Proceedings of the Seventh International Conference on
Railway Technology: Research, Development and Maintenance
Edited by: J. Pombo
Civil-Comp Conferences, Volume 15, Paper 20.1
Civil-Comp Press, Edinburgh, United Kingdom, 2026
ISSN: 2753-3239, doi: 10.4203/ccc.15.20.1

SynDRA-Gen: A Configurable Pipeline for Controlled Synthetic Multi-Modal Data Generation in a Railway Shunting Yard

G. Damico[1,2], F. Nesti[1,2], M. Marinoni[1,2], S. Sabina[1,2] and G. Buttazzo[1]

[1] Tecip Institute, Scuola Superiore Sant'Anna, Pisa, Italy
[2] Simulatrix MV s.r.l., Pisa, Italy

Abstract

The development of perception systems for railway applications relies on large volumes of annotated multimodal data. However, acquiring such data in real-world railway environments is inherently challenging due to safety constraints, high operational costs, and limited access to infrastructure. This paper introduces SynDRA-Gen, a standalone application designed for controlled synthetic data generation in railway yard scenarios. The system builds upon a previously developed simulation framework used for railway perception datasets and extends it into a fully configurable and user-friendly pipeline. SynDRA-Gen enables users to customize scene layouts, environmental conditions, and sensor configurations through a structured interface. The application supports a synchronized sensor suite that generates both camera and LiDAR data, along with automatically produced annotations such as depth maps, semantic segmentation, and 2D/3D bounding boxes. To facilitate reproducibility and accessibility, the system is distributed as a precompiled executable, allowing consistent dataset generation without requiring direct access to the underlying simulation framework. The tool is publicly available at: `https://github.com/DAmicoGlc/SynDRA-Gen`.

Keywords: synthetic data generation, configurable data generation pipeline, simulation framework, railway shunting yard, virtual environments, perception systems.

1 Introduction

Perception systems are becoming a key component of modern railway applications, supporting tasks such as obstacle detection, infrastructure monitoring, and safety-critical decision making. These systems typically rely on machine learning models that require large volumes of annotated data acquired from multiple sensing modalities, including cameras and LiDAR sensors.

In railway environments, the acquisition of such data is particularly challenging, due to safety-critical and strict operational constraints, which limit large-scale data collection campaigns. Simulation has emerged as an effective approach to overcome these limitations. Modern game engines provide photorealistic rendering and flexible environment design, enabling the generation of synthetic datasets with full ground-truth annotations. However, most existing simulation frameworks focus on road scenarios, and only limited attention has been given to tools specifically designed for railway perception.

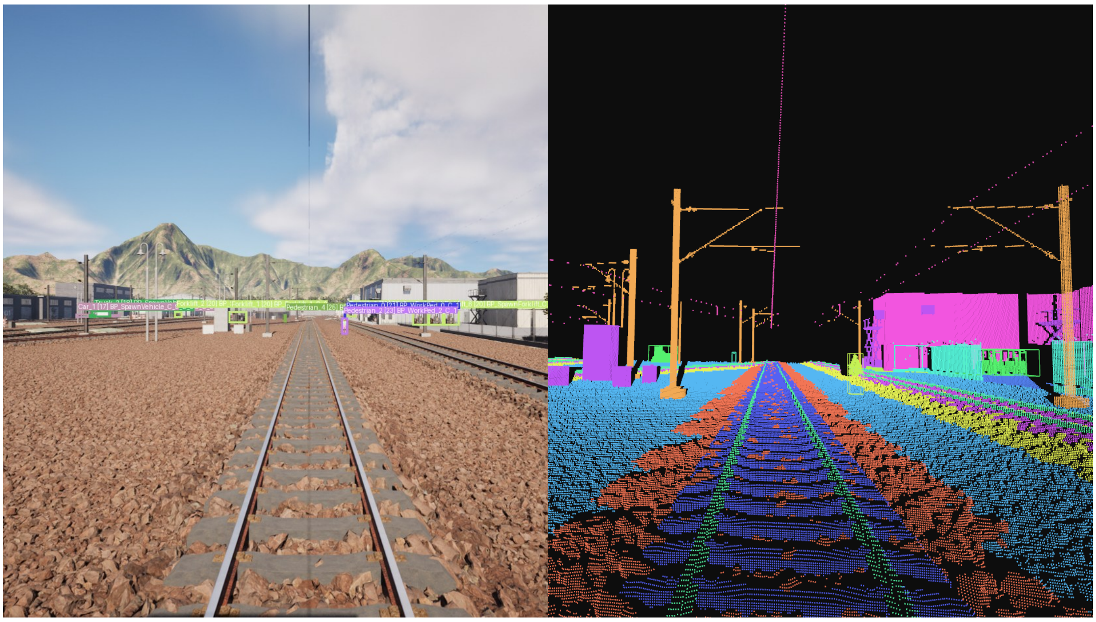

Figure 1: Generated RGB image (on the left) and point cloud (on the right) along with boxes and semantic annotations.

