Digital Twin Journey: Computer Vision AI Model Enhancement with Dell Technologies Solutions & NVIDIA Omniverse

H20019

Abstract

This paper provides practical guidance on starting a Digital Twins AI development journey. It leverages Dell Technologies' Computer Vision and AI solutions expertise, using Dell PowerEdge servers and NVIDIA Omniverse Enterprise as foundational building blocks.

Dell Technologies Solutions

April 2024

Notes, cautions, and warnings

NOTE: A NOTE indicates important information that helps you make better use of your product.

CAUTION: A CAUTION indicates either potential damage to hardware or loss of data and tells you how to avoid the problem.

WARNING: A WARNING indicates a potential for property damage, personal injury, or death.

Contents

1

Introduction

Topics:

White paper focus

Digital Twin is a rapidly evolving, complex arena in AI, and requires careful planning. This paper provides practical guidance and recommendations on how to begin a 3D Digital Twin development journey.

Dell Technologies is building on our rich tradition and experience of providing validated designs with enterprise-grade Computer Vision (CV) and AI solutions. In this validated design, we apply Dell PowerEdge R760xa servers and NVIDIA Omniverse Enterprise suite as the foundational building blocks to demonstrate how a 3D Digital Twin can be practically applied to improve the object detection capabilities of CV solutions.

The first three chapters of this white paper are intended for a broad Digital Twin audience, including those with both business and technical interests.

The final three chapters are relevant to technical staff and managers who are currently working with 3D Digital Twin technologies; we detail several practical implementation considerations when working with 3D Digital Twin workflows.

Digital Twin

A Digital Twin is a true-to-reality, physically accurate virtual representation of a real-world physical object, process, or environment. This representation is ideally updated in real time.

Digital Twins have been applied within various domains such as:

Manufacturing and assembly	Factory production line workflows accurately simulated by virtual models.
Transportation	Smart City virtual simulations of transportation assets that are linked to real-time telemetry from physical vehicles.
Construction industry	CAD drawings and plans accurately mirror the real-time progress of physical build.
Aerospace industry	Flight telemetry data mirrored to aviation simulators.

While the concept of Digital Twins has existed for more than 20 years, the domain and scope of Digital Twin implementation continues to evolve. This has resulted in a dynamic Digital Twin concept, which can often cause confusion.

Digital Twin Consortium

Digital Twin Consortium is respected for their Digital Twin expertise. It brings together industry, government, and academia to promote consistency in Digital Twin technology, particularly regarding vocabulary, architecture, security, and interoperability. The goal of the consortium is to advance the use of Digital Twin technology across all industries. Dell Technologies is a committee member of the Digital Twin Consortium.

3D Digital Twins

The advent of a 3D graphic paradigm has introduced a new class of Digital Twin simulations based on true-to-reality 3D asset attributes such as physical movement, material, light, and rendering.

These Digital Twins can be `AI-enabled` and/or `AI enabling`, enhancing the capabilities of real-world solutions based on decisions and actions accurately modeled within virtual 3D worlds.

While data requirements can vary by use case, a significant amount of the data can be 3D in nature. Digital Twin datasets will most likely be constituted from a blend of datatypes, such as architecture, product, process, real-world telemetry, and geospatial. These datasets will also likely be sourced from several sources, such as CAD, CAE, BIM, MES, PLM, GIS, IoT, and edge computing.

With the integration of technologies such as edge computing, industrial AI, and 3D Digital Twin, the nascent concept of the Industrial Metaverse (IM) is beginning to emerge. IM solutions can optimize processes and drive sustainable practices, helping to shape the future beyond simulations.

2

Business considerations

Topics:

- Business benefits
- Business challenges

Business benefits

Digital Twin solutions offer significant benefits for businesses, such as:

Increased efficiency and productivity	Deliver greater agility and optimize asset performance, enhancing business decisions and increasing efficiency. For example, spot mistakes before they happen in the real world.
Cost savings	Save money by modeling ways to reduce waste, save energy, and avoid downtime.
Improved safety and risk management	Simulate asset and/or human behavior in different scenarios, in a safe and cost-effective manner that might not be feasible in the real world.
Better decision making and problem solving	Study and adjust these virtual models to simulate different scenarios, helping businesses make better decisions by providing insights and new data.
Predictive maintenance	Enable businesses to identify and prevent potential equipment problems. Anomalies can be detected earlier or ideally avoided before occurring in real-world assets, reducing downtimes and extending equipment uptimes.

Business challenges

While Digital Twin presents many benefits for businesses, their implementation can also bring significant challenges.

Businesses should keep several key considerations in mind when adopting Digital Twins:

Data privacy and security	Generation of vast amounts of data, potentially containing sensitive and confidential information relating to business operations, customers, or products. Companies must design their Digital Twin solutions with data privacy and security in mind.
Cost of implementation	Implementing a Digital Twin can be a significant investment for businesses. Regardless of the particular use case or domain, the implementation will likely require a multidiscipline team spanning many functions (such as information technology, engineering, data science, and business).
Organizational alignment	Developing and maintaining Digital Twins requires specialized expertise and knowledge, such as process engineering, mechanical engineering, planners and managers, software engineering, data integration, data analytics, simulation modeling, and 3D artists and modelers.
Integration with existing systems	Digital Twins are typically implemented to mirror existing real-world use cases where existing systems and infrastructure are already in place. The integration of these data pipelines is a critical requirement for an effective Digital Twin solution.

3

Solution overview

Before outlining the solution overview, we need to establish some key Digital Twin concepts and terms.

Topics:

Characteristics of a Digital Twin

- A Digital Twin is a virtual, true-to-life replica of a specific real-world asset.
- Digital Twins enable timely, optimized, real-world outcomes by using synchronized virtualized models.
- The Digital Twin paradigm can be applied to most real-world scenarios once the relevant elements are available.
- There is no "one size fits all" approach when it comes to Digital Twins.
- Greenfield implementations of Digital Twins are rare and are nearly always integrated into existing solutions (processes, workflows, software, infrastructure).
- The implementation and integration challenges are heavily dependent on the context and domain of the twinned physical real-world assets.
- For new Digital Twin initiatives, best practice is to start small, with a building block approach for each step of the Digital Twin workflow.

Foundational building blocks of a Digital Twin

Users often make pragmatic initial design decisions when developing their Digital Twin solutions. These instances might not initially contain all elements to realize the full benefits of or meet the strict definition of a Digital Twin in itself. In the absence of a single implementation approach, unfortunately the discussion around Digital Twins can be clouded in confusion.

What is clear however is that undertaking a Digital Twin solution requires several core components.

Core Digital Twin components

These core components form the nucleus around which a Digital Twin is formed.

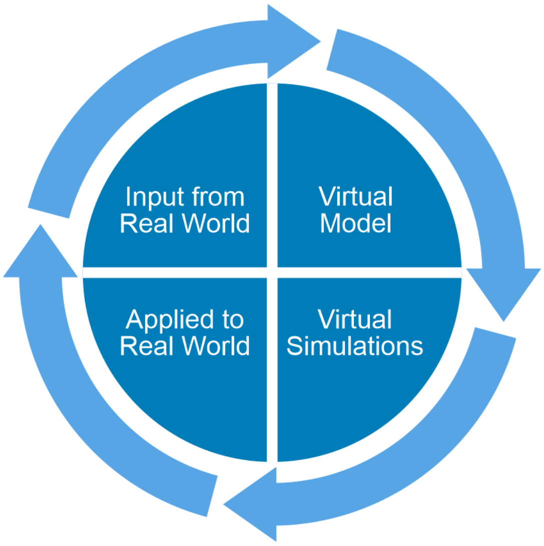


Figure 1. Core components of a Digital Twin workflow

- Input from Real-world - data ingestion
- Virtual model- software model of the real-world
- Virtual simulations- Ability to simulate within a virtual world
- Applied to real-world- Feedback mechanism from virtual to real-world

Real-world data input

At the heart of all Digital Twin implementations is data. Real-world objects, processes, or environments all have their own data characteristics and challenges, such as availability, accessibility, and formats.

