INTELLIRISK2: RISK-AWARE EXPLORATION THROUGH ACTIVE LEARNING, ANOMALY DETECTION AND EPISODIC MEMORIES

Lennart Puck, Tristan Schnell, Timothée Buettner, Arne Roennau, and Ruediger Dillmann

FZI Forschungszentrum Informatik, 76137 Karlsruhe, Germany

ABSTRACT

Exploring different planets yields a multitude of hazards for the robotic explorers. To mitigate the risks imposed on the robots, intelliRISK2 proposes a risk-aware exploration framework. This framework includes different learning strategies to both gain a quick understanding of the surroundings, as well as learning from previous experiences. To adapt and understand the surroundings, active learning strategies and external anomaly detection are applied. During the runtime, the health of the robot is continuously monitored through an internal anomaly detection. Furthermore the robot learns from previous experiences by incorporating the corresponding knowledge into an episodic memory. The combined strategies allow the autonomous robots to perform risk-aware exploration missions and adapt to various situations

Key words: Planetary Exploration, Machine Learning, Walking Robots.

1. INTRODUCTION

Exploration missions into the unknown are a complex and daunting task for robots. To safely navigate in unknown environments, the robots need to both evaluate their surroundings as well as their own current health status. In the context of planetary exploration, the missions can only be partially simulated on earth due to different constraints, such as gravitational force and material composition of the surroundings. To minimize the potential hazards to the robot, mainly safe and conservative decision are taken, this however limits the potential of the robot. Previously in intelliRISK [1], the knowledge of the system was modeled by hand and only known surroundings and robot behavior were encoded. To enhance this, three different learning strategies are applied to adapt to novel situations and learn from previous experiences: Active Learning, Anomaly Detection and Episodic Memory. The anomaly detection is applied both to the internal assessment of the robot as well as the classification of the surroundings. The active learning component is utilized to improve the segmentation of the surroundings, while the episodic memory is applied to learn from previous

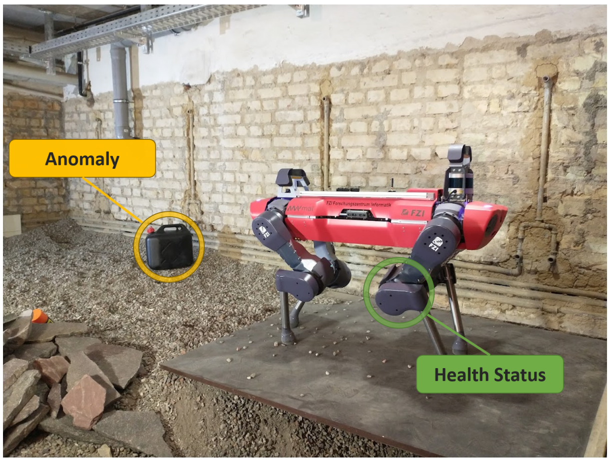

Figure 1. ANYmal C [15] in the test environment. The robot detects anomalies in the environment and monitors its current status during the mission. The learned experiences are encoded in the episodic memory of the robot.

situations and states of the robot.

To determine hazards from the surroundings a twofold approach is used, classification of known ground types and novel impressions. Through this, the robot can gain an in depth understanding of its surroundings. To continuously improve the semantic segmentation an active learning approach is applied. The robot can independently, select images which it thinks would benefit the learned model and is provided by a human operator on a base station with annotations. This is further enhanced through an ensemble-based anomaly detection: when the robot detects a novelty, it potentially has detected a hazard. Furthermore, anomalies can be regarded as interesting exploration goals, since they would lead to new insights. Thereby, exploring them would increase the knowledge gained throughout the mission.

The safe and robust execution requires a self-assessment of the robot, where it can react to known situations as well as novel circumstances. An unsupervised anomaly detection approach is utilized to determine defects or novel behavior of the robot. With this approach the robot can react to changes through damages or different payloads. This gained knowledge can then further be incorporated into an episodic memory, which allows the robot to recollect experiences from the past to gain insights on the best behavior in the current circumstances. Furthermore, by extending the members of the mission to a heterogeneous team, the robots could negotiate which robot is currently best suited to carry out a specific mission goal.

