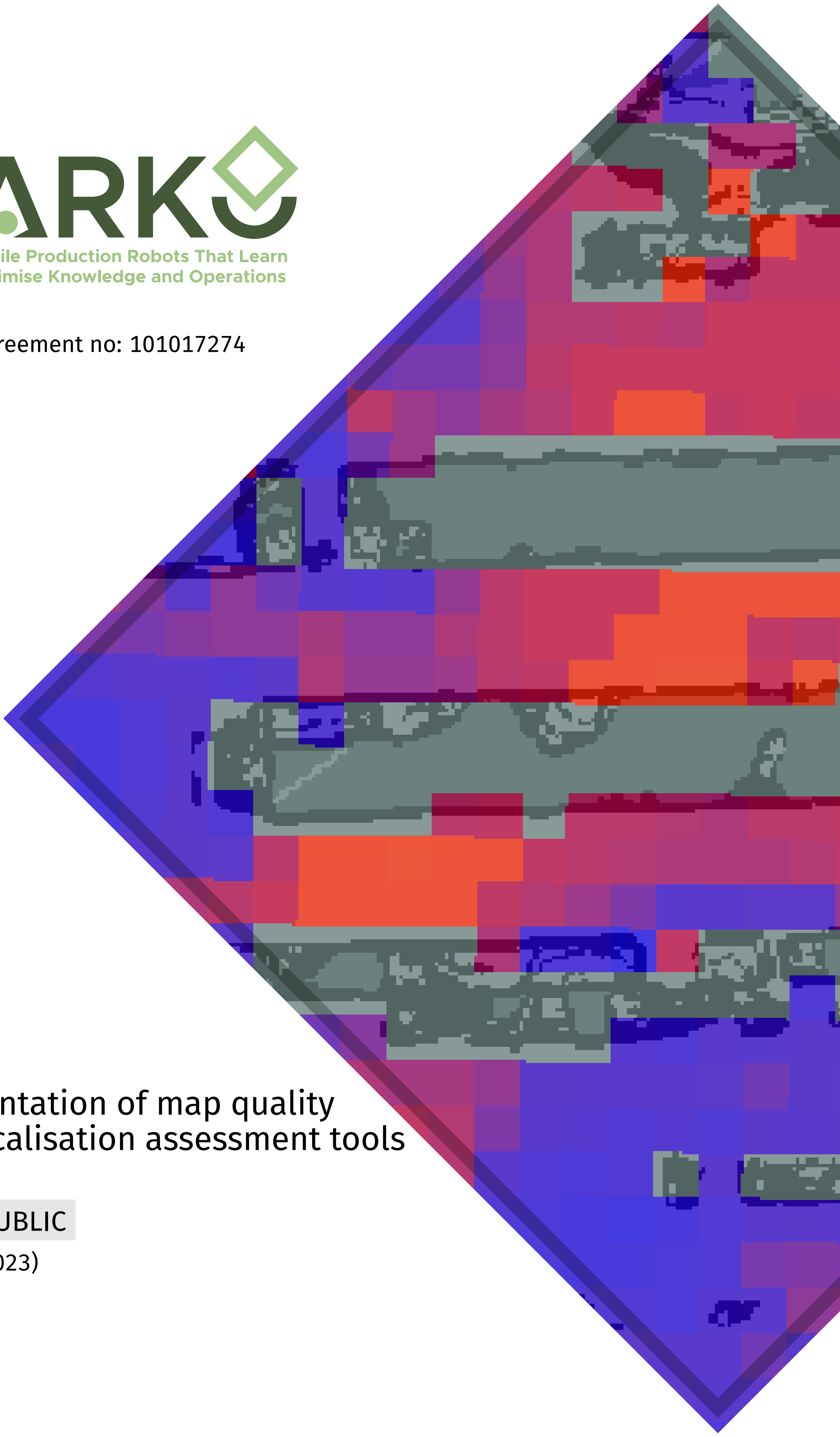
DARKO

Dynamic Agile Production Robots That Learn and Optimise Knowledge and Operations

H2020-ICT-2020-2 Grant agreement no: 101017274



# DELIVERABLE 3.2

## Prototype implementation of map quality assessment and localisation assessment tools

Dissemination Level: PUBLIC

Due date: month 29 (May 2023)

Deliverable type: Software

Lead beneficiary: ORU

# Contents

# 1 Introduction

This report describes the software deliverable D3.2: the first implementation of map quality assessment and localisation assessment functionalities for task T3.4. Our prototype framework for reliability-aware mapping and safe localisation at this point in the project includes software for scan registration quality assessment (section 2), localisation quality assessment for anticipating and mitigating inaccurate localisation (section 3), as well as a more generic map quality assessment module (section 4) that labels "implausible" areas of a map.

