

Smart Devices, Smart Manufacturing: Pegatron Taps AI, Digital Twins

One of the world's largest manufacturers of smartphones, notebooks and more created an AI platform based on NVIDIA technologies. Now, it's simulating operations in the metaverse with Omniverse.

Author: Rick Merritt

In the fast-paced field of making the world's tech devices, Pegatron Corp. initially harnessed AI to gain an edge. Now, it's on the cusp of creating digital twins to further streamline its efficiency.

Whether or not they're familiar with the name, most people have probably used smartphones, tablets, Wi-Fi routers or other products that Taiwan-based Pegatron makes in nearly a dozen factories across seven countries. Last year, it made more than 10 million notebook computers.

Andrew Hsiao, associate vice president of Pegatron's software R&D; division, is leading the company's move into machine learning and the 3D internet known as the metaverse .

"We've been collecting factory data since 2012 to find patterns and insights that enhance operations," said Hsiao, a veteran tech manager who's been with the company for 14 years, since it spun out of ASUS, one of the world's largest PC makers.

In 2016, Pegatron's COO, Denese Yao, launched a task force to apply new technology to improve operations. Hsiao's team of AI experts collaborated with factory workers to find use cases for AI. One of their first pilot projects used deep learning to detect anomalies in products as they came down the line.

It got solid results using modified versions of neural network models like ResNet, so they stepped on the gas.

Today, Pegatron uses Cambrian, an AI platform it built for automated inspection, deployed in most of its factories. It maintains hundreds of AI models, trained and running in production on NVIDIA GPUs.

The new platform catches up to 60% more defects with 30% fewer variations than human inspectors, and factory employees appreciate it.

"Manual inspection is a boring, repetitive job, so it's not surprising employees don't like it," he said. "Now, we're seeing employees motivated to learn about the new technology, so it's empowering people to do more value-added work."

The system may also improve throughput as factories adjust workflows on assembly and packing stations to account for faster inspection lines.

Pegatron's system uses NVIDIA A100 Tensor Core GPUs to deploy AI models up to 50x faster than when it trained them on workstations, cutting weeks of work down to a few hours.

"With our unified platform based on DGX, we have our data lake, datasets and training all in one place, so we can deploy a model in one click," Hsiao said.

Using the Multi-Instance GPU capability in A100 GPUs, Pegatron cut developers' wait time for access to an accelerator from nearly an hour to 30 seconds. "That lets us dynamically schedule jobs like AI inference and lightweight model training," he said.

As part of its AI inference work, the system analyzes more than 10 million images a day using NVIDIA A40 and other GPUs.

Pegatron uses NVIDIA Triton Inference Server , open-source software that helps deploy, run and scale AI models across all types of processors, and frameworks. It works hand-in-hand with NVIDIA

TensorRT , software that simplifies neural networks to reduce latency.

“Triton and TensorRT make it easy to serve multiple clients and convert jobs to the most cost-effective precision levels,” he said.

Hsiao’s team optimizes pretrained AI models it downloads in integrated Kubernetes containers from the NVIDIA NGC hub for GPU-optimized software.

“NGC is very helpful because we get with one click the deep learning frameworks and all the other software components we need, stuff that used to take us a lot of time to pull together,” he said.

Taking another step in smarter manufacturing, Pegatron is piloting NVIDIA Omniverse , a platform for developing digital twins

It has two use cases so far. First, testing Omniverse Replicator to generate synthetic data of what products coming down the inspection line might look like under different lighting conditions or orientations. This information will make its perception models smarter.

Second, it’s creating digital twins of inspection machines. That lets remote workers manage them remotely, have better insight into predictive maintenance and simulate software updates before deploying them to a physical machine.

“Today, when a system goes down, we can only check logs that might be incomplete, but with Omniverse, we can replay what happened to understand how to fix it, or, run simulations to predict how it will behave in the future,” he said.

What’s more, industrial engineers who care about throughput, automation engineers responsible for downtime, and equipment engineers who handle maintenance can work together on the same virtual system at the same time, even when logging in from different countries.

If all goes well, Pegatron could have Omniverse available on its inspection machines before the end of the year.

Meanwhile, Hsiao is looking for partners who can help build virtual versions of a whole production line in Omniverse. Longer term, his vision is to create a digital twin of an entire factory.

“In my opinion, the greatest impact will come from building a full virtual factory so we can try out things like new ways to route products through the plant,” he said. “When you just build it out without a simulation first, your mistakes are very costly.”

Original URL: <https://blogs.nvidia.com/blog/2022/08/16/pegatron-ai-omniverse/>