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

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

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

LEARNER

Project No

015339

Deliverable No

D2.2

Deliverable Title

Hybrid map model

Deliverable Completion Date

19 January 2025

Table of Contents

Document Revision History

Version	Date	Notes
1.0	19/01/2025	First version document describing the developed Hybrid map model for supporting SLAM and Path Planning approaches in dynamic environments.

List of Acronyms

Acronym	Meaning
LEARNER	SLAM and Path Planning Middleware Package for Robots in Challenging Environments
SLAM	Simultaneous Localization and Mapping
PP	Path Planning
ROS	Robot Operating System

1. Introduction

LEARNER aims to develop a Simultaneous Localization and Mapping (SLAM) and Path Planning (PP) middleware package for robots operating in complex and dynamic environments. The hybrid map model, which provides a structured representation of the environment by providing occupancy information, semantic classification, and historical state tracking, is one of the fundamental components of our system. This deliverable documents the implemented map model, detailing its structure, data representation, and interconnection with the broader system's architecture.

The hybrid map is built upon a ROS2-based occupancy grid, which serves as the foundation for mapping the environment. Unlike standard occupancy maps that only indicate whether a region is occupied or free, the hybrid map extends this functionality by classifying occupied regions into different semantic categories and retaining a history of changes over time. This additional information layer allows for a context-aware representation of the environment, supporting the decision-making aspects of autonomous robots in dynamic and human-populated sites.

To achieve the above, inputs from both SLAM and semantic segmentation algorithms are required, which are parts of deliverables D2.3 and D2.4. The SLAM module continuously updates the occupancy grid in order to maintain an up-to-date representation of the robot's surroundings. Furthermore, the semantic segmentation module processes the RGB images in real time, and it classifies entities in the environment, which are projected onto the occupancy grid. With this information, the map can better separate stationary structures, moveable objects, and humans, which in turn leads to better localization and enhanced social-aware PP properties. Additionally, the hybrid map provides a mechanism to track the history of changes to the content of the mapped environment. This essentially means that any change in the environment's state (such as when a door opens or a box is moved) is recorded and tracked in time, allowing the system to retain a time record of each occupancy grid's cell states.

This deliverable presents the implementation of the developed hybrid map and describes the way it stores, updates, and utilizes information to support the core functionalities of LEARNER. It also details the structure and interconnection of its individual components (occupancy data, semantic classification, and history tracking).

2. ROS2 and the Standard Occupancy Grid

The Robot Operating System 2 (ROS2) is a flexible framework for developing robot software. It provides tools, libraries, and conventions to simplify the creation of complex and robust robot behavior across a wide variety of robotic platforms. A central piece in ROS2 for modeling an environment is known as the occupancy grid: a two-dimensional structure where each cell stores the likelihood of containing an obstacle. This grid-based approach equips robots with a clear way to interpret and navigate their surroundings effectively. The grid is typically structured as follows:

- **Header**: Contains metadata such as a timestamp and coordinate frame information.
- **MapMetaData**: Includes details like the map's resolution (the real-world size of each cell) and the dimensions of the grid.
- **DataArray**: A one-dimensional array storing occupancy probabilities for each cell in a row-major order.

In this structure, occupancy probabilities are usually expressed as integer values within the range of 0 to 100, where:

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

- 0 indicates the cell is completely free (no obstacle),
- 100 signifies the cell is fully occupied (presence of an obstacle), and
- -1 denotes an unknown state, meaning the cell's occupancy is not determined.

The *nav_msgs/OccupancyGrid* message type in ROS2 standardizes how environmental information is structured and exchanged. It allows robots to easily share and interpret data about their surroundings. However, although the classic occupancy grid effectively denotes which areas are free or blocked, it does not include semantic details about obstacles nor does it keep a record of how the environment has changed over time. This lack of information can limit a robot's capacity to respond intelligently in complex, rapidly evolving settings.

To overcome these drawbacks, the occupancy grid was extended in this deliverable to support both semantic categorization and temporal information. By additional data, the system gains a richer understanding of the environment, leading to improved map awareness and better interactions in dynamic scenarios.

