IMAGE PROCESSING AND ANALYSIS OF HISTOPATHOLOGICAL IMAGES RELATING TO HIRSCHSPRUNG'S DISEASE

Jinu Kurian[a], Marco T.K. Law[b], Victoria Madge[c],
Adrian D.C. Chan[a], Dina El Demellawy[d], Eranga Ukwatta[a]
Carleton University[a], University of British Columbia[b], McGill University[c], CHEO[d]

INTRODUCTION

Hirschsprung's disease (HD) is a congenital gut motility disorder, affecting approximately one in 5000 live births [1]. HD is suspected after an infant has not passed meconium within 48 hours. The inability for segments of the distal colon to relax is attributed to the stunted growth of neurons along the large intestine [1]. These neurons, termed *ganglia*, innervate the *muscularis propria* (the muscle region of the large bowel) to permit colonic movement. Treatment for HD involves a surgical pull-through procedure to connect healthy colon directly to the rectum, removing estimated portions of the aganglionic colon (as determined by suction and seromuscular biopsies). To ensure the remaining colon is sufficiently ganglionated, a pathologist examines sections of the excised colon searching for adequate *ganglion* cell presence in the proximal margin. However, due to a lack of evaluation standards regarding *ganglion* identification and count, this current practice is subjective; conclusions rely on physicians' experience and time constraints. Unfortunately, a subpopulation of patients exists for whom symptoms persist regardless of a satisfactory histopathological evaluation [2]. Reoperation may not be possible due to its severe morbid nature, including increased risks for enterocolitis, permanent colostomy, and other associated complications [3].

The proportion of patients with pathology indications and persistent HD symptoms post-surgery [2], incites the query into whether subjectivity in the proximal margin assessment, regarding the adequate presence of ganglia, may contribute to improper diagnoses. An objective evaluation method for *ganglion* identification may improve the outcome of surgical treatment, reduce the need for re-operation, and aid in the further understanding of HD. For instance, structural data relating to the colon, such as *ganglion* count and spread, may be characteristic of and correlate to the severity of HD. Computer-aided approaches can avoid manual count, which is time-consuming, tedious and subject to high observer variability.

This paper describes the development of image analysis methods for extracting structural data from digital histopathological images of excised colon specimens, from which objective measures for universal scoring rubrics for HD may be derived. It expands upon the preliminary results obtained by Law M. *et al.* in [4].

METHODS

As described in [4], the image-processing pipeline to identify ganglion cells is to: 1) isolate the *muscularis propria*, 2) identify *myenteric plexus* regions (pockets between muscle layers populated by *ganglia*), and 3) quantify *ganglia* within each *myenteric plexus*.

Study Subjects and Image Acquisition

This study included ten cases from patients diagnosed with HD at CHEO (formerly known as the Children's Hospital of Eastern Ontario). The Research Ethics Board at CHEO approved the study protocol.

Using immunohistochemistry, the Leica BOND-MAX from Leica Microsystems and the Leica BOND-III from Leica Biosystems stained histopathological paraffin blocks of ganglionated large intestine segments for the HD cases. Calretinin stain was preferred due to its increased ganglion discrimination with digital histopathology [5]. Aperio ScanScope CS (Model SC) by Aperio Technologies scanned the prepared slides at an effective magnification of

200x for image resolution of 0.5μm/pixel. These images were compressed in SVS-file format. Figure 1 is a thumbnail of one such image, with a window zoomed in on myenteric plexi.
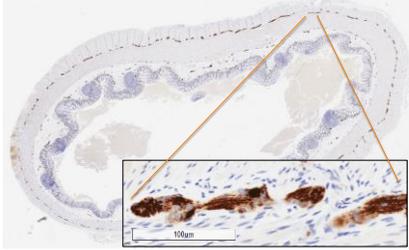

Figure 1: *Myenteric plexi at 20x magnification within the muscularis propria of a colon section*

Segmentation of the *Muscularis Propria*

To avoid high computational costs, the *muscularis propria* is segmented on a down sampled image (ratio of 16:1). This also serves as a primary low-pass filter, allowing for the effective use of k-means clustering in the RGB colour space. Further preprocessing steps remove noise and artefacts, but also allow large structures to appear as an averaged colour. After applying a median filter and an averaging filter, the blurred image is separated into four clusters (k = 4). In general, these clusters correspond to the white background; the violet mucosa; the near-white submucosa; and the *muscularis propria,* which tends towards lavender tones. Experimentally, the cluster associated with the *muscularis propria* is often the second largest cluster in terms of pixel area. Finally, morphological operations help fill in holes from the binary image of the cluster.

