Automatic Liver Tissue Section Image Characterization

Dr I Suneetha[1], Chandra Vani D[2], Neelakantesh P[3], Chandana M[4], Narendra D[5]
[1]*Professor, Dept of ECE, Annamacharya Institute of Technology & Science, Tirupati*
[2,3,4,5]*UG Students, Dept. of ECE, Annamacharya Institute of Technology & Sciences, Tirupati*

Abstract **- Analyzing liver cross-section images is a manual and laborious task. Liver transplantation is currently the only cure for patients with end-stage liver disease. Alternating treatment options are critically needed for these patients. Understanding the molecular mechanisms by which the proliferation of an important type of liver cell called hepatocytes is regulated. The methods to automatically count hepatocytes cells, count nuclei and classify liver vessel types, using input images of cell boundary and cell nuclei based on a dataset of 21 images. Compared to a trained researcher, the methods are able to count cells, segment overlapping nuclei, and classify vessel types, including portal vein, central vein, and bile duct with reasonable precision and accuracy and detecting other cell types. All the methods are categorized into three main classes including gray level-based method, structure based-method and texture based-method.**

Index Terms **- Microscopy image processing, cell counting, nuclei segmentation, vessel classification, SVM, canny edge detection, Hough transform.**

INTRODUCTION

Liver transplantation is currently the only cure for patients with end-stage liver disease, yet currently over 14,000 patients are on the liver transplant waitlist awaiting a donor [1]. Alternative treatment options are critically needed for these patients.

Physicians and scientists contemplating new treatment methods face the same major obstacle: understanding the molecular mechanisms by which the proliferation of an important type of liver cell called hepatocytes is regulated. The Nusse Lab of the Stanford Institute of Stem Cell Biology and Regenerative Medicine studies this class of specialized cells and recently published work in identifying a population of hepatocyte stem cells in the liver [2]. A major obstacle in scaling the research is the laborious task of manual cell counting from tissue section images.

This project leverages image processing and machine learning techniques to tackle the unique challenges presented by tissue section images. First, compared to tissue culture images, cells in a tissue section image are highly heterogenous with complex structures and ambiguous boundaries, making cell counting a more difficult task. Second, tissue section images may contain different types of cells, including hepatocytes, bile duct epithelial cells, endothelial cells, and immune cells. In this project, we focus exclusively on the hepatocytes cells which is the focus of the research. Third, in contrast to most other cells in our body, hepatocytes can have more than one nucleus. A final challenge is that in addition to cells, the liver section images also show important features of a liver, including central veins, portal veins, and bile ducts.

There are three distinct goals in the project: accurately count the number of hepatocyte cells irrespective to their stain, accurately segmented clustered nuclei, and accurately classify liver features, including central vein, portal vein and bile duct. This work focuses on developing algorithms to accomplish each goal and analyzing their performances. The algorithm is applied to a dataset provided by the Nusse Lab, consisting of 21 liver tissue section images, which are manually labeled by a trained researcher.

This paper is structured as the following: first we will present the implementation pipelines for cell boundary seg- mentation, cell nuclei segmentation, and liver vessel type classification. Next, the performance of each task using various techniques is analyzed and compared.

BACKGROUND

Microscopy has been the tool used by biomedical scientists to advance our fundamental understanding of life at the cellular level. Advances in modern microscopic techniques, such as the use of fluorescent markers and microscopic systems with sub-light diffraction limits, lowered the cost of acquiring images

and enhanced the level of details of these images [3]. On the flip side, the abundance, heterogeneity, dimensionality, and complexity make manual image analysis a laborious and expensive task. Consequently, automating the analysis of these images can help accelerate the progress in cell biology and other related biomedical fields.

There are numerous benefits associated with potentially automating these tasks, based on a survey of existing literature. Automating these tasks using image processing and machine learning may result in significant time savings as well as reduce measurement variability due to operator- dependent and parameter-sensitive conditions [3]. Additionally, automation has the potential of quantifying numerous cell topology characteristics that are difficult or expensive to do so manually [4].

There have been many papers published on developing automated methods for segmenting and counting cells in microscopy images. Common approaches range from classical image processing techniques, such intensity thresholding and morphological operations, to modern approaches such as tensor voting schemes, neural networks, and Markov random fields [3]. Due to the astounding variety of the different types of cell analysis problems, there are just as many cell segmentation methods that combines existing techniques and apply it to a new problem, and our project is no exception.