A Digital Twin solution should be able to ingest (import and convert) data from particular domains such as:

- 3D Graphics Systems, for example Universal Scene Description (USD) format
- Computer Vision, for example image and video formats (png, RTSP)
- Computer Aided Design (CAD), for example 2D-3D drawings (DWG, DXF)
- Edge and robotics, for example LiDAR sensor E57 format (.bag)
- Geographic Information Systems (G.I.S), for example GeoJason format

This data can then be leveraged to form the basis of virtualized digitized assets. The ability to easily ingest a myriad of data formats from multiple different sources with live synchronization is a critical consideration when developing Digital Twin solutions.

Virtual model of the real world

Building digitized virtual models involves data transfers, data conversions, and data aggregations into a central environment before these digitized assets can be presented to a modeling framework. Subsequent changes in the real-world data sources need to be constantly synchronized and mirrored to these virtualized models.

The ability to manage a common data format, unified asset pipeline, and central collaboration are key capabilities of Digital Twin solutions. These capabilities enable easy visualizations and interactions with digitized assets within virtual models.

Simulations within the virtual world

Simulation offers the possibility for faster innovation, rare scenario modeling, and higher-quality outcomes. Anomalies can be detected earlier or ideally avoided before they occur in real-world instances.

The Digital Twin can become a single source of truth in which scenario modeling can take place after the following components are in place:

- Data Pipelines
- Assets ingestion
- Data aggregation into digitized models
- Models synchronization

The ability to easily collaborate and run simulations is therefore a critical consideration when developing Digital Twin solutions.

Feedback mechanism from virtual to Real-World

The final core component of a Digital Twin is the ability to leverage what you have modeled and learned from the twinned virtual assets to positively impact real-world outcomes. Hence, the feedback mechanism from the virtual to the real world is critical for any instance of Digital Twin.

Real-world objects, processes, or environments will all face feedback-related challenges such as feasibility, synchronization, and safety. It might not be prudent, practical, or desirable to build real-time or automated feedback into a Digital Twin workflow, and this should be considered on a case-by-case basis. For example, automated action in a life-threatening scenario requires careful consideration.

Starting point for a Digital Twin journey

Practical Digital Twin solutions are often cast in the context of an evolving journey, initially starting with small scale building blocks interconnecting adjacent elements before full end-to-end pipeline integration is achieved.

It might be natural to assume that all Digital Twin journeys start when all the core components are in place, or that they begin with data ingestion and follow a standard flow, but in practice this may not be the case. Businesses will likely already have existing objects, processes, environments, and data in place, so they might choose to embark upon their Digital Twin journey where it makes practical sense within their existing data workflows.

Solution

This paper explores the practical implications and considerations for developing Digital Twin solutions using DellPowerEdgeR760xa servers and NVIDIA Omniverse Enterprise suite as the foundational building blocks.

NOTE: We leverage a CV object detection use case as an illustration instrument for this paper.

Computer Vision

Computer Vision (CV) is a field of Artificial Intelligence (AI) that provides computers and systems with the ability to extract relevant information from digital images, videos, and other digital sensor inputs.

Object detection is a specific application of CV AI technology using algorithms to detect, identify, and track objects in images and videos. It involves finding the coordinates of an object within an image and video, then updating the scene with the specific object enclosed with a bounding box. Typically, CV object detection models are trained on datasets containing sufficient examples of the required object(s) including object labels and coordinates information.

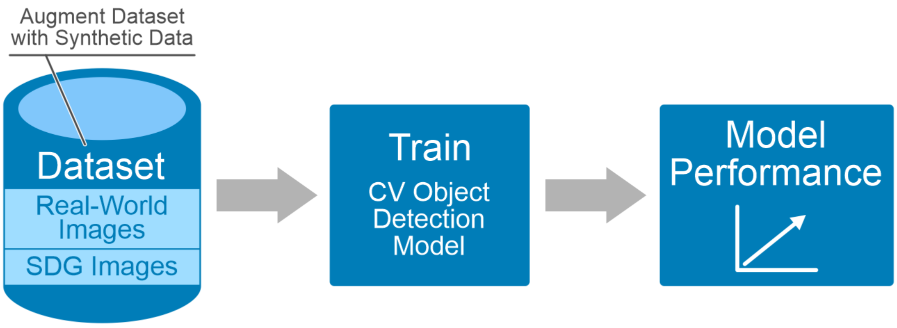


Figure 2. Example Computer Vision workflow utilizing synthetic data

This solution leverages 3D modeling, simulation, and synthetic data generation (SDG) technologies to enhance the performance capabilities of a typical CV object detection model.

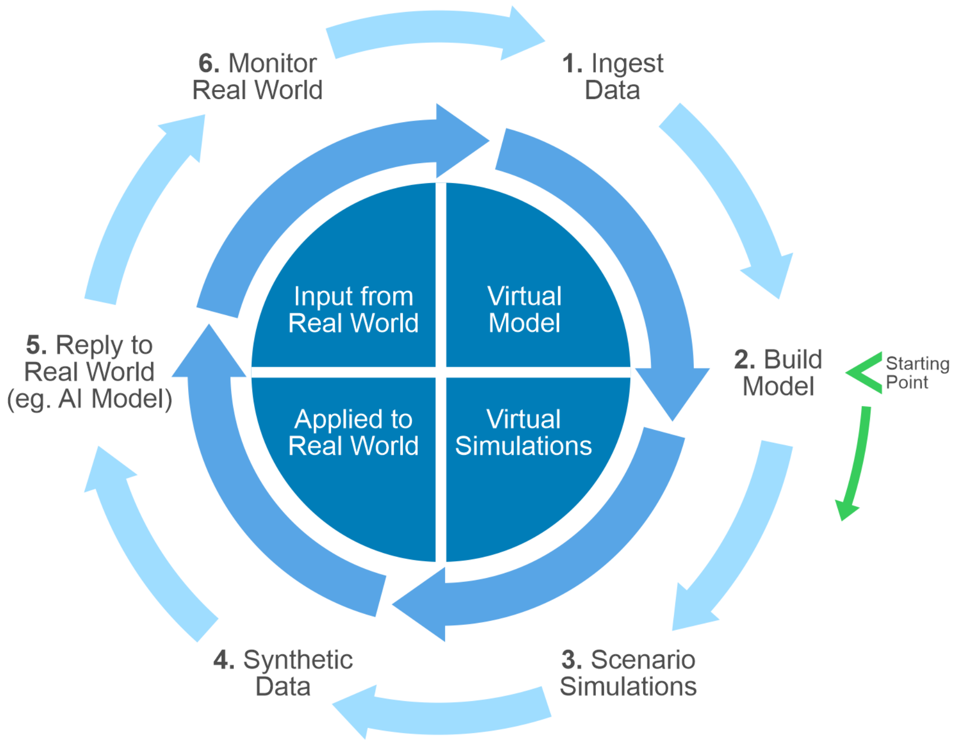


Figure 3. Applied Computer Vision Digital Twin workflow

This paper focuses on the following Digital Twin CV workflow categories (2, 3, 4, and 5 in the previous figure:

2. Build Model
3. Scenario Simulations
4. Synthetic Data Generation
5. Apply to the Real World

3D Digital Twin model/simulation framework

To create, manage, manipulate, collaborate, and simulate Digital Twin assets, a suitable modeling framework is required. In the context of the selected CV use case, the solutions modeling framework needs to provide capabilities such as:

- Accurately model 2D and 3D visual environments
- Create additional 2D and 3D assets
- Physically and visually simulate different scenarios such as different scene perspectives, domain conditions, location, lighting, color, texture, and background
- Generate synthetic data based on modeled scenarios, including metadata such as object labeling and location coordinates
- Integrate with Digital Twin workflows and data pipelines

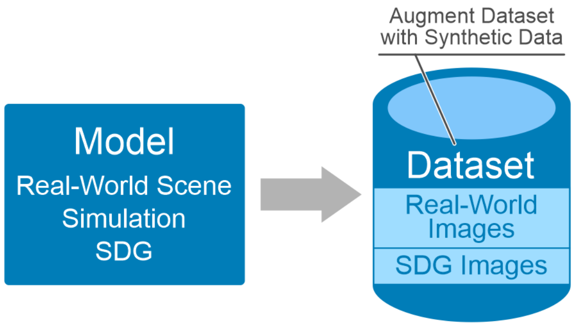


Figure 4. 3D simulation framework used to model suitable scene and corresponding data

This solution leverages 3D SimReady assets (visually and physically accurate digital building blocks) to model numerous scene scenarios and generate corresponding synthetic data, which is then used to enhance the training of a CV object detection model.