Segmentation of the *Myenteric Plexus* Regions

Using a 3x3 median filter on the 16:1 low resolution colour image, high-frequency noise components are reduced while preserving edges. Each colour channel is contrast-adjusted, such that one percent of the pixels is saturated at the lowest and highest intensities of the original image. To isolate the brown-stained plexus regions, Digital Image Subtraction Blue Enhancement (DISBE) is applied to the image, as described in [6]. DISBE highlights regions where the blue channel has a lower intensity relative to the rest of the image (such as brown pixels). Masking the resulting image with the binary image of the *muscularis propria*, help differentiate *myenteric plexus* regions from *submucosal plexus* regions. Otsu thresholding is applied to preserve regions of high intensity and morphological operators are used to refine the mask by removing speckles.

Evaluation of the Developed Algorithm

We used manual segmentations as surrogate of ground truth to evaluate the algorithm-generated results. Manual annotations, outlining the *muscularis propria* and *myenteric plexus* regions, were drawn by co-author Kurian (biomedical engineering master's student) and reviewed by co-author El Demellawy (expert pathologist at CHEO) utilizing Leica Biosystems' Aperio ImageScope. The binary images generated from these annotations serve as the standard to validate segmentation algorithms.

Using images resized to match the down-sampled automated segmentation results, the ten algorithm-generated segmentations results were evaluated based on similarity to the manual segmentation of the *muscularis propria*, as well as inclusion of the *myenteric plexus* regions. The second evaluation metric is crucial to the performance for the remainder of the image-processing pipeline.

Region-based similarity for the *muscularis propria* was determined using the Sørensen–Dice index. It evaluates the proportion of overlapping regions between samples, as follows:

$$\frac{2 \, |S_{manual} \, \cap \, S_{automated}|}{|S_{manual}| + |S_{automated}|} \quad (1)$$

where S_{manual} and $S_{automated}$ are the number of white pixels in the binary masks of manual and algorithm-based segmentations, respectively. A Dice coefficient of 100% indicates a perfect match between two images, whereas a value of zero will represent a lack of overlap.

Automated muscularis propria segmentation was also evaluated using plexus inclusion (PI)

$$\frac{\text{Area of Plexus Regions}_{in \, automated \, muscularis}}{\text{Area of Plexus Regions}_{in \, manual \, muscularis}} \times 100\% \quad (2)$$

which examines the percentage of manually segmented *myenteric plexus* regions included in the automated segmentation of the *muscularis propria*. This metric accounts for discrepancies

introduced from converting manual contours into region-based masks, such as edge distortions.

The automated identification of *myenteric plexus* was evaluated against the manually segmented plexus regions. Determining the intersection between objects in the two binary masks allowed us to calculate precision and recall, respectively, as

$$precision = \frac{TP}{TP+FP} \tag{3}$$

$$recall = \frac{TP}{TP+FN} \tag{4}$$

where TP refers to true positives (overlap of objects), FP to false positives (objects misidentified as *myenteric plexi*), and FN to false negatives (unidentified plexus regions).

RESULTS

The results as summarized in Table 1, reveal an average dice coefficient of 71.22% ± 20.44% for the *muscularis propria*. Likewise, the average total area of manually segmented plexus region included in the algorithm-generated *muscularis propria* is 72.51% ± 35.64%. Figure 2 depicts the results of the automated segmentation for patient case three.

Continuing with the identification *of myenteric plexus* regions, the average precision and recall values are 54.48% ± 29.97% and 37.71% ± 25.78%, respectively. Masking the plexus region binary mask with the manual segmentation of the *muscularis propria* improved these metrics by approximately 20%.

DISCUSSION

Eight out of the ten algorithm-generated *muscularis* segmentations have displayed satisfactory spatial overlap with the manual segmentations, exhibiting Dice coefficients greater than 70%. The morphological operations prevent detailed edges of the *muscularis propria* in the binary mask, as seen in Figure 2(c). However, for this clinical application, the inclusion of *myenteric plexus* regions is more pertinent for the segmentation of the *muscularis propria*. Excluding two cases, the Dice coefficient positively correlated with the inclusion of *myenteric plexus* regions.
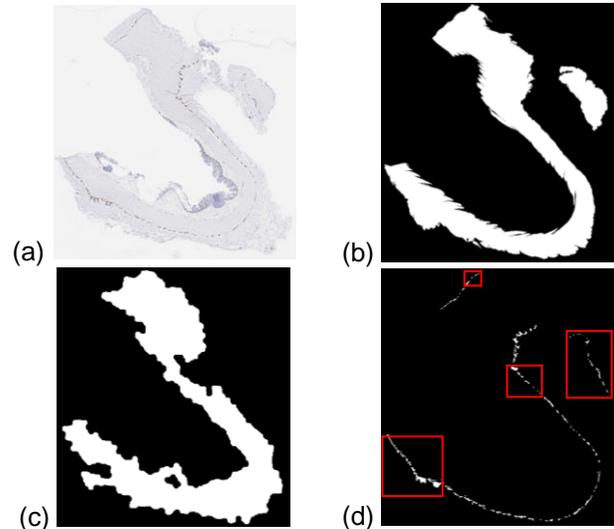

Figure 2: *Case 3 (a) Calretinin-stained thumbnail (b) Manual muscularis propria segmentation (b) Algorithm-generated muscularis propria mask (c) Manual segmentation of myenteric plexus regions. The red boxes highlight the excluded plexus regions after the muscularis propria is segmented using the automated method*

Further investigation into the two patient cases with poor Dice similarity revealed areas for algorithm improvement. In patient case six, the morphological operations failed to provide robust results. Interestingly, in patient case seven, the k-means operation divided the *muscularis propria* between two different clusters. Additional blurring or pre-processing may prevent reoccurrence of similar results.