In this work, we present SynDRA-Gen, a standalone simulation application developed in Unreal Engine 5 for the controlled generation of synthetic multimodal perception datasets in a railway yard environment. An example of synthetic data generated by SynDRA-Gen is shown in Figure 1. The proposed system is designed around two complementary components: a scenario-generation module, which defines the structure and variability of the railway environment, and a synthetic data generation module, which produces synchronized sensor outputs and annotations from the configured scene.

SynDRA-Gen supports configurable environmental conditions, railway layouts, dynamic pedestrian activity, and a flexible sensor suite including camera and LiDAR configurations. All sensors are mounted on a moving acquisition platform, ensuring synchronized data capture during train motion. The system also provides automatic generation of multiple annotation types, including pixel-level labels and 2D/3D bounding boxes. The simulation framework used in this work has been previously adopted to generate synthetic datasets for railway perception. In particular, the SynDRA [1] and SynDRA-BBox [2] datasets were produced using the same framework.

This work provides the following main contributions:

- The design and implementation of SynDRA-Gen, a standalone application for synthetic data generation in railway yard.
- A configurable pipeline for multimodal dataset generation, supporting synchronized camera and LiDAR data with automatic annotations.
- A reproducible workflow that enables controlled variation of scenario and sensor parameters without requiring access to the underlying simulation framework.

The rest of the paper is organized as follow: Section 2 introduces the related works; Section 3 describes the SynDRA-Gen architecture; Section 4 details the application workflow, configurable parameters, and presents performance experiments; and Section 5 concludes the paper.

2 Related Work

Railway perception datasets are often harder to collect and publish than road datasets due to safety constraints, restricted access to infrastructure, and limited opportunities to capture rare hazardous events. Synthetic data aims to reduce this bottleneck by enabling scalable generation of labeled scenes under controlled variability (including rare corner cases), at a fraction of the manual labeling cost. Classic synthetic dataset efforts in vision demonstrated how virtual worlds can provide dense ground truth at scale and how this data can support training and evaluation of semantic understanding pipelines [3,4].

Real-world railway datasets. Real-world railway datasets provide the most valuable anchor for evaluation and for validating synthetic-to-real transfer. Recent multi-sensor initiatives such as OSDaR23 provide synchronized sensor data and railway-focused labels that support automated train operation and perception benchmarking [5]. Earlier datasets such as RailSem19 targeted semantic understanding in rail scenes and offer baseline benchmarks for rail-specific semantics and operating conditions [6]. In practice, such datasets help justify the need for synthetic data (scarcity and limited coverage of adverse conditions) and also provide a sanity-check route for validating whether synthetic scenarios and labels are aligned with real railway content distributions.

Railway-specific synthetic data. Railway-focused simulation frameworks have been developed to generate synthetic datasets tailored to trackside and yard environments. TrainSim explicitly supports LiDAR and camera dataset generation in railway scenarios [7]. More recently, SynDRA introduced a dedicated synthetic railway dataset for vision tasks, strengthening the case that railway-specific synthetic data pipelines can fill gaps left by limited public real data [1]. Complementary work such as SynDRA-BBox extends the railway synthetic ecosystem toward LiDAR-centric 3D detection and sim-to-real adaptation, aligning with scenarios in which synthetic data is combined with real data for transfer [2]. General-purpose autonomous driving simulators provide widely referenced design patterns for reproducible experiments, sensor configuration, and scenario variability. CARLA is a prominent open simulator supporting configurable environments and sensor suites, frequently used as a baseline reference for reproducible synthetic-data experiments [8].

A central limitation of synthetic training data is the *reality gap*. Domain randomization addresses this by training across controlled variability in appearance, geometry, and illumination so that real data becomes a *sample* within a broader distribution [9]. In parallel, domain adaptation methods explicitly align features between source (simulated) and target (real) domains. In railway perception, these strategies are relevant both when synthetic data is used as a pretraining source and when a small real dataset is available for adaptation or fine-tuning.