Synthetic data generation

Typically, training CV AI models requires labeled and diverse datasets which might contain countless elements (not only visual in nature). The collection, curation, and labeling of real-world data is inherently time-consuming, expensive, and might not be feasible. This can impede the development of CV models, slowing time to market/solution.

SDG can be used to augment existing datasets and/or bootstrap initial CV solution developments, resulting in significant time and cost reductions.

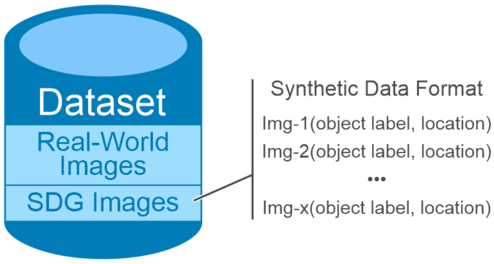


Figure 5. Synthetically generated annotated images

This solution leverages SD generated by previous virtual simulations, consisting of 2D and 3D assets and associated metadata (object labels and coordinates), which is then used to collect data to train a CV model.

Apply virtual model to real world

After sufficient training and validation cycles, CV AI models are typically published and released to production environments. New or existing CV solutions update their models and make subsequent inferences and predictions on real-world image and video inputs.

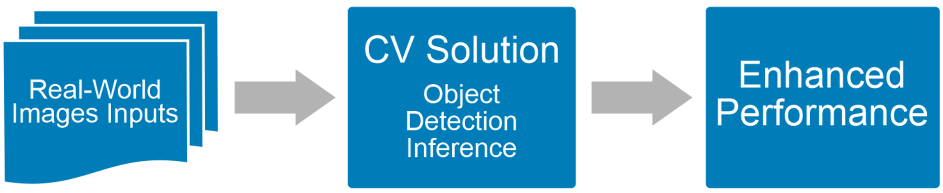


Figure 6. Real-world object detection validation

This solution takes a CV object detection model trained on synthetic data and applies it to real-world images.

4

Partner overview

Since its founding in 1993, NVIDIA Corporation has evolved from developing graphics processing units (GPUs) for gaming into a broad technology company with a diverse product portfolio in parallel computing, AI, and data center technology. NVIDIA has pivoted strongly towards AI, leveraging its GPU technology for deep learning and AI workloads. The company's CUDA programming model and GPUs are widely used in data centers for AI research and cloud computing and by enterprises for AI inference and training tasks.

NVIDIA has expanded its offerings beyond hardware and its proprietary CUDA development environment to include AI with deep learning libraries (like cuDNN), AI software platforms (like NVIDIA AI Enterprise), and 3D collaboration and simulation platforms (like Omniverse). These software offerings enhance the value of NVIDIA hardware by enabling developers to build and deploy applications on the NVIDIA computing platform efficiently.

NVIDIA's go-to-market model reflects its transition from originally focusing on graphics technology to becoming a central player in the broader field of parallel computing and AI. NVIDIA is positioned as a critical enabler of the growing AI market by leveraging its GPU technology across various industry sectors, including edge computing. This broad market approach and an emphasis on software and ecosystem development help explain NVIDIA's growth and influence in the AI technology sector.

Topics:

- Omniverse
- Omniverse Digital Twin Technology
- GPUs for Omniverse

Omniverse

NVIDIA Omniverse is a platform of APIs, SDKs, and services that enable developers to easily integrate Universal Scene Description (OpenUSD) and RTX rendering technologies into existing software tools and simulation workflows for building AI systems.

Applications built on Omniverse core technologies fundamentally transform complex 3D workflows, allowing individuals and teams to build unified tool and data pipelines and simulate large-scale, physically accurate virtual worlds for industrial and scientific use cases.

Omniverse enables creators, designers, researchers, and engineers to enhance workflows within a common 3D framework that enables seamless collaboration and interoperability across applications.

Omniverse attributes

Omniverse is aligned to four key attributes, designed to enhance the interoperability of 3D workflows.

Table 1. Omniverse attributes

Attributes	Description
Scalable Visualization & Simulation	Create physically accurate visualizations and simulations of products and environments, reducing the time and costs of physical testing.
Easy-to-Use Suite of Developer Tools	Quickly create and deploy custom workflows and apps without extensive programming knowledge. Use sample reference applications to get started and focus on design and building.
Ecosystem of Application Building Blocks	Connect to a rich catalog of industry-leading third-party extensions to integrate your tools seamlessly with proven applications and connectors.

Table 1. Omniverse attributes (continued)

Attributes	Description
Portal to AI	Embed AI into your tools and applications to automate repetitive tasks for users with predictive capabilities and natural language processing.

See NVIDIA Omniverse Enterprise for more information.

Omniverse platform

Omniverse platform consists of several foundational elements including Connect, Nucleus, Kit, Simulation, and RTX Renderer.

Figure 7. Omniverse foundational components

Table 2. Omniverse platform components

Component	Description
Connect	The libraries that connect popular content creation tools to Omniverse. With Connectors, you can continue to work in your favorite software applications while benefitting from other Omniverse tooling.
Nucleus	The central database and collaboration engine of Omniverse. With Nucleus, users share and modify representations of virtual worlds.
Kit	The toolkit for building native Omniverse Apps, Extensions, and Microservices.
Simulation	A powerful suite of tools and SDKs that simulate a physically accurate world.
Renderer	An advanced, multi-GPU renderer based on NVIDIA RTX™ that powers both real-time ray tracing and referenced path tracing.

See NVIDIA Omniverse Platform Elements for more information.

Omniverse foundation app templates

Omniverse platform also provides several foundational application templates such as:

Table 3. Omniverse sample application templates

Foundation Template	Description
USD Composer	A foundation application for developing and composing USD scenes.
USD Explorer	A foundation application for exploring and reviewing USD scenes.
Omniverse Code	IDE for building Omniverse extensions and applications.
Isaac Sim	Simulation toolkit to design, test, train, and simulate AI-based CV application and robots in photorealistic and physically accurate virtual environments.
Audio2Face	Audio-driven foundation application for face to lip sync.

Omniverse foundation app templates are best-practice example implementations and configurations of Kit extensions. They can be used out of the box or as a generic template that developers and customers can customize, extend, and personalize according to their workflow.

Omniverse applications cover a broad range of use cases and vertical markets including:

- Industry, architecture, engineering, and construction
- Film and TV special effects
- Robotics and computer vision
- Game development

Omniverse Enterprise is architected with interoperability in mind. Based on Universal Scene Description (USD), the platform fundamentally transforms complex 3D workflows. You can easily connect your 3D product pipelines and unlock full design-fidelity visualization of your 3D datasets using purpose-built connector plugins for the most common 3D applications and datatypes. NVIDIA is continually adding Omniverse interoperability for third-party accelerated applications.

Once you have established your data pipelines and your 3D datasets are ingested and aggregated into Omniverse, your 3D models can become a single source of truth that can be easily shared with, collaborated on, and refined by key contributors and stakeholders in real time. Working this way streamlines decision-making and accelerates time to production.

By leveraging the flexibility, interoperability and collaboration capabilities NVIDIA Omniverse can facilitate the development and use of digital twins technologies and solutions.

Omniverse Digital Twin Technology

NVIDIA Omniverse Digital Twin technology is being used to solve the world's most significant challenges, ranging from factory scale to planetary scale and everything in between. Creating digital twins with NVIDIA Omniverse allows developers to create physically accurate virtual replicas of unique objects, processes, and environments, all constantly synchronized with real-world data inputs and powered by AI.

This technology enables companies to innovate, optimize, and troubleshoot in a virtual environment, thereby reducing costs, improving efficiency, and accelerating the development of new products and services. See Omniverse Digital Twin Extensions

At the heart of Digital Twin solutions is the ability to perform simulations. Omniverse offers several simulation technologies including:

- SimReady assets
- Isaac Sim and Replicator

SimReady assets

Simulation environments, such as Omniverse, can leverage additional model information and metadata to make that content far more useful for research and training of a particular product or ecosystem. In order to achieve this level of fidelity from both a visual and simulation perspective, NVIDIA is working to create a new 3D standard specification called SimReady. At a fundamental level, you can think of SimReady assets as building blocks for industrial virtual worlds. SimReady assets are supposed to be more than just visual 3D objects.

They are built on top of the USD platform, and they are designed to be modular to take advantage of the flexibility that the USD framework provides while including embedded physical properties, semantic labeling, behaviors, and connected data streams so that they are immediately useful within simulation environments.

