An Augmented Reality Microscope for Real-time Automated Detection of Cancer

Authors

Po-Hsuan (Cameron) Chen[+]
Krishna Gadepalli[+]
Robert MacDonald[+]
Yun Liu,
Kunal Nagpal,
Timo Kohlberger,
Greg S. Corrado
Jason D. Hipp
Martin C. Stumpe*

Google AI Healthcare, Google Inc, Mountain View, CA, USA

[+] contributed equally
* corresponding author: mstumpe@google.com

Abstract

Purpose:
Microscopic review of tissues on glass slides is the gold standard for cancer diagnostics and provides fundamental insights into cancer biology. Using digital images, artificial intelligence (AI) algorithms have been shown to improve diagnostic accuracy and provide quantitative metrics that can not be provided on glass tissue slides alone; however, they require investment in a digital pathology workflow. In this paper, we describe the integration of modern AI into a standard microscope that is routinely used to review cancer specimens and perform basic and translational research.

Experimental Design:
We modified a standard optical microscope to have built-in real-time AI capabilities. This required integrating an augmented reality display into the optics, training a deep neural network for cancer detection, and running the AI algorithms connected to the microscope with low latency. As sample applications, we trained AI algorithms to detect metastatic breast cancer in sentinel lymph nodes and prostate cancer in prostatectomy specimens.

Results:
The integration of projection optics into an existing microscope provided a parallax-free, high resolution display of information derived by machine learning overlaying the sample. The AI algorithm was running in real-time using the live image feed from the camera, and was generating high accuracy cancer detection results. Pathologists testing the device reported a seamless experience that provided immediately useful information.

Conclusion:
The augmented microscope has applications in research and clinical usage, and could provide a complementary technology to digital pathology based on whole slide imagery. Its main promise lies in expanding availability of modern AI algorithms.

Statement of Translational Relevance

Microscopic analysis of tissue provides critical information for cancer diagnosis and staging in the clinic. It also enables fundamental insights into tumor biology via reviewing animal model tissue, human explant tissues and tissue microarrays for biomarker expression. Recent advances in artificial intelligence (AI) have shown that accurate automated tissue analysis is possible. However, significant financial and workflow barriers exist for pathologists and cancer researchers to utilize these algorithms, such as the purchase of a whole-slide imaging scanner, IT infrastructure, and operating personnel.

In this paper, we propose an alternative solution: an augmented microscope that integrates real-time AI capabilities, allowing these AI algorithms to be applied with seamless integration into existing workflows and comparatively low equipment cost. By leveraging the latest AI algorithms, our microscope modification has the potential to improve speed, quality and accuracy of quantitative tissue assessments for clinical practice and translational tissue experiments.

Introduction

Microscopes are the gold standard for cancer diagnostics in anatomic pathology. These have proven to be effective for more than 100 years across a wide variety of specimens and stains. These analog microscopes are critical for patient's American Joint Committee on Cancer (AJCC) cancer staging that enable treatment decisions and prognostic determination to be made. For the pathologic staging, a pathologist examines tissue sections to determine the type and grade of cancer, lymph node spread and metastasis to other organs. Recent editions of the AJCC and clinical studies require pathologists to provide more quantitative assessments of their tissue or specimen reviews. For example, breast cancer staging requires pathologists to quantify the amount of tumor in sentinel lymph nodes by measuring maximal dimensions of tumor foci and counting tumor cells. Additional studies have shown the importance of counting tumor infiltrating lymphocytes[1][2]. Although these tasks are time consuming, tedious, and demonstrate considerable inter- and intra-rater variability, they are helpful to clinicians in determining immunotherapy treatment plans. Lastly, quantitative measurements of human epidermal growth factor receptor 2 (HER2), Estrogen Receptor (ER), and Progesterone Receptor (PR) immunohistochemistry (IHC) biomarker expression is critical for breast cancer therapy selection. These advances in cancer classification, treatment and associated companion diagnostics have driven increasing complexity and reporting requirements for the pathologist, in some cases stretching human capability for efficient assessment.

As a potential solution, recent advances in artificial intelligence, specifically deep learning[3], have demonstrated that computer systems are able to perform human intelligence level tasks in medical image analysis[4]–[12]. Research has also shown the potential to improve diagnostic accuracy, quantitation and efficiency by applying deep learning algorithms to whole-slide pathology images for cancer classification and detection[6], [7], [9], [11], [13]. However, because of the large upfront cost for the scanner, the infrastructure required for transmitting and storing images, and the need for trained personnel to operate the equipment, digital pathology adoption has been slow in both clinical and translational research settings.

To bridge this gap and bring the advantages of deep learning into the analog world, we present an augmented optical light microscope that enables real-time usage and presentation of the results of AI algorithms. As in a traditional analog microscope, the user views the sample through the eyepiece, but the predictions of an AI algorithm, which sees the same view of the sample as the user, can be projected back into the optical path in real-time. This projection is visually superimposed on the original (analog) image of the specimen. Because the user still sees the actual sample, as opposed to a digitized version, this process is termed augmented reality, and we term our microscope the Augmented Reality Microscope (ARM). Although we focus on the application of this novel technology to cancer diagnosis in this paper, the ARM itself is application-agnostic and equally lends itself to other microscopy applications in the life sciences or beyond.

Materials and Methods

To provide a useful assistant technology, the ARM needs to demonstrate high accuracy, speed and clarity. First, it must provide digital information, e.g. classification results from deep learning algorithms, with high quality diagnostic performance. Second, the system must keep pace with the rapid, semi-continuous movement of the slide under review. Finally, the digital content must have sufficient resolution and brightness to effectively display the information to the human operator, and the digital features must be precisely aligned with the analog image of the specimen in the field of view. The alignment of digital and analog images should be insensitive to small changes in the operator viewing position (parallax-free). A system capable of delivering on these core capabilities can seamlessly fit into a traditional pathologist workstation, eliminating important adoption barriers.

We designed and developed the ARM with three major components integrated end-to-end to achieve real-time assisted diagnostic functionality. The system consists of an augmented brightfield microscope; a computer to collect the microscope camera images, run the deep learning algorithm and drive a display module attached to the microscope;and a set of trained deep learning algorithms. Figure 1 shows a schematic sketch of the ARM and a photo of the actual implementation.