Some of the contents of this report related to registration quality assessment and localisation quality assessment were reported already in D3.1, although they were not planned to be ready until this D3.2. This report includes some more detail and information about the integration into the complete DARKO software/hardware implementation. Notably, the section on map quality assessment (section 4) is completely new for this deliverable.

# 2 Registration quality assessment

Scan registration is a central part of both the mapping and the localisation pipeline described in D3.1. As such, automatic assessment of the result of scan matching is important for detecting and mitigating errors and to improve mapping and localisation. Our current version of registration quality assessment is the CorAl method of Adolfsson et al. [1]. The current implementation is a refactoring that, compared to the original version [2], includes extensions to 2D range data, which is applicable to the 2D range data from the Sick safety scanners on the DARKO robot platform.

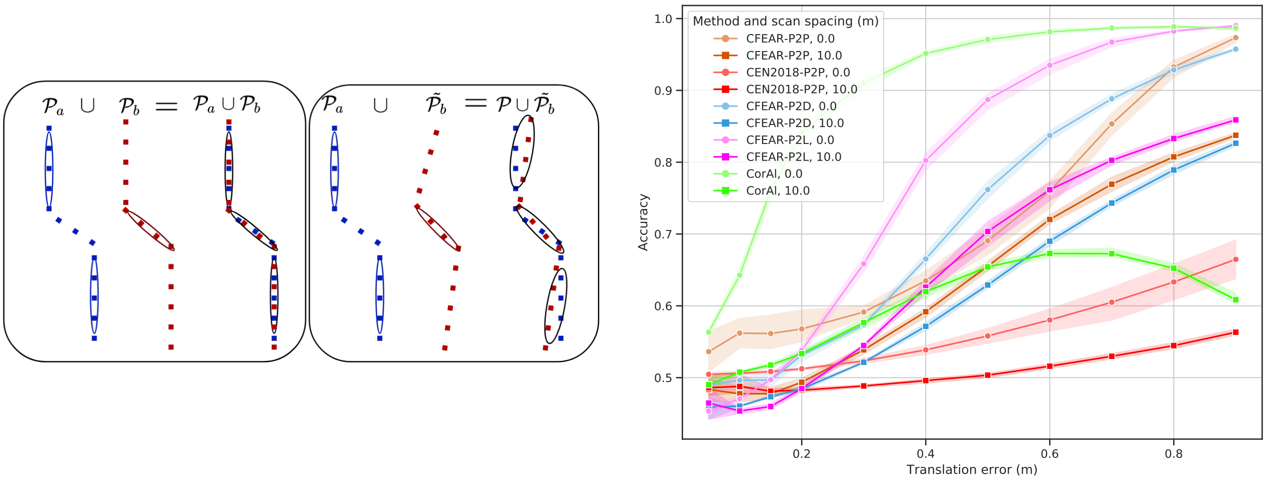
CorAl computes the average differential entropy in two point clouds, comparing the local point entropy in each point cloud separately to the union of the point clouds. See fig. 1a. A key idea is to estimate the entropy inherent in the scene from the entropy in the separate point clouds, which enables CorAl to accurately assess quality in a range of different environments. The decision boundaries between aligned vs non-aligned point clouds can be learned in a self-supervised fashion from accurately aligned scans with poses.

The key features of the CorAl method, compared to alternative methods for scan alignment assessment, is that it (1) provides a simple and intuitive measure of alignment correctness between point cloud pairs that generalize well, and as such (2) can highlight regions that indicate misalignment, and (3) employs self-supervised learning of decision boundaries. Our recent publication [1] contributes further systematic evaluations, including an ablation study of parameter importance, comparisons to four aditional baselines, and a new cross-environment study. Figure 1b shows results as classification accuracy vs the alignment error, compared to a set of baselines on the Oxford RobotCar dataset. Small errors are generally the hardest ones to detect.

