

# Vietnam's VinBrain Deploys Healthcare AI Models to 100+ Hospitals

NVIDIA DGX SuperPOD and NVIDIA Clara enable VinBrain to detect abnormalities and accelerate time to diagnosis with AI.

Author: Renee Yao

Doctors rarely make diagnoses based on a single factor — they look at a mix of data types, such as a patient's symptoms, laboratory and radiology reports, and medical history.

VinBrain, a Vietnam-based health-tech startup, is ensuring that AI diagnostics can take a similarly holistic view across vital signs, blood tests, medical images and more.

"Multimodal data is key to delivering precision care that can improve patient outcomes," said Steven Truong, CEO of VinBrain. "Our medical imaging models, for instance, can analyze chest X-rays and make automated observations about abnormal findings in a patient's heart, lungs and bones."

If a medical-imaging AI model reports that a patient's scan shows lung consolidation, Truong explained, doctors could combine the X-ray analysis with a large language model that reads health records to learn the patient has a fever — helping clinicians more quickly determine a more specific diagnosis of pneumonia.

Funded by Vingroup — one of Vietnam's largest public companies — VinBrain is the creator of DrAid, which is the only AI software for automated X-ray diagnostics in Southeast Asia, and among the first AI platforms to be cleared by the FDA to detect features suggestive of collapsed lungs from chest X-rays.

Trained on a dataset of more than 2.5 million images, DrAid is deployed in more than 100 hospitals in Vietnam, Myanmar, New Zealand and the U.S. The software applies AI analysis to medical images for more than 120,000 patients each month. VinBrain is also building a host of other AI applications, including a telehealth product that analyzes lab test results, medical reports and other electronic health records.

The company is part of NVIDIA Inception, a global program designed to offer cutting-edge startups expertise, technology and go-to-market support. The VinBrain team has also collaborated with Microsoft and with academic researchers at Stanford University, Harvard University, the University of Toronto and the University of California, San Diego to develop its core AI technology and submit research publications to top conferences.

The VinBrain team has developed more than 300 AI models that process speech, text, video and images — including X-ray, CT and MRI data.

"Healthcare is complex, so the pipeline requires hundreds of models for each step, such as preprocessing, segmentation, object detection and post-processing," Truong said. "We aim to package these models together so everything runs on GPU servers at the hospital — like a refrigerator or household appliance."

VinBrain recently launched DrAid Appliance, an on-premises, NVIDIA GPU-powered device for automatic screening of medical imaging studies that could improve doctors' productivity by up to 80%, the team estimates.

The company also offers a hybrid solution, where images are preprocessed at the edge with DrAid Appliance, then sent to NVIDIA GPUs in the cloud for more demanding computational workloads.

Another way to access VinBrain's DrAid software is through Ferrum Health, an NVIDIA Inception company that has developed a secure platform to help healthcare organizations deploy AI applications

across therapeutic areas.

VinBrain trains its AI models — which include medical imaging, intelligent video analytics, automatic speech recognition, natural language processing and text-to-speech — using NVIDIA DGX SuperPOD . Adopting DGX SuperPOD enabled Vinbrain to achieve near-linear-level speedups for model training, achieving 100x faster training compared with CPU-only training and significantly shortening the turnaround time for model development.

The team is using software from NVIDIA AI Enterprise , an end-to-end solution for production AI, which includes the NVIDIA Clara platform, the MONAI open-source framework for medical imaging development and the NVIDIA NeMo conversational AI toolkit for its transcription model.

“To develop good AI models, you can’t just train once and be done,” said Truong. “It’s an evolving process to refine the neural networks.”

VinBrain has set up an early validation pipeline for its AI projects: The company tests its early-stage models across a couple dozen hospitals in Vietnam to collect performance data, gather feedback and fine-tune its neural networks.

In addition to using NVIDIA DGX SuperPOD for AI training, the company has adopted NVIDIA GPUs to improve run-time efficiency and deployment. It uses the NVIDIA Triton inference server and NVIDIA TensorRT to streamline inference for more than hundreds of AI models on cloud-based NVIDIA Tensor Core GPUs .

“We shifted to NVIDIA GPUs for inference because of the higher throughput, faster response time and, most importantly, the cost ratio,” Truong said.

After switching from CPUs to NVIDIA Tensor Core GPUs, the team was able to accelerate inference for medical imaging AI by more than 3x, and video streaming by more than 30x.

“In the coming years, we want to become the top company solving the problem of multimodality in healthcare data,” said Truong. “Using AI and edge computing, we aim to improve the quality and accessibility of healthcare, making intelligent insights accessible to patients and doctors across countries.”

Register for NVIDIA GTC , taking place online March 20-23, to learn more about AI in healthcare .

Original URL: <https://blogs.nvidia.com/blog/2023/02/07/vietnam-vinbrain-deploys-healthcare-ai/>