3 SynDRA-Gen Architecture

SynDRA-Gen is a standalone application for generating multimodal synthetic perception datasets in a railway yard environment. It is built on top of a railway simulation framework previously used for synthetic dataset generation in rail scenarios [1, 2]. SynDRA-Gen exposes the underlying engine via a structured and configurable pipeline so that users can generate synchronized camera and LiDAR datasets without modifying simulator internals.

The architecture separates two layers: (i) scenario configuration, which defines the yard environment and its variability, and (ii) data acquisition, which produces synchronized sensor outputs and annotations from the configured scene. This separation enables controlled studies in which scene variability and sensing variability can be manipulated independently.

3.1 Scenario generation

The simulated environment reproduces a shunting yard for railway inspection and perception experiments, including parallel tracks and switch lanes surrounded by service infrastructure (e.g., facility buildings, parking areas, signals/lights). The scenario is depicted in Figure 2.

Figure 2: The virtual scenario of SynDRA-Gen.

The yard is populated with work-related objects and maintenance props (e.g., stacks of sleepers, tool bags, forklifts, lamps) to increase geometric and semantic variability near the tracks.

Environmental variability is introduced in a controlled and repeatable manner. Vegetation around the facility is generated procedurally using mesh-family selection and deterministic seeds that regulate placement of trees, lower vegetation, and grass. This design supports *visually different yet repeatable* configurations and aligns with established sim-to-real motivation for controlled variability [9].

3.2 Weather, illumination, and low-visibility sensing

SynDRA-Gen provides an environmental module controlling both weather and lighting conditions, enabling acquisitions under clear daytime, nighttime, and low-visibility fog regimes. For RGB data, fog reduces visibility and contrast through scattering and distance-dependent attenuation [10]. For LiDAR data, fog is modeled through range-dependent degradation, including a probabilistic reduction of valid returns as a function of distance and noise injection. These choices are consistent with degradation mechanisms reported in adverse-weather LiDAR perception literature and support evaluation of robustness under low visibility [11, 12].

3.3 Railway configuration and dynamic actors

The application includes a train-motion module that moves the acquisition platform along a user-selected rail using a configurable speed profile. Since the sensors are rigidly attached to the train reference frame, all outputs are acquired during train travel and remain spatially synchronized with the selected rail path. The application also supports controlled variation of railway surface appearance (e.g., ballast and surrounding terrain presets) while preserving deterministic reconstruction.

To introduce controlled moving actors, pedestrian workers are enabled in the selected track area. Workers are assigned to predefined paths surrounding and crossing the rails, and aspect parameters can be varied across capture configurations to create different visual patterns.

3.4 Synthetic data generation: sensors and annotations

After the scenario is configured, synthetic perception data are generated through a configurable sensor suite mounted on the moving train. The camera configuration includes at least one RGB camera and can be set to monocular, stereo, or dual-view camera. Camera placement and intrinsic parameters are controlled by mounting and acquisition presets, enabling systematic comparison of observation geometries in the same environment.

Optionally, a LiDAR sensor can be mounted on the same platform. Two LiDAR families are supported, flash LiDAR and beam-stacked (scanning) LiDAR, via configurable sensor presets. Mounting parameters adjust the sensor pose relative to the rails. All sensors share the same moving reference frame, supporting synchronized acquisition. Internally, platform motion is updated at a fixed simulation step, while sensor outputs are captured at a lower fixed acquisition rate, preserving temporal alignment across modalities.

The application supports multiple annotation modalities for both image and point-cloud data. For camera sensors, the system can export RGB images together with depth, semantic segmentation, and 2D object-level bounding boxes. For LiDAR, it produces point clouds enriched with semantic information and can additionally export box annotations aligned with the range data. Annotation type are configurable, enabling both lightweight (raw sensing) and richer (dense labels and object-level annotations) datasets.

4 Application Flow and Experimental Setup

This section describes how SynDRA-Gen is used in practice and summarizes a workflow for synthetic dataset generation. Instead of assessing downstream perception methods (addressed in prior SynDRA works [1,2]), the focus is on the data-generation process and on practical aspects of producing datasets.