We believe that CorAl has great potential to serve as an alignment quality tool for point clouds in general and can improve localisation robustness by equipping odometry, relocalisation, and loop closure systems with the capability of introspectively detecting small errors in diverse environments. For example, the method has since been included in the TBV-SLAM framework [3].

# 3 Localisation quality assessment

Precise localisation is key to most mobile-robot systems, not least those that are deployed in industrial settings. However, even state-of-the art lidar-based systems may fail or lose




**(a)** Uncertainty (entropy) is preserved when joining aligned point clouds $\mathscr{P}_a \cup \mathscr{P}_b$ *(left)*, but increases when joining misaligned point clouds *(right)*. The entropy for aligned point clouds should be similar to the entropy in the separate point clouds and can be used when quantifying alignment quality.

**(b)** Classification accuracy vs translation error for consecutive scans (0.0) and scans taken at least 10 m apart (10.0), respectiveley. CorAl is the most accurate of the evaluated methods in particular for consecutive scans and can detect small alignment errors of 0.3 m with > 90 % accuracy. (CFEAR-P2L requires 0.2 m larger errors to reach similar level of accuracy.)

**Figure 1:** The CorAl method for registration quality assessment (from Adolfsson et al. [1]).

accuracy, in particular in feature-sparse environments (e. g., fully stacked warehouse aisles or transport corridors).

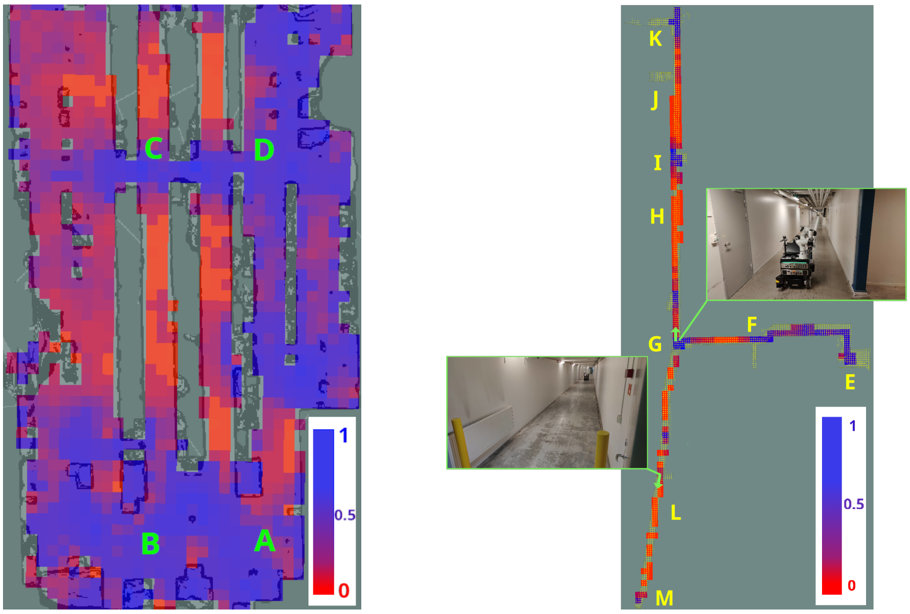
Our aim is to be able to predict localisation risk (i.e., the risk of generating inaccurate pose estimates) and account for it by taking preemptive measures; e.g., such that a planner can generate "risk-aware" paths that takes both the risk of inaccurate localisation and the path length into account.

Some recent methods attempt to determine whether the current sensor view contains sufficient features to be matched to a map [4]. However, the potential loss of localisation may be overestimated by such measures as they do not consider the temporal filtering aspects inherent in localisation methods like Monte Carlo localisation. Our proposal [5] is based on the concept of *alignability*, which represents the estimated capacity of a given range scan to be aligned other ones – prior to attempting registration, in contrast to CorAl (section 2) which assesses pairwise alignment "after the fact". CorAl and related methods are useful for *detecting* localisation and mapping issues but cannot be used useful for *proactively* avoiding inaccurate localisation.

