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 polyclonal 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.

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.

**Tissue Segmentation Algorithm**

This algorithm is based on a segmentation CNN. To train this algorithm, hundreds of slides were annotated with tissue vs. background and a segmentation CNN was trained to segment the tissue from 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 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**

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.

**Network Architecture**

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.

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

**Results**

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.

**Table 1 - Result summary**

<table>
<thead>
<tr>
<th>Overall concordance</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>91.7%</td>
<td>97.9%</td>
<td>79.9%</td>
</tr>
</tbody>
</table>

**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 H&E 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**