Isaac Sim and Replicator

Isaac Sim makes the most of the Omniverse platform's powerful simulation technologies. These include advanced GPU-enabled physics simulation with NVIDIA PhysX 5, photo realism with real-time ray and path tracing, MDL material definition support for physically based rendering, and synthetic data generation.

Omniverse Replicator is a framework for developing custom synthetic data generation pipelines and services. Replicator is exposed as a set of extensions within Isaac Sim and Omniverse Code. Replicator is designed to easily integrate with existing pipelines using open-source standards like USD, PhysX, and MDL.

You can use Omniverse Replicator to easily generate diverse datasets by varying many parameters such as types of defects, locations, ambient lighting, SimReady objects and assets and more to speed up training and iteration of the AI model development. Replicator has the ability to annotate these synthetic images with precisely labeled metadata and to produce specific data formats required for various AI frameworks.

Developers can customize these Omniverse Tools to:

- Vary scene generation (scenario modeling)
- Generate annotated and labeled data
- Write data in various formats

These tools enable flexible integration with existing AI workflows.

GPUs for Omniverse

The Omniverse platform is optimized for NVIDIA RTX-powered professional workstations, mobile workstations, and NVIDIA Certified Systems. NVIDIA Certified Systems enable enterprises to confidently deploy hardware solutions that securely and optimally run their modern accelerated workloads.

For virtualized deployments, the NVIDIA L4, NVIDIA L40 , or NVIDIA L40S, Ada™ Architecture GPUs with 20GB or 48GB of graphics memory are the best options for unprecedented graphics, rendering, and AI performance.

NVIDIA vGPU

NVIDIA vGPU software empowers virtualized environments by providing GPU performance to VMs, enhancing productivity, and enabling a wide range of workloads across different industries. It allows multiple VMs to have simultaneous, direct access to a single physical GPU.

Key benefits include:

- Bare metal performance: Achieve performance virtually indistinguishable from a bare metal environment.
- Management and monitoring: Leverage common data center management tools for tasks like live migration.
- Optimal resource utilization: Provision GPU resources with fractional or multi-GPU VM instances.
- Improved business continuity: Respond to changing business requirements and remote teams.

vGPU Editions:

- NVIDIA RTX Virtual Workstation (vWS): Designed for creative and technical professionals.
- NVIDIA Virtual PC (vPC): Ideal for knowledge workers using office productivity applications.
- NVIDIA Virtual Applications (vApps): Supports application streaming with RDSH solutions

Transfer Learning ToolKit for AI Optimization (TAO)

Transfer learning is a powerful technique that instantly transfers learned features from an existing neural network model to a new customized one. The open-source NVIDIA TAO Toolkit, built on TensorFlow and PyTorch, uses the power of transfer learning while simultaneously simplifying the model training process and optimizing the model for inference throughput on practically any platform. The result is an ultra-streamlined workflow. You can adapt your own models or pre-trained models to your own real or synthetic data, and then optimize them for inference throughput, all without needing AI expertise or large training datasets.

5

Solution architecture overview

This assessment used a Dell PowerEdge 760xa server and the NVIDIA Omniverse Enterprise Platform to examine the flexibility and performance requirements for enhancing CV workflows within the context of a Digital Twin development journey.

Omniverse Enterprise Platform was deployed as a virtualized instance enabling a flexible infrastructure configuration that can be tailored to individual requirements, such as splitting physical GPUs resources into vGPU partitions. This flexibility can prove immensely beneficial when Digital Twin or AI workload needs are likely to change during development.

The Dell PowerEdge 760xa is positioned to meet the diverse needs of Digital Twin requirements such as 3D modeling, physics simulations, image rendering, computer vision, robotics, edge computing, and AI training and inferencing.

The functionality and validation discussed below was tested on a nonproduction CV system configuration and deployed in a Dell engineering lab.

See Dell Technologies Validated Designs hub for a full list of Computer Vision platforms.

Topics:

Physical architecture

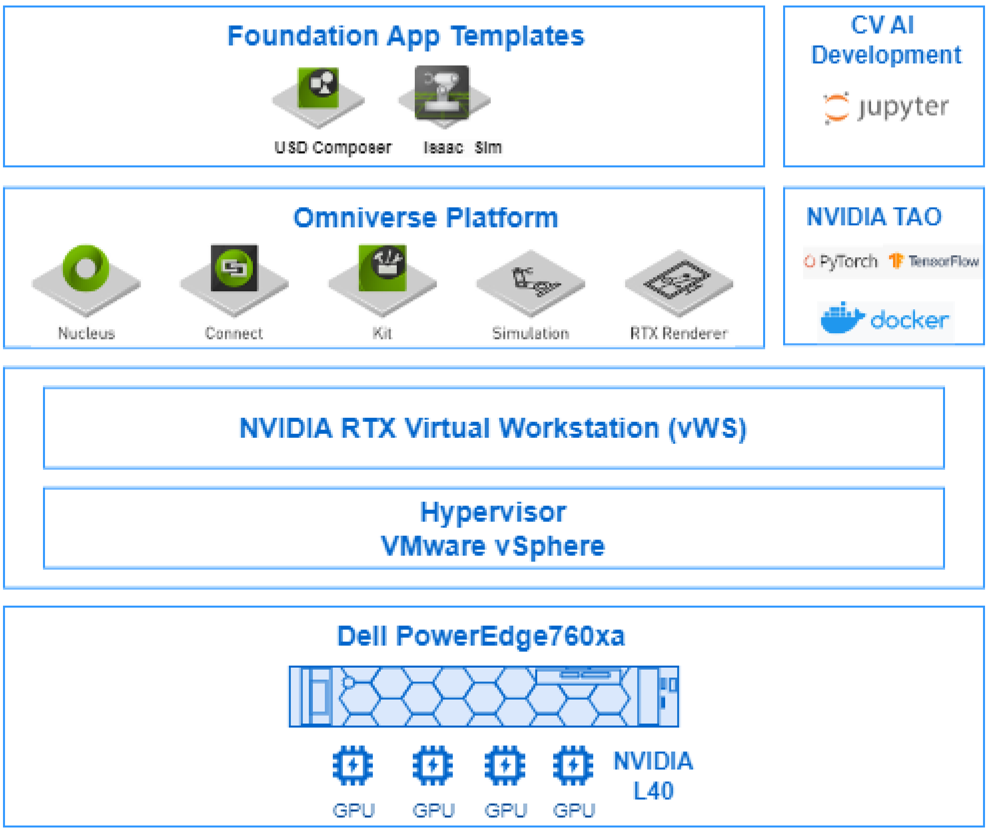


Figure 8. Digital Twin foundation Architecture for Computer Vision Workloads

In the above diagram, we deployed a single Dell Power Edge 760xa node with 4x NVIDIA L40 GPUs that were configured and managed by VMware vSphere suite. Physical GPU resources were partitioned with NVIDIA vGPU software and assigned to VMs that were configured to provide various functionalities, such as 3D modeling framework, 3D simulations, SDG, and CV AI model development such as NVIDIA Omniverse and NVIDIA TAO.

Computing

Dell PowerEdge 760xa

The solution was validated on a Dell PowerEdge 760xa and managed with an ESXi cluster.

The following table lists the hardware specifications for our test setup:

Table 4. Case setup: hardware specification

Component	Details
CPU	2x Intel(R) Xeon(R) Gold 6438M CPU @ 2.20 GHz (128 vCPU)
Memory	512 GB (16x 32 GB DDR-5 ECC)
Storage	Local Raw 6,400 TB (8x 800 GB SSD SAS 12Gbps)
GPU	4x L40 (FP32 TF 90) Memory: 48 GB GDDR6 w/EEC Media Engines: 3 Video Encoder 3 Video Decoder 4 JPG Decoder Power: 300 Watts

Applications and Virtualized Infrastructure

This solution leverages NVIDIA vGPU technology, NVIDIA Omniverse Enterprise Platform, NVIDIA TAO, within a virtualized VMware deployment.