In order to predict the quality of localisation, we contribute here with *localisation risk maps*, the first implementation of which being an *alignability map*. (For future iterations we intend to also include dynamics as a localisation risk factor.) Our approach aims to capture localisation risk spatially and serves to predict where in a map localisation may be less accurate. Example alignability maps are shown in figs. 2a and 2b.

We have presented these alignability maps and how they can be used for risk-aware motion planning in two IROS workshops [8, 5]. The main contributions are summarised below.

An alignability map in our case is a 2D grid map in which each cell represents the expected alignability that can be obtained from different scans within that area. The definition of our alignability map is based on different measures of alignability: the one proposed in Nobili et al. [6] and another set of metrics based on the 3D-NDT represen-




**(a)** Warehouse environment. Aisles that lack geometric features are marked as less alignable (red shades) indicating higher risk of localisation error.

**(b)** Corridor environment. As expected, alignability is low in the featureless corridors, but higher near the end of corridors and when passing other passages.

**Figure 2:** Two examples of *alignability maps* for prediction and mitigation of localisation error. Red regions indicate feature sparsity, which entails a higher risk of inaccurate localisation

tation [7]. The former measure aims to quantify the variety of surface normals present in a given scan: the higher this variety is, the higher the alignability and, therefore, the lower the localisation risk. On the other hand, NDT-based metrics are defined on the covariance of the pose estimate from the scan registration optimization problem based on the D2D-NDT function. The lower this uncertainty is, the better the expected alignment between point clouds and, intuitively, the lower the risk of localisation error.

Our quantitative experiments show that alignability can be used as an indicator of localisation error and we validate, with Granger causality tests, that it also serves to *anticipate* the occurrence of errors. Figure 3 shows time series of Nobili alignability and localisation error when driving several rounds in the warehouse environment shown in fig. 2a. We have marked in this plot every time the error gets noticeably high and also all events of extremely poor alignability. Each peak of localisation error always takes place some time after the events of poor alignability. Thus, we can visually assess that the alignability metrics are useful to predict localisation errors. We have also demonstrated this more formally through statistical Granger causality tests. In all Granger tests (each based on a Pearson $\chi^2$ test with a significance level of 0.05) the null hypothesis $H_0$ was confidently rejected ($p$-values are around 0.02 for Nobili's metric and below $2 \cdot 10^{-18}$ for the NDT-based ones) meaning that all the alignability metrics *Granger-cause* the localisation error. More details can be found in Castellano-Quero et al. [5].

Figure 4 demonstrates the use of alignability maps in motion planning. In these examples, we have constructed a cost map from the alignability map in fig. 2a in order to influence the motion planner to avoid regions with extremely low alignability. We shwo the results of four experiments in which the robot follows a path between two waypoints. Each experiment has been repeated two times, leading to different paths (see figs. 4a and 4b). The paths that ignore alignability costs (red) are the shortest possible, while the

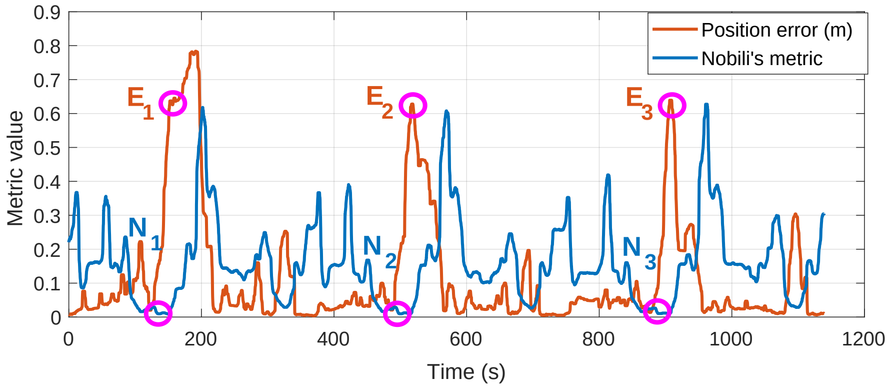



**Figure 3:** Localisation error vs alignability. The three events marked as E correspond to localisation errors higher than 0.6 m, the ones labeled as N to Nobili's alignability values lower than 0.01. There is a clear visual correspondence between low-alignability events and high-error events, as also quantified by Granger causality tests (described in the text). (The time series have been filtered by using a moving median filter with 6 seconds of window size.)

other ones (blue) tend to avoid the two central corridors, which exhibit lower alignability (see fig. 2a). We have also measured the localisation errors for each path (see fig. 4c), finding that those generated by considering alignability costs always suffer from lower positioning errors.

