Results from CAMELYON17: Using HALO AI™ to Identify and Stage Breast Cancer Metastases

INTRODUCTION

HALO AI™ integrates a revolutionary deep learning convolutional neural network (CNN), ideally suited for image classification, within the intuitive HALO® image analysis platform. Using training annotations provided by a pathologist, HALO AI can 'learn' to decipher the complex patterns of histologically-stained tissues in a similar way to the pathologist eye. Unlike predecessor machine learning algorithms, the neural network behind HALO AI can manage the high staining and morphological variability seen in diagnostic pathology and continually improves with interaction and training.

The metastasis of breast cancer to lymph nodes has significant clinical implications for breast cancer patients with regard to therapy selection and prognosis. The identification and staging of metastases currently involves thorough microscopic evaluation of lymph node sections by a pathologist, a laborious and time consuming task. This application note describes how HALO AI was trained to automate identification and staging of metastases in breast cancer and how the final classifier performed in the **CAMELYON17 Challenge**. The results are highly encouraging for the transfer of this technology from research application to routine clinical use.

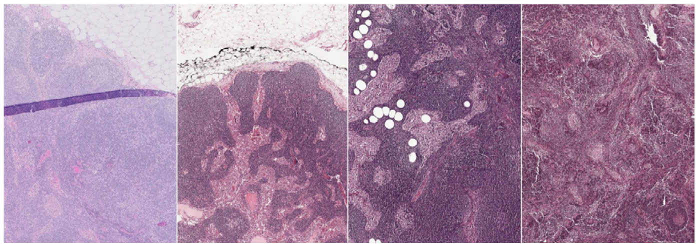

Figure 1. Four representative lymph node images from the training data set with considerable variability in stain intensity/color, tissue quality, and overall morphology.

BACKGROUND

CAMELYON17, organized by Diagnostic Image Analysis Group (DIAG) and Department of Pathology of the Radboud University Medical Center, challenged academic and commercial teams to develop a fully-automated method for identifying and staging breast cancer metastases (Geessink, O. *et al.*, 2017). Challengers were given five digital slides of lymph nodes from each of 100 breast cancer patients (500 slides total). The first objective was to identify a method to find all metastatic tumor cells in each slide and, based on the size of the tumor area, assign each lymph node to one of four categories: negative, isolated tumor cells (ITCs), micro-metastases, or macro-metastases. The final objective was to combine the results for the five slides and assign a pN tumor stage at the patient level as follows Pathologic lymph node classification (pN-stage) according to the CAMELYON17 guidelines: pN0: No micro-metastases or macro-metastases or ITCs found, pN0(i+): Only ITCs found, pN1mi: Micro-metastases found, but no macro-metastases found, pN1: Metastases found in 1–3 lymph nodes, of which at least one is a macro-metastasis, and pN2: Metastases found in 4–9 lymph nodes, of which at least one is a macro-metastasis.

METHODS

GENERATION OF TRAINING DATA. The "ground truth" training data used in for the CAMELYON17 study consisted of approximately 160 images with tumor annotations and 370 negative images from a total of 770, including 270 from the previous CAMELYON16 challenge (Bejnordi, B.E. *et al.*, 2016). The data set contained considerable variability in stain intensity, tissue/section quality, and morphology, as demonstrated in **(Figure 1)**, similar to what a pathologist may need to evaluate.

Prior to training, a pre-processing step was used to detect tissue on the slide and glass. These negative annotations and the ground-truth annotations provided by the organizers were imported into the HALO AI platform and assigned to one of two classes, tumor and non-tumor.

TRAINING. HALO AI was trained for $1.3x10^6$ iterations using the CAMELYON16 data set, and then fine-tuned in a second round on the combined CAMELYON16 and CAMELYON17 data sets for a further $7x10^5$ iterations. All training was performed at a resolution of 0.25 µm/px. HALO AI automatically balances and augments the training data according to published methods (Liu, Y. et *al.*, 2017), so there is no need for users to intervene at this step in the training process.