The following table lists the solutions software and versions that we tested:

Table 5. Case setup: software specifications

Component	Details
Hypervisor	VMware ESXi 8.0 U2 Build-22380479(A04)
GPU	vGPU Driver Grid 16.3 535.154.02, vGPU Profiles "nvidia_l40-48q"
VM Operating Systems	Windows 10 Enterprise 10.0.19045, Ubuntu 20.04.6 LTS
Omniverse Platform	Launcher 1.9.10, Nucleus 2023.1.0
Omniverse Apps	USD Composer 2023.2.3, Isaac SIM 2023.1.1
NVIDIA TAO	TAO Tool Kit 5.1.0

6

Use case validation

The use cases validated during this testing focused on three main areas:

- 3D Simulations
- Synthetic data generation
- Computer Vision AI model training on synthetic data

NOTE: A single L40 GPU resource with an l40-48q vGPU profile was assigned to each virtual machine for each use case validation.

Topics:

3D simulations

Our 3D simulations investigated how to generate a physically accurate 3D scene and how to perform realistic scenario simulations with Omniverse USD Composer. We broke the test down into the following subscenarios:

- Simple scene interaction with SimReady Assets
- Complex Scene interaction with several SimReady Assets

Simple scene interaction with SimReady assets

Figure 9. Default Omniverse stage containing two SimReady assets

The previous figure shows the Default Omniverse stage containing two SimReady assets: a pallet and cardboard boxes.

Apart from the ability to generally interact with a 3D scene and change its parameters (such as perspective, lighting, and materials), the addition of SimReady assets enables the capability to interact with these assets in a physically accurate manner.

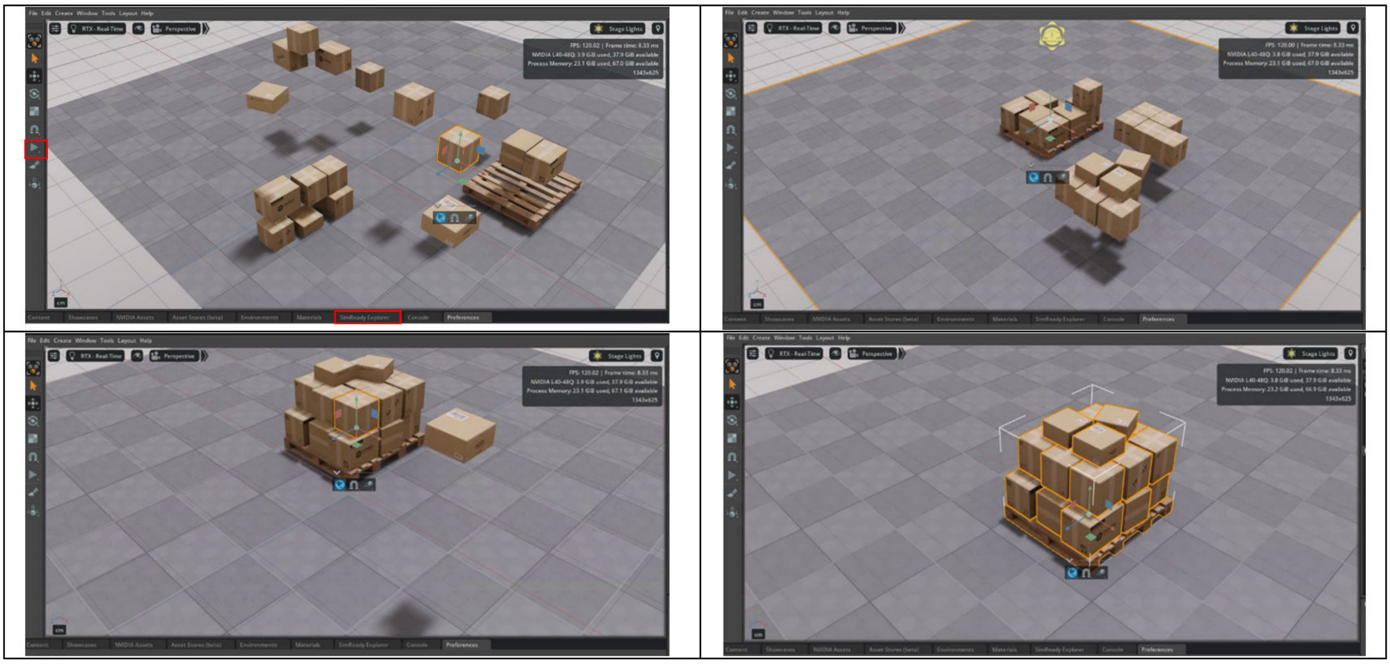


Figure 10. Default Omniverse stage containing two SimReady assets

The **Play** button enables simulation capabilities within the scene.

See Omniverse Isaac warehouse tutorial.

Results

In this example, the user can model a range of simple GUI interactions with both SimReady assets (pallet or cardboard box) within the scene, such as picking up and moving boxes and modeling drop box scenarios.

- The ability to select wanted asset(s) and model their behavior and or interactions.
- Average CPU and GPU utilization never rose above 5% during this simple scenario interaction.

Findings

Users can manually interact within the stage and scene or automate these interactions. Omniverse Enterprise suite contains various SimReady categories with numerous assets. They are available from various locations, including Omniverse USD Composer on the SimReady Explorer tab. Additional categories and assets are available within Omniverse Nucleus.

Complex scene interaction with Several SimReady assets

Complex warehouse scene with physics enabled animation workflow for racks, boxes, shelves, and hanging lamps.

Enable the **PhysX FlatCache** Omniverse extension before loading the warehouse scene. *Recommended when working with a large number of rigid body assets.*

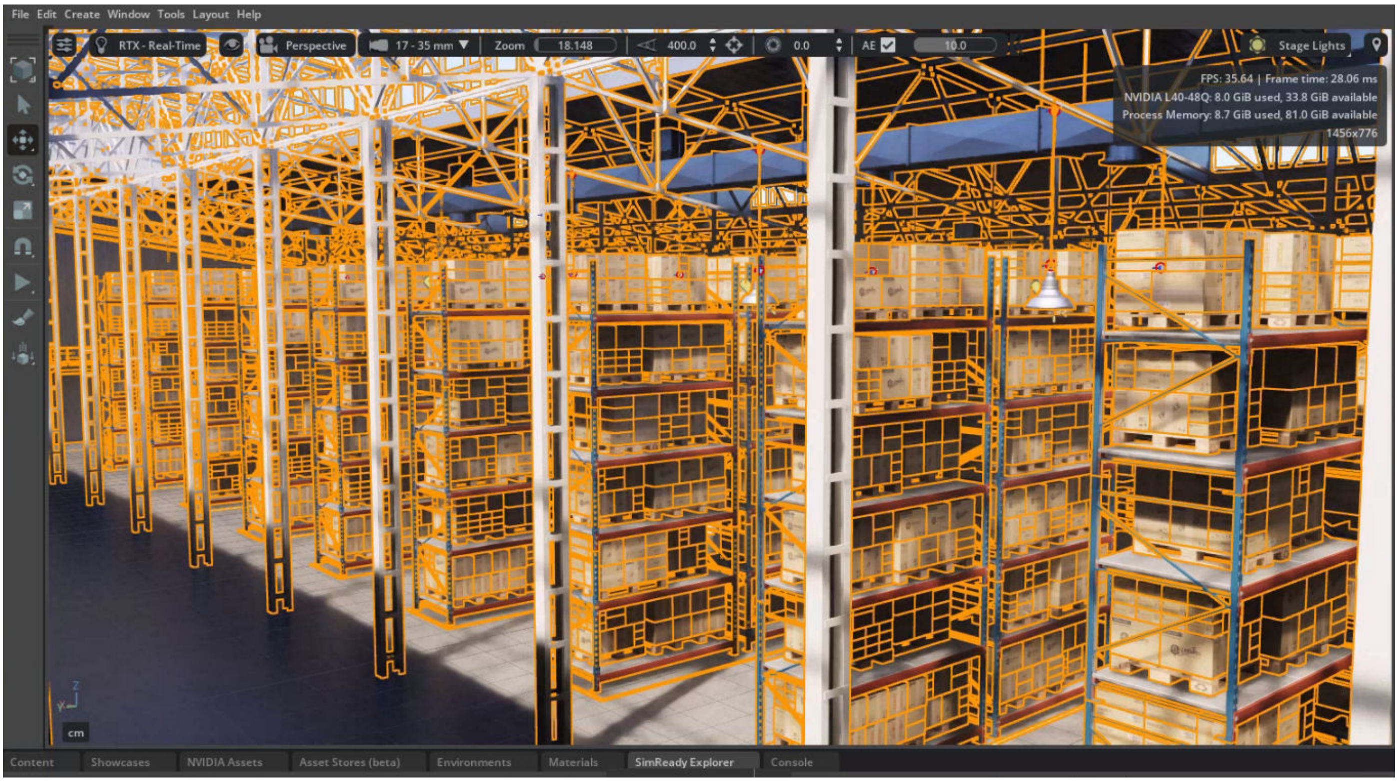


Figure 11. Omniverse scene set within a warehouse environment

The previous figure shows the Omniverse scene set within a warehouse environment containing multiple assets both static and SimReady including pallets, Racks, Wooden Crate, Shelf, Metal Fencing, Lamps, Marking Lines, Cardboard box.