Optico-Mechanical Design

The augmented microscope, shown in Figure 1, includes a commercial off-the-shelf brightfield microscope (Nikon Eclipse Ni-U) with custom modules designed to capture high resolution images and superimpose the digital display content, free of parallax. The standard upright microscope illuminates the specimen (S) from behind and captures the image rays with a conventional objective. These rays propagate upward, in a collimated state, towards the oculars. A teaching module (Nikon Y-IDP) with a beam splitter (BS1) was inserted into the optical pathway in the collimated light space. This module was modified to accept a microscope camera (C1), so that the specimen image relayed from BS1 was in focus at the camera sensor when the specimen was also in focus to the observer. A second customized teaching module (Nikon T-THM) was inserted between the oculars and the first teaching module. The beam splitter in this module (BS2) was rotated 90 degrees to combine light from the specimen image (SI) with that from the projected image (PI). Collimating optics between the microdisplay and BS2 were chosen to match the display size with the ocular size (22 mm). The relay optic was adjusted such that the microdisplay was positioned in a virtual focal plane of the specimen. In this way the viewer observes SI and PI in the same plane, which minimizes relative motion when the observer moves, a phenomenon known as parallax. Note that BS1 needs to be before BS2 in the optical pathway from objective to ocular, so that camera C1 sees a view of the specimen without the projection PI. The observer looking through the eyepiece sees PI superimposed onto SI.

Component design and selection was driven by final performance requirements. Camera and display devices were chosen for effective cell and gland level feature representation. The camera (Adimec S25A80) included a 5120x5120 pixel color sensor with high sensitivity and global shutter capable of running up to 80 frames/sec. Camera images were captured by an industrial frame-grabber board (Cyton CXP-4) with PCI-E interface to the workstation. The microdisplay (eMagin SXGA096, 1292 x 1036 pixels) was mounted on the side of the microscope and imaged with an achromatic condenser (Nikon MBL71305) at a location tuned to minimize parallax and ensure that the specimen and display image are simultaneously in focus. The microdisplay includes an HDMI interface for receiving images from the workstation. Due to the limited brightness of this display, BS2 was chosen to transmit 90% of the light from the display and 10% from the sample, which resulted in adequate contrast between PI and SI when operating the microscope light source near half of its maximum intensity.

Software and Hardware System

The application driving the entire system runs on a standard off-the-shelf PC with a BitFlow frame grabber connected to a camera (C1) for live image capture and an NVidia Titan Xp GPU for running deep learning algorithm. The end-to-end process from frame grabbing to the final display is shown in Figure 2. To improve responsiveness, the system is implemented as a highly optimized, pipelined, multi-threaded process, resulting in low overall latency. The software is written in C++ and TensorFlow[14].

The primary pipeline consists of a set of threads that continuously grab an image frame from the camera, debayer it (i.e. convert the raw sensor output into a color image), prepare the data, run algorithm, process the results, and finally display the output. Other preprocessing steps such as flat-field correction and white balancing can be done in this thread as well for cameras which can not do them directly on-chip. To reduce the overall latency, these steps run in parallel for a sequence of successive frames, i.e. the display of frame ‘N’, generation of heatmap of frame ‘N+1’, and running algorithm on frame ‘N+2’ all happen in parallel.

In addition to this primary pipeline, the system also runs a background control thread. One purpose of this thread is to determine whether the camera image is sufficiently in focus to yield accurate deep learning algorithm results. The system uses an out-of-focus detection algorithm to assess focus quality. A second purpose of this thread is to determine the currently used microscope objective (see Supplemental Material for details), so that the deep learning algorithm tuned for the respective magnification is used. Additionally, settings for white balance and exposure time on camera C1 can be set to optimal profiles for the respective lens.

Data Acquisition

In this paper, we demonstrate two clinical applications: lymph node metastasis detection and prostate cancer detection. For both applications, the system assists via a diagnosis outline highlighting the tumor region.

For lymph node metastasis detection, we obtained training data from the Cancer Metastases in Lymph Nodes (Camelyon) 2016 challenge data set[6]. This data set comprises 399 whole slide images from H&E slides digitized by one of two whole-slide scanners: a 3DHISTECH Pannoramic 250 Flash II (pixel size 0.243×0.243μm) or a Hamamatsu XR C12000 (pixel size 0.226×0.226μm). The dataset contains pixel-level ground truth diagnoses of tumor vs benign. For prostate Gleason grading, we collected 285 prostatectomy slides and digitized them with an Aperio AT2 scanner (pixel size 0.218×0.218μm). These whole slide images were annotated by board-certified pathologists outlining regions as: benign, Gleason grade 3, Gleason grade 4 or Gleason grade 5.

To evaluate the deep learning algorithm performance, we obtained 8 slides for each tissue type as testing data. To establish the ground truth of tumor localization, we also obtained consecutive serial sections of the testing data. These slides are IHC stained, pancytokeratin (AE1/AE3) for lymph nodes and PIN4 for prostatectomies, and reviewed by at least two board-certified pathologists.

Deep Learning Algorithm Development

We developed a new convolutional neural network based on Inception v3[15], which has been used in various medical imaging applications, including pathology[9], dermatology[8], [10], and ophthalmology[4]. The variant of Inception v3, which we call Inception-FCN, combines inception v3 with the concept of fully convolutional networks[16]. It is optimized for real-time prediction and generalization between different image sizes.

The deep learning image analysis workflow includes two phases: algorithm development and algorithm application, illustrated in Figure 2. For algorithm development, we train the neural networks on digitized pathology slides with patches of size 911×911 pixels at each of the magnifications (4X, 10X, 20X, 40X for lymph node and 4X, 10X, 20X for prostate), and the corresponding labels indicating the diagnosis, e.g. tumor/benign or Gleason grades. By changing the weights of the neural network to reduce the difference between the predicted results and the corresponding labels, the neural network learned to recognize patterns and distinguish between different diagnoses. During neural network training, we also scaled the whole slide images to match the pixel resolution from the scanner images (~0.21μm/pixel) to the pixel resolution from the microscope camera (~0.11μm/pixel). Although pixel scaling can be done in real-time on the ARM, we chose to do it in the neural network training step to shift the required computation into the training phase. For algorithm application, the neural network is

provided with images of size 5120×5120 pixels captured from the microscope camera. The output from the network is a heatmap depicting the likelihood of cancer at each pixel location. The heatmap can be displayed directly using a colormap, or thresholded to get an outline that is then displayed as an overlay on the sample. Microscope users found that outlines are usually favored over heatmaps in these applications since the latter occludes the underlying sample. However, heatmaps could be useful in some applications, and the system is capable of displaying either visualization mode.