Both the external and internal assessment lead to a robotic system which has the capability to adapt to novel situations in an exploration mission. The results of the work can lead towards a robust and risk-aware path-planning which continuously reassesses the robots health. Thereby, allowing an autonomous, yet safe exploration in unknown areas.

The paper is structured as follows, first an overview on the state of the art is given. Afterwards the core concepts of intellIRISK2 are detailed. We are going to highlight specifically the active learning, anomaly detection and episodic memory components of the presented stack Finally, the concept of orchestration is provided before the conclusion and outlook are highlighted.

2. RELATED WORK

State of the art as presented here will be subdivided according to the three categories active learning, anomaly detection and episodic memory.

In the context of active learning, the foundations are highlighted by Settles in [3]. The author, provides a detailed summary of different active learning strategies, query generation and data collection. Most approaches rely on pool-based strategies in contrast to the stream-based approach in intelliRISK2. Yang and Loog [4] show in their overview of different active learning strategies that even simple and even random query selection can compete with more complicated selection methods for pool-based queries in the context of logistic regression. For semantic segmentation tasks Mackowiak et al. [5] propose an approach based on estimated information gain and annotation cost to find interesting pixels in an image. The queries are based on patches of an image, an approach on complete images is proposed by Tan et al. in [6]. The authors apply the information gained from an edge detection algorithm on the image to the query strategy. Both Mackowiak et al. [5], and Tan et al. [6] take approximately 17% of the images of their respective pool to achieve an accuracy of 95% compared to the baseline. However, most approaches are tackling pool-based methods and are missing the transfer to stream-based policies.

Anomaly detection is the task of finding unusual patterns in data. One of the first studies is by Silverman [7] and is based on density estimation. A detailed review of existing approaches and key studies was conducted by Pimentel et al. in [8]. An early approach for novelty detection in the surroundings is proposed by Marsland [9], the author uses self organising networks in an inspection task. A risk minimizing approach is shown by Sofman et al. in [10]. The authors create models of the environment to detect changes in repeated execution. Through encoding knowledge of successful missions and constant monitoring they are able to minimize risks. For internal fault detection a health estimation needs to be performed. Buderath et al. [11] show the requirements for structural health assessment in UAVs. A data based unsupervised machine learning approach for fault detection is proposed by Amruthnath and Gupta in [12].

A formal design for episodic memory is proposed by Stachowicz et al. in [13]. The recall of such episodic memories for robot action execution is detailed by Rothfuss et al. in [14]. The authors utilize a deep neural network to implement the episodic memory. With this they are able to encode and recall previously learned visual episodes. The integration of the active learning and anomaly detection into an episodic memory and an overall mission orchestration is currently missing. IntelliRISK2 tries to solve the open questions and provide an overall risk-aware exploration framework.

3. RISK-AWARE EXPLORATION

IntelliRISK2 can be roughly divided into two major parts: the introspection and the external perception. The complete framework is shown in Figure 2. Both areas are tackled with different approaches. For the introspection fault detections through anomaly detection and reflection through episodic memories are key components. The perception utilizes active learning components and anomaly detections on a distributed setup with a base station to adapt throughout the mission. Through the setup with both self and environment awareness the capabilities of the robot are vastly increased and risk aware missions are possible.

In the following sections first the external perception through active learning and anomaly detection is described before the introspection through episodic memories is detailed. Finally, the whole interacting system and mission control via behavior trees is explained.

3.1. Setup

The main robot which is utilized in intelliRISK2 is ANYmal C [15]. The four legged robot is capable of navigating across rough terrain and tackle difficult challenges. Through the open setup using ROS [16] as an underlying framework, it is possible to get both external and internal states of the robot. Especially the internal states are utilized to analyze the robots own current health status.