We have integrated these alignability maps as cost maps for the DARKO risk-aware motion planner. The integration of alignability maps into DARKO's planning framework is further described in D6.3. Prototype software for alignability mapping is available below. First, the implementation of Nobili's alignability metric can be found in:

```
mapping/alignability:
type: git
url: https://gitsvn-nt.oru.se/software/alignability.git
version: master
```

The implementation of NDT-based alignability metrics is available in:

```
ndt_core_public:
type: git
url: https://gitsvn-nt.oru.se/software/ndt_core_public.git
version: nice-devel
```

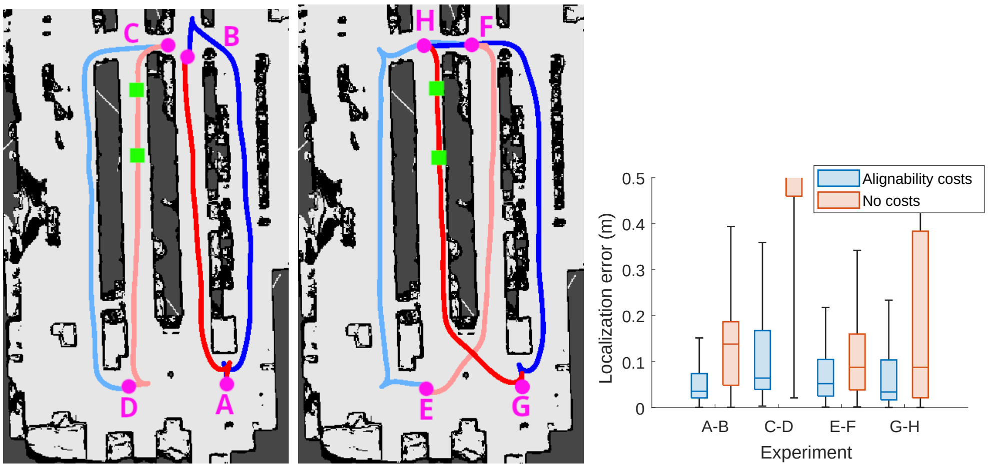
Finally, the implementation of our alignability map is available in:

```
risk_map:
type: git
url: https://gitsvn-nt.oru.se/darko/software/risk_map.git
version: master
```

which relies in the previous two packages and that will be updated in the remainder of DARKO to integrate further layers of sources of localisation risk, e. g., dynamics.

## 4 Map quality assessment

Our prototype implementation for *reference-free map quality assessment* is based on a variational autoencoder that can assess 2-D occupancy grid maps. Occupancy grid maps are widely used in the robotics landscape, including DARKO, and offer a convenient way




**(a)** Planning A–B and C–D. **(b)** Planning E–F and G–H. **(c)** Localisation errors obtained while executing the paths in figs. 4a and 4b.

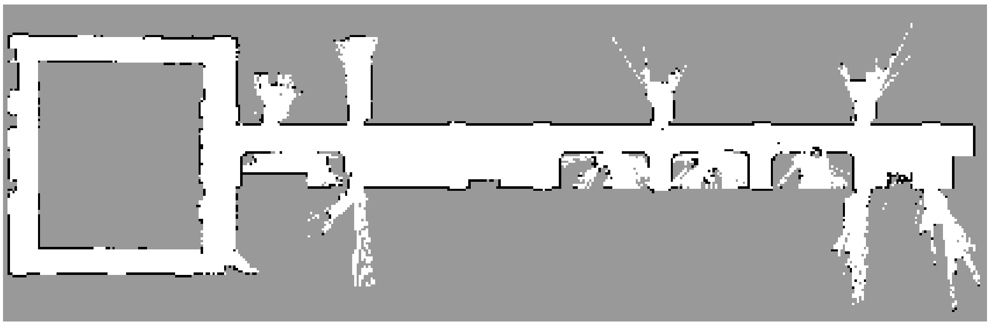
**Figure 4:** Motion planning experiments with localisation risk maps, in the environment shown in fig. 2a. *Blue* trajectories represent paths generated by considering alignability costs, while *red* ones do not consider them. Green squares delimit a region with very low alignability. Please note that the sensor range has been cropped to 6 m here in order to provoke larger localisation errors.

