CUSTOMER SUCCESS STORY | NETHERLANDS CANCER INSTITUTE

CANCER RESEARCH ADVANCES WITH NVIDIA AI ENTERPRISE

Image courtesy of Netherlands Cancer Institute

NVIDIA

WITH NVIDIA AI ENTERPRISE, DOCTORS COULD START TREATING TUMORS MORE EFFECTIVELY WITH LESS RADIATION.

INTRODUCTION

The Netherlands Cancer Institute (NKI) has been at the forefront of cancer research and treatment since 1913. Comprised of an internationally acclaimed research center and a dedicated cancer clinic, NKI puts innovative ideas into action for the benefit of patients.

Recently, NKI started conducting research to see if trained artificial intelligence (AI) models, running on virtualized mainstream servers, could deliver the performance needed to enable more accurate tumor targeting and reduced radiation exposure using cone beam computed tomography (CBCT). The solution, enabled by NVIDIA's AI Enterprise software suite, promises to make cancer treatments more effective for more people.

CUSTOMER PROFILE

ANTONI VAN LEEUWENHOEK
NEDERLANDS KANKER INSTITUUT

Organization	Industry	Founded	Location	Size	Website
Netherlands Cancer Institute (NKI)	Healthcare	1913	Amsterdam, the Netherlands	783 researchers, 3,000 clinicians	www.nki.nl

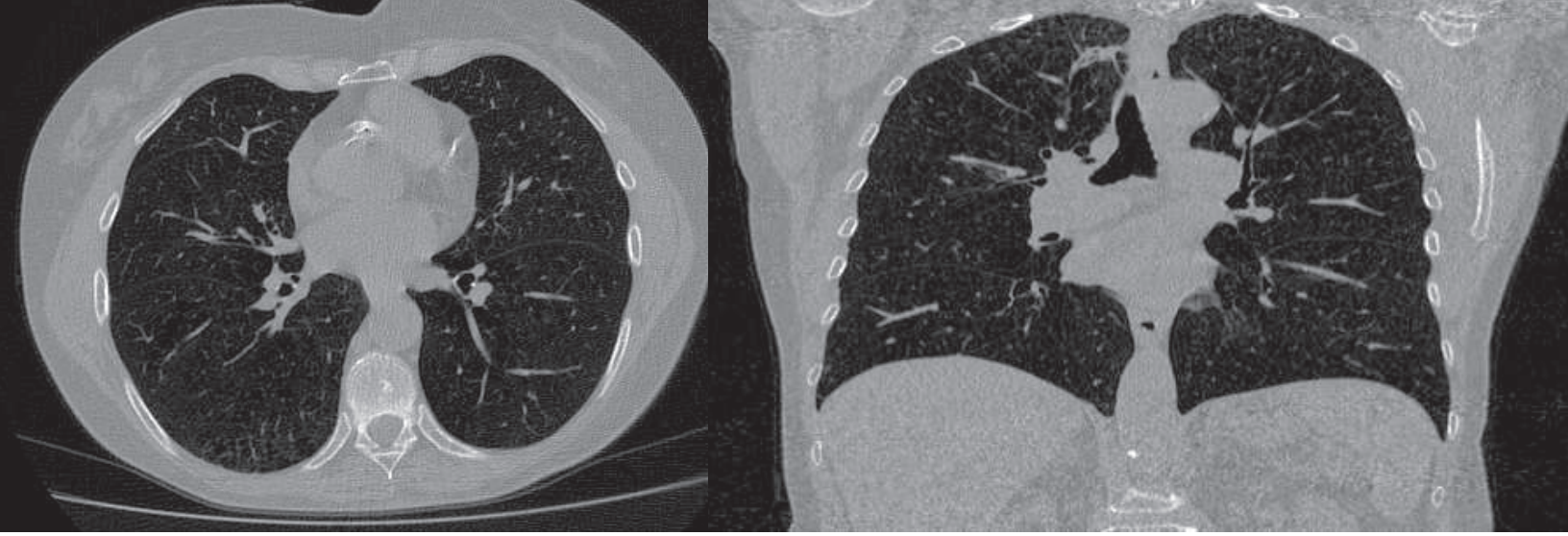

Image courtesy of Netherlands Cancer Institute

SUMMARY

- NKI researchers wanted a way to adapt the radiotherapy treatment to the current anatomy for precision tumor targeting and better treatment planning.
- Using NVIDIA AI Enterprise software and NVIDIA A100 GPUs in a VMware vSphere environment, cone beam computed tomography (CBCT) scans could be rendered in as little as 60 seconds at 2mm resolution and in as little as five minutes at 1mm resolution.
- AI models were trained to reconstruct 3D volumes directly from projection data, removing artifacts caused by photon noise.
- Researchers are hopeful that the technology can one day enable high-resolution 4D CBCT reconstruction, enabling accurate tracking of tumors as they move during breathing cycles.
- Faster, more robust computing power is resulting in more accurate and efficient radiation treatments for patients with cancer.
- From an IT perspective, the NVIDIA AI Enterprise platform makes it easy to isolate and manage large AI workloads at scale, while simplifying infrastructure maintenance related to consolidation, expansion, and upgrades.

