Identification of Helicobacter pylori Organisms by Digital Image Analysis

Mark A. Valasek, MD, PhD[1], Eli Brazowski[2], Yaron Anavi[2], Ifat Rotbein[2], Lotan Chorev[2] and Eliron Amir[2]
University of California San Diego, Department of Pathology, San Diego, CA, USA[1]; Nucleai, Tel Aviv, Israel[2]

Background

Clinical laboratory testing, including serum antibody and stool antigen testing, is commonly utilized for the evaluation and diagnosis of Helicobacter pylori (HP) infection. In addition, detection of HP often occurs via visual identification of organisms during histopathological evaluation of endoscopic stomach biopsies or surgical resection specimens. Although routine H&E stains are often sufficient, various special stains (Warthin-Starry, Giemsa, etc.) or immunohistochemistry (usually polyclonocal anti-HP antibodies) can be utilized to enhance sensitivity and specificity of detection. In part because such ancillary methods can be costly and a potential source of test overutilization, their precise role is controversial with particular interest placed on whether all gastric biopsies should be stained up-front or only selected cases should be stained (1, 2, 3). Here, we explore the potential of digital image analysis (DIA) utilizing convolutional neural networks (CNN) to aid in the identification of HP in H&E stained pathological tissue specimens.

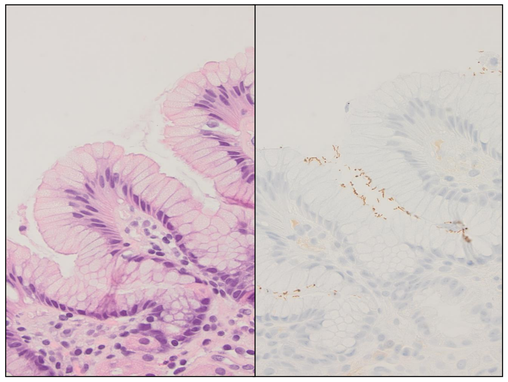

Figure 1 – Visual Helicobacter Detection on Routine H&E (A) and Immunohistochemical Stains (B) (600X)

Design

We trained an artificial intelligence DIA algorithm based on convolutional neural networks (Nucleai-HP-Algorithm) using 35 H&E slides to identify Helicobacter pylori (HP) organisms. A test image set of 80 cases (including both gastric biopsies and surgical resection specimens) was utilized and included the following diagnoses: 1) normal gastric mucosa or chronic gastritis/reactive gastropathy, HP negative (n=24), 2) chronic gastritis, HP positive (n=48), and 3) HP difficult to determine (n=8). Two methods were used to evaluate HP: 1) whole slide (tissue biopsy) HP detection, or 2) within slide (within tissue biopsy) patch HP detection (1 patch = 22.5 um x 22.5 um) to assist in locating the HP within a slide.

HP Classification Algorithm - High Level Design

The algorithm is based on 3 main sub-components:

- Tissue Segmentation Algorithm
- Tile Segmentation Algorithm
- Slide level decision and detections

This was performed for all test slides and the results were compared the diagnosis provided by pathologists.

HP Network Architecture and Classification Algorithm

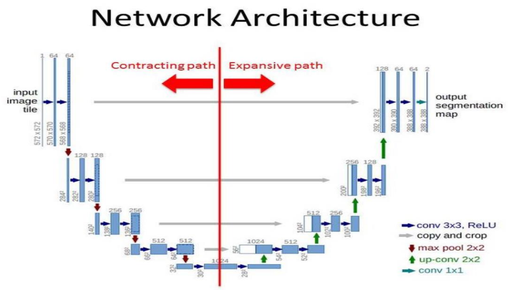

Figure 2 - The architecture of the segmentation convolutional neural network used by Nucleai-HP-Algorithm

Tissue Segmentation Algorithm

This algorithm is based on a segmentation CNN. To train this algorithm, hundreds of slides were annotated with tissue and background, and a segmentation CNN was trained to segment the tissue vs. the background and any noise (including markers, scratches on the surface etc.). Therefore, this algorithm is utilized to identify tissue from a slide image.