Apart from the ability to model different physics behaviors, SimReady parameters can also be modified to allow more granular control. We experimented with the parameter properties of Linear Velocity and Angular Velocity to change the momentum of the cubes that trigger the initial collisions within the scene or stage.

Figure 12. Result of physics collision simulations set within a warehouse layout

See NVIDIA Warehouse Physics Workflow tutorial , which details how to configure the above scene to model smart collision for rigid body dynamics.

Results

To move objects around the scene while the scene or stage is playing back, use Shift + Left-Click+Mouse Drag. You can use this to pull individual or multiple assets around during playback. To stop simulation and reset the scene or stage, press the Stop button.

Findings

- Average CPU and GPU utilization of approximately 50% and approximately 35% respectively during this complex multi-asset, PhysX-enabled simulation.
- The duration of initially loading the scene was reduced by enabling and configuring suitable Omniverse cache settings. See Omniverse cache options for Workstation and Enterprise configurations.
- The performance of the Omniverse USD Composer VM for this PhysX-enabled scene was noticeable. NVIDIA suggests separating rendering and PhysX tasks to dedicated GPU resources if available. However, this PhysX configuration was not available using the NVIDIA control panel within the virtualized Windows instance.

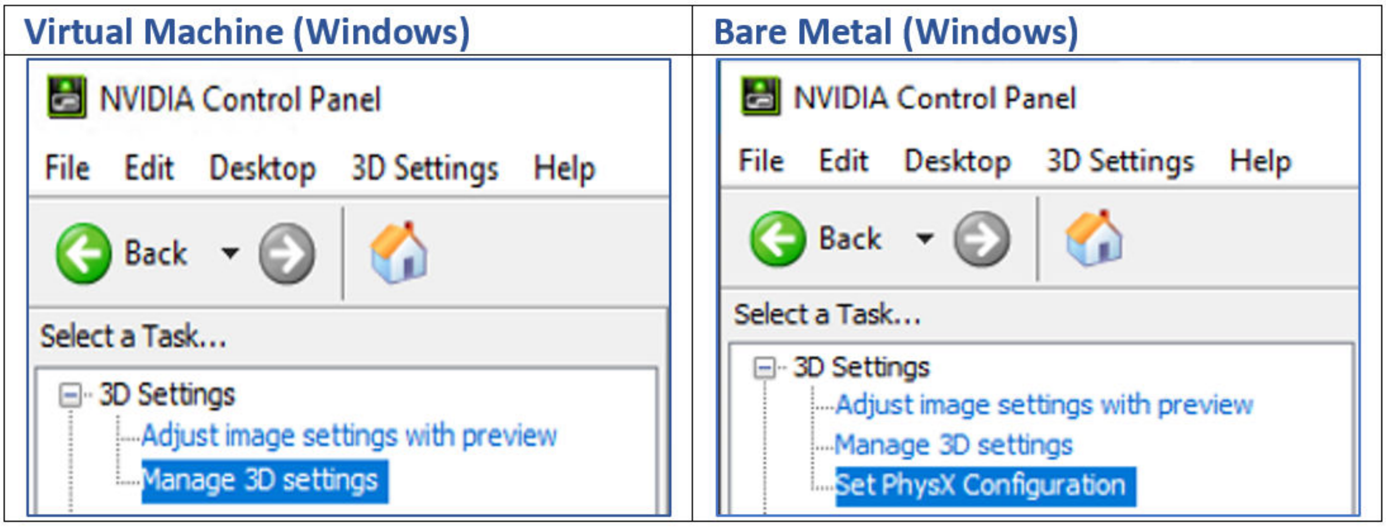


Figure 13. NVIDIA control panel properties exposed to Bare metal and virtualized Windows instance

Synthetic data generation

We investigated how to generate synthetic 3D data with NVIDIA Omniverse Isaac Sim. This test explored how to generate suitable synthetic data that is required to train a CV AI model to detect *"Pallet Jack'"* objects within a warehouse (logistics) setting.

The Isaac Sim environment configurations are based on NVIDIA-Omniverse synthetic-data repository examples . For strategies to improve the quality of synthetic data such as domain randomization (texture and lighting changes, scene distractors), see How to Train Autonomous Mobile Robots to Detect Warehouse Pallet Jacks Using Synthetic Data documentation.

High-level simulation and synthetic data generation steps:

- Load warehouse scene
- Load SimReady assets
- Simulate domain randomization
- Generate data

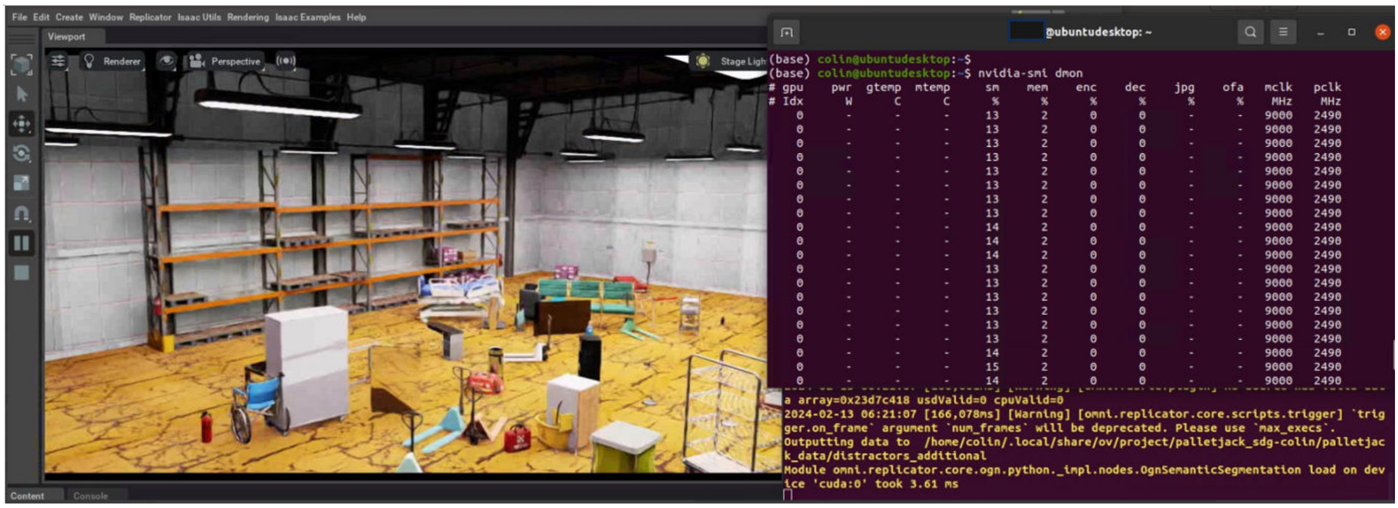


Figure 14. Warehouse scene containing several instances of pallet jacks and other assets

Results

Several Python scripts are provided to configure both Omniverse Isaac Sim and Replicator to automatically generate scene simulations and associated metadata (labeling and annotations) required for subsequent AI model training.

- Running these scripts generates approximately 5,000 pallet jack images and associated metadata.
- The associated metadata is written in standard CV object detection KITTI format.
- The images generated vary based on the domain randomization configurations.

