intel.

Breakthrough in Cancer Screening Using Deep Learning

Researchers have created an innovative AI solution for analyzing breast cancer images that achieves accuracy of 98 percent with only 20 percent of the data labeled[1]. While previous solutions could detect cancer, this one can grade it too.

At a Glance:

- NAS-SGAN uses AI to detect and classify breast cancer tumors, which helps physicians making healthcare decisions.

- The innovative approach uses Intel® Xeon® Scalable processors for training. GPUs could not hold the entire AI model in memory.

- NAS-SGAN achieved 98 percent accuracy with only 20 percent of the training data annotated[1].

AI can improve the consistency and repeatability of breast cancer tumor grading, but it's difficult to build the AI solution. Annotating images for training is time-consuming and laborious, and there aren't many labeled images already available. Two researchers have pioneered a novel approach. It uses labeled and unlabeled images together to achieve high accuracy while minimizing the annotation workload. The solution failed on GPUs, so Intel helped with a technology architecture based on 2nd Generation Intel® Xeon® Scalable processors.

Challenge

- Deep learning solutions can help with cancer diagnosis, but they require labeled images for training.
- Labeling histopathological images is time-consuming and labor-intensive.
- For best results, the deep learning solution needs to be able to process high resolution images, and GPUs cannot hold the entire AI model in memory.

Solution

- The NAS-SGAN approach uses unlabeled images to understand the data distribution and labeled images to grade the cancer.
- Four servers based on 2nd gen Intel® Xeon® Scalable processors train the solution in parallel, with 192GB of memory per server.
- The Intel® Optimization for TensorFlow makes it easy to use acceleration features in the processors.

Results

- NAS-SGAN achieved 98 percent accuracy with only 20 percent of the data annotated[1].
- The new solution can not only detect cancer, but also classify it, something previous solutions based on Generative Adversarial Networks (GANs) could not do.
- The solution streamlines diagnosis, with physicians reviewing the images and classifications to make their treatment decisions.

Limited Labeled Data Makes Cancer AI a Challenge

Breast cancer is the world's most prevalent cancer, with 2.3 million women diagnosed in 2020[2]. It occurs in every country, and affects mostly women, at any age after puberty.

To diagnose and monitor the disease, histopathological images are used. These are images of tissue specimens at a microscopic scale. Nuclear atypia scoring (NAS) grades a tumor according to how much the tumor cells differ from normal tissue. The World Health Organization (WHO) has adopted the Nottingham grading system (NGS) as the standard for breast cancer grading. The NGS is associated with the survival chances of the patient and can be used to guide individualized treatment plans.

Grading the images manually is difficult. Physicians analyzing images might not consistently make the same grading decision, and opinions can vary between those who perform the analysis. It's a time-consuming process, too.

Automatic screening using deep learning can overcome the limitations of manual analysis. However, it's difficult to generate a deep learning model because of insufficient training data. While the raw images can be created cost-effectively in minutes, the process of labeling them is time-consuming and laborious. Labeled histopathological images are scarce.

In addition, the deep learning model needs to be able to handle large images (1024 x 1024 pixels) for best results. "To grade the images accurately, it is important that the morphological features are well extracted," said Dr Madhu Nair from the Artificial Intelligence & Computer Vision Lab, Cochin University of Science and Technology. "The grade and stage of the cancer depend on the morphological features. We cannot simply reduce the size of the image to fit the size of the model. We need to use high-resolution images to extract all the differences."

Dr Asha Das (Union Christian College, India) worked with Nair as they took on the challenge. "Our question was: Is it possible to develop a model using less labeled data, and still get high accuracy?" said Nair.

They needed a hardware architecture and software architecture that could meet the demanding performance requirements, as they built a pioneering new approach to cancer screening.

Generative Adversarial Networks Help AI Generalize

Generative Adversarial Networks (GANs) are a type of neural network solution that can be used to generate new images, and judge whether images are genuine or fake. GANs are made of two neural networks that work together: a generator, which creates images; and a discriminator, which judges whether the images match a sample set or not.

GANs can be used to generate histopathological images. These can be added to a training data set of genuine histopathological images, so that a deep learning model can better generalize when shown previously unseen images.

GANs have also been used in the past to detect tumors and other anomalies, but previous implementations have not been able to grade the images.