METHOD

The high-level image processing pipeline is illustrated in figure 1. Each step in the pipeline is described in detail in the following sections.

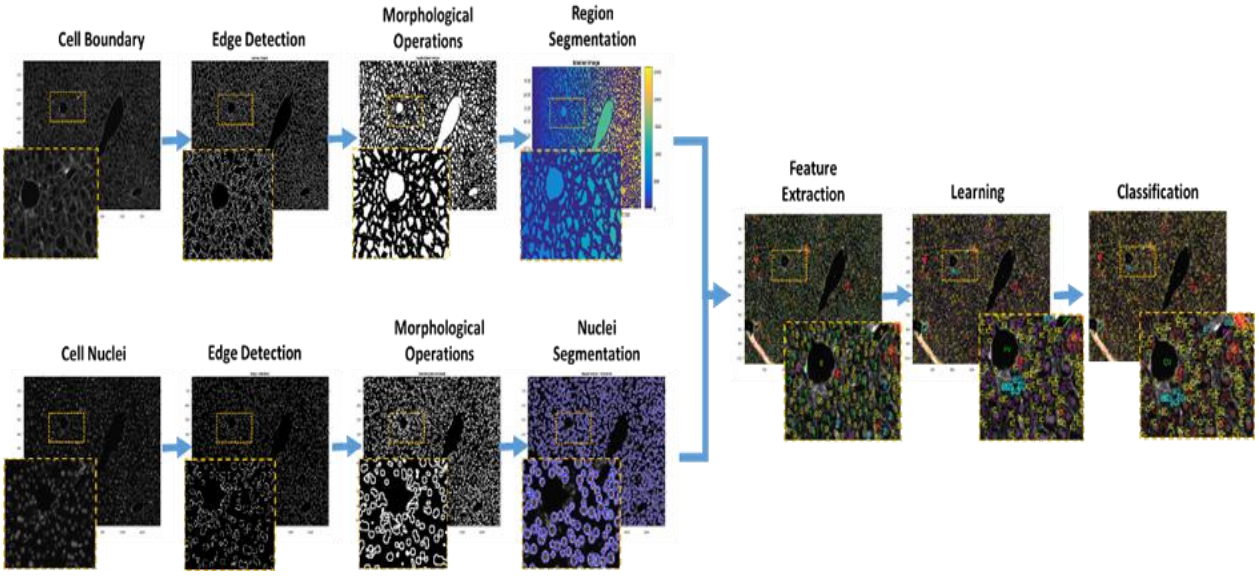

Fig.1. Block diagram of proposed method cell boundary, cell nuclei segmentation, and liver feature classification.

A. Cell Boundary Segmentation

a) Edge Detection: The cell boundaries of a gray-scale cell boundary input image are detected using the Canny edge gradient. The Canny edge detector is chosen over other edge detection methods such as Prewitt and Sobel edge detection because the Canny edge detector calculates gradients based on derivative of a Gaussian filter and therefore is less susceptible to errors due to ambiguous cell boundaries. An image resulting from Canny Edge Detection is shown in fig 2.

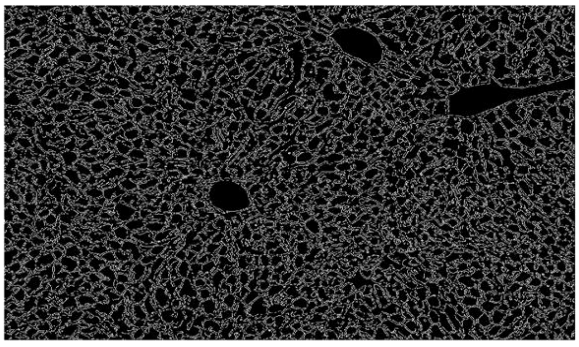

Fig.2. A cell boundary output image using Canny edge detection

b) Morphological Operations: A few morphological operations are performed to the resulting edge image to enhance the salient cell boundaries so that distinct cell regions can be identified. First, the image is dilated using a line structure of 2-pixel length in both the horizontal and vertical direction. Second, the image is negated to so that only the cell boundary information is present. Next, small regions are flood filled to reduce noise are flood filled to reduce noise. Finally, the cell regions are smoothed using opening and closing operations with a 1-pixel diamond structuring element. The process of morphological operations shown in fig.3.

