HELLENIC REPUBLIC
MINISTRY OF DEVELOPMENT AND INVESTMENT
GENERAL SECRETARIAT FOR RESEARCH AND INNOVATION
HELLENIC FOUNDATION FOR RESEARCH AND INNOVATION

H.F.R.I.
Hellenic Foundation for Research & Innovation

Greece 2.0
NATIONAL RECOVERY AND RESILIENCE PLAN

Funded by the European Union
NextGenerationEU

Greece 2.0
Basic Research Financing Action
(Horizontal support of all Sciences)
Sub-action 1
Funding New Researchers

LEARNER

Project Title

SLAM AND PATH PLANNING MIDDLEWARE PACKAGE FOR ROBOTS IN CHALLENGING ENVIRONMENTS

Project Duration

20 November 2023 – 19 November 2025
24 Months

Project Acronym	Project No
LEARNER	**015339**

Deliverable No	Deliverable Title
D2.1	**Simulated environments and datasets formulation**

Deliverable Completion Date

19 June 2024

Table of Contents

Table of Figures

Document Revision History

Version	Date	Notes
1.0	19/06/2024	First version document containing the formulated simulation environments for acquiring the required training and evaluation datasets

List of Acronyms

Acronym	Meaning
SLAM	Simultaneous Localization and Mapping
USD	Universal Scene Description
MDL	Material Definition Language

1. Introduction

The objective of this deliverable is to present a comprehensive report on the development of simulated environments and the associated sequences produced for LEARNER project. These simulations are designed to facilitate the training and evaluation of our Simultaneous Localization and Mapping (SLAM) algorithm and Path Planning systems, while taking into consideration obstacle avoidance and the development of the social awareness models. The produced environments will serve as a basis for developing our techniques, which will then be adjusted accordingly to be applied in real world conditions. This document also includes information on a publicly available repository that other beneficiaries can use for their own purposes.

For this task, we utilized two advanced simulation tools: Unreal Engine 5 and NVIDIA Omniverse. Both platforms were employed separately, each bringing unique strengths and weaknesses to the table. Unreal Engine 5, with its Lumen rendering engine, excels in creating highly detailed and dynamic environments. On the other hand, NVIDIA Omniverse offers robust real-time collaboration and simulation capabilities with its proprietary rendering engine. By leveraging both platforms, we were able to create comprehensive and varied synthetic worlds essential for our project's objectives.

From these virtual worlds, we rendered photorealistic sequences of images that we split into two datasets. One targeting the training process and one the evaluation of the developed algorithms.

1.1. Unreal Engine 5 Simulation

Unreal Engine 5 was utilized to develop a sophisticated simulation environment representing office spaces. This simulation includes several critical features to challenge and train our navigation systems:

- **Dynamic Lighting**: The office environment includes varying lighting conditions to test the system's adaptability to different visibility scenarios. Lumen, the real-time global illumination and reflections system in Unreal Engine 5, enhances the realism of these lighting conditions.
- **Smoke and Fire Effects**: To simulate emergency conditions, the environment includes smoke and dynamic fire particles. These elements are crucial for testing the system's ability to navigate and operate under impaired visibility.

- **Falling Obstacles and Explosions**: The inclusion of dynamic elements, such as falling objects, adds complexity to the environment, challenging the robustness and responsiveness of SLAM and path planning systems.

1.2. NVIDIA Omniverse Simulation

NVIDIA Omniverse was employed to create a detailed and interactive industrial warehouse environment. This platform's strength relies on the Omniverse RTX Renderer, which is a physics-based real-time ray-tracing renderer built on NVIDIA's RTX technology, Pixar's Universal Scene Description (USD), and NVIDIA's Material Definition Language (MDL). It provides two render modes supporting fully dynamic lighting (without any light baking) with thousands of lights, millions of objects, and the flexible MDL material representation. With these render modes it enables a variety of workflows not possible before, and it provides a reference rendering solution for USD-based content in Omniverse leveraging the power of NVIDIA GPUs with RTX, making it ideal for creating dynamic and collaborative scenarios.