TUMOR STAGING. The 500 slides in the test set were classified using the trained classifier and probability maps were generated such that different shades of red indicate probability of being 'tumor' as shown in (**Figure 2**). Regions classified as 'tumor' were converted to annotations at three different probability thresholds (50%, 95% and 99%). The total annotated area, area of the largest single annotation, and the longest major axis length of a single annotation were measured across all three probability thresholds and an average value was calculated for each slide. A slide was categorized as 1) macro-metastases if an annotation had a major axis greater than 2mm^2 or if the area of annotation was greater than 1mm^2, 2) micro-metastases if an annotation had a major axis greater than 0.2mm, 3) ITC if it was neither micro nor macro and had an area greater than 3000 μm^2 or 4) negative if none of the other categories apply. For each patient, the results were combined for the five sections to assign a final pN stage as described in the background section.

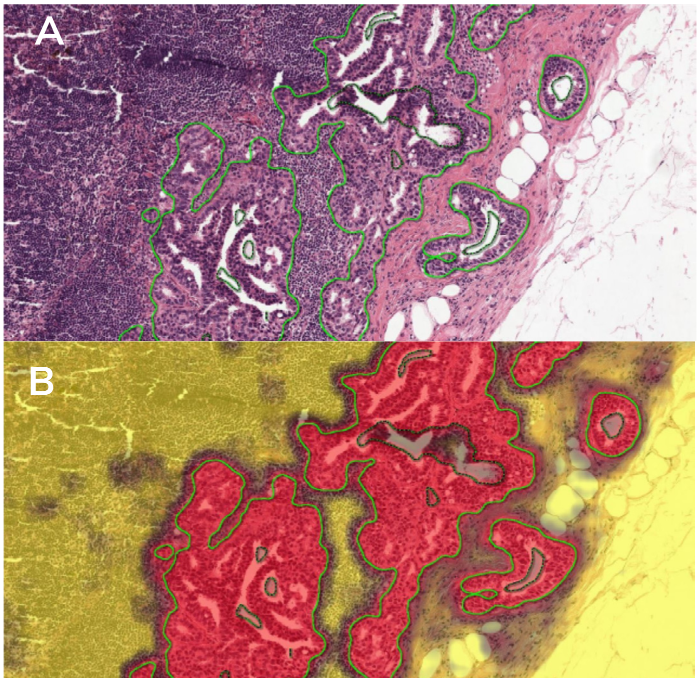

Figure 2. Example result of classifier training. Annotated training regions (A), and correct identification of tumor positive regions by classifier (B). Red = tumor (more intense red = higher probability of tumor) and yellow = non-tumor.

RESULTS

All applications to the CAMELYON17 challenge were evaluated and scored using five class quadratic weighted kappa, where the classes were the tumor level (pN) stages. HALO AI achieved a kappa score of 0.8554, which was **the highest ranking of all commercial entries presented within the CAMELYON17 challenge deadline** (April 1, 2017), demonstrating the power of HALO AI in this application. Full results and methods can be reviewed at the CAMELYON17 website **https://camelyon17.grand-challenge.org/.**

In total, HALO AI successfully classified 500 slides from the equivalent of 100 breast cancer patients, generating very few false negatives. Some false positives were evident, likely caused by overrepresentation of lymphocytes in the negative training class. Additional training or a two-step classification is expected to reduce the likelihood of false-positives and is currently being tested.

CONCLUSIONS

Deep learning has considerable potential in digital pathology, as demonstrated by the performance of HALO AI in this challenge. This approach took the established CNN-style neural network and enhanced it by integrating with the user-friendly HALO platform to create HALO AI, a complete workflow solution for histopathology.

REFERENCES

1. Geessink, O. B´andi, P., Litjens, G., and van der Laak, J. Camelyon17: Grand challenge on cancer metastasis detection and classification in lymph nodes (2017). *Accessible from: https://camelyon17.grand-challenge.org/*
2. Bejnordi, B.E. and van der Laak, J. Camelyon16: Grand challenge on cancer metastasis detection in lymph nodes (2016). Accessible from: *https://camelyon16.grand-challenge.org/*
3. Yun Liu, Krishna Gadepalli, Mohammad Norouzi, George E Dahl, Timo Kohlberger, Aleksey Boyko, Subhashini Venugopalan, Aleksei Timofeev, Philip Q Nelson, Greg S Corrado, *et al.*, "Detecting cancer metastases on gigapixel pathology images," arXiv preprint arXiv:1703.02442, 2017.

For more information on HALO AI or any other Indica Labs' products or services, visit us at http://indicalab.com or send an email to info@indicalab.com.