Algorithm Evaluation

We evaluate the algorithm performance of tumor detection within the field of view with the following metrics: receiver operating characteristic (ROC) curves (the true positive rate against the false positive rate) , area under the ROC curve (AUC), accuracy, precision, and recall (TP: true positive; FP: false positive; FN: false negative):

$$\text{Precision} = TP/(TP+FP),$$
$$\text{Recall/True Positive Rate} = TP/(TP+FN),$$
$$\text{False Positive Rate} = FP/(FP+TN).$$

Results

Using modern deep learning algorithms with off-the-shelf graphics card for accelerated computation, the system achieved a total latency of about 100 ms (10 frames per second) which is fast enough for most workflows. The projection of the deep learning predictions into the optics was high enough contrast to be clearly visible on top of the tissue sample using common background illumination levels and was parallax-free. Operating the ARM was seamless for first-time users (pathologists that were not part of the study) who tested it, with almost no learning curve.

We demonstrate two sample clinical applications: the detection of metastatic breast cancer in sentinel lymph nodes and the identification of prostate cancer in prostatectomy slides. Figure 3 shows the view through the eyepieces for these two applications with different magnifications. The green outlines highlight the tumor regions in both cases. The outlines were a natural visualization for most users, although some preferred heatmaps or semi-transparent filled outlines which we offered as an alternative visualization options.

Deep Learning Algorithm Performance

Figure 4 shows several sample fields of view on lymph node specimens for metastasis detection. To validate the algorithm performance, we also show the digitally scanned H&E images and IHC images on the side for comparison. Despite the differences in image quality and color distribution between microscope images and digitally scanned images, our algorithms correctly identify the tumor region in the fields of view.

To quantify the algorithm's performance, we collected 6 fields of view enriched for tumor and 6 fields of view that contained only benign tissue from all lymph node and prostate testing slides at 10X (i.e. a total of 96 fields of view for each application). The fields of view were collected with coverage across a diverse range of cell types. For the lymph node evaluation, the images contained lymphoid cells, connective tissue, blood and lymphatic vessels, fat, and metastatic breast cancer. For the prostate evaluation, the images contained benign prostatic glands, blood and lymphatic vessels, fat, inflammation, and prostate cancer. Note that the image quality from the microscope and whole slide scanner differ significantly with respect to the level of focus, exposure time, etc. In this study, we collect images with proper focus and exposure level for evaluation as in the regular microscopic tissue review workflow.

Next, we evaluated the ARM across these 96 fields-of-view. For the lymph node metastasis detection algorithm, the ARM achieved a field-of-view-level AUC of 98% on the test dataset. For the prostate cancer detection algorithm, the ARM achieved a field-of-view-level AUC of 96% on the test dataset. We also report accuracy, precision and recall with three different operating thresholds: a high accuracy operating threshold, a high precision operating threshold for a diagnostic use case, and a high recall threshold for a screening use case. The ROC curve and the metrics are shown in Figure 5.

Discussion

In this paper, we present a novel type of optical light microscope with built-in real-time deep learning capabilities for assisted diagnosis, and the ability to project the deep learning results into the optics. As a proof of concept, we developed deep learning algorithms for two applications: detection of breast cancer lymph node metastasis and detection of prostate cancer. We demonstrate the feasibility of the end-to-end integrated system to be used in a seamless workflow and with a high cancer detection accuracy, comparable to algorithms run on WSI. This bridges the gap between AI algorithms and analog microscope workflow, enabling the integration of existing AI research on pathology to current clinical workflow.

To enable AI algorithms on analog microscope requires three novel technologies working in unison. First, state-of-the-art convolutional neural networks for high accuracy detection and

classification. High accuracy neural networks have been shown possible in the literature on digitally scanned images alone or images from microscope alone. A contribution of this work is the demonstration of successful generalization of deep learning algorithms from digitally scanned images to microscope images. The feasibility of overcoming the differences in image modality allows us to use digitally scanned images for deep learning algorithm development for the ARM. Second, the ability to run these algorithms in real-time to provide an interactive user experience. This is achieved by a tightly integrated software, hardware and AI algorithm system with a real-time performance for live algorithm predictions. Third, a parallax-free head-up display in the microscope to project high-resolution heatmaps, contours, or textual information onto the sample. This is made possible by the novel configuration of off-the-shelf optic components that are generally available.

Even though this study details two clinical diagnostic applications, the platform is application agnostic and can be used for any kind of image analysis task for which a deep learning algorithm has been trained. We envision the usage of this device for many other clinical and research applications, such as mitotic rate estimation, IHC quantification and positive margin detection in frozen sections. Figure 6 shows several images captured from the microscope with examples of suitable information overlays including biomarker/IHC quantification (HER2, ER, PR, Ki67, P53, CD8), infectious disease detection, mitotic figure counting. Other applications include translational experiments that explore the quantification of immune cells, immunotherapy biomarkers such as PDL1, and cell cycle markers such as P53 and Ki67 using IHC. Additional applications include the analysis of fluorescent (FISH) images, live cell culture images, and even applications outside of life sciences such as in semiconductor quality control.

Aside from the core novelty of having AI integrated into the microscope, the ARM can also be used for projection of any auxiliary information, such as case data or annotations, into the optics. Furthermore, still images or video streams, with or without the augmented display, can be captured for inclusion in reports, or directly transmitted to external displays for usage in tumor boards, educational settings, or live telepathology. Furthermore, using off-the-shelf components, a robotic stage could be added to control horizontal slide motion and focus adjustment, effectively enabling automated exhaustive ‘scanning’ of the slide for applications such as metastasis or tumor detection.