Omniverse RTX Renderer provides the RTX-Real-Time ray tracing mode which allows rendering more geometry than traditional rasterization methods, as well as physical-based materials at a high fidelity, in real-time. In RTX - Real-Time mode, the renderer performs a series of separate passes that compute the different lighting contributions (for example: ray-traced ambient occlusion, direct lighting with ray-traced shadows, ray-traced indirect diffuse global illumination, ray-traced reflections, ray-traced translucency and subsurface scattering). Each pass is separately denoised, and the results are composited.

The reason why an industrial warehouse environment was chosen is due to its complexity. Studying and simulating an industrial facility is crucial for ensuring the safety of a robotic vehicle navigating through such environments, particularly when performing SLAM. Industrial settings are inherently complex and hazardous, presenting a variety of dynamic factors such as smoke, fire, human presence, machinery, moving objects, and varying lighting conditions. Each of these elements can significantly impact the robot's sensors and its ability to accurately perceive and map its surroundings. Potential risks can be identified and robust algorithms that enhance the robot's navigational accuracy and safety to be developed.

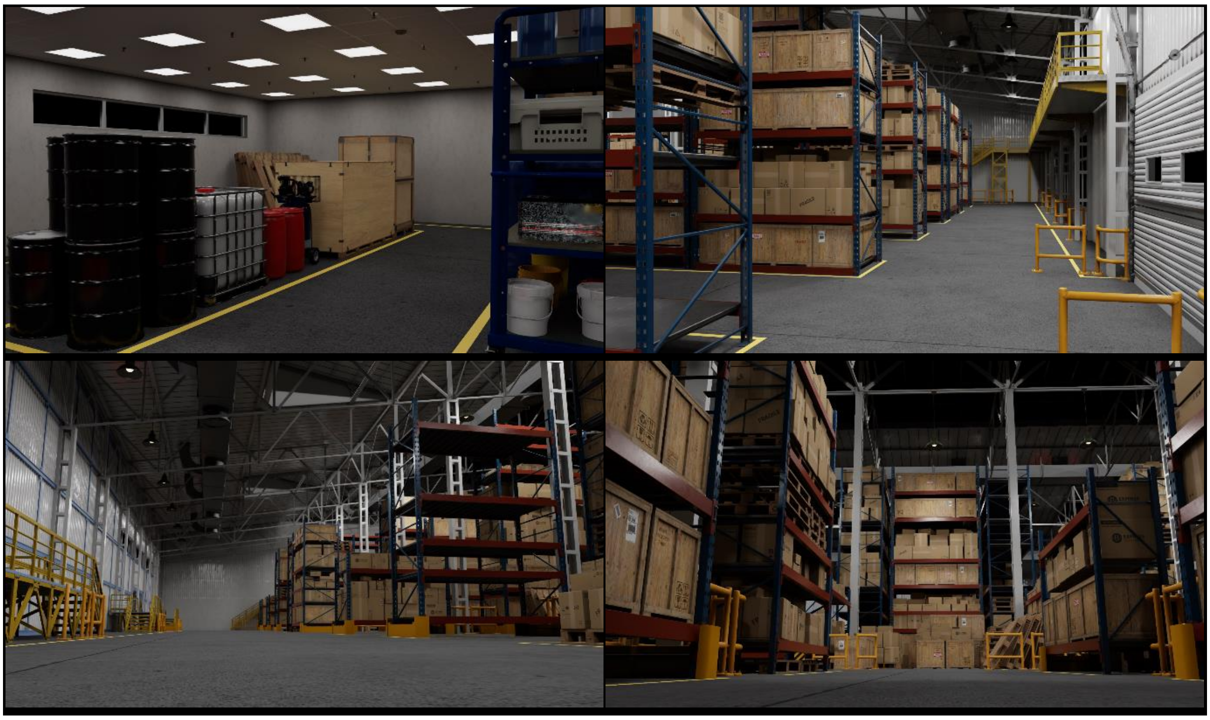


Figure 1. Representative examples of the industrial warehouse in NVIDIA Omniverse, where the simulation takes place.

