



Bootstrapping Computer Vision Models With Synthetic Data

NVIDIA Deep Learning Institute

Workshop Overview

Training computer vision models is complex, iterative, and requires a vast amount of high-quality, relevant visual data. Traditionally, this process relies on visual data gathered from the real world with cameras and sensors, often manually labeled, to represent the scenarios and situations that the model needs to learn.

NVIDIA Omniverse™ Replicator is a powerful synthetic data generation (SDG) engine that produces physically simulated synthetic data for training deep neural networks (DNNs). It augments costly, laborious human-labeled data, which can be error-prone and incomplete, with diverse physically accurate data tailored to the needs of developers.

In this course, you'll use Omniverse Replicator and the Omniverse Defects Generation Extension to generate synthetic data. Next, you'll iterate on the dataset to train a DNN to find target objects (scratches) in a scene.



Learning Objectives

By participating in this workshop, you'll learn how to:

- > Create a synthetic training dataset for later processing using NVIDIA Omniverse Replicator
- > Customize and refine existing tools to match your dataset feature and format requirements
- > Parameterize data generation offline for faster iteration when creating new or refined datasets
- > Import a synthetic dataset into your workflow, train it, iterate on the design, and export a model to be used for inference

Overview

Duration	8 hours
Price	Contact us for pricing.
Prerequisites	> Intermediate understanding of Python (including classes, objects, and decorators) > Basic understanding of machine learning and deep learning concepts and pipelines Suggested materials to satisfy prerequisites: Python tutorial , Deep Learning in a Nutshell , Deep Learning Demystified
Tools, libraries, and frameworks	NVIDIA Omniverse Replicator, NVIDIA TAO Toolkit, Python
Assessment type	Skills-based coding assessments evaluate the learner's ability to generate synthetic data for training computer vision models.
Certificate	Upon successful completion of the assessment, participants will receive an NVIDIA DLI certificate to recognize their subject matter competency and support professional career growth.
Hardware Requirements	Desktop or laptop computer capable of running the latest version of Chrome. Each participant will be provided with dedicated access to a fully configured, GPU-accelerated workstation in the cloud.
Language	English

Workshop Outline

Introduction

(15 minutes)

Meet the instructor.

- > Create an account at courses.nvidia.com/join

Introduction to Synthetic Data Generation (SDG) With Omniverse Replicator

(150 minutes)

Learn how to create a synthetic training dataset for later processing:

- > Discuss the case for synthetic data.
- > Learn the basics of the Replicator Python API for SDG.
- > Create example datasets using Python scripts using an NVIDIA Omniverse application interface.
- > Create a defects dataset using the Omniverse Defects Generation Extension and the Omniverse Defects demo pack.
- > Modify the extension code to change the dataset generated.

Break (60 minutes)

Headless SDG and Replicator YAML Extension

(120 minutes)

Learn to parameterize data generation offline using the Replicator YAML extension for faster iteration when creating new or refined datasets:

- > Discuss the advantages and disadvantages of running Omniverse Replicator in headless mode.
- > Learn to run Omniverse Replicator in headless mode using a configuration file.
- > Iterate on changes to the configuration file to generate new datasets.

Break (15 minutes)

Integrating Dataset Iteration Into the Training Workflow

(75 minutes)

Learn how to import a synthetic dataset into your workflow, train it, iterate on the dataset design, and export a model to be used for inference:

- > Discuss practical guidelines and examples for training a perception dataset to find a target object.
- > Fine-tune a visual transformer (ViT) model using NVIDIA TAO as the example workflow.
- > Iterate on the model by improving the data to solve accuracy issues.
- > Export the model for later deployment.

Assessment and Q&A

(45 minutes)

- > Review key learnings.
- > Take a code-based assessment to earn a certificate.

Why Choose NVIDIA Deep Learning Institute for Hands-On Training?

- > Access workshops from anywhere with just your desktop/laptop and an internet connection. Each participant will have access to a fully configured, GPU-accelerated server in the cloud.
- > Obtain hands-on experience with the most widely used, industry-standard software, tools, and frameworks.
- > Learn to build deep learning and accelerated computing applications for industries, such as healthcare, robotics, manufacturing, accelerated computing, and more.
- > Gain real-world experience through content designed in collaboration with industry leaders, such as the Children's Hospital of Los Angeles, Mayo Clinic, and PwC.
- > Earn an NVIDIA DLI certificate to demonstrate your subject matter competency and support your career growth.

Ready to Get Started?

For the latest DLI workshops and trainings, visit

www.nvidia.com/dli

For questions, contact us at nvdli@nvidia.com

© NVIDIA Corporation. All rights reserved. NVIDIA, the NVIDIA logo, and Omniverse are trademarks and/or registered trademarks of NVIDIA Corporation in the U.S. and other countries. All other trademarks and copyrights are the property of their respective owners. 2970454. OCT23