In this study, we have shown high cancer detection accuracy in a controlled setup as a proof of concept. A larger study will be required to evaluate the algorithm performance with higher statistical significance, and to assess failure modes. Furthermore, even though we have shown the successful generalization of deep learning algorithms from digitally scanned images to microscope images, an alternative approach will be to train deep learning algorithms directly on microscope images. It is expected that this would lead to additional improvement to the algorithm performance. Most importantly, a reader study will be needed to verify the feasibility on actual clinical usage.

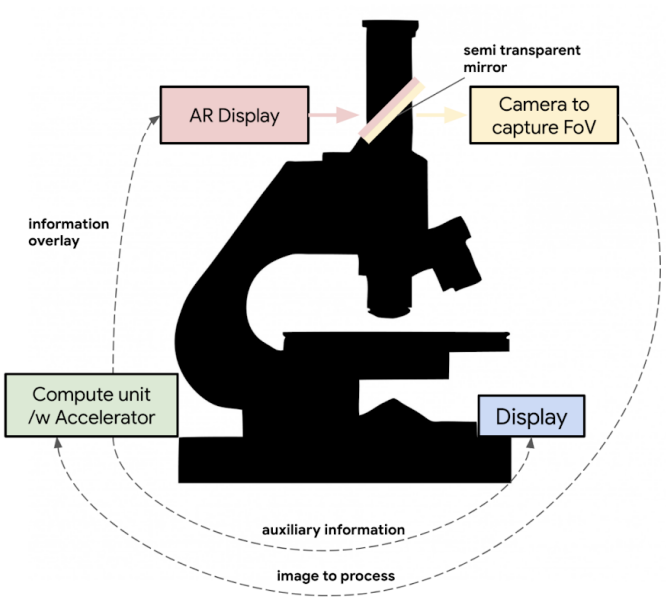
Comparison to Digital Pathology Approach, Limitations and Potential Impact

The ARM preserves many benefits of conventional pathology workflows while while integrating digital tools of potential utility to clinical practice. It enables pathologists to use their familiar microscope rather than change the way they view and diagnose cases. Also, it retains the superior image quality of analog microscopes, gives the user the ability to adjust the focus, pan and zoom quickly around the image, and provides an interactive experience for reading the sample. Furthermore, it can be adapted to microscopes in cancer research laboratories performing various types of tissue analyses. However, it also has limitations compared to a fully digital solution. In particular, it does not provide a fast way to exhaustively scan the entire slide and create a whole slide image. A human in the loop is required, and the system does not provide a scalable way to digitize slides at scale. Consequently, it is not well suited for digital archiving or for running AI to detect tiny features at higher power if the user has not already located the region of interest.

Conclusion

Given these advantages and limitations, we believe the main potential of the ARM is in expanding access to AI to users who are not likely to adopt a digital workflow in the near future. These include areas where the integration or economic barrier for a digital workflow is too high, e.g. small labs and clinics, and in particular, developing countries where a microscope can easily be deployed whereas a standard digital pathology solution is impractical. In the developing world, applications such as detection of infectious diseases (e.g. tuberculosis, malaria) stand out in particular. Even in places that will adopt a standard digital pathology workflow for the bulk of their work, there are domains where digital scanners still face major challenges (e.g. cytology or fluorescent imaging), or where low latency is required (e.g. frozen sections). In those places, the ARM could be used in combination with the digital workflow.

Figures



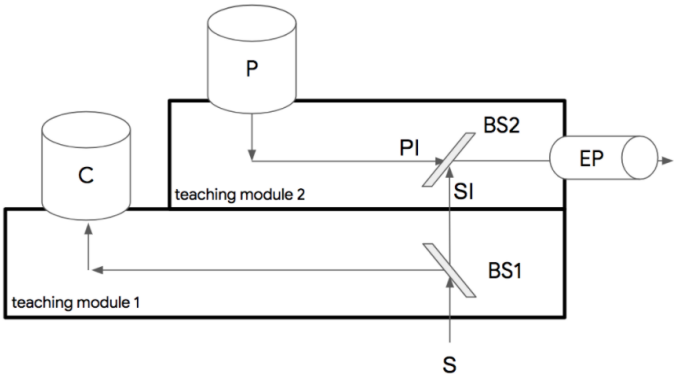


Figure 1: System overview.
1: Schematic sketch of the whole device.
2: A photo of the actual implementation.
3: Schematic of the optic pathway.
4: Photos with indication of corresponding modules.