2. Datasets

To develop robust Neural Network models for real-world applications, it is essential to train them on diverse and high-quality datasets. In this study, we generated unique runs with the aim to capture images under various environmental and lighting conditions, ensuring comprehensive training for improved performance and evaluation in challenging scenarios.

2.1. Training Subset

For the training part of the dataset, distinct runs were created where the robot's navigation was designed using fixed paths and transformations, with the aim to capture high-quality images. Specifically, the robot maintained the same pose and movement across all runs, ensuring consistency in the structural environment. Varying elements for each dataset subset were the environmental and lighting conditions, such as fog density and light levels, but also environmental and non-environmental barriers, such as people, vehicles, boxes, and more. Recording, understanding, and reacting to an environment through model-based and Deep Learning techniques works best in bright light conditions with high contrast, but accuracy typically decreases in challenging conditions, like dense fog or low light. To improve performance in these difficult scenarios, we generated a series of image sequences where the conditions were changed, while keeping the camera movement and environment constant. This method allows the Deep Learning techniques to learn from the light sequences and apply that knowledge to dim light or other challenging conditions, enhancing their ability to adapt and maintain accuracy across extreme variations.

2.1.1. Unreal Engine 5 Training Dataset

The conditions under which the Unreal Engine 5 training datasets were extracted, are as follows:

- **Optimal Conditions**: Bright light with high contrast for feature detection.

- **Challenging Conditions**: Dark lighting to simulate poor visibility scenarios.
- **Fog Conditions**: Simulating fog to test feature detection under impaired visibility.

Figure 2. Indicative snapshots from the Unreal Engine 5 Training Dataset showing areas of the office environment in different conditions. In each group, the samples on the left show the area in ideal conditions and on the right under darkness.

2.1.2. NVIDIA Omniverse Training Dataset

The conditions under which the NVIDIA Omniverse training datasets were extracted, are as follows:

- **Optimal Conditions**: Bright light with high contrast is essential for effective Feature Detection. Clear illumination allows the developed models to distinguish objects and details with greater accuracy. Furthermore, it enhances the visibility of edges and textures, enabling the models to identify and classify features with precision. These conditions provide a reliable baseline for training, ensuring that the models can perform fundamental detection tasks efficiently before being tested in more challenging environments.
- **Low-Lighting Conditions**: Dark or ambient levels of lighting are used to simulate poor visibility scenarios. The reduced illumination will make it difficult for the model to discern objects and features accurately, mimicking real-world situations such as nighttime operations or dimly lit indoor spaces. These conditions will test the robot's ability to adapt and maintain functionality

when visual information is limited, ensuring they can still perform essential tasks like navigation and obstacle detection reliably even under adverse lighting conditions.

- **Smoke Conditions:** Smoke introduces visual noise and reduces clarity, making it difficult for the visual sensors and cognitive system of the robots to identify and track objects accurately. By inserting such conditions into the training data, the Deep Learning models learn to handle impaired visibility, increasing their applicability in hazardous environments.