3. Hybrid Map Model

3.1. Data Structure and Storage

Fundamentally, the hybrid map builds upon the classic grid structure found in ROS2 occupancy grid, dividing the environment into distinct cells that each represent a specific physical area. Unlike the traditional grid, where a cell only holds a value indicating the chance of occupancy, this enhanced map enriches each cell with several layers of data. In particular, every cell now includes:

- **Occupancy Status**: A Boolean indicator or probability measure that shows if the cell is free, occupied, or in an indeterminate state, following the standard format of ROS2 *nav_msgs/OccupancyGrid* message type.
- **Semantic Label**: A category assigned to the cell based on the type of object it contains, derived from deep learning-based semantic segmentation.
- **Historical Tracking**: A record with time stamps that tracks previous states of the cell, allowing the robot to monitor how the environment has changed over time.

As it can be seen, the hybrid map's base remains the standard occupancy grid, while additional layers are added to store semantic and temporal information. This design supports quick data retrieval and updates while keeping memory usage under control.

3.2. Semantic Classification of Environmental Elements

A key advancement of the proposed hybrid map over the original occupancy grid is the organization of its cells based on the kind of objects they contain. Rather than treating every entity as the same, the system recognizes structural features, movable items, and human agents separately, enabling the robot to fine-tune its navigation strategies. Deep-learning-based segmentation is employed to analyze RGB images from the robot's camera and identify objects, which are then mapped to the corresponding cells in the occupancy grid. The classification framework is divided into specific categories, each reflecting different types of objects:

1. **Stable Structural Elements**
   Structural walls, columns, and other permanent features belong in this category. Since these objects are regarded as unchanging, they serve as reliable references for localization and PP.

2. **Unstable Structural Elements**
   This group covers elements like doors, retractable barriers, or storage racks. While these fixtures tend to remain in one place, they can alter their state, for instance by opening or closing, which influences the PP and mapping processes.
3. **Dynamic Objects**
   Items such as boxes, carts, or pallets fall under this label because they often shift in location or may even vanish from the environment. Although their positions are logged for obstacle avoidance, data originating from them is not utilized for SLAM.
4. **Humans**
   Human agents receive a unique classification due to the unpredictable nature of their movements. Within the hybrid map, human activity is further broken down into subcategories, such as standing, sitting, running, gesturing, working intensely, or working lightly, to enable socially-aware PP.

3.3. History Tracking and Environment Evolution

Standard occupancy grids lack internal memory because they mainly record a single, real-time snapshot of the surroundings. The hybrid map fills that gap by including a history record in every cell, which allows the system to track changes in the environment over time. In particular, this memory system keeps:

- **timestamps** that indicate when a cell is classified as occupied or when it is no longer occupied,
- **previous classifications** for each occupied cell to identify patterns and recurring shifts in the environment, and
- **transition states**, such as a door initially recognized as closed but later recorded as open.

This temporal data is particularly useful in situations where objects frequently move but may eventually return to their original positions or when an entity's classification state (Section 3.2) changes. For instance, once a robot discovers an unobstructed passage, it can use historical data to determine whether the obstruction has a tendency to resurface in the same location in the past or if this clearance is permanent.

3.4. Integration with SLAM and Semantic Segmentation

The hybrid map is continuously updated by integrating two main sources of information. On the one hand, the mapping module of SLAM updates the grid's occupancy status by identifying which regions are occupied, free, or unknown. On the other hand, the semantic segmentation module identifies objects in the recorded images and classifies them according to the semantic categories defined earlier. This information is then projected on the 3D map and, consequently, on the enhanced occupancy grid.

A pipeline architecture governs how these two elements interact with the created hybrid map model. Firstly, the SLAM module provides updates on the occupancy grid based on real-time estimations. After processing the RGB images, the semantic segmentation module classifies the items it has found and updates the hybrid map's semantic layer. Finally, the historical tracking system records changes over time based on the previous stage of the occupancy map, allowing the robot to retain knowledge of past states.