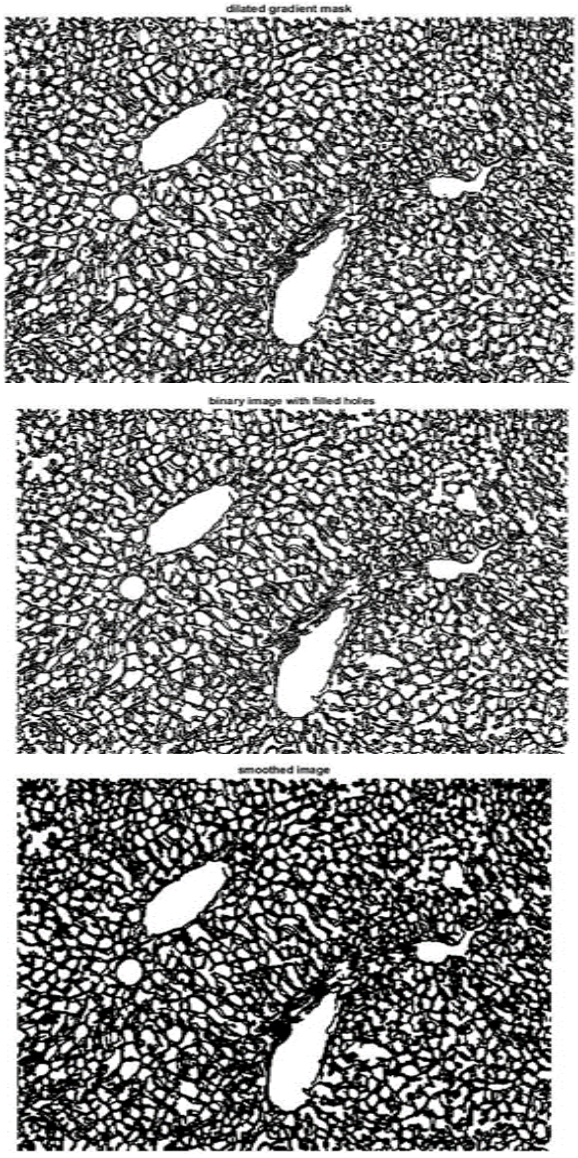

Fig.3. Morphological operations of a) dilated gradient mask, b) binary image with small holes filled in, c) smoothed cell boundaries using opening and closing.

c) Region Segmentation: Following edge detection and morphological operation that enhance the boundary of cell structures, regions are segmented by getting 8-connected neighborhoods. A sample image of segmented regions, including potential cells, non-cellular regions, and veins, are distinctively colored in a gradient in figure 4.

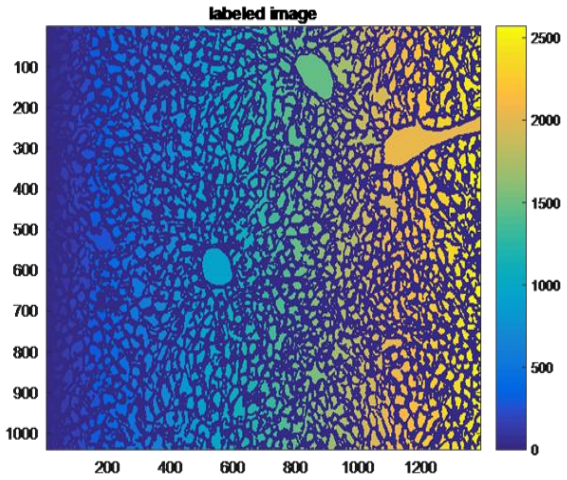

Fig.4. Image of segmented regions, including potential cells, non-cellular, and veins, each distinctively colored in a gradient.

B. Cell Nuclei Segmentation

a) Pre-processing: An input gray-scale image of the nuclei is preprocessed using global image thresholding via Otsu's method that is adjusted by a small tuning parameter to reject ambiguous regions in the image.

b) Edge Detection: Following the binarization of the image, Canny edge detection is used to detect nuclei edges in the image. Canny edge detection also works well for nuclei segmentation due to its robustness to noise.

c) Morphological Operation: Following edge detection, the nuclei boundaries are enhanced by dilating the image using a line structuring element of 3-pixel length in the horizontal and vertical directions.

d) Hough Transform: To detect the circular shaped hepatocyte cell nuclei, circular Hough Transforms emerged as the best option among other techniques attempted which include SIFT key point, SURF key point, and morphological operation. The performances of each operation are compared in the results section. MATLAB implements the circular Hough Transform via the function imfindcircles.