to leverage image based learning methods. In occupancy grid maps there can be regions that misrepresent the environment, for example where a wall is too thick, or regions that are only partially explored and the map is incomplete in those areas. In practice, maps are usually assessed qualitatively by a human expert, but these assessments are not easily reproducible and tend to vary greatly between people. Furthermore, assessing all of the regions of a map by a human is a time consuming task. With this variability in mind, a data driven method is chosen to limit potential sources of errors and to ensure a reproducible and repeatable process to map quality assessment.

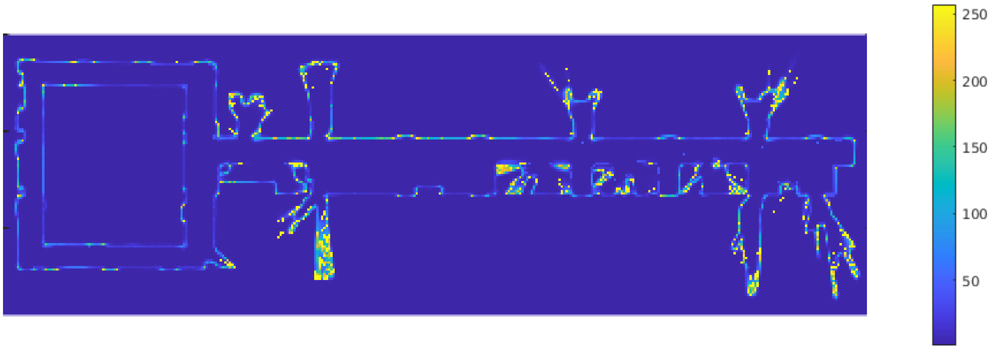
Occupancy grid maps are composed of cells that represent the probability of being occupied where the value of a cell is in the range of $[0,1]$ where a value of 0 or 1 represents a high probability of the cell being free or occupied, respectively. The proposed method remaps the cells of the map such that the cell can be one of three classes, **U**nknown, **O**ccupied and **F**ree with corresponding value 1, 0.25 and 0, respectively.

One common issue with building models is the availability of ground truth or labelled data which can be time consuming to generate (as also noted by, e. g., Andersone [9]). Our proposed approached uses an autoencoder, as a type of unsupervised learning, to learn an encoding of maps. An autoencoder first encodes the input data by transforming the data, usually to reduce the dimensionality, and then decodes the data, attempting to recreate the input data from the encoding. For occupancy grid maps, the input data would then be patches of the map with size $N \times N$, along with tunable patch constraints such as the minimum number of occupied cells $\tau_O$ per patch and the maximum number of unknown cells $\tau_U$ per patch. In the current prototype implementation, prior knowledge of the working environment (such as corridor width, expected wall thickness and cell size) should be considered when setting these parameters, in order to feed the encoder patches that have enough structure that it should learn.

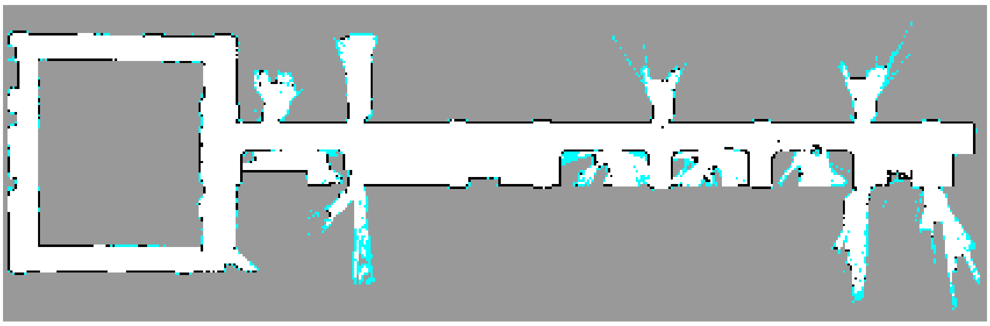
The data driven aspect of the proposed approach hypothesises that for any given map, the majority of the map accurately represents the environment and that only a small portion poorly represents the environment. An autoencoder trained on patches of the map