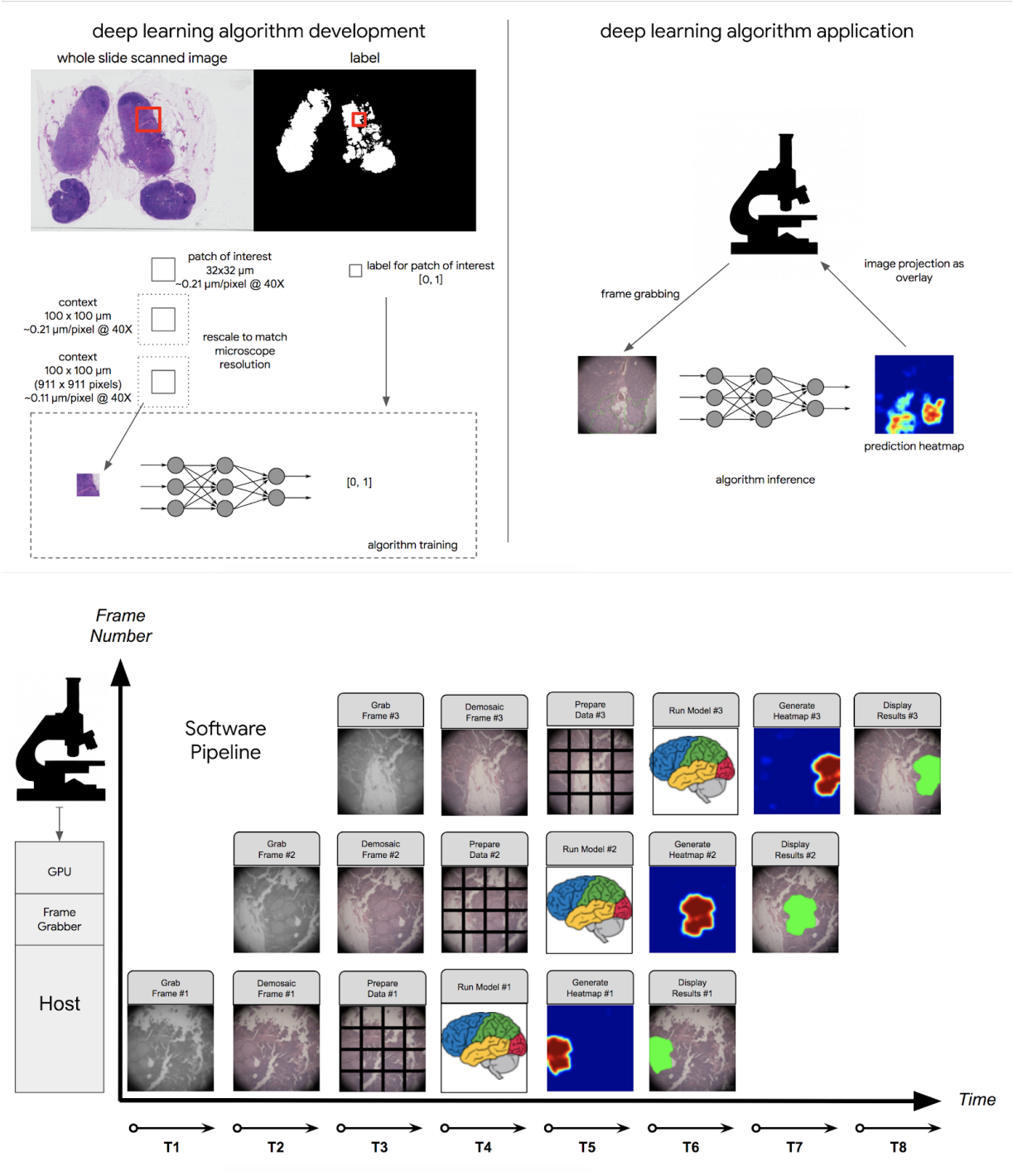


Figure 2: Deep learning, software and hardware integration.

1. Deep learning algorithm development and application.
2. Schematic of the software pipeline.

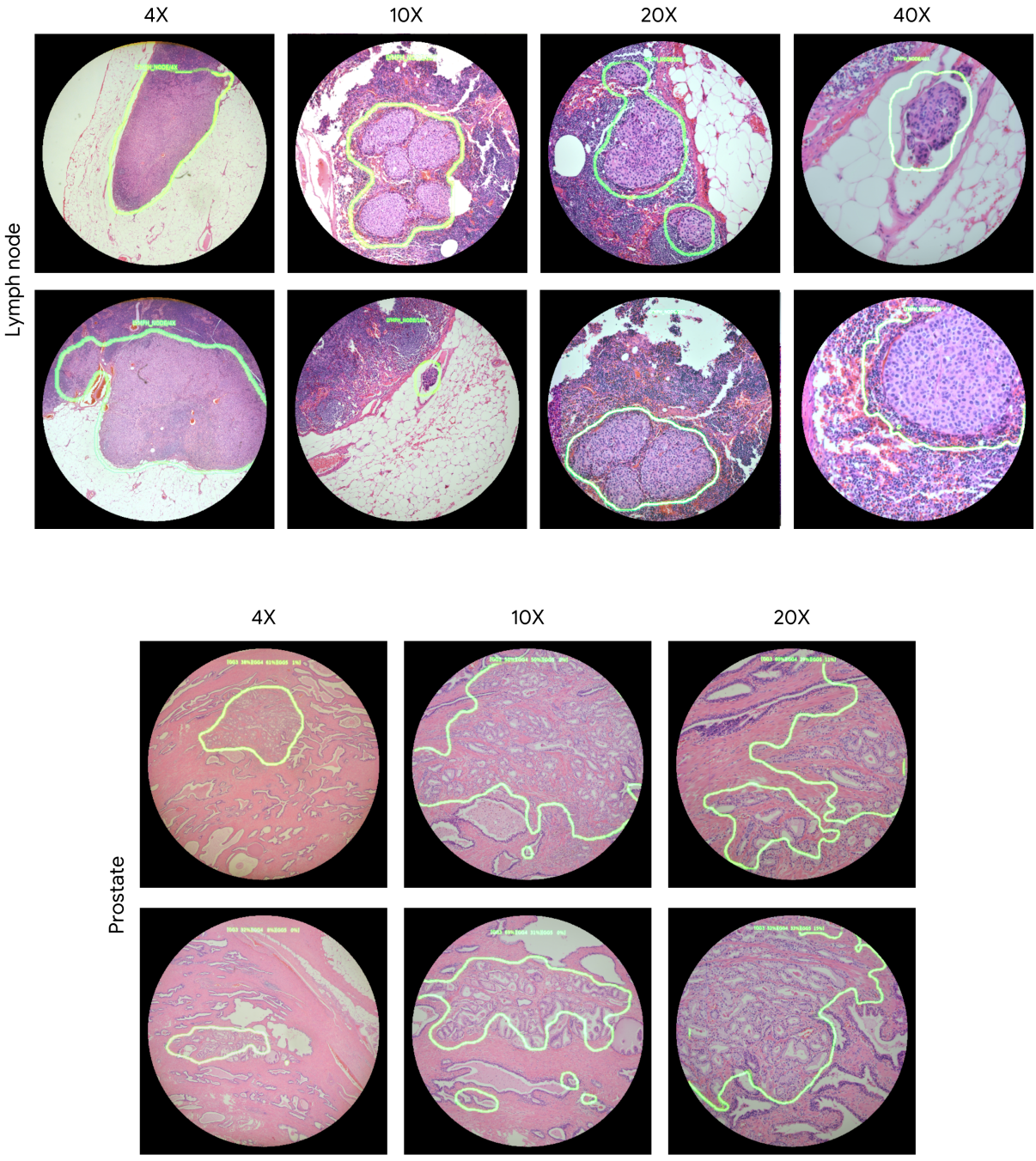


Figure 3: Sample views through the lens.
Top: Lymph node metastasis detection at 4X, 10X, 20X, and 40X.
Bottom: Prostate cancer detection at 4X, 10X, and 20X.