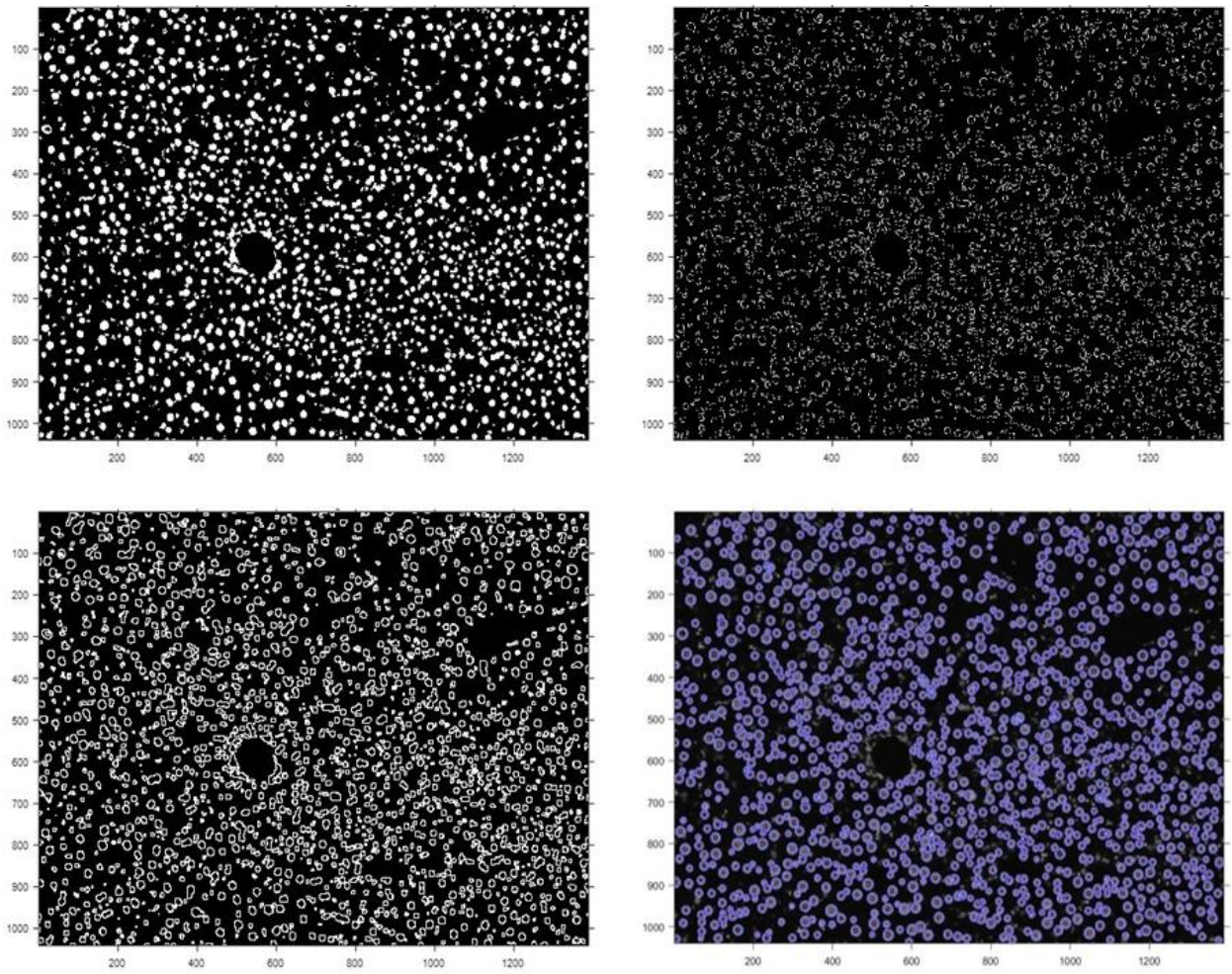

Fig.5. Images illustrating the steps of nuclei segmentation a) binarized image using global thresholding, b) edge image using canny edge detection, c) enhanced edge image using dilation, d) detected nuclei using Hough transform.

