

# Case Study

Healthcare  
Medical Imaging with AI



## Ningbo Konfoong Bioinformation Technology (KFBIO) Accelerates M. Tuberculosis Detection with Intel® AI

**PyTorch 1.6 with integrated Intel MKL-DNN delivered up to 11.4X<sup>1</sup> faster inferencing on Intel Xeon® Gold 6252 processors**

Intel®  
AI Builders  
Member

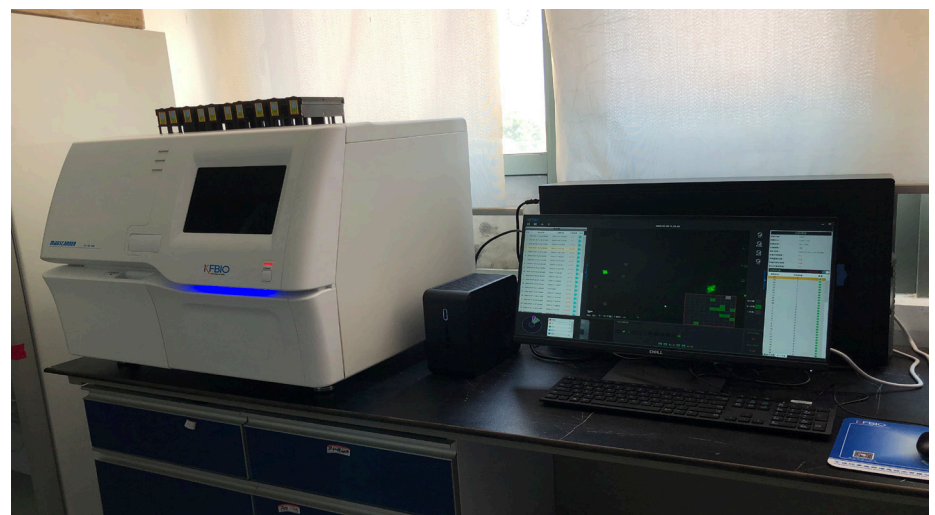


According to the World Health Organization (WHO), tuberculosis (TB) is one of the top ten causes of death and the leading cause from a single infectious agent (above HIV/AIDS). One of the WHO's Sustainable Development Goals is to end the TB epidemic by 2030. Some 58 million lives were saved through TB diagnosis and treatment between 2000 and 2018,<sup>2</sup> thus rapid and accurate detection remains critical to continuing to reduce the number of cases each year.

Sputum smear microscopy (SSM) has been the primary method of diagnosis of tuberculosis because of its simplicity and associated low costs. However, manual reading of stained slides is time-consuming and exposes doctors to potential risk of infection. KFBIO uses its pathology scanning systems to digitize SSM slides as high resolution images and then apply deep learning (DL) algorithms to support the diagnosis of M. TB infection.

### DL-based Screening Assists Diagnosis of M. Tuberculosis

Ningbo Konfoong Bioinformation Co., Ltd (KFBIO) develops pathology-related solutions. These include sample processing equipment, digital pathology scanning systems, pathology information systems, and DL algorithms and models for detecting and classifying precancerous changes and abnormalities. KFBIO's capability in scanning traditional pathological sections into digital images allows scientists to apply DL techniques to assist medical diagnosis. KFBIO has developed a suite of AI-powered pathological solutions based on DL technology, including cervical cancer and M. tuberculosis detection.



Digital imaging makes it possible for medical data to be stored, shared, and analyzed collaboratively over a network. Digital pathology with AI helps address the shortages in medical resources and provides better accessibility to diagnosis for people in remote and rural areas.

The performance demonstrated in this paper underscores the potential of AI-based M. tuberculosis screening and diagnosis through collaboration between KFBIO and Intel.

### Intel AI Accelerates Automated Screening 8.40X<sup>1</sup>

While KFBIO had an effective DL solution for scanning M. Tuberculosis specimens using GPUs, their engineers needed higher performance in order to complete scanning and diagnosis faster. Collaborating with Intel AI engineers, KFBIO optimized their code based on PyTorch and the detectron2 model to take advantage of Intel® Xeon® Scalable processor architecture enhancements, including Intel Deep Learning Boost (Intel DL Boost) technology and the processor's large memory capacity.