The images show actual views through the lens with overlaid algorithm prediction results. Green outlines are the predicted tumor region.

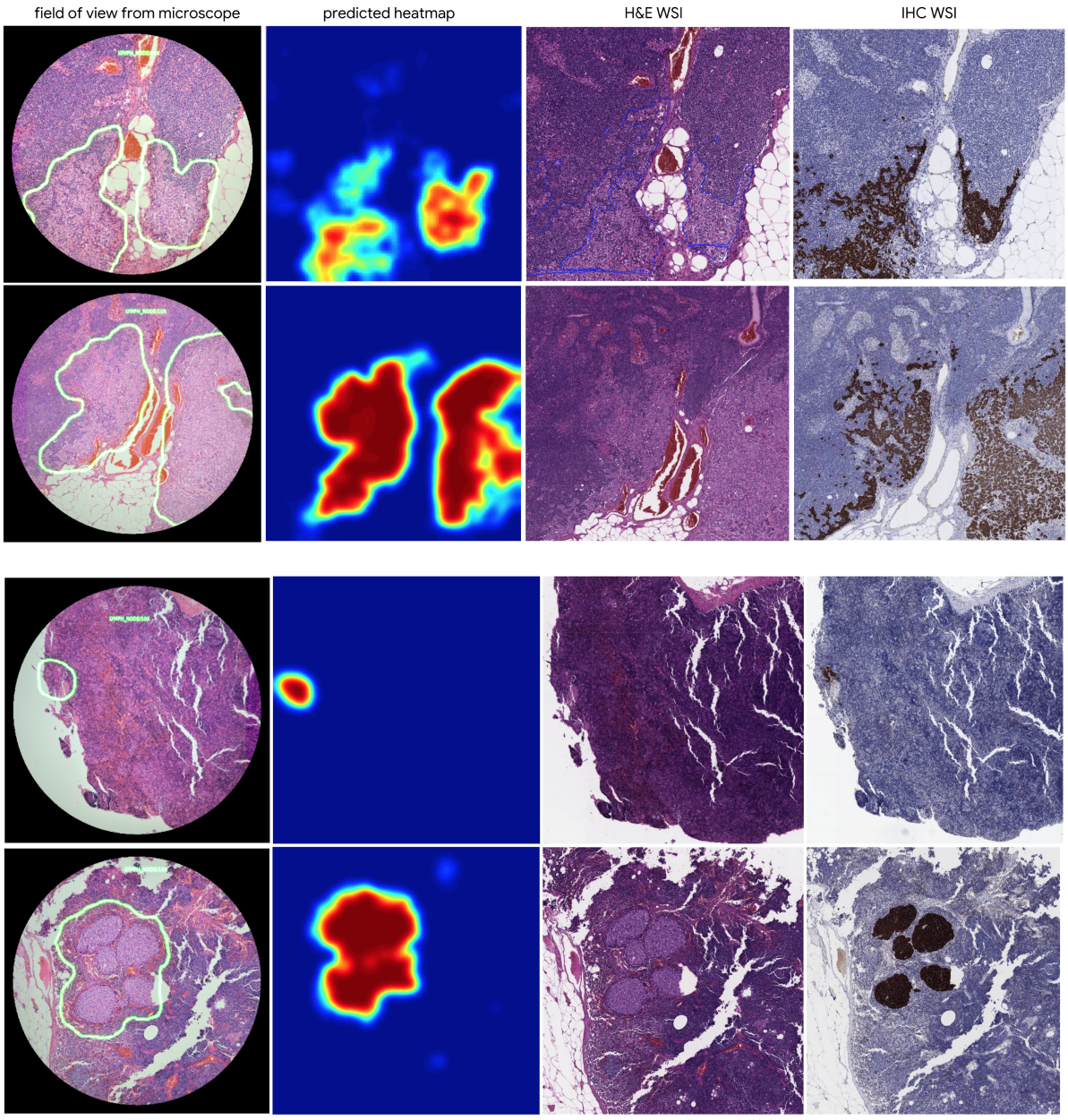
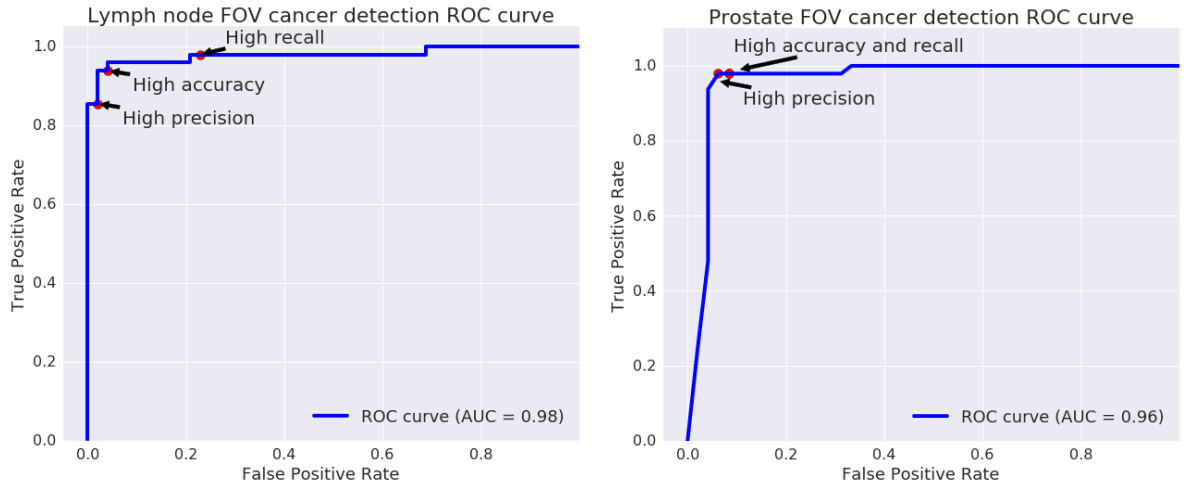


Figure 4: Lymph node cancer detection at 10X.
Left to right: Microscope fields of view, predicted heatmap, WSI H&E, WSI IHC.