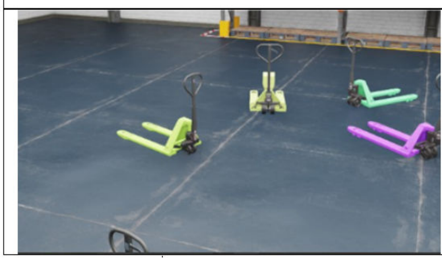
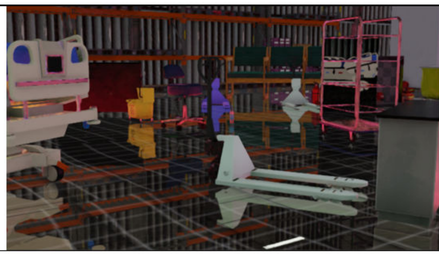
No Distractors *Images only contain Pallet Jacks objects*	**Distractors** *Images with numerous objects including Pallet Jacks*	**Distractors Additional** *Same as Distractors category but including challenging textures and lighting.*
		

Figure 15. Examples of pallet jack images based on domain randomization scene distractor techniques

Findings

- Generation of approximately 5,000 annotated images was seamless and took less than five minutes to complete.
- Average CPU and GPU utilization of approximately 7% and approximately 33% respectively during this SDG test.
- Setup of this predefined scene was slightly complex for first-time use by a novice.
- Developing a new scene from scratch would be a significant undertaking. Domain randomization would require a suitable heuristic approach that includes different scenes. Indoor and outdoor scenes might require additional considerations.

Developing a Computer Vision AI object detection model with synthetic data

This test uses the synthetic data generated in the previous test ("Pallet Jack" images in a logistics setting) to train a CV AI object detection model.

The AI development environment configurations are based on NVIDIA-Omniverse synthetic-data repository examples , which leverage the NVIDIA open-source Pre-Trained Tool Kit (TAO).

High-level training, validation, and real-world inference steps:

1. Convert synthetic data metadata records into a suitable format for an AI model.
   - Use an NVIDIA TAO detectnet-v2 container to convert KITTI to Tensorflow format.
   - Download a pre-trained CV AI model (ResNet-18 model).
2. Train the model on synthetic data.
3. Validate model on synthetic data.
4. Perform training and build validation model with synthetic data.

1. Set up TAO via Docker Container

- We will follow the pre-requisites section of instructions for using TAO toolkit. Make sure that the pre-requisite steps are completed (installing `docker`, `nvidia container toolkit` and `docker login nvcr.io`)
- The docker container being used for training will be pulled in the cells below, make sure you have completed the pre-requisite steps and `docker login nvcr.io` to allow pulling of the container from NGC

In [2]:

```
import os
%env DOCKER_REGISTRY=nvcr.io
%env DOCKER_NAME=nvidia/tao/tao-toolkit
%env DOCKER_TAG=4.0.0-tf1.15.5 ## for TensorFlow docker

%env DOCKER_CONTAINER=nvcr.io/nvidia/tao/tao-toolkit:4.0.0-tf1.15.5
```

```
env: DOCKER_REGISTRY=nvcr.io
env: DOCKER_NAME=nvidia/tao/tao-toolkit
env: DOCKER_TAG=4.0.0-tf1.15.5 ## for TensorFlow docker
env: DOCKER_CONTAINER=nvcr.io/nvidia/tao/tao-toolkit:4.0.0-tf1.15.5
```

2. Download Pretrained Model

- We will use the `detectnet_v2` Object Detection model with a `resnet18` backbone
- Make sure the `LOCAL_PROJECT_DIR` environment variable has the path of this cloned repository in the cell below

In [ ]:

```
# os.environ["LOCAL_PROJECT_DIR"] = "<LOCAL_PATH_OF_CLONED_REPO>"
os.environ["LOCAL_PROJECT_DIR"] = os.path.dirname(os.getcwd()) # This is the location of the root of the cloned repo
print(os.environ["LOCAL_PROJECT_DIR"])
```

In [ ]:

```
!wget --quiet --show-progress --progress=bar:force:noscroll --auth-no-challenge --no-check-certificate \
        https://api.ngc.nvidia.com/v2/models/nvidia/tao/pretrained_detectnet_v2/versions/resnet18/files/resnet18.hdf5 \
        -P  $LOCAL_PROJECT_DIR/local/training/tao/pretrained_model/
```

3. Convert Dataset to TFRecords for TAO

- The `Detectnet_v2` model in TAO expects data in the form of TFRecords for training.
- We can convert the KITTI Format Dataset generated from Part 1 with the `detectnet_v2 dataset_convert` tool provided with TAO toolkit

Figure 16. Extract from the jupyter notebook from the NVIDIA synthetic-data-example repository

Results

The provided juypter notebook contains the code required to run training, validation, and inferencing.

Training and validation

We conducted training and validation with various numbers of synthetic images to categorize model performance. The following table shows the validation accuracy based on the number of synthetic images used in training.

Table 6. Validation accuracy with synthetic images

Number of Synthetic Images	Validation Accuracy
0 Pre-Trained Model Base Line	2%
100	25%
500	62%
5000	83%

Inferencing

The pre-trained CV AI Object Detection model was tested with images of "Pallet Jacks" from the Logistic Objects in Context (LOCO) dataset to visualize real-world model performance. We tested inference on real-world images from the LOCO dataset with the model accurately detecting "Pallet Jacks" objects. The following figure shows "Pallet Jacks" objects detected (delineated with red bounding boxes) in real-world images.