**(a)** Input occupancy grid map.




**(b)** Reconstruction loss $L$, using the trained encoder to reconstruct the occupancy grid map from patches extracted from the map used for training.



**(c)** Grid map where cells with a loss $L > \tau_L$ are highlighted with blue, indicating cells that are of concern.

**Figure 5:** Example occupancy grid map for map quality assessment. Source: Pre-2014 Robotics 2D-Laser Datasets (http://www.ipb.uni-bonn.de/datasets/): Seattle UW (D. Haehnel).

that meet certain criteria should converge on an encoding that captures the majority of the map that represents the environment accurately and the poor sections will be poorly represented by the encoding. We can then analyse the reconstruction loss $L$ of a patch by comparing the input patch cell values with the ones generated using the encoder. Cell values with a higher reconstruction loss should then correspond to regions with potential map errors.

The proposed method has been evaluated using three occupancy grid maps for this deliverable, which are given in figs. 5 and 6. Patches extracted from the map shown in fig. 5a with size $N = 28$ and meeting the requirements $\tau_O = 5\%$ and $\tau_U = 45\%$ were used to train an autoencoder. In other words, patches with at least 5% occupied cells and at most 45% unknown cells were chosen to be of interest to train on, in order to avoid training on a large number of all-free or all-unknown map areas.

The resulting reconstruction loss of the entire map, generated by applying *all* patches of the map that do not have all the cell values with the same value, is shown in fig. 5b. A cutoff value of $\tau_L = 132$ was chosen by assessing the reconstruction loss. Any grid cell with a value larger than $\tau_L$ would be deemed to be incorrect. Figure 5c illustrates which cells of the trained map would then be considered incorrect.

Figure 6 shows two more examples of applying the encoding learned on fig. 5a with the same threshold $\tau_L$, highlighting the ability of the developed method to successfully encode a map and the ability to transfer to other occupancy grid maps of different environments but with similar structure.

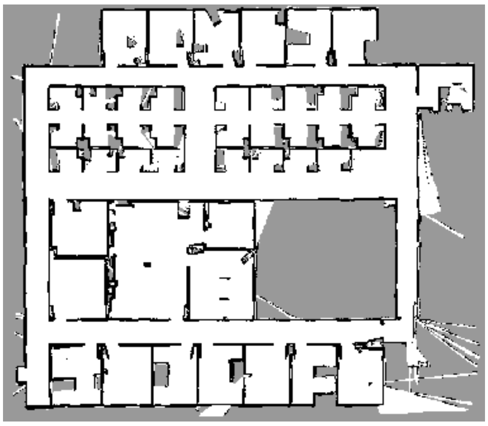
In future work, we aim to include self-supervised learning in new environments by incorporating online sensor data, to learn suitable values for the tunable hyperparameters, and to extend the method to three-dimensional data.

## 5 Summary

In this report we have outlined the current implementation of the prototype software used for reference-free map quality assessment and localisation quality estimation contributing to DARKO's Objective 4 ("risk-aware operation for safety and efficiency").

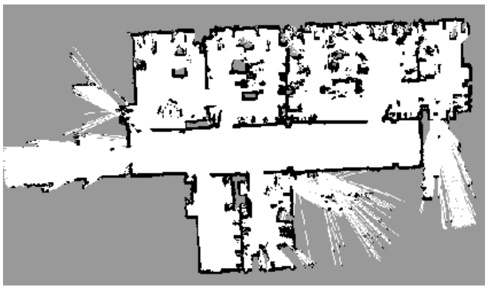
For localisation quality assurance we have implemented a first version of (on-line) *localisation quality assessment* by means of the CorAl scan alignment measure (section 2) and (off-line) *localisation risk maps* based on alignability (section 3). We have shown how these localisation risk maps can be used to predict and to mitigate localisation error.