Tissue Type	Operating Threshold	Accuracy (CI_{95})	Precision (CI_{95})	Recall (CI_{95})
lymph node	High Accuracy	**0.96 (0.92-0.99)**	0.98 (0.93-1.00)	0.94 (0.86-1.00)
	High Precision	0.92 (0.86-0.97)	**0.98 (0.92-1.00)**	0.85 (0.75-0.95)
	High Recall	0.89 (0.81-0.95)	0.82 (0.72-0.92)	**0.98 (0.93-1.00)**
prostate	High Accuracy	**0.96 (0.92-0.99)**	0.94 (0.87-1.00)	0.98 (0.93-1.00)
	High Precision	0.95 (0.90-0.99)	**0.96 (0.89-1.00)**	0.94 (0.86-1.00)
	High Recall	0.96 (0.92-0.99)	0.94 (0.87-1.00)	**0.98 (0.93-1.00)**

Figure 5: ROC curve and performance metrics for lymph node metastases detection and prostate cancer detection.

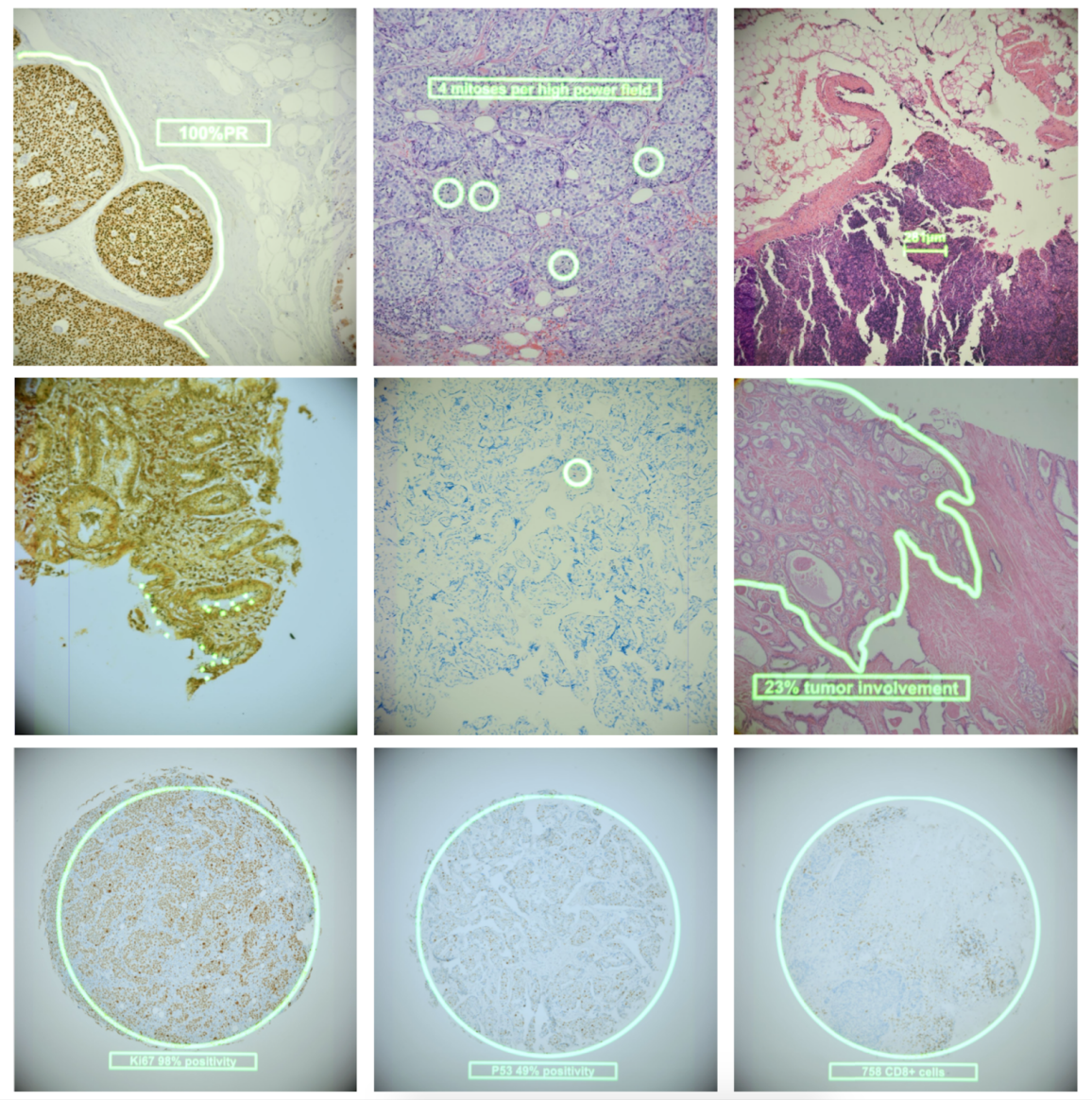


Figure 6: Possible applications on the platform.

1. Progesterone Receptor (PR) quantification.
2. Mitosis Counting on H&E slide.
3. Measurement of tumor size.
4. Identification of Helicobacter pylori (H. pylori).
5. Identification of Mycobacterium.
6. Identification of prostate cancer region with estimation of percentage tumor involvement.
7. Ki67 quantification.
8. P53 quantification.
9. CD8 quantification.

Note that these are illustrations how relevant predictions for these tasks could look like in the ARM, we have not trained the respective algorithms yet.

Disclosure of Potential Conflicts of Interest

P. Chen, R. MacDonald, K. Gadepalli, K. Nagpal, Y. Liu, T. Kohlberger, G. Corrado and M. Stumpe are employees of Google and own Alphabet stock.

Acknowledgments

We would like to thank the pathologists that provided initial user feedback: Dr. Mahul Amin, Dr. Michael Emmert-Buck, Dr. Jenny Smith, Dr. Niels Holger Olson, Dr. Trissia Brown, Dr. Scott Binder, Dr. Sangoi Ankur, as well as colleagues who provided assistance with engineering components and paper writing: Toby Boyd, Lina Dong, Craig Mermel.