Das and Nair created a model, called NAS-SGAN, which can discriminate the different cancer grades (see Figure 1). Its name is short for nuclear atypia scoring (NAS) semi-supervised generative adversarial network (SGAN). NAS-SGAN uses the unlabeled images to understand the data distribution, and the labeled images to grade the cancer.

Technical Components of the Solution

- **Intel® Xeon® Gold 6248 processors.** These two-socket processors offer 20 cores. The processors are housed in servers with 192GB of memory for the deep learning solution.

- **Intel® Optimization for TensorFlow.** This software enhances the performance of TensorFlow on Intel® architecture, by taking advantage of acceleration features in Intel® processors.

- **Horovod.** This open-source solution enabled the researchers to run their training across multiple servers, working in a single cluster. Each server processed 16 images from a 64-image batch simultaneously.

It works in two phases:

1. A GAN is used to create images that are indistinguishable from genuine histopathological images. The GAN is trained using unlabeled images, which are relatively easy to obtain. The new images are used to help the solution understand the data distribution.

2. The GAN discriminator is then trained with the labeled images to predict the cancer grades.

Attempts to implement the solution using GPUs failed. "It took several days to execute and would sometimes stop," said Nair. "We couldn't complete the project using those machines."

Das and Nair worked with Intel to implement the solution using 2nd gen Intel® Xeon® Scalable processors. Four servers were organized as a compute cluster, without any deep learning accelerators. The servers and storage were connected using a 25G Ethernet network.

"The Intel® architecture was amazing," said Nair. "We were able to complete the training in a few hours. Because the servers had 192GB of memory, more than the 40GB or 80GB available on graphics cards, we were able to use high-resolution images and fit the whole model in memory."

The software stack used the Intel® Optimization for TensorFlow, which is designed to use acceleration features of Intel® processors. They include Intel® Deep Learning Boost (Intel® DL Boost), which accelerates matrix operations often used in deep learning training. "The Intel Optimization for TensorFlow worked the same as the main version of TensorFlow," said Nair.

The open-source Horovod training framework was used to enable distributed training across the server cluster.
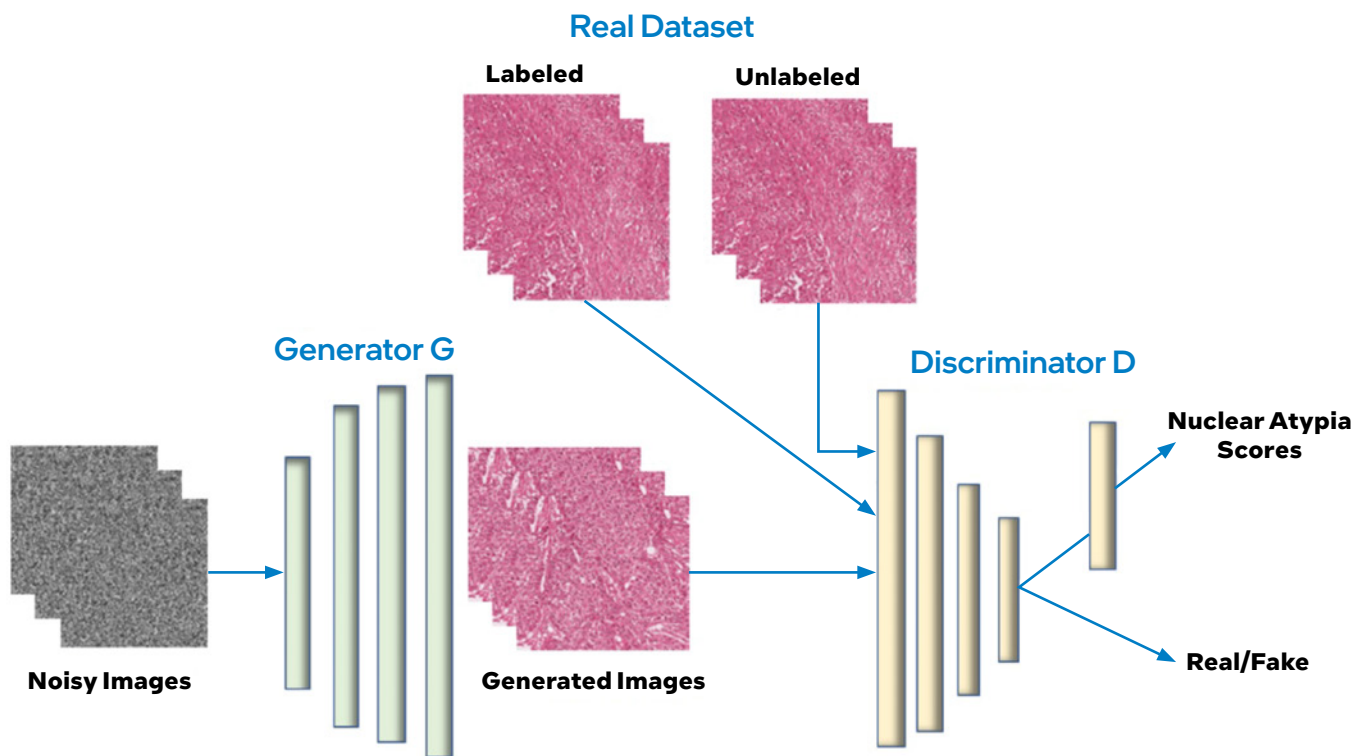
2