For map quality assessment we have implemented a variational autoencoder (section 4) that learns the expected appearance of 2D grid maps in an unsupervised fashion and thus is able to mark "suspicious" regions of the map. Our preliminary experiments indicate that this approach is useful for highlighting low-quality areas of a map, due to mapping errors (e. g., registration mistakes or missed loop closures) or due to missing coverage during mapping.

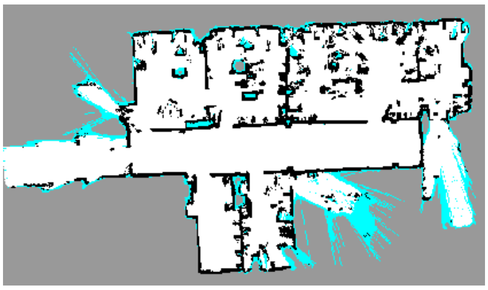


**(a)** Second occupancy grid map used to test the propose methodology. Scale unknown. Source: Radish: Robotics Research Datasets (Wagan et al., 2008): sdr_site_b (A. Howard) https://dspace.mit.edu/handle/1721.1/62245.

**(b)** Second occupancy grid map with regions highlighted with the colour blue indicating cells that are of concern.



**(c)** Third occupancy grid map used to test the propose methodology. Scale unknown. Source: Radish: Robotics Research Datasets Albert B (C. Stachniss, 2010) https://dspace.mit.edu/handle/1721.1/62291.



**(d)** Third occupancy grid map with regions highlighted with the colour blue indicating cells that are of concern.

**Figure 6:** Further examples of map quality assessment using the variational autoencoder proposed in section 4.

# References


[1] Daniel Adolfsson, Manuel Castellano-Quero, Martin Magnusson, Achim J. Lilienthal, and Henrik Andreasson. "CorAl: Introspection for robust radar and lidar perception in diverse environments using differential entropy". In: *Robotics and Autonomous Systems* (2022), p. 104136.

[2] Daniel Adolfsson, Martin Magnusson, Qianfang Liao, Achim J. Lilienthal, and Henrik Andreasson. "CorAl – Are the point clouds Correctly Aligned?" In: *Proc. of the Europ. Conf. on Mobile Robotics (ECMR)*. Sept. 2021.

[3] Daniel Adolfsson, Mattias Karlsson, Vladimír Kubelka, Martin Magnusson, and Henrik Andreasson. "TBV Radar SLAM – trust but verify loop candidates". In: *IEEE Robotics and Automation Letters (RA-L)* 8.6 (June 2023), pp. 3613–3620. arXiv: 2301.04397 [cs.RO].

[4] Julian Nubert, Etienne Walther, Shehryar Khattak, and Marco Hutter. "Learning-based Localizability Estimation for Robust LiDAR Localization". In: *Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*. 2022.

[5] Manuel Castellano-Quero, Tomasz Piotr Kucner, and Martin Magnusson. "Alignability maps for the prediction and mitigation of localization error". In: *IROS 2023 Workshop on "Closing the Loop on Localization"*.

[6] Simona Nobili, Georgi Tinchev, and Maurice Fallon. "Predicting Alignment Risk to Prevent Localization Failure". In: *Proc. of the IEEE Int. Conf. on Robotics and Automation (ICRA)*. 2018, pp. 1003–1010.

[7] Todor Stoyanov. "Reliable Autonomous Navigation in Semi-Structured Environments using the Three-Dimensional Normal Distributions Transform (3D-NDT)". PhD thesis. Örebro University, 2012.

[8] Manuel Castellano-Quero, Tomasz Piotr Kucner, and Martin Magnusson. "Alignability maps for ensuring high-precision localization". In: *13th IROS Workshop on Planning, Perception, Navigation for Intelligent Vehicles*.

[9] Ilze Andersone. "Quality Evaluation of the Occupancy Grids without Ground Truth Maps". In: *Proceedings of the 12th International Conference on Agents and Artificial Intelligence - Volume 1: ICAART,* INSTICC. SciTePress, 2020, pp. 319–326.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101017274