Authors' contributions

P. Chen led the deep learning algorithm development and evaluation, K. Gadepalli led the software integration, R. MacDonald led the optics development, K. Nagpal prepared data for the prostate model, Y. Liu prepared data for the lymph node model, T. Kohlberger prepared data for the focus quality model, G. Corrado provided strategic guidance, J. Hipp provided clinical guidance, M. Stumpe conceived the idea and led the overall development. All authors contributed to writing the manuscript.

References

[1] M. V. Dieci *et al.*, “Update on tumor-infiltrating lymphocytes (TILs) in breast cancer, including recommendations to assess TILs in residual disease after neoadjuvant therapy and in carcinoma in situ: A report of the International Immuno-Oncology Biomarker Working Group on Breast Cancer,” *Semin. Cancer Biol.*, Oct. 2017.

[2] R. Salgado *et al.*, “The evaluation of tumor-infiltrating lymphocytes (TILs) in breast cancer: recommendations by an International TILs Working Group 2014,” *Ann. Oncol.*, vol. 26, no. 2, pp. 259–271, Feb. 2015.

[3] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015.

[4] V. Gulshan *et al.*, “Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs,” *JAMA*, vol. 316, no. 22, pp. 2402–2410, Dec. 2016.

[5] J. G. Elmore *et al.*, “Diagnostic concordance among pathologists interpreting breast biopsy specimens,” *JAMA*, vol. 313, no. 11, pp. 1122–1132, Mar. 2015.

[6] B. Ehteshami Bejnordi *et al.*, “Diagnostic Assessment of Deep Learning Algorithms for Detection of Lymph Node Metastases in Women With Breast Cancer,” *JAMA*, vol. 318, no. 22, pp. 2199–2210, Dec. 2017.

[7] G. Litjens *et al.*, “Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis,” *Sci. Rep.*, vol. 6, p. 26286, May 2016.

[8] A. Esteva *et al.*, “Dermatologist-level classification of skin cancer with deep neural networks,” *Nature*, vol. 542, no. 7639, pp. 115–118, Feb. 2017.

[9] Y. Liu *et al.*, “Detecting Cancer Metastases on Gigapixel Pathology Images,” *arXiv [cs.CV]*, 03-Mar-2017.

[10] D. S. W. Ting *et al.*, “Development and Validation of a Deep Learning System for Diabetic Retinopathy and Related Eye Diseases Using Retinal Images From Multiethnic Populations With Diabetes,” *JAMA*, vol. 318, no. 22, pp. 2211–2223, Dec. 2017.

[11] A. Janowczyk and A. Madabhushi, “Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases,” *J. Pathol. Inform.*, vol. 7, p. 29, Jul. 2016.

[12] R. Poplin *et al.*, “Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning,” *Nature Biomedical Engineering*, vol. 2, no. 3, pp. 158–164, Mar. 2018.

[13] D. Wang, A. Khosla, R. Gargeya, H. Irshad, and A. H. Beck, “Deep Learning for Identifying Metastatic Breast Cancer,” *arXiv [q-bio.QM]*, 18-Jun-2016.

[14] M. Abadi *et al.*, “TensorFlow: A system for large-scale machine learning,” *arXiv [cs.DC]*, 27-May-2016.

[15] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, “Rethinking the inception architecture for computer vision,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 2818–2826.

[16] E. Shelhamer, J. Long, and T. Darrell, “Fully Convolutional Networks for Semantic Segmentation,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 4, pp. 640–651, Apr. 2017.

[17] C. W. Pirnstill and G. L. Coté, “Malaria Diagnosis Using a Mobile Phone Polarized Microscope,” *Sci. Rep.*, vol. 5, p. 13368, Aug. 2015.

Supplementary Data

Video

A video describing the Augmented Reality Microscope is available at
https://www.youtube.com/watch?v=9Mz84cwVmS0

Detection of current objective power

Identification of the currently active objective could be done via an intelligent nose piece that transmits the current objective ring position to a computer. However, that solution would not work for all microscopes, requires additional drivers for the nose piece hardware (e.g. no Linux drivers exist for the Nikon nose piece we used here), and increases to the overall hardware cost. As an alternative, we outfitted the nosepiece with color codings next to each objective, and attached a small camera to the microscope frame (Figure S1). The camera is held in a six degrees of freedom mechanical mount to point at the color codings and transmits its image via USB connection to the workstation. The control software determines the currently active color coding via simple thresholding of the hue channel. This approach worked as an instantaneous and highly reliable detection method in our setup.

Figure S1: Close-up of the objective detection mechanism: A USB camera is pointed at the nose piece ring for detection of color coded markers that are unique for each objective ring position.

Alternative designs considered and related technologies

An alternative design for the heads-up display was considered where the projection P via BS2 would be integrated into the eye pieces instead of its current position in the collimated light column. However, the advantages of the current design are (1) it is much more easy to achieve parallax-free display and simultaneous focus for both S and P, (2) the size of the eye pieces

would have imposed a considerable size restriction on the optics for the projection and made the form factor less desirable (e.g. the projection optics would have been close to the eye piece), and (3) that the current design can easily be fitted into most common microscopes.

The projection of the image into the optical path could be done using an LCD display instead of an active display and a beamsplitter. However, since the LCD would be blocking light from the underlying sample to create the projection image (instead of adding light into it), this approach would come with the disadvantages of reduced brightness and contrast, distortions and obscurations of the underlying tissue, and inherently high parallax.

For identifying the current objective power, an alternative to the intelligent nosepiece or the color-coding approach used here would be to detect the magnification of the tissue sample directly from the image using a deep learning approach, for instance a convolutional neural network. However, while in many anatomical pathology cases this classification should be relatively straightforward due to the inherent scale of typical morphologies (e.g. nuclei, glands), this approach would be limited to applications in which the scale can be clearly discerned. And even within those, feature-less regions of the sample (e.g. fat tissue or stroma) could cause errors and lead to a less robust solution.

A fundamentally different approach for displaying auxiliary information along with a sample would be to simply show a digitized image of the camera C1 plus the projection P in a digital eyepiece. Similar systems exist already[17]. The main advantage of that approach is that it eliminates the need for the custom optics for a parallax-free projection, and can reduce the cost of the overall system. However, losing the “real” analog view onto the sample would significantly decrease the image quality and sacrifice many of the advantages of an optical microscope.