Patch Segmentation Algorithm

Tens of slides (positive and negative) were scanned, and annotated in detail. The annotation included marking individual points at the location of HPs within the positive slides. In total, tens of thousands of individual HPs were annotated. Patches (as defined above) were extracted from the slides. Positive patches were taken in areas where we had HP annotations, and negative patches were taken from negative slides. Each HP annotated point was expanded to a small circle around it (to mark the HP area). A segmentation Convolutional Neural Network (CNN) was trained, where the input was the extracted patch, and the ground truth for the segmentation training was a binary mask where positive was the expanded circles derived from annotated HPs, and negative was the remainder tissue area outside of the expanded circle regions. Negative patches had a segmentation ground truth that was all negative. After a first CNN was trained there were several iterations of training that included of adding annotations and hard negative training with negative patches that had HP detections. This was performed until satisfying results were achieved in the patch level segmentation accuracy.

Patch Level Detection and Slide Level Decision

Patch detection and slide level decision are based on a thresholds from the segmentation outputs of the different patches. After the patch segmentation there is a threshold per patch, where the patch is marked as HP positive or HP negative. Nearby, or adjacent patches are united to be one detected area in the slide. Finally, a per-slide threshold is determined, wherein the slide is marked as either HP Negative or HP Positive and given a confidence measurement which is based on the amount of tissue with HP detections.

Inference & Testing Phase

During inference phase, the first step is the tissue segmentation algorithm, then patches are extracted from the tissue, each patch is then segmented using the patch segmentation algorithm. The patch segmentation output is used to mark the HP positive areas in the slide. The patch segmentation output also serves as an input to the final step of the slide level decision. The slide level decision algorithm outputs the result and confidence for the slide.

Results

Figure 3 – Tile level segmentation masks as output by the algorithm

Overall concordance	91.7%
Sensitivity	97.9%
Specificity	79.9%

Table 1 – Result summary

The overall concordance between DIA and the reference pathologist was 91.7%. Slide level sensitivity is 97.9% and specificity is 79.9%. When the patch level is set at 80% sensitivity this leads to an average of 7 false positive areas (less than 0.01% of tissue surface area) for a gastric biopsy slide (where an area is defined as a group of patches with a distance up to 11.25um). Discrepancies include surgical specimens where there is a large amount of tissue or non-gastric tissue types.

Conclusions

The findings demonstrate the capability of DIA using CNN to successfully identify HP organisms in gastric tissue samples. Therefore, DIA may have a potential role to reduce time spent by pathologists on HP detection, reduce the need for use of ancillary testing (including IHC stains) or to be utilized in quality assurance/quality control measures.

Future improvements of the algorithm can include training with larger data sets with more variety, which will improve performance, and/or incorporation of inflammatory milieu evaluation for risk stratification of images. It is important to note that similar algorithms using this type of approach can be used to identify signet ring cells in gastric samples.

References

1. Batts KP, Ketover S, Kakar S, Krasinskas AM, Mitchell KA, Wilcox R, Westerhoff M, Rank J, Gibson J, Mattia AR, Cummings OW, Davison JM, Naini BV, Dry SM, Yantiss RK; Rodger C Haggitt Gastrointestinal Pathology Society. Appropriate use of special stains for identifying Helicobacter pylori: Recommendations from the Rodger C. Haggitt Gastrointestinal Pathology Society. Am J Surg Pathol. 2013 Nov;37(11):e12-22.

2. Lash RH, Genta RM. Routine Anti-Helicobacter Immunohistochemical Staining is Significantly Superior to Reflex Staining Protocols for the Detection of Helicobacter in Gastric Biopsy Specimens. Helicobacter. 2016 Dec;21(6):581-585.

3. Pittman ME, Kharazjian A, Wood LD, Montgomery EA, Voltaggio L. Prospective identification of Helicobacter pylori in routine gastric biopsies without reflex ancillary stains is cost-efficient for our health care system. Hum Pathol. 2016 Dec;58:90-96.