SOFTWARE

- **AI Software:** NVIDIA AI Enterprise
- **Server Virtualization:** VMware vSphere

HARDWARE

- **Server:** Dell R740
- **GPU:** NVIDIA A100 80GB

CHALLENGE STATEMENT

As medical imaging advances, so does the need for increased compute, bandwidth, and storage capabilities. This is especially true in the treatment of cancer, which relies on body imaging with Conebeam CT (CBCT) for accurate planning of targeted radiation therapies.

Unlike higher-quality CT scans, CBCT uses an X-ray tube and a large flat detector panel that rotates around the patient, capturing data with a cone-shaped X-ray beam instead of the "slices" that CT scanners are typically known for. The data received from these systems is used to reconstruct 3D images for a variety of specialties, including thorax, dental, oral/maxillofacial (mouth, jaw, and neck), and ears, nose, and throat (ENT). CBCT is essential in radiation therapy, where a daily CBCT scan is used for adapting the treatment plan to the current anatomy of the patient, which can change due to weight loss or eating a big meal prior to treatment.

The more accurate the CBCT images, the easier it is for clinicians to localize smaller tumors during radiation therapy and provide more targeted treatments. Accurate CBCT reconstruction would simplify treatment planning, potentially reducing the need for a dedicated planning CT scan.

CBCT images are computed by a reconstruction algorithm using the projection data acquired by the scanner. Reconstructing CBCT images, however, is more challenging compared to CT reconstruction, and the image quality given by classical reconstruction methods is poor.

Deep learning reconstruction methods have received a lot of attention from the medical imaging community recently for modalities such as CT and MRI, but the applications to CBCT remain very limited. Training models to reconstruct 3D volumes directly from projection data at clinically relevant resolution and projection count is an extremely memory-hungry task, since a single CBCT volume at 1 mm resolution would already occupy 1 GB of memory taking gradient information into account.

WHY NVIDIA AI ENTERPRISE WITH VMWARE VSPHERE?

Clinician Benefits

- Get higher-resolution scans much faster
- Detect tumors more reliably
- Simplify planning for adaptive radiotherapy
- Potentially reduce radiation doses for CBCT outside radiotherapy

IT Admin Benefits

- Utilize GPU resources more easily
- Isolate and manage large AI workloads
- Test the platform/solution remotely

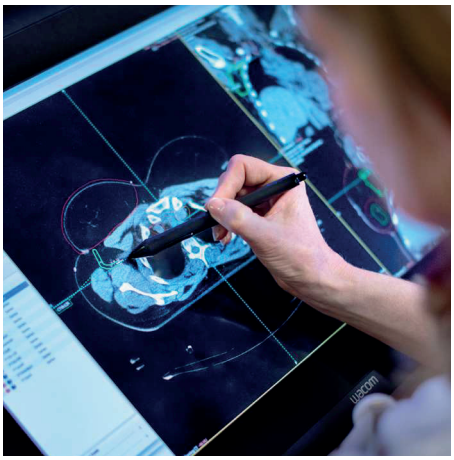

Image courtesy of Netherlands Cancer Institute

"We can get native resolution from a CBCT scan using just 55 gigabytes of GPU memory out of 80. This leaves us some extra room to correct for other image distortions like scatter and potentially do 4D reconstructions to account for movement during a patient's breathing cycle."

Nikita Moriakov
Post Doctoral Researcher, NKI

SOLUTION STATEMENT

NKI leveraged a powerful solution developed by NVIDIA, in partnership with VMware, to deploy and manage AI workloads on mainstream 1U/2U servers in a virtualized data center environment, without compromising performance.

NKI researchers were given access to best-in-class AI software—the NVIDIA AI Enterprise suite—certified for the industry's leading virtualization platform, VMware vSphere. Designed to achieve near bare-metal performance for AI training and inference, the solution makes larger, more complex AI training and data analytics possible with new scale-out capabilities.

The solution enables high availability for AI workloads, while simplifying infrastructure maintenance such as consolidation, expansion, and upgrades. VMware vSphere transforms bare-metal servers (including CPU- and GPU-based resources) into centrally managed AI and machine learning infrastructure pools that can quickly provision virtual machines (VMs) and containers that can be accessed from anywhere on demand.