4.1 Application workflow

SynDRA-Gen is designed as a standalone application for repeated scenario configuration and data acquisition. At startup, the main module loads the railway yard scene and presents a configuration menu. Through this menu, the user selects scenario and sensor settings, including environmental conditions, railway parameters, pedestrian aspects, camera presets, and LiDAR presets. After the configuration is applied, the application instantiates or updates configurable scene elements, prepares the sensor suite, and initializes output directories for data export. The user then starts acquisition; once a run finishes, the tool returns to the menu to enable repeated runs under modified configurations.

The train trajectory is updated at 100 Hz (i.e., platform motion updates every 0.01 s), while sensor acquisition is performed at a fixed rate of 10 Hz. Consequently, one synchronized sensor sample is produced every 10 simulation steps. This decoupling between motion update and acquisition rate ensures smooth platform motion and preserves temporal coherence among cameras, LiDAR, and annotations by producing all modalities from the same world state at each acquisition instant.

4.2 Customization parameters

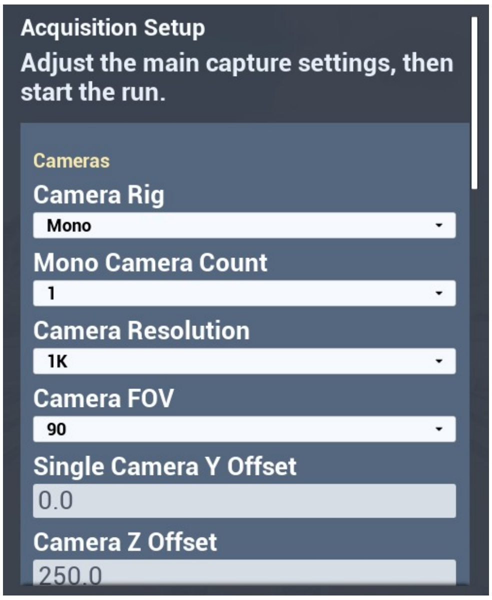

Figure 3: SynDRA-Gen configuration Menu.

SynDRA-Gen exposes a focused set of parameters that control both scenario variability and sensing while keeping the interface oriented toward dataset generation, an image of the user menu is presented in Figure 3.

Railway configuration: acquisition is performed along a selected railway track, using configurable constant speed profiles. The visual context can be varied through predefined ballast/terrain combinations, enabling appearance changes while preserving deterministic reconstruction.

Environmental and pedestrian configuration: weather and illumination presets (clear, fog, night) can be selected, and pedestrian workers that are enabled along predefined paths in the selected track area, introduce dynamic actors in a controlled manner.

Camera configuration: the system supports monocular, stereo, and dual-camera setups. These options control view coverage (single forward view versus wider multi-view) and baseline disparity (stereo). Camera intrinsics and mounting offsets can be varied systematically to explore different sensing geometries.

LiDAR configuration: two LiDAR modalities are supported (beam-stacked and flash LiDAR) via configurable presets and mounting parameters, enabling different sampling strategies and point-cloud structures.

4.3 Experimental Setup and Reporting Metrics

To evaluate the practical usability of the proposed application, we design a representative experimental protocol aimed at quantifying both computational cost and storage requirements during a complete data acquisition session.

All experiments are conducted on a fully specified hardware and software platform to ensure reproducibility and to contextualize the reported performance. The system configuration consists of a Windows 11 operating system and Unreal Engine 5.4. The hardware setup includes a 13th Gen Intel(R) Core(TM) i9-13900K CPU (3.00 GHz), an NVIDIA GeForce RTX 5080 GPU with 16 GB of dedicated memory, and 64 GB of system RAM. Rather than targeting an average operating condition, we deliberately adopt a worst-case configuration, intended to stress the system and provide an upper-bound estimate of resource consumption.

In particular, it includes two cameras for high-resolution image acquisition, the LiDAR configuration with the highest number of rays, and the export of all available annotation modalities. Such a setup reflects a scenario in which maximum data fidelity is prioritized over efficiency.

Table 1: Performance metrics for dataset acquisition under a high-demand configuration, where the train run at 10 km/h and generates 1029 data samples.

Metric	**Value**
Total acquisition time	4 min
Average time per sample	20 ms
Peak memory usage	19 GB
Dataset size	21 GB

Performance is assessed through a set of aggregate metrics that capture both temporal and resource-related aspects of the acquisition process. Specifically, we measure the total time required to generate the dataset, the average time per sample, the peak memory usage observed during execution, and the overall storage footprint of the generated data.