C. Liver Vessel Type Classification

a) Feature extraction: From the results of nuclei and cell boundary segmentation, relevant features are extracted from each image in the dataset, including geometric proper- ties, boundary properties, cell properties, neighboring region properties. While all the geometric properties and a couple of boundary properties are obtainable via the MATLAB function Region Props, the interesting features including SIFT key point density, distance from bile duct, and average cell size vs. relative distance are intended to quantify the domain knowledge used by an expert researcher classifying vessels present in a tissue as portal vein or central vein.

b) Learning and Classification: The features extracted above are used to create feature vectors for bile duct, central vein and portal vein training examples in 12 images. These training examples are passed into the following classifiers and the models are tested using the remaining 9 images.

RESULTS

A. Cell Counting

The performance of the cell counting method based on the cell boundary and nuclei segmentation pipelines is analyzed in this section. An overlay image of cell boundary and nuclei shows stained hepatocyte cells, which may have more than 1 nucleus. The positions of regions segmented via cell boundary are registered with the positions of segmented nuclei, and a region is counted as a cell if it contains at least one nucleus. image showing regions labeled with the number of nuclei present in that region is shown in figure 6.

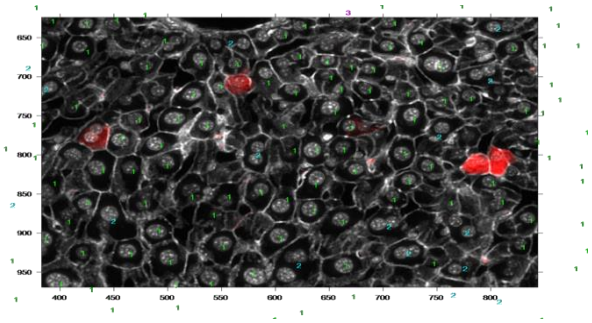

Fig.6. image showing cells labeled with the number of nuclei present.

The automatic cell counting method tends to overcount the number of 1 nuclei cell as the 1 nuclei cell have lower accuracy. On the other hand, the automatic method is fairly accurate counting the number of 2 nuclei cells. Overall, the automatic algorithm achieves a respectable accuracy of 78% for all cells.

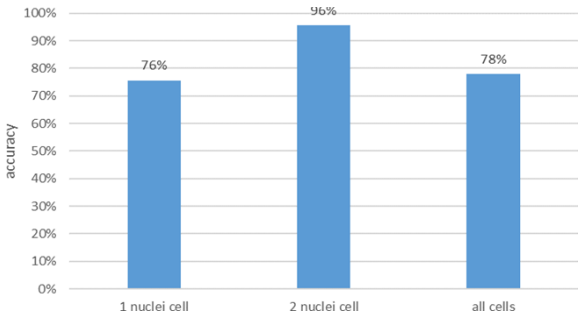

Fig. 7. A graph showing the comparison between the number of cells manually counted by a trained researcher and these automatically counted.

B. Nuclei Segmentation

The performance of the nuclei segmentation method is analyzed in this section. The output of the nuclei segmentation pipeline is the count of hepatocyte cell nuclei that have a characteristic large, circular shape. A sample output of each of the four methods is shown in the figure 8, where a detected nucleus is traced by a circle outline.

The Hough transform not only detects nuclei of different size but also distinguishes between two overlapping nuclei. SIFT key point detection fails to positively identify many nuclei perhaps due to the uniform nature of nuclei that lack corners for SIFT to positively identify as key points. SURF key point detection does a much better job than SIFT key point detection, however SURF still creates many false positives, i.e., nuclei identified where none exists. Finally, nuclei counting based on connected regions post morphological erosion and dilation is not robust enough because it cannot be finetuned to work across a range of images, and it detects many false positives.