The robot is equipped with three CPUs, however it is missing a GPU. Therefore, all deployed systems have to able to run with low computational burden. To train the according networks a base station is employed. There,

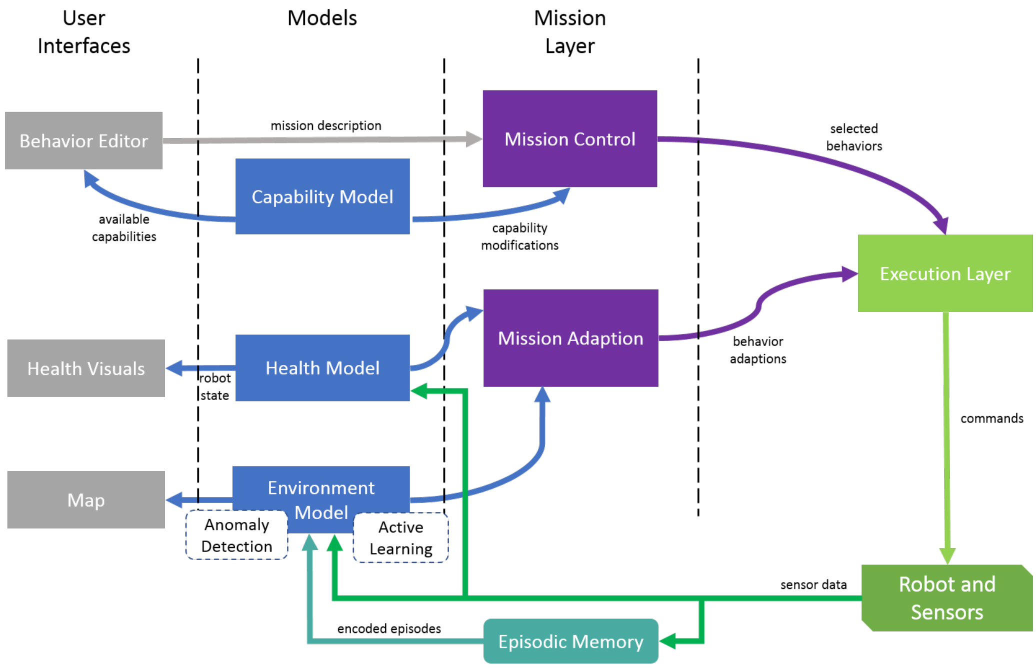

Figure 2. The general setup of the intelliRISK2 stack. On the left the different user interfaces are specified, these include a behavior editor for the behavior trees, visuals for the health and the map. The internal models of the robot can be divided into the capability model, the health model and the environment model. On the environment model the external anomaly detection and the active learning is executed. In the mission layer, both the mission control and adaptation is performed, here the behavior trees are generated and can be adapted during runtime. Afterwards the missions are executed and the live data is returned into the health and environment models.

the data acquired by the robot is stored and used in different training scenarios before the models are deployed onto the robot.

The sensors used for the perception are the Velodyne Puck Lidar for the mapping and Intel Realsense D435 for the segmentation and external anomaly detection. Google Cartographer is used for SLAM applications [17]. Since Cartographer only provides a 2D output (including 3D localization), a 3D mapping solution was built on top of it. To this end, the VDB Mapping [18] approach is used. This allows for a real-time capable mapping of the sensor data into a voxel representation. The map enables additional data to be stored alongside the occupancy values and can be utilized for 3D navigation. In the context of intelliRISK2, a risk-aware 3D Navigation will be developed.

Furthermore the software stack, even though primarily used on ANYmal C, can be easily transferred to different robotic systems. For this, sensors providing the similar ROS topics containing lidar and image data are required.

3.2. Active Learning

Understanding the surroundings is vital to find safe paths through the robots surroundings. However, exploring unknown areas with only sparse previous knowledge is a difficult task for any autonomous system. To help the robot in these situations, an active learning environment was set up, where a human operator can collaborate with the robot through a base station.

The active learning framework is distributed between the robot and a base station. On the robot, the continuous camera stream is passed into a semantic segmentation. The results of the segmentation are then analyzed by an active learning node. Once this node deems the segmented image as promising for further training the original image alongside the segmentation is passed into a buffer for storage [2]. The human operator can request batches of images from the buffer for annotation from the base station. As soon as the images are annotated, they are incorporated into the training data and a new model can be trained and then deployed onto the robot. This communication cycle can be utilized throughout the whole mission to continuously improve the model of the semantic segmentation. To account for changing selection thresholds with new models an adaptive threshold is introduced. Furthermore, since the images are stored in the buffer and come from a continuous data source, a filter asserts that these images in the buffer have less similarity. A detailed review on a small custom dataset and existing drone datasets is performed in [2]

By using the human for annotation in the loop, the robot is guided to a better understanding of the surroundings, while simultaneously the annotation overhead is reduced. This approach can further be extended to utilize different

Figure 3. The anomaly detection with live data from the robot. On the left side, the input image is shown. On the right, the corresponding segmentation mask of the ensemble based anomaly detection is highlighted.

sensors to gain a robust segmentation. It is even possible to have a team of robots, trained with different datasets, cooperate to label the data of each other to increase the overall results.