2nd Gen Intel Xeon Scalable processors (formerly Cascade Lake) include new Intel DL Boost technology to accelerate inferencing performance. Fast math has been added to Intel Math Kernel Library for Deep Neural Networks (Intel MKL-DNN) to take advantage of the processor's hardware advances. The Intel MKL-DNN optimizations are abstracted and integrated directly into the PyTorch framework,<sup>3</sup> so end users can take advantage of this technology without special code modifications. The latest official release of PyTorch integrates Intel MKL-DNN by default. Thus, users can achieve performance benefits on 2nd Gen Intel Xeon Scalable processor-based platforms without additional installation steps.

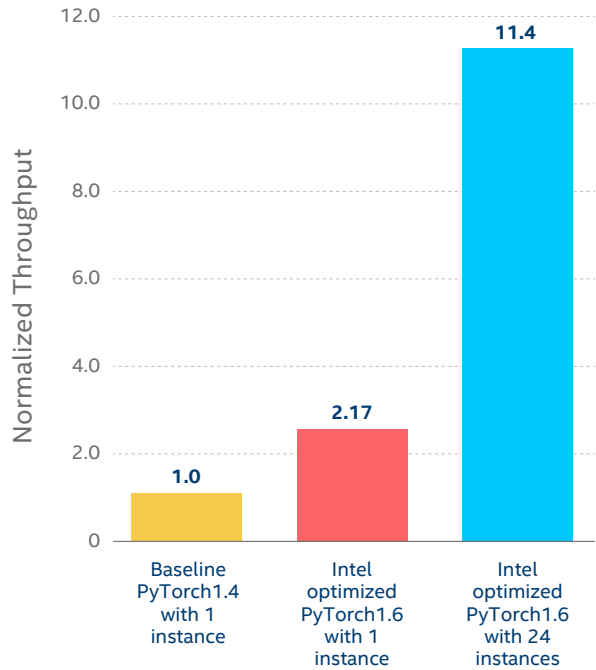
Additionally, 2nd Gen Intel Xeon Scalable processors support many terabytes of memory enabling fast access to large datasets. Digital pathology images tend to be large in size, which can strain GPU memory capacity. With larger memory capacity, the scanning solution could run multiple, simultaneous inferencing instances on the same configuration.

Optimizations for KFBIO's M. Tuberculosis scanning algorithm were straightforward by simply using PyTorch 1.6 optimized for Intel architecture instead of the previous version of PyTorch 1.4. Additional optimizations were completed by profiling the detectron2 model and optimizing minimal operations.

Benchmark testing on Intel Xeon Gold 6252 processors revealed improved inference performance compared to a baseline model without optimizations running on the same processor.

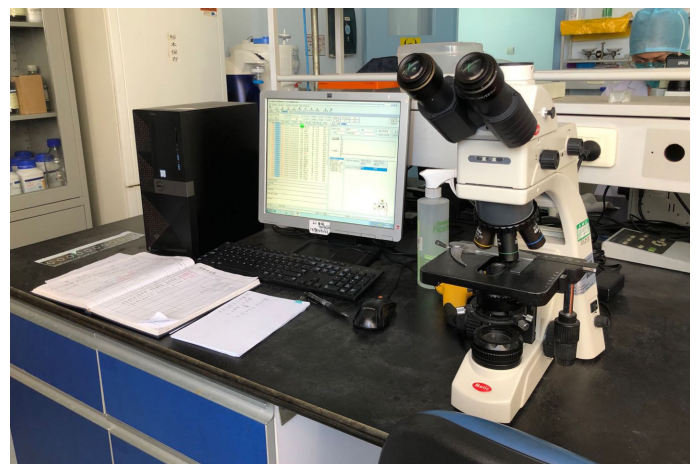
Two optimized benchmarks returned performance boosts of 2.17X for a single instance and 11.4X for 24 instances (see chart at top right). The solution benefited from large memory capacity when running 24 instances.

### Detectron2 Throughput Performance (Higher is Better)



### A Customer Deployment

Recently, KFBIO successfully deployed their DL-based TB screening solution at Shanghai Public Health Clinical Center in China in an integrated workstation running on 2nd Gen Intel Xeon Scalable processors. The center will use it to accelerate processing of Tuberculosis diagnosis caseloads. The deployment integrates the KFBIO digital pathology scanner and AI software to provide end-to-end capability: digitizing specimen slides, performing AI inferencing for detection, visualizing results, and producing diagnostic reports. According to customer evaluation, the system delivers 86.8 percent average precision (AP) in detection and 88.9 percent accuracy in classification.<sup>4</sup> The entire workflow pipeline can be completed in less than 80 seconds for a single-case diagnosis from input of specimen to output of report. The customer is satisfied with the performance and expects to further shorten the process to 60 seconds.