VMware vSphere increases availability, tightens security, streamlines maintenance, and reduces costs to create an agile, efficient, resilient, and secure infrastructure platform that supports existing workloads and next-generation applications.

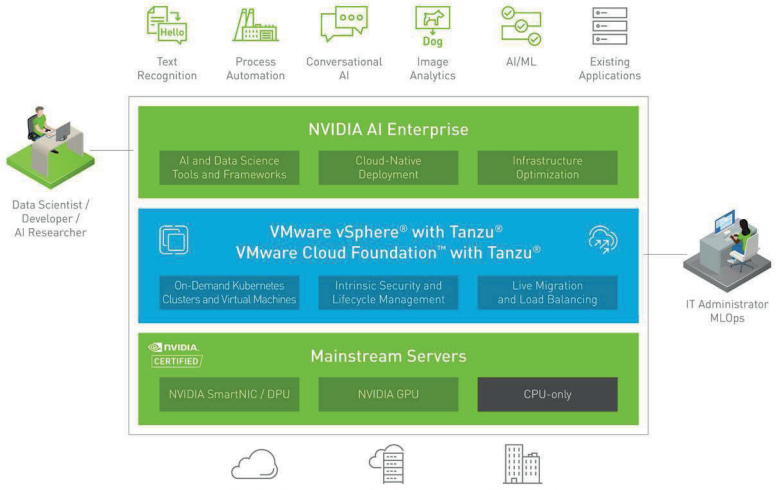

Figure 1. With the AI-Ready Platform, end users can access the software they need to build successful AI projects, and IT admins can support the projects with tools they're familiar with.

To evaluate the benefits of updating their infrastructure to the latest GPU architecture, Ampere, the NKI research team developed two different AI models. One model was trained using eight NVIDIA RTX 8000 graphics cards, based on the previous generation NVIDIA Turing™ architecture, and NVIDIA RTX 8000 graphics cards in a bare metal environment, and the other was trained using two NVIDIA A100 (80GB) graphics cards, based on the newer NVIDIA Ampere GPU architecture, graphics cards in a virtualized environment with NVIDIA AI Enterprise.

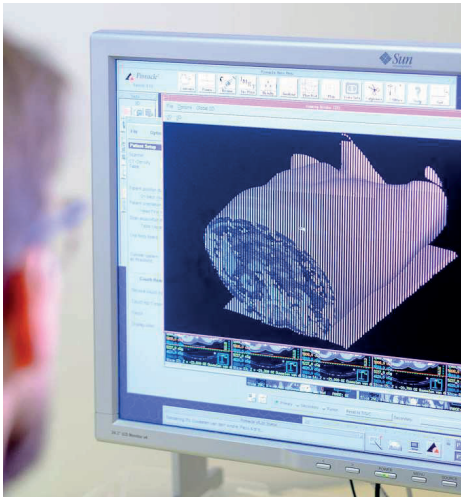

Image courtesy of Netherlands Cancer Institute

"With AI models and hardware from two or three years ago, we'd have to down-step the resolution to free up memory for training. Now, with the power of current GPUs, we can develop end-to-end systems that reconstruct high-resolution volumes directly from the projection data."

Nikita Moriakov
Post Doctoral Researcher, NKI

RESULTS STATEMENT

Using eight NVIDIA RTX 8000 graphics cards available on premise, NKI researchers were able to train a low-resolution model that enabled CBCT reconstruction at clinically acceptable 2 mm isotropic resolution. This resolution is sufficient for radiotherapy purposes, but it is behind the 1 mm resolution given by a dedicated CT scan.

By moving up to two NVIDIA A100 (80GB) graphics cards, the research team was able to fine-tune the low-resolution model and reconstruct full diagnostic images at 1 mm resolution. The high-resolution model was initialized from the low-resolution version, and the process of fine-tuning took approximately one month.

The inference speed at 1 mm resolution was also greatly improved by moving to the A100: A full volume at 1 mm resolution took just five minutes to reconstruct on a single GPU, compared to 14 minutes on the bare metal RTX 8000 machine.

WHAT THIS MEANS FOR PATIENTS

The ability to achieve native resolution in near-real time from a CBCT scan makes it possible to provide more accurate treatments to patients, based on their current anatomy when they come in for an appointment.

With a better image, smaller tumors are more likely to be localized reliably on the day of treatment, taking into account displacement of lesions due to weight fluctuations or recent food consumption, and the prescribed radiation dosage can be administered more precisely. In the future, it may be possible to increase treatment accuracy even more by improving tracking of the tumor motion, caused by body movements during respiratory cycles.

To learn more about NVIDIA AI Enterprise, visit:
www.pny.com/nvidia-ai-enterprise