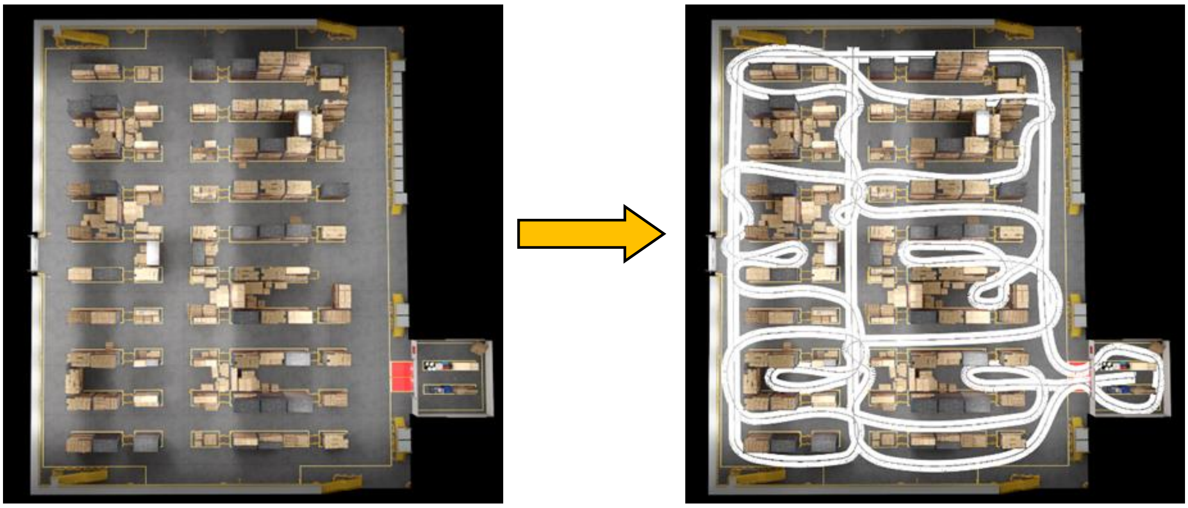
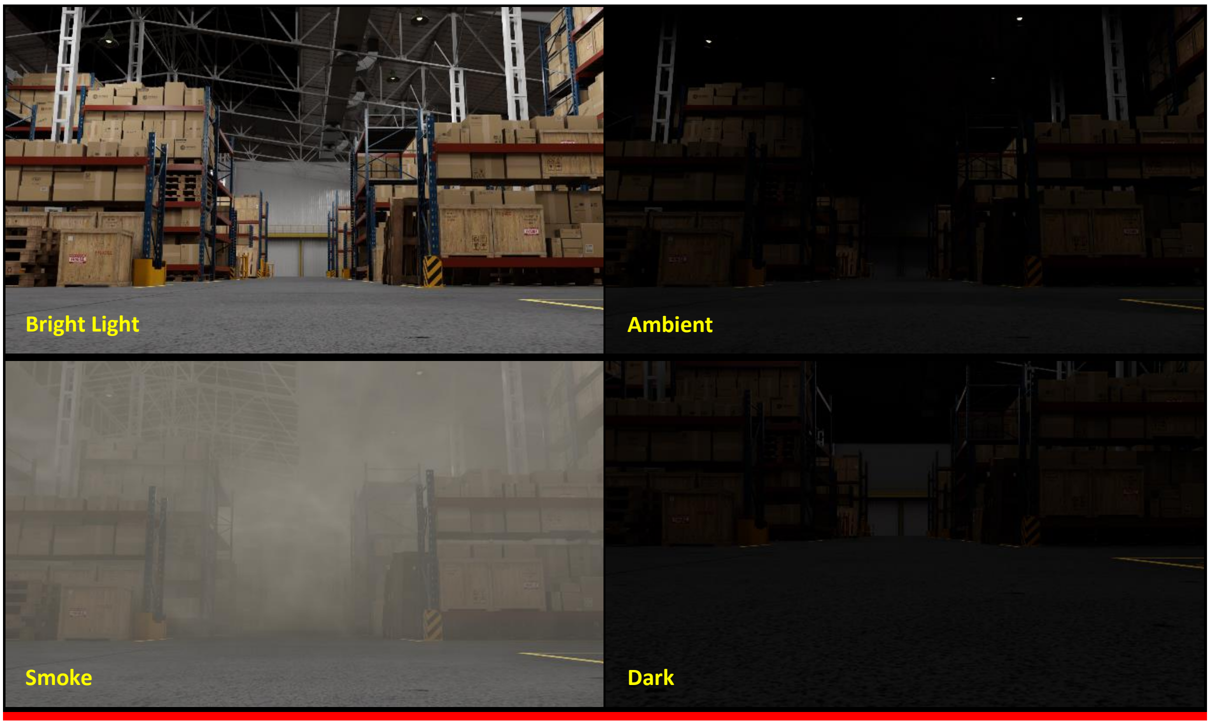