The unrefined edges of the algorithm-generated *muscularis* segmentation and the frequency of plexus regions found at the border of the *muscularis propria* and the adjoining *submucosal* region, contribute to higher probabilities of identifying plexus regions outside of the muscle layer (i.e. false positives). However, as shown in Table 1, the segmentation of the *myenteric plexus* regions is relatively successful when considered independently of the automated *muscularis propria* segmentation. Nevertheless, brown staining errors in the image promote the number of false positives detected, thereby lowering precision. In addition, the non-discriminatory nature of morphological operators removes many of the smaller plexus regions in

the final mask, adding to the false negatives. This negatively affects the recall scores and inadvertently lowers precision.

CONCLUSION

In some cases, the segmentation results of the *muscularis propria* limit the performance of identifying *myenteric plexus* regions. Although preliminary results are promising, improvements can certainly be made. Considerations towards convolutional neural networks (CNN), as well as other texture-based techniques, may result in enhanced *muscularis propria* segmentations. This, in turn, would inevitably improve *myenteric plexus* region identification, and thereby, ganglion classification as well. Given the encouraging results presented in this paper, further exploration into the improvement of plexus region identification will be conducted.

ACKNOWLEDGEMENTS

REFERENCES

[1] J. Kessmann, "Hirschsprung disease: diagnosis and management," *American Family Physician, vol. 74, no. 8,* pp. 1319-1322, 2006.

[2] T. A. Lawal, K. Chatoorgoon, M. H. Collins, A. Coe, A. Pe˜na, and M. A. Levitt, "Redo pull-through in Hirschprung's disease for obstructive symptoms due to residual aganglionosis and transition zone bowel," *Journal of Pediatric Surgery,* vol. 46, no. 2, p. 342–347, 2011.

[3] A. Arshad, C. Powell, and M. P. Tighe, "Hirschsprung's disease," *BMJ,* vol. 345, no. 7883, pp. 47-49, 2012.

[4] M. T. K. Law, A. D. C. Chan and D. El Demellawy, "Color image processing in Hirschsprung's disease diagnosis," in *2016 IEEE EMBS International Student Conference (ISC)*, Ottawa, 2016.

[5] D. H. Gonzalo and T. Plesec, "Hirschsprung disease and use of calretinin in inadequate rectal suction biopsies," *Archives of pathology & laboratory medicine,* vol. 137, no. 8, p. 1099–102, 2013.

[6] S V. Bernardo, S. Lourenço, R. Cruz, L. Monteiro-Leal, L. Silva, D. Camisasca, M. Farina and U. Lins, "Reproducibility of immunostaining quantification and description of a new digital image processing procedure for quantitative evaluation of immunohistochemistry in pathology," *Microscopy and Microanalysis,* vol. 15, no. 04, pp. 353-365, 2009.

Table 1: *Evaluation of Muscularis Propria and Myenteric Plexus Segmentation Results*

Images	Muscularis Propria Evaluation		Myenteric Plexus Identification			
			Precision		Recall	
	Dice Values	Inclusion of Plexus Regions	Using Algorithm-Generated Muscularis	Using Manually Segmented Muscularis	Using Algorithm-Generated Muscularis	Using Manually Segmented Muscularis
Case 1	87.20%	99.90%	45.00%	51.33%	69.90%	74.76%
Case 2	83.32%	98.94%	93.33%	97.56%	13.33%	38.10%
Case 3	77.03%	57.20%	26.36%	43.62%	53.13%	82.81%
Case 4	77.85%	100.00%	27.37%	70.27%	50.00%	50.00%
Case 5	89.99%	99.76%	70.13%	96.43%	73.97%	73.97%
Case 6	27.93%	13.75%	75.00%	100.00%	16.07%	68.75%
Case 7	40.07%	19.28%	91.18%	91.09%	32.29%	47.92%
Case 8	78.42%	96.26%	77.17%	87.10%	54.20%	61.83%
Case 9	74.00%	43.66%	18.75%	20.54%	8.74%	81.55%
Case 10	76.34%	96.37%	20.51%	47.37%	5.48%	6.16%
Mean	71.22%	72.51%	54.48%	70.53%	37.71%	58.59%
Std. Dev.	20.44%	35.64%	29.97%	28.08%	25.78%	23.76%