The results reported in Table 1 highlight the computational and storage demands of the proposed pipeline under extreme operating conditions. The measured acquisition times and memory footprint emphasize the cost of high-fidelity sensing configurations, while the dataset size illustrates the scalability challenges inherent to multimodal data generation. Overall, these findings provide practical insight into the trade-offs between data quality and resource consumption, offering guidance for selecting appropriate configurations in real-world applications.

5 Conclusions

This paper presented SynDRA-Gen, a standalone application for synthetic data generation in railway environments. Built on top of an existing simulation framework, the application provides a configurable and reproducible pipeline for generating multimodal perception datasets, including camera images, LiDAR point clouds, and associated annotations.

The proposed system enables users to control both scenario configuration and sensing parameters through a structured interface, supporting systematic dataset generation without requiring access to the underlying simulation implementation. By focusing on the data-generation process rather than downstream perception tasks, this work introduced a practical tool for producing consistent and scalable datasets tailored to railway applications.

Future work will extend the capabilities of SynDRA-Gen along several directions. First, additional environmental effects such as rain will be incorporated to further increase the realism and variability of the generated data. Second, the railway dynamics will be enhanced by enabling train traversal along switch rails, allowing more complex motion patterns and scene coverage. Finally, the inclusion of moving vehicles within the environment will further enrich scene dynamics and enable the generation of more diverse and challenging perception scenarios.

These extensions will further strengthen the applicability of the proposed system as a comprehensive tool for synthetic data generation in railway perception research.

The application will be publicly released at The tool is publicly available at: `https://github.com/DAmicoGlc/SynDRA-Gen` to support reproducible research and facilitate the adoption of synthetic data generation techniques in the railway domain.

References

[1] G. D'Amico, F. Nesti, G. Rossolini, M. Marinoni, S. Sabina, and G. Buttazzo, "SynDRA: Synthetic Dataset for Railway Applications," in *Winter Conference*

on Applications of Computer Vision (WACV), 2025.

[2] X. Diaz, G. D'Amico, R. Dominguez-Sanchez, F. Nesti, M. Ronecker, and G. Buttazzo, "Towards Railway Domain Adaptation for LiDAR-based 3D Detection: Road-to-Rail and Sim-to-Real via SynDRA-BBox," arXiv:2507.16413, 2025.

[3] G. Ros, L. Sellart, J. Materzynska, D. Vazquez, and A. M. Lopez, "The SYNTHIA Dataset: A Large Collection of Synthetic Images for Semantic Segmentation of Urban Scenes," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.

[4] T. Hofmeier, M. Chichon, "Vir2Rail – Modular LiDAR Simulation Framework for ATO Research in Railway Environments," Zenodo, Mar. 2026. DOI: 10.5281/zenodo.18769046.

[5] R. Tagiew, M. Koeppel, K. Schwalbe, P. Denzler, P. Neumaier, T. Klockau, M. Boekhoff, P. Klasek, and R. Tilly, "OSDaR23: Open Sensor Data for Rail 2023," in *International Conference on Robotics and Automation Engineering (ICRAE)*, 2023.

[6] O. Zendel, M. Murschitz, M. Zeilinger, D. Steininger, S. Abbasi, and C. Beleznai, "RailSem19: A Dataset for Semantic Rail Scene Understanding," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2019.

[7] G. D'Amico, M. Marinoni, F. Nesti, G. Rossolini, G. Buttazzo, S. Sabina, and G. Lauro, "TrainSim: A Railway Simulation Framework for LiDAR and Camera Dataset Generation," *IEEE Transactions on Intelligent Transportation Systems*, 2023.

[8] A. Dosovitskiy, G. Ros, F. Codevilla, A. Lopez, and V. Koltun, "CARLA: An Open Urban Driving Simulator," in *Conference on Robot Learning (CoRL)*, 2017.

[9] J. Tobin, R. Fong, A. Ray, J. Schneider, W. Zaremba, and P. Abbeel, "Domain Randomization for Transferring Deep Neural Networks from Simulation to the Real World," arXiv:1703.06907, 2017.

[10] C. Sakaridis, D. Dai, and L. Van Gool, "Semantic Foggy Scene Understanding with Foggy Cityscapes," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018.

[11] M. Dreissig, D. Scheuble, F. Piewak, and J. Boedecker, "Survey on LiDAR Perception in Adverse Weather Conditions," arXiv:2304.06312, 2023.

[12] A. Haider et al., "A Methodology to Model the Rain and Fog Effect on the Performance of Automotive LiDAR Sensors," *Sensors*, 2023.