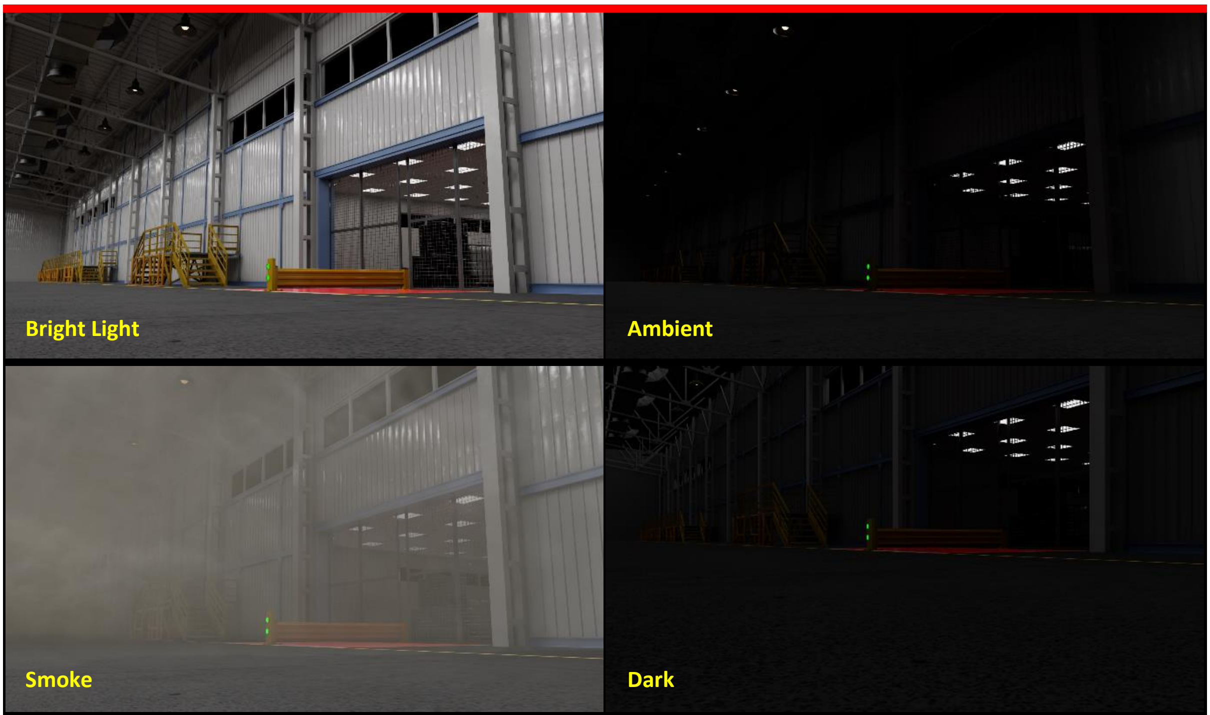