3.3. Anomaly Detection

The anomaly detection can be divided into two parts: external novelty detection and internal fault detection. For the novelty detection, the camera stream is analyzed by an ensemble of neural networks to find previously unseen structures and environments in the surroundings. Each network is an auto-encoder trained on a different dataset. By combining different input modalities such as RGB and depth information and utilizing the versatile training data, the ensemble can outperform the individual networks.

To achieve the best results at the beginning of a new mission, data from aerial images or previous missions can be used. By combining the different models where each focusses on different aspects and features of the data, the overall precision of the anomaly detection can be increased. Furthermore, to enhance the system, data can be acquired during the runtime of the mission and after training on the base station, the new model can be added to the ensemble. This allows for a continuous learning and improvement of the system, similar to the active learning approach.

The applied loss function for the training of the individual autoencoders is a pixelwise mean squared error over the processed image patch. For ease of use and to keep the network parameter size small, images are divided into patches of 64x64 px to utilize the higher resolution of the complete image. On each patch the squared error is used as reconstruction error to determine if a pixel is anomalous. After the detection the patches are stitched together again and the results are given on the complete image. In a post-processing step only areas with pixel clusters are kept, thereby eroding single, anomalous pixels without classifying them as anomaly.

By combining multiple neural networks with different learned weights the results can be improved. Depending on the context, the different networks can be combined with a weighted mean, as described in Equation 1.

$$\bar{p} = \frac{1}{E} \sum_{i=1}^{E} w_i s_i \tag{1}$$

Where $\bar{p}$ describes the anomaly score of the pixel overall different ensembles. E is the number of networks in the current ensemble, and i describes the individual networks. Therefore, the individual score(s) of a network is calculated and weighted (w). The weights can be adapted according to the certainty of the individual networks depending on each context. With this setup the robot is capable of detection novel surroundings and thereby, find interesting sites worth exploring.

The internal fault detection is implemented in two different approaches: the first relies on Gaussian Mixture Models (GMMs), whereas the second utilizes a bidirectional Wasserstein GAN. Both approaches are used to estimate the health of the robot. The core idea behind both systems is to accurately asses faults in the internal behavior of the robot, both implement an unsupervised fault detection system.

The GMMs are fitted to the healthy recorded data (representing nominal operation) using the Expectation Maximization algorithm. Afterwards the continuous data stream can be scored using the GMMs. Here, the average logarithmic probabilities of the data points are computed. When the score decreases, a deviation from the normal state is detected, thereby showing a decrease of the health of the robot [19].

The updated fault detection approach uses a bidirectional Wasserstein GAN [20]. Here a small amount of the non-anomalous data is utilized to train the model. Since the multivariate data is highly dependent the latent space of the bidirectional Wasserstein GAN can be of a low dimension. In the work different encoding ranging from 2 to 20 latent variables are applied. Furthermore, different variants of the system are compared against each other. The first consists of fully connected linear layers with LeakyReLu activation function. The second is based on a long short-term memory (LSTM) network. As a third variant, a convolutional approach is implement. Finally the second and third approach are combined to a more complex approach. The results show that especially the LSTM based approaches show a stable anomaly scoring over time. However, the other approaches show a clear distinction of anomalies as well. Thereby, showing a robust detection of anomalies with different models.

Both internal fault detection approaches are evaluated on ANYmal C. The introduced faults are weights attached to a single leg. Hereby, the GAN framework can consistently detect and localize even small anomalies during the robots runtime.

3.4. Episodic Memory

A memory function was designed to assess the risks encountered by the robot in earlier episodes or missions. This allows the robot, to scan for similar episodes in its past and remember the outcome and performed actions.

