EDITORIAL   OPEN

Check for updates

Efficient cellular annotation of histopathology slides with real-time AI augmentation

In recent years, a steady swell of biological image data has driven rapid progress in healthcare applications of computer vision and machine learning. To make sense of this data, scientists often rely on detailed annotations from domain experts for training artificial intelligence (AI) algorithms. The time-consuming and costly process of collecting annotations presents a sizable bottleneck for AI research and development. HALS (Human-Augmenting Labeling System) is a collaborative human-AI labeling workflow that uses an iterative "review-and-revise" model to improve the efficiency of this critical process in computational pathology.

Recent advances in computer vision and the rapid digitization of histology slides have escalated interest in artificial intelligence (AI) applications for pathology. Modern algorithms are not only capable of predicting regions of cancer,[1] but also driver mutations,[2] metastatic origins,[3] and patient prognosis.[4] Many algorithms rely only on the slide image and metadata, but others also require annotations of cells, tissues, and other entities within the slide. This additional information helps connect histology images to their underlying biology and may improve the interpretability and generalizability of resulting algorithms.[5] Still, acquiring sufficient labels to train data-hungry models is a significant challenge, potentially requiring millions of annotations across thousands of slides.[5] To address this bottleneck, van der Wal et al.[6] introduce HALS (Human-Augmenting Labeling System) to enable efficient, high-quality annotations at scale.

HALS consists of three software components—a segmentation model, a classifier model, and an active learner model—that jointly support a human annotator. First, the segmentation model identifies the location of all cells in a small region. Second, the annotator, usually a board-certified pathologist, begins labeling cells. These labels are used to train a classifier model that identifies and labels other cells in the region that the annotator may correct. As the classifier learns from these corrections, its predictions progressively improve and require fewer corrections. When the first region is sufficiently labeled (20–30 annotations per class), the active learner model identifies the next most informative region for annotation, and the process repeats.

The advantages are two-fold. First, HALS's classifier model reduces workload by replacing the laborious task of primary annotations with the simpler task of correction annotations. Second, HALS's active learner model improves data efficiency by directing the annotator to more informative slide regions. To demonstrate these advantages, the authors measure workload (proportion of AI suggestions requiring correction) and data efficiency (validation accuracy vs. sample size) for seven pathologists using HALS to label tumor cells, tumor-infiltrating lymphocytes, eosinophils, and Ki-67[+] cells. HALS resulted in significant workload reductions for all tasks, ranging from 96% for tumor cells (the easiest task) to 83% for eosinophils. In other words, a pathologist armed with HALS is expected to produce approximately 10 times more annotations with the same number of clicks. Algorithms trained using HALS were also 1.4–6.4% more data efficient than those trained without AI augmentation, suggesting higher quality labels. Improvements were robust across the four cell types and two stains, lending credence to the system's generalizability. In addition to these objective improvements, AI augmentation also allows pathologists to monitor classifier improvement and identify failure modes throughout the labeling process. This may enable closer collaboration with engineers and contribute insights for understanding model deficiencies and allocating labeling resources.

Similar "review-and-revise" systems are widely commercialized and have been utilized in areas such as object recognition,[7] autonomous driving research,[8] and dermatologic classification.[9] HALS offers several unique contributions for pathology applications. The active learning component[10] is particularly well-suited for digital pathology, where gigapixel-resolution images preclude exhaustive annotation. In addition, HALS synthesizes several pathology-specific segmentation,[11] classification,[12] and interface[13] systems into a single modular framework that simplifies application and adaptation. Future work should investigate scalability to large slide repositories, region selection heuristics that prioritize cell classes in greatest need of improvement, and application areas beyond microscopy.

In past decades, the decreasing cost of sequencing has powered explosive progress in genetics.[14] Overcoming the high cost of expert annotations may unlock similar possibilities for the continuing deluge of biological image data. In their article, Van der Wal and colleagues present compelling results that labeling augmentation with HALS may decrease annotation cost up to 10-fold, indicating substantial progress toward this goal. For academic investigators, AI augmented labels could enable more rapid and granular study of biological hypotheses. For industry teams, newly freed resources and higher quality data may accelerate development of AI-based software-as-a-medical-device (SaMD). As systems like HALS continue to improve, we anticipate that AI augmentation may soon become an ubiquitous aid for embedding human knowledge into data.

James A. Diao [ID]*1,4✉, Richard J. Chen*[2,4] *and Joseph C. Kvedar* [ID]*3*
1Program in Health Sciences and Technology, Harvard Medical School, Boston, MA, USA. 2Department of Biomedical Informatics, Harvard Medical School, Boston, MA, USA. 3Department of Dermatology, Massachusetts General Hospital, Boston, MA, USA. 4These authors contributed equally: James A. Diao, Richard J. Chen.
✉email: james_diao@hms.harvard.edu

REFERENCES

1. Wang, D., Khosla, A., Gargeya, R., Irshad, H. & Beck, A. H. Deep learning for identifying metastatic breast cancer. Preprint at https://arxiv.org/abs/1606.05718 (2016).

2. Fu, Y. et al. Pan-cancer computational histopathology reveals mutations, tumor composition and prognosis. *Nature Cancer* **1**, 800–810 (2020).

3. Lu, M. Y. et al. AI-based pathology predicts origins for cancers of unknown primary. *Nature* **594**, 106–110 (2021).

4. Skrede, O.-J. et al. Deep learning for prediction of colorectal cancer outcome: a discovery and validation study. *Lancet* **395**, 350–360 (2020).

5. Diao, J. A. et al. Human-interpretable image features derived from densely mapped cancer pathology slides predict diverse molecular phenotypes. *Nat. Commun.* **12**, 1613 (2021).

6. van der Wal, D. et al. Biological data annotation via a human-augmenting AI-based labeling system. *NPJ Digit. Med.* **4**, 145 (2021).

7. Yu, F. et al. LSUN: construction of a large-scale image dataset using deep learning with humans in the loop. Preprint at https://arxiv.org/abs/1506.03365 (2015).

8. Segal, S. et al. Just label what you need: fine-grained active selection for perception and prediction through partially labeled scenes. Preprint at https://arxiv.org/abs/2104.03956 (2021).

9. Groh, M. et al. Evaluating deep neural networks trained on clinical images in dermatology with the fitzpatrick 17k dataset. Preprint at https://arxiv.org/abs/2104.09957 (2021).

10. Sener, O. & Savarese, S. Active learning for convolutional neural networks: a core-set approach. Preprint at https://arxiv.org/abs/1708.00489 (2017).

11. Graham, S. et al. Hover-net: simultaneous segmentation and classification of nuclei in multi-tissue histology images. *Med. Image Anal.* **58**, 101563 (2019).

12. Gamper, J., Alemi Koohbanani, N., Benet, K., Khuram, A. & Rajpoot, N. PanNuke: an open pan-cancer histology dataset for nuclei instance segmentation and classification. in *Digital Pathology* 11–19 (Springer International Publishing, 2019). https://doi.org/10.1007/978-3-030-23937-4_2.

13. Aubreville, M., Bertram, C., Klopfleisch, R. & Maier, A. SlideRunner - a tool for massive cell annotations in whole slide images. Preprint at https://arxiv.org/abs/1802.02347 (2018).

14. Koboldt, D. C., Steinberg, K. M., Larson, D. E., Wilson, R. K. & Mardis, E. R. The next-generation sequencing revolution and its impact on genomics. *Cell* **155**, 27–38 (2013).

AUTHOR CONTRIBUTIONS

COMPETING INTERESTS