Figure 3. Top view of the industrial plant (left) and the steady path of the robot during the simulations (right).




Bright Light
Ambient
Smoke
Dark


Bright Light
Ambient
Smoke
Dark


Bright Light
Ambient

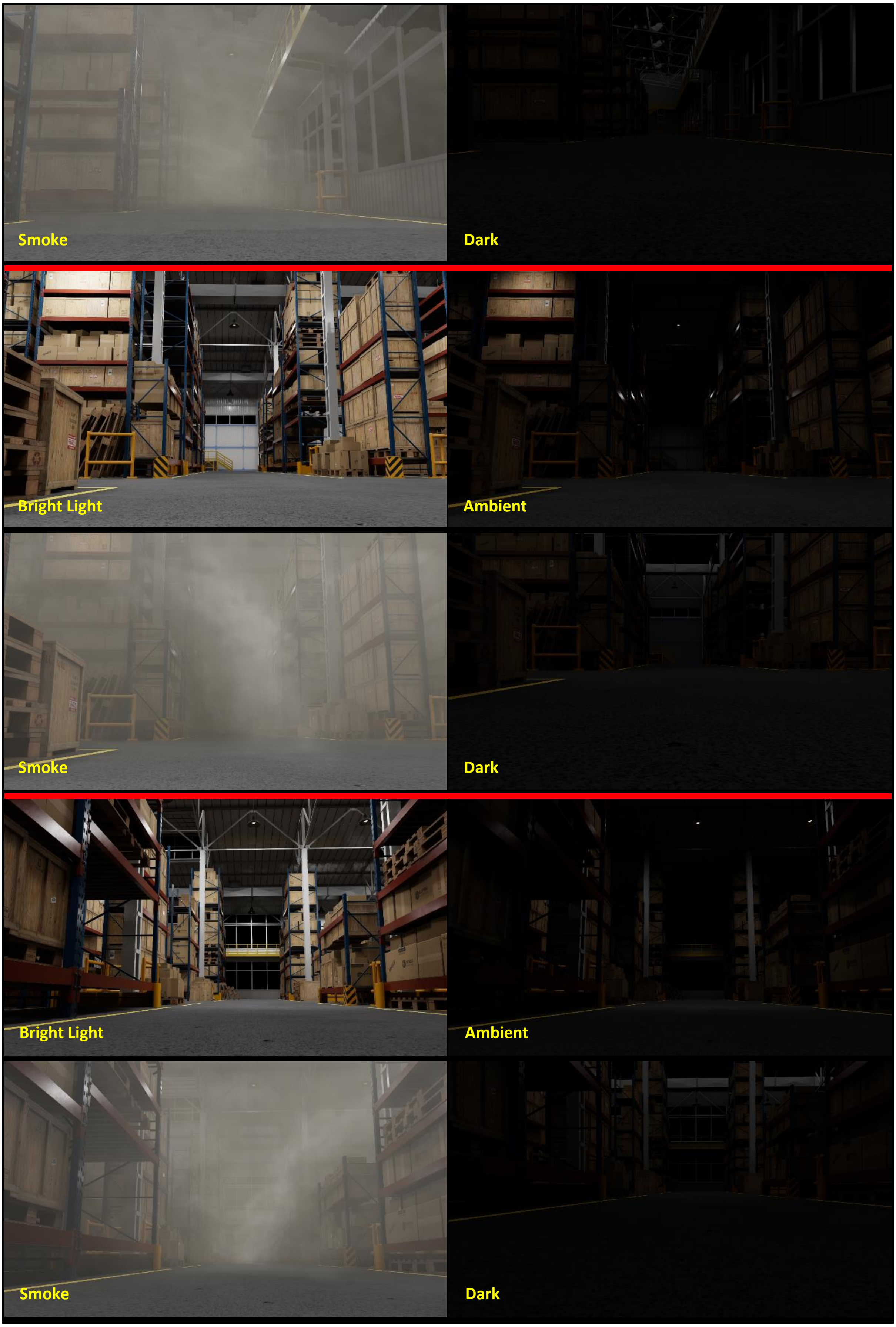

Smoke
Dark
Bright Light
Ambient
Smoke
Dark
Bright Light
Ambient
Smoke
Dark

Figure 4. Indicative group snapshots from the dataset showing one area of the warehouse in different environmental conditions. In each group, the top left sample shows the area in ideal conditions, the bottom left with smoke, the top right with ambient lighting, and the bottom right in darkness.