4. Functional Impact of the Hybrid Map

4.1. Adaptive Safety Bounds

One of the most significant advantages of the proposed hybrid map model is its ability to dynamically adjust safety margins based on the classification and movement likelihood of obstacles. The hybrid map distinguishes between stable and unstable structures, dynamic objects, and human beings, in contrast to standard navigation methods that treat all obstacles as static and equally dangerous. By making use of the additional available information the robot is able to implement context-sensitive safety precautions. For instance, a wall or other structural part is not expected to shift, allowing for tight safety margins. However, a human worker or a moving cart adds a degree of uncertainty and should force the robot to increase its safety distance in order to avoid collision. This concept is further expanded due to the subcategorization of human activity which is detailed in Section 4.3.

4.2. Dynamic Path Replanning

Due to the static representations used by traditional occupancy grids, obstacles are viewed as permanent fixtures until they are manually modified. By identifying and adjusting to environmental changes, the hybrid map allows for the robot's PP system to always reassess its pre-planned paths. As occupancy grid cells may undergo classification changes over time, the system continuously tracks updates from the historical tracking layer and semantic segmentation module. To provide real-time movement optimization, the navigation module initiates a new PP event whenever an occupancy grid cell is eliminated or reclassified. For instance, a door that was closed and is now open drastically changes the amount of space that can be used for navigation.

4.3. Enhanced Human-Robot Interaction

In addition to addressing dynamic obstacles, the hybrid map permits social-aware robot behavior, which enables the system to recognize and react to human presence. By classifying their actions and affective state, the robot can modify its motion patterns, thus avoiding interferences with ongoing work, making human collaborators uncomfortable, or endangering their presence. For example, the system can differentiate human-robot interaction conditions, such as a worker sitting that imposes minimum clearance restrictions because they are unlikely to move quickly, a person performing a demanding task that should cause the robot to slow down or alter its course to keep a safe distance, or a human pointing at the robot, possibly indicating an intention to engage.

4.4. Improved SLAM and Localization Accuracy

The hybrid map model additionally introduces key improvements in SLAM, particularly in the localization process, by providing additional context regarding the stability and predictability of environmental features. The SLAM module incorporated in LEARNER relies on detecting and tracking local visual features (such as keypoints, corners, and edges) to estimate the robot’s pose over time. However, in dynamic environments where objects frequently move, this can lead to erroneous localization due to features suddenly disappearing or shifting position.

By combining historical tracking and semantic classification, the hybrid map minimizes this issue and enables the robot to discriminate between stable, dependable features and unstable, temporary ones. In particular, the algorithm can mark a feature as inappropriate for long-term localization when it is found on an object that is categorized as a dynamic element, such as a moving crate or a human. Therefore, the SLAM method prioritizes features detected on stable structural elements (e.g., walls, pillars, or permanent installations) rather than depending on these unstable landmarks. Furthermore, loop closure detection, a critical SLAM procedure that fixes long-term localization issues by identifying previously visited sites, is strengthened when it is based on stable landmarks. Since the hybrid map records past environmental changes, the system can avoid false loop closure detections that could occur from objects being present in different parts of the environment than the ones they were originally recorded. Moreover, an additional prevalent challenge in SLAM is map corruption, which occurs when dynamic factors (such as people moving or objects shifting) cause irregularities in the map that is produced. By classifying these elements and tracking their past states, the hybrid map ensures that only reliable, persistent structures contribute to the long-term environment representation. Lastly, the created hybrid map enables the system to differentiate between genuinely structural alterations and momentary changes in situations where a robotic agent operates in the same area for several sessions. This prohibits transient objects that have already been registered from misleading the localization system in subsequent sessions.

5. Conclusions

LEARNER's proposed hybrid map model is an essential component since it offers a structured approach for combining occupancy information, semantic classification, and historical tracking. In this manner, it paves the way for adaptive navigation and better localization in dynamic environments by differentiating between various object categories and preserving knowledge about environmental changes over time.

Our future steps involve fully integrating the hybrid map with the social-aware PP and SLAM modules. To improve pose estimation in dynamic situations, the SLAM system will filter out local features originating from dynamic objects using the semantic and historical layers of the map. In the meantime, semantic segmentation and human activity classification will be incorporated into the PP module to dynamically modify navigation behavior.