Figure 1. The NAS-SGAN model uses unlabeled images to generate new images that broaden the training data set and uses labeled images to train the discriminator to classify the different cancer grades.

Intel Supports the Study

Intel worked closely with Das and Nair on the technology aspects of the project. "We had the idea but were worried about whether it would work or not," said Nair. "I shared our problem with the Intel team, and was extremely happy that they immediately understood the importance of this work. They gave us the opportunity to use this distributed architecture."

He added: "They also helped us to improve the model, and shared optimizations with us to get it working. That's the reason why we were able to succeed. We're grateful to Intel for its support. It was a really nice experience working with the Intel team, and we look forward to continuing this collaboration."

High Accuracy with Limited Labeled Training Data

Das and Nair compared the performance of NAS-SGAN with ten other GAN algorithms used to detect breast cancer. NAS-SGAN achieved an accuracy of 98 percent, approximately 10 percent higher than the next-rated GAN (WGAN-GP) [1]. Precision was 97 percent, 18 percent higher than WGAN-GP[1]. NAS-SGAN also excelled on the F1-score, which represents a harmonic mean1. It includes (among other things) false positives and negatives. NAS-GAN had an F1-score of 97 percent, 15 percent higher than WGAN-GP[1].

"We were really happy with the results," said Das. "It's remarkable that we were able to achieve 98 percent accuracy with only 20 percent of the data annotated. That's really exciting."

She added: "Our model can better discriminate the features that separate the grades because of the way it uses two GANs. We were able to get almost comparable results using medium sized images, too."

Lessons Learned

The key lessons from this project are:

- The NAS-SGAN algorithm addresses the shortcomings of other GAN models for breast cancer screening by adding the ability to grade cancer images.

- NAS-SGAN achieves high accuracy results even when using a limited amount of annotated data[1]. This helps to minimize the time-consuming and labor-intensive process of classifying images.

- The researchers were unable to realize this project using GPUs. Using CPUs, it was possible to train the AI model using high-resolution training images of 1024 x 1024 pixels and retain the entire model in memory.

- The Intel® Optimization for TensorFlow includes features to accelerate the performance of TensorFlow on Intel® architecture.

Using the other GANs, a physician had to study the image to grade it. The automated scoring of NAS-SGAN helps to streamline the diagnosis process and analysis process, and helps to improve consistency and accuracy in grading. Physicians can review the image and grading to make their treatment decisions.

As for the future? The researchers are looking at how a similar approach could be used for mortalities resulting from cerebral aneurysms and classifying polyps from endoscopies.

Learn More

- Intel® Optimization for TensorFlow
- Intel® Xeon® Scalable processors
- NAS-SGAN: A Semi-supervised Generative Adversarial Network Model for Atypia Scoring of Breast Cancer Histopathological Images

Find the solution that is right for your organization. Contact your Intel representative or visit **intel.com/healthcare**.

intel.

Intel technologies' features and benefits depend on system configuration and may require enabled hardware, software, or service activation. Performance varies depending on system configuration. No computer system can be absolutely secure. Check with your system manufacturer or retailer, or learn more at intel.com

[2] https://www.who.int/news-room/fact-sheets/detail/breast-cancer