2.2. Evaluation Subset

Evaluation is a critical step in the development of model-based and Deep Learning models, as it assesses their performance and generalization capabilities in real-world scenarios. This process involves testing the developed architectures on a separate set of data that includes samples not encountered during the development and training stages. By evaluating the models in such conditions, we can measure their accuracy, robustness, and adaptability. Additionally, evaluation helps identify any weaknesses or limitations in the total system, providing guidance for further refinement and optimization to enhance their overall performance and reliability.

2.2.1. Unreal Engine 5 Evaluation Dataset

For the case of evaluation dataset, highly dynamic events are crucial for evaluating the final system algorithms and testing the overall effectiveness of our methods in real-world-like conditions. Towards that end, the selected conditions adopted are as follows:

- **Lighting Changes**: Simulating various lighting conditions to evaluate adaptability.
- **Fog Appearance and Disappearance**: Testing the system's response to changing visibility.
- **Moving Obstacles and Passages**: Evaluating the navigation system's ability to adapt to sudden changes.

Figure 5. Example instances showing the dynamic events and conditions that could occur in an office environment.

2.2.2. NVIDIA Omniverse Evaluation Dataset

The evaluation process involved meticulously chosen features to thoroughly assess the models across a spectrum of scenarios. This approach ensured that the models could be tested for their accuracy and navigational capabilities in diverse and challenging industrial warehouse environments. The evaluation included features such as:

- **Human Interaction:** The warehouse simulation includes both static individuals and dynamic crowds, providing a realistic and challenging environment for testing social navigation and obstacle avoidance. By involving stationary humans to the simulation, the robot can practice maneuvering around fixed obstacles, while the presence of moving crowds introduces unpredictability and requires the robot to dynamically adapt its path. This comprehensive approach ensures that the robot can effectively navigate complex human-populated environments, maintaining safety and efficiency in real-world warehouse operations where human interaction is frequent and varied.
- **Dynamic Elements:** The environment also includes both falling and stationary obstacles, as well as other variable elements such as forklifts and doors, to evaluate the system's performance under unpredictable conditions. Falling obstacles are there to challenge the robot to respond to sudden changes in its surroundings, requiring quick and accurate adjustments to its navigation paths. Stationary obstacles, on the other hand, test the robot's ability to plan and execute efficient routes in a cluttered environment. Additionally, the inclusion of moving forklifts introduces dynamic interactions that further assess the robot's anticipation and reaction to moving objects. This comprehensive testing helps in thoroughly evaluating the robustness of the SLAM algorithm and the overall navigation system, ensuring that the complexities and uncertainties of real-world industrial settings can be handled and overcome.
- **Shifts of Lighting Conditions:** To thoroughly evaluate the model's adaptability and performance, we created a simulation which divides the warehouse into several imaginary zones, each featuring different events. In some places, there are optimal lighting conditions in order to observe if the model is able to find points in the space which are similar and even identical to those of the testing ones. Other zones feature ambient lighting to simulate typical warehouse environments with consistent but not always optimal illumination. Additionally, some zones are designed with low visibility due to smoke or darkness, challenging the model to maintain functionality and accuracy under adverse conditions. By evaluating the models across these varied lighting environments and events, we want to ensure the robust adaptation of the model to different illumination levels and its reliable performance in both ideal and challenging situations.

H.F.R.I. Hellenic Foundation for Research & Innovation
Greece 2.0 NATIONAL RECOVERY AND RESILIENCE PLAN
Funded by the European Union NextGenerationEU

Figure 6. Representative snapshots from the evaluation dataset, showing the dynamic changes that could occur in an industrial warehouse.

3. Public Repository

Following the objectives of LEARNER project, the above datasets together with the source files for reproducing the simulated environments are published under the EU Open Research Repository: Zenodo. All the available material can be found at the following link, which was also incorporated at the project's official website: https://doi.org/10.5281/zenodo.14823070.

H.F.R.I. Hellenic Foundation for Research & Innovation
Greece 2.0 NATIONAL RECOVERY AND RESILIENCE PLAN
Funded by the European Union NextGenerationEU