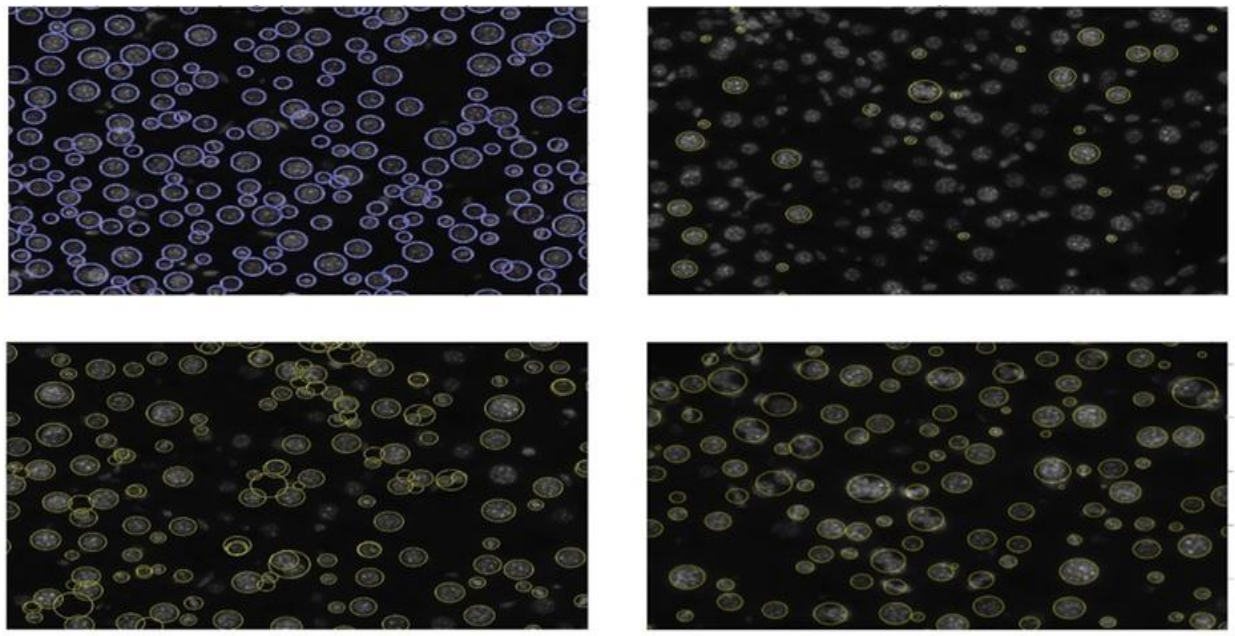

Fig .8. Nuclei segmentation output using the following techniques: a) Hough Transform, b) SIFT key point, c) SURF key point, and d) morphological connected neighbors.

C. Vessel Type Classification

The classification results shown in figures.9 validates the choice of features used, especially those extracted based on domain knowledge The SVM model unsurprisingly performs the best out of the three learning models tested. The features extracted capture the characteristics of bile ducts as all three methods gave great prediction for them. However, the lack of very clear distinctive characteristics between central veins and portal vein image regions leads to some misclassifications.