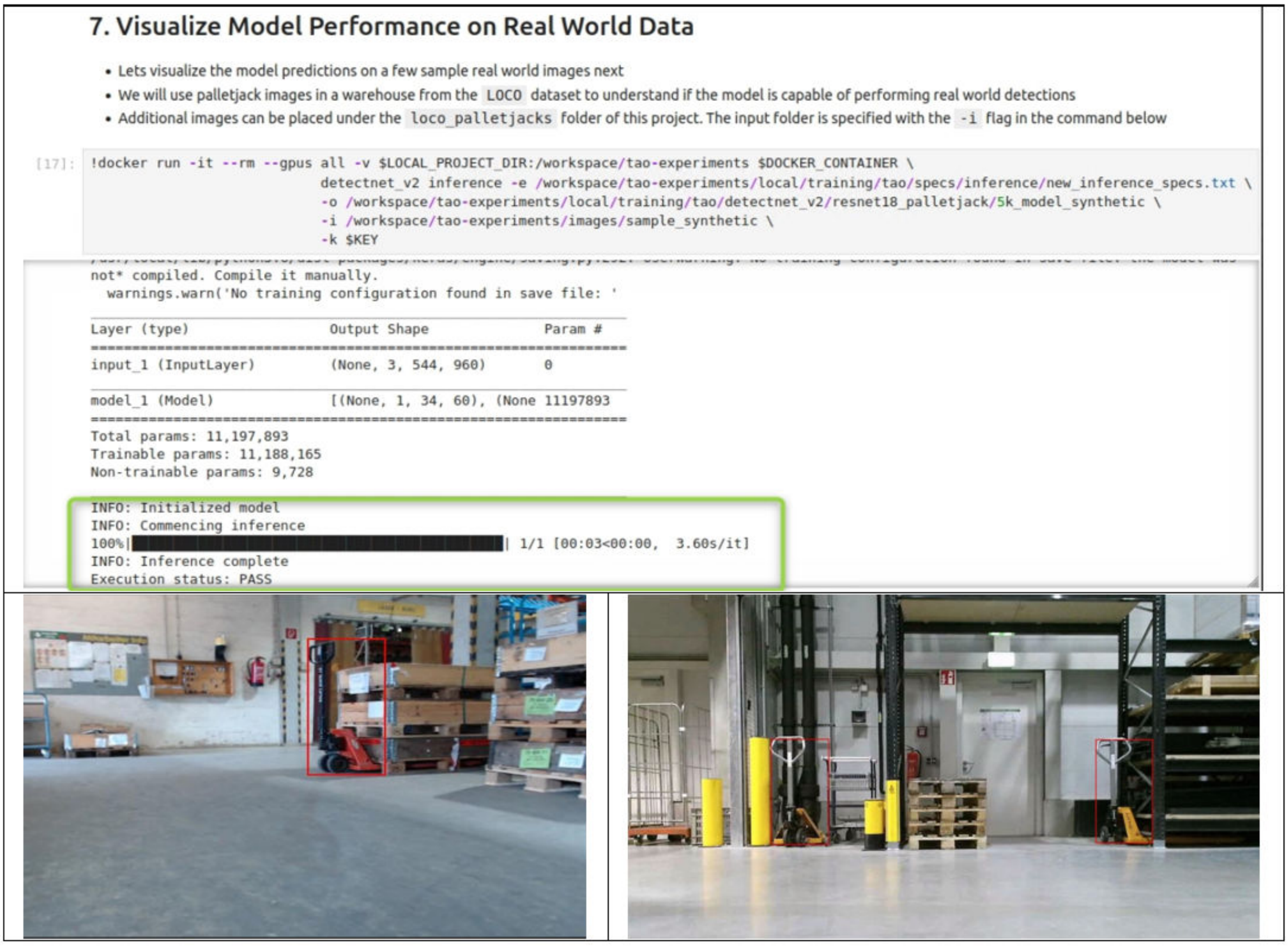


Figure 17. "Pallet Jacks" objects detected

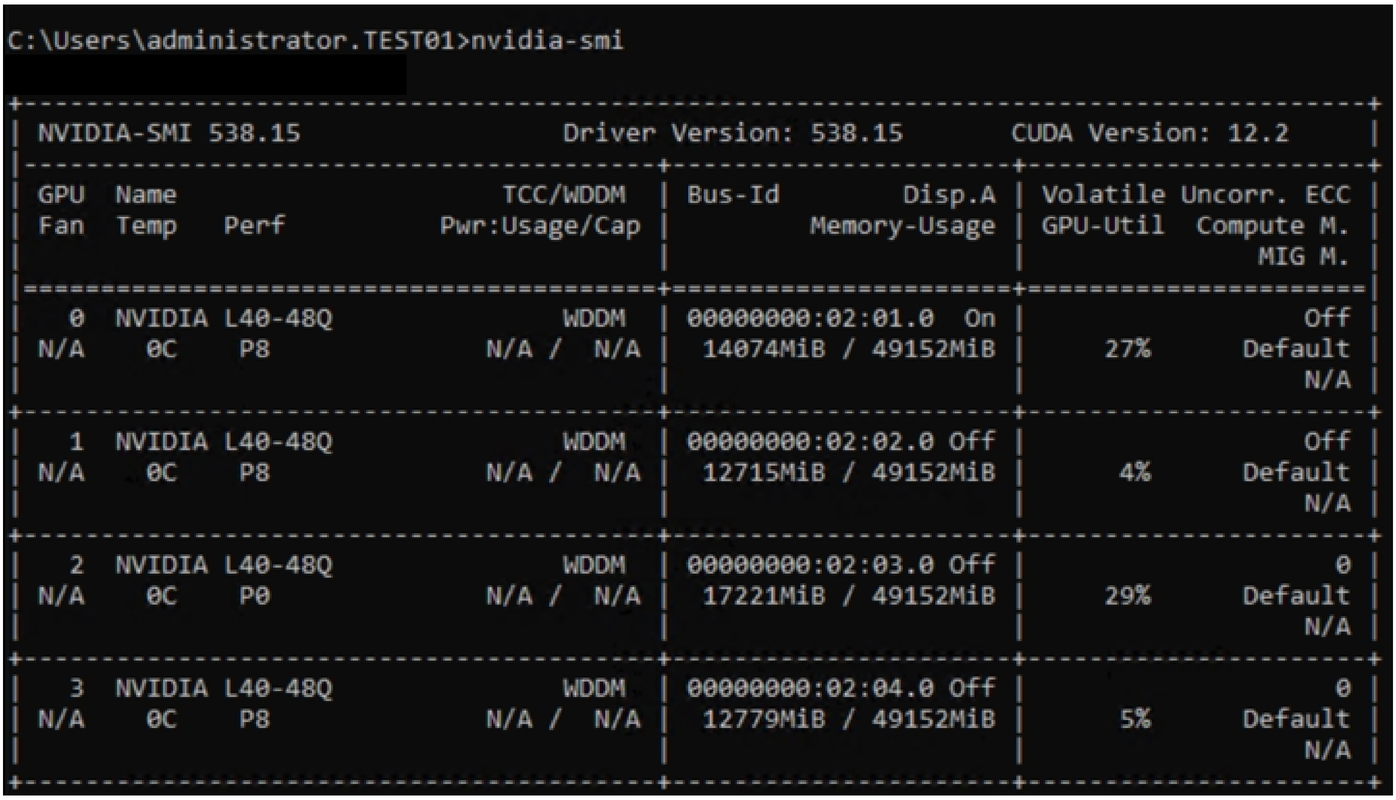
Findings

- Average CPU and GPU utilization of approximately 25% and approximately 72% respectively during model training.
- The CV AI model was solely trained on synthetic data. For further model optimization, it would be interesting to conduct training with both synthetic and user-curated real images.

Note on resource assignment

The above test cases (3D Simulations, Synthetic Data Generation, and Computer Vision AI Model Training) were run sequentially. In practice during Digital Twin development, these workflows might be required to run concurrently. We found that a virtualized NVIDIA Omniverse Enterprise deployment allowed for flexible resource assignments.

The following figure shows a snapshot of L40 utilization for concurrent workflows (3D model development, SDG generation, and CV AI model training) running on a single Dell PowerEdge 760xa.



```
C:\Users\administrator.TEST01>nvidia-smi
[redacted]
+---------------------------------------------------------------------------------------+
| NVIDIA-SMI 538.15                 Driver Version: 538.15       CUDA Version: 12.2     |
|-----------------------------------------+----------------------+----------------------+
| GPU  Name                     TCC/WDDM  | Bus-Id        Disp.A | Volatile Uncorr. ECC |
| Fan  Temp   Perf          Pwr:Usage/Cap |         Memory-Usage | GPU-Util  Compute M. |
|                                         |                      |               MIG M. |
|=========================================+======================+======================|
|   0  NVIDIA L40-48Q               WDDM  | 00000000:02:01.0  On |                  Off |
| N/A    0C    P8             N/A /  N/A  |  14074MiB / 49152MiB |     27%      Default |
|                                         |                      |                  N/A |
+-----------------------------------------+----------------------+----------------------+
|   1  NVIDIA L40-48Q               WDDM  | 00000000:02:02.0 Off |                  Off |
| N/A    0C    P8             N/A /  N/A  |  12715MiB / 49152MiB |      4%      Default |
|                                         |                      |                  N/A |
+-----------------------------------------+----------------------+----------------------+
|   2  NVIDIA L40-48Q               WDDM  | 00000000:02:03.0 Off |                    0 |
| N/A    0C    P0             N/A /  N/A  |  17221MiB / 49152MiB |     29%      Default |
|                                         |                      |                  N/A |
+-----------------------------------------+----------------------+----------------------+
|   3  NVIDIA L40-48Q               WDDM  | 00000000:02:04.0 Off |                    0 |
| N/A    0C    P8             N/A /  N/A  |  12779MiB / 49152MiB |      5%      Default |
|                                         |                      |                  N/A |
+-----------------------------------------+----------------------+----------------------+
```

Figure 18. Snapshot of L40 utilization for concurrent workflows

7

Conclusions

The Digital Twin paradigm can be applied to numerous real-world scenarios, but there is no one size fits all approach. Greenfield implementations of Digital Twins are rare, and practical Digital Twin solutions nearly always follow a step-by-step approach of integrations into existing processes, workflows, infrastructure, and software. For new Digital Twin initiatives, best practice is to start small, with a modular approach for each part of the Digital Twin workflow.

Digital Twin implementation and integration challenges are heavily dependent on the context and domain of the "twinned" physical real-world assets, which may only become apparent as development evolves, hence the need for flexible Digital Twin ecosystems.

This paper explored the practical implications and considerations for developing Digital Twin solutions using Dell PowerEdge R760xa servers and NVIDIA Omniverse Enterprise Platform as foundational building blocks.

We demonstrated:

- The ability to easily customize and assign hardware for various workflows, with accelerated compute and graphics.
- A virtualized deployment of the NVIDIA Omniverse Enterprise Platform.
- 3D Modeling Framework in a scenario simulation.
- Synthetic data generation with automatic image annotation and labeling.
- AI object detection model development:
    - Training and validation with SDG images
    - Model performance on real-world images

These demonstrations combine to offer the flexibility, interoperability, and collaborative capabilities needed to begin a Digital Twin journey with CV solutions.

8

References

Dell Technologies

- Dell Technologies Validated Designs hub
- Dell PowerEdge R760xa servers

NVIDIA

- NVIDIA Omniverse Enterprise Platform
- NVIDIA TAO Toolkit
- NVIDIA-Omniverse synthetic-data repository examples

Digital Twin

- Digital Twin Consortium