## Conclusion

Accelerating machine screening of specimens using AI technologies can increase the number of specimens screened and potentially alert physicians to patient concerns more quickly than manual methods alone. Using PyTorch 1.6 with integrated Intel MKL-DNN library allowed KFBIO's detectron2 model to take advantage of Intel DL boost in 2nd Gen Intel Xeon Scalable processors. The large memory support of the processor architecture also enabled multiple inferencing instances to run simultaneously directly from memory. With PyTorch optimized for Intel DL boost and large memory, the detectron2 model ran as much as 11.4X faster than the baseline model.

For more information about KFBIO, visit [en.kfbio.cn/index.php](http://en.kfbio.cn/index.php)

Learn more about the Intel AI Builders program at [builders.intel.com/ai/membership](http://builders.intel.com/ai/membership)

Intel®  
AI Builders  
Member

**Konfoong Biotech International Co., Ltd.** is a high-tech biological information technology enterprise specializing in digital pathology system development and production. Their integrated digital pathology diagnosis system provides an ultimate solution platform, integrating health care, education, scientific research and information services.

<sup>1</sup> KFBIO m. tuberculosis screening detectron2 model throughput performance on 2nd Intel® Xeon® Gold 6252 processor:

NEW:

Test 1 (single instance with PyTorch 1.6: Tested by Intel as of 5/22/2020. 2-socket 2nd Gen Intel® Xeon® Gold 6252 Processor, 24 cores, HT On, Turbo ON, Total Memory 192 GB (12 slots/16 GB/2666 MHz), BIOS: SSE5C620.86B.02.01.0008.031920191559 (ucode: 0x500002c), Ubuntu 18.04.4 LTS, kernel 5.3.0-51-generic, mitigated

Test 2 (24 instances with PyTorch 1.6: Tested by Intel as of 5/22/2020. 2-socket 2nd Gen Intel Xeon Gold 6252 Processor, 24 cores, HT On, Turbo ON, Total Memory 192 GB (12 slots/16 GB/2666 MHz), BIOS: SSE5C620.86B.02.01.0008.031920191559 (ucode: 0x500002c), Ubuntu 18.04.4 LTS, kernel 5.3.0-51-generic, mitigated

BASELINE:

(single instance with PyTorch 1.4): Tested by Intel as of 5/22/2020. 2-socket 2nd Gen Intel Xeon Gold 6252 Processor, 24 cores, HT On, Turbo ON, Total Memory 192 GB (12 slots/16 GB/2666 MHz), BIOS: SSE5C620.86B.02.01.0008.031920191559 (ucode: 0x500002c), Ubuntu 18.04.4 LTS, kernel 5.3.0-51-generic, mitigated.

<sup>2</sup> <https://www.who.int/news-room/fact-sheets/detail/tuberculosis>

<sup>3</sup> See the article Intel and Facebook collaborate to Boost PyTorch CPU Performance for more details on recent performance accelerations.

<sup>4</sup> Testing done by Shanghai Public Health Clinical Center and KFBIO. Intel does not control or audit third-party data. You should consult other sources to evaluate accuracy.

Performance results are based on testing as of dates shown in configuration and may not reflect all publicly available security updates. No product can be absolutely secure. See configuration disclosure for details.

Intel's compilers may or may not optimize to the same degree for non-Intel microprocessors for optimizations that are not unique to Intel microprocessors. These optimizations include SSE2, SSE3, and SSSE3 instruction sets and other optimizations. Intel does not guarantee the availability, functionality, or effectiveness of any optimization on microprocessors not manufactured by Intel. Microprocessor-dependent optimizations in this product are intended for use with Intel microprocessors. Certain optimizations not specific to Intel microarchitecture are reserved for Intel microprocessors.

Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products. For more complete information visit: <http://www.intel.com/performance>

© Intel Corporation. Intel, the Intel logo, and other Intel marks are trademarks of Intel Corporation or its subsidiaries. Other names and brands may be claimed as the property of others.