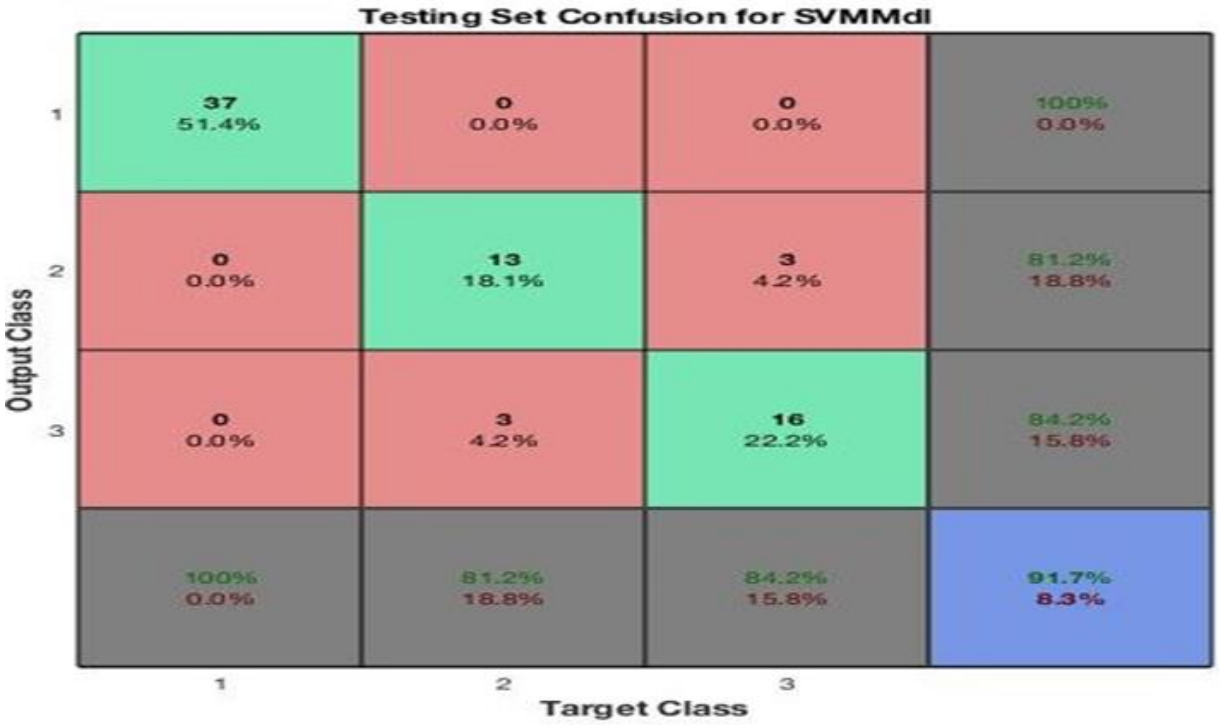

Fig. 9. Multiclass SVM model test confusion matrix.

CONCLUSION

The goals of this projects are threefold: 1) accurately detect cell counts, 2) correctly classify liver vessel types, and 3) accurately segment overlapping nuclei. Given the limited timeframe and resources available, this project successfully developed methods that automatically count cells with about 80% accuracy and segment overlapping nuclei with about 60% precision and 100% recall. Also, portal vein, central vein, and bile duct in our very limited dataset were classified with 91% precision.

There were many challenges encountered that were re-solved to varying degrees in the project. Ambiguous cell boundary is the biggest factor that contributes to over- counting 1-nuclei cells, because the 2-nuclei cells with ambiguous boundaries may detected as two 1-nuclei cells. Another major challenge is extracting the relevant features for classification of vessel types based on domain knowledge. For example, average cell size relative to distance from vessel boundary is computed based on the observation that cells around portal veins tend to be more compressed than these around central veins. A final major challenge is the expensive nature of generating labeled datasets, limiting the scope of supervised learning algorithms that are feasible.

FUTURE SCOPE

Future works could focus on improving the classification accuracy of the vessel types, by building up a dataset with thousands of labeled images. This could present the opportunity of using autoencoders to extract classification features for the models used here and perhaps a convolutional neural network as a classification model. Based on a robust feature classification algorithm, certain expensive research tasks can also be automated, such as measuring the number of cells closer to certain types of vessels. Finally, the project scope can be expanded to include

detecting other cell types, such as endothethial cells and bile duct epithelial cells.

REFERENCES

[1] Data from United Network for Organ Sharing (UNOS), http://www.unos.org, 2014.

[2] W. B, Z. L, F. M, L. CY, and N. R, "Self-renewing diploid axin2+ cells fuel homeostatic renewal of the liver.," Nature, 2015.

[3] E. Meijering, "Cell segmentation: 50 years down the road," IEEE Signal Processing Magazine, vol. 29, no. 5, September 2012, pp. 140145.

[4] I. Grishagin, "Automatic cell counting with imagej," Analytical Bio- chemistry, 2015.

[5] C. G. Loukas, G. D. Wilson, B. Vojnovic, and A. Linney, "An image analysis-based approach for automated counting of cancer cell nuclei in tissue sections," Cytometry Part A, vol. 55A, no. 1, pp. 30–42, 2003.

[6] P. Wuttisarnwattana, M. Gargesha, W. van't Hof, K. R. Cooke, and L. Wilson, "Automatic stem cell detection in microscopic whole mouse cryo-imaging," IEEE Transactions on Medical Imaging, vol. 35, pp. 819–829, March 2016.

[7] G. M. Faustino, M. Gattass, S. Rehen, and C. J. P. de Lucena, "Automatic embryonic stem cells detection and counting method in fluorescence microscopy images," IEEE, 2009.

[8] J. Cheng and J. C. R. ast, "Segmentation of clustered nuclei with shape markers and marking function," IEEE Transactions on Biomedical Engineering, vol. 56, pp. 741–748, March 2009.

[9] F. Maruhashi, S. Murakami, and K. Baba, "Automated monitoring of cell concentration and viability using an image analysis system," Cytotechnology, vol. 15, no. 1, pp. 281–289, 1994.