The episodic memory is implemented as a short term memory ring buffer which gets persistently stored when one of the previous mentioned internal anomaly detections finds a fault. Thereby generating a long term memory of the sensor data before an error. Through a Fast Fourier Transform (FFT) the current sensor data can be compared with the previously stored episodes. Therefore, the robot can learn through its memory to avoid potentially dangerous episodes in future scenarios. Potentially a single episode which did impose harm to the robot can lead to the robot avoiding similar situations in the future. In the episodic memory, both internal sensor data as well as external perception can be included. Therefore, a multitude of sensory data can be processed to evaluate the episodes and choose an action accordingly.

3.5. Risk Map

Risk-aware navigation requires in-depth knowledge of the current surroundings. To find paths for safe navigation, an understanding has to be provided to the respective path planners. To this end, a map with a risk understanding is of utter importance. Since the approaches in intelliRISK2 should account for all possibilities of unknown environments a complete 3D Mapping is desired. Therefore, the VDB Mapping approach [18] is enriched with additional information.

In the extended map, different kinds of layers can be included for either semantic understanding of the surroundings or different properties of the ground as can be seen in Figure 4. With this information a detailed robot and task specific planning can be conducted. This means, depending on the surroundings of the robot and its current risk affinity, different plans could be generated. The more risks the robot would be willing to take, the more direct the paths would be. If the risk parameter is low, the robot would prefer detours through better traversable environments. On the other hand, a high risk affinity would tip the robot towards using shortcuts through risky terrain.

Additionally, depending on the robots current health status as estimated by the internal anomaly detection, the robot would adapt its behavior, i.e. changing its locomotion mode. Therefore, utilizing the internal state and the episodic memory to assert a risk-aware execution of the path.

3.6. Risk-aware Exploration

The overall orchestration of the risk-aware exploration will be done by a behavior tree. The behavior trees are

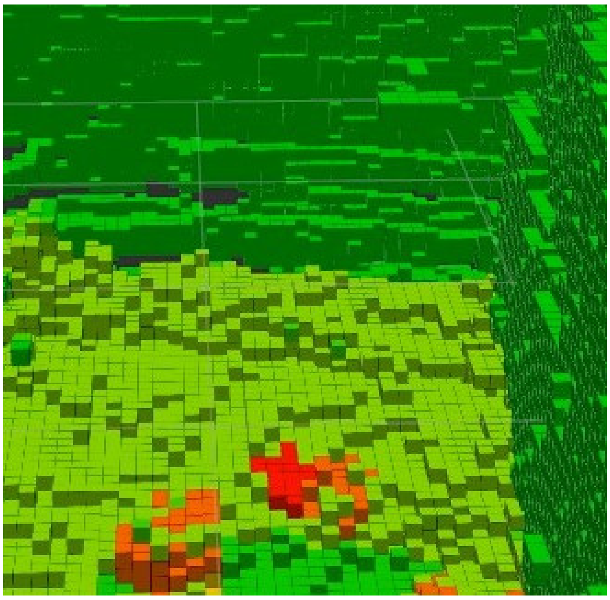

Figure 4. Voxel representation of the surroundings as it is seen by the robot. In the view of the robot different hazardous objects are highlighted in red. The more danger the robot anticipates the darker the red hue is. Green denotes safe environments. In this context the gravel testbed from Figure 1 is visible with added additional stones in front of the robot.

capable of triggering the mission planning, its execution and can organize multiple agents.

The main mission planning will be conducted on a base station and can generate behavior trees which, in return, can be passed to an individual robot. However, each robot will be able to plan its own risk-aware path according to its current status and estimate the health costs and risks based on their own trained models, capabilities and experiences. Once assigned an exploration target, the robot would start its path towards said goal. During the execution the internal and external assessment is active and once the robot finds anomalous behavior a replanning will be started. This replanning will contain both the overall mission and the path planning, and can incorporate the current models of the robot's health and environment. Therefore, the mission will always be assessed with the latest insights the robot gained in the unfamiliar surroundings.

4. CONCLUSION AND FUTURE WORK

This paper highlights the overall structure of intelliRISK2 and its main components. It is shown how the robots can both evaluate the hazards from their surroundings, as well as their own current health status. To achieve the risk assessment of the surroundings, an active learning approach for semantic segmentation is applied. This allows an operator to aid the robot from a base station to quickly gain an understanding of the environment. If there are unforeseen and unknown structures, a novelty detection based on ensembles provides insight. This enables the robot to generate interesting points which are worth exploring. The monitoring and recollection of the robots health is based on an anomaly detection of the internal sensor readings and an episodic memory. The anomaly detection monitors the current state of all readings and through a bidirectional Wasserstein GAN the sensor data can be transformed into a low dimensional latent space. When an anomaly is found, the episodic memory saves the corresponding episode. Thereby, the robot can continuously compare its current execution with previously collected negative experiences. This allows the robot to avoid doing actions which have proven to have a harmful outcome in similar situations. Through an overall mission planning using behavior trees, the robot can perform risk-aware exploration missions.

The next steps are improving the autonomy of the behavior tree generation and bringing the whole setup into an analog mission. Thereby, going from a controlled lab environment into more complex environments.

ACKNOWLEDGMENTS

The research leading to these results has received funding from the DLR Space Administration under grant agreement No. 50RA2026 by the German Bundestag.

REFERENCES

[1] Schnell, T., Büttner, T., Puck, L., Plasberg, C., Heppner, G., Rönnau, A. and Dillmann, R., 2018. Risikobewusste roboter für erhöhte autonomie in planetaren explorationsmissionen. Deutsche Gesellschaft für Luft- und Raumfahrt-Lilienthal-Oberth eV.

[2] Puck, L., Spielbauer, N., Schäble, D., Schnell, T., Büttner, T., Roennau, A. and Dillmann, R., 2021, December. Distributed Active Learning for Semantic Segmentation on Walking Robots. In 2021 20th International Conference on Advanced Robotics (ICAR)

[3] Settles B., 2009, Active Learning Literature Survey, Computer Science Technical Report

[4] Yang Y., Loog M., 2018, A benchmark and comparison of active learning for logistic regresssion, Pattern Recognition

[5] Mackowiak R., Lenz P., Ghori O., Diego F., Lange O., Rother C., 2018, Cereals-cost-effective region-based active learning for semantic segmentation, arXiv preprint arXiv:1810.09726

[6] Tan Y., Yang L., Hu Q., Du Z., 2019, Batch Mode Active Learning for Semantic Segmentation Based on Multi-Clue Sample Selection, Proceedings of the 28th ACM International Conference of Information and Knowledge Management

[7] Silverman B., 1986, Density Estiamtion for Statistics and Data Analysis, CRC Press

[8] Pimentel M., Clifton D., Clifton L., Tarassenko L., 2014, A review of novelty detection, Signal Processing

[9] Marsland S. 2001, On-line novelty detection through self-organisation, with application to inspection robotics, The University of Manchester(United Kingdom)

[10] Sofman B., Neuman B., Stentz A., Bagnell J. A., Anytime online novelty and change detection for mobile robots, Journal of Field Robotics

[11] Buderath M., Neumair M., 2007, Operational Risk Assessment for Unamnned Aircraft Vehicles by Using Structural Health and Event Management, European Aeronautic Defence and Space Deutschland GMBH Muenchen (Germany) Technical Report

[12] Amruthnath N., Gupta T., 2018, A research study on unsupervised machine learning algorithms for ealry fault detection in predicitive maintenance, 5th Internation Conference on Industrial Engineering and Applications(ICIEA)

[13] Stachowicz D., Kruijff G., 2018, Episodic-Like Memory for Cognitive Robots, Third Conference of Electircal Engineering (SCEE)

[14] Rothfuss, J., Ferreira, F., Aksoy, E. E., Zhou, Y., Asfour, T. (2018). Deep episodic memory: Encoding, recalling, and predicting episodic experiences for robot action execution. IEEE Robotics and Automation Letters, 3(4), 4007-4014.

[15] Hutter, M., Gehring, C., Jud, D., Lauber, A., Bellicoso, C.D., Tsounis, V., Hwangbo, J., Bodie, K., Fankhauser, P., Bloesch, M. and Diethelm, R., 2016, October. Anymal-a highly mobile and dynamic quadrupedal robot. In 2016 IEEE/RSJ international conference on intelligent robots and systems (IROS)

[16] Quigley, M., Conley, K., Gerkey, B., Faust, J., Foote, T., Leibs, J., Wheeler, R. and Ng, A.Y., 2009, May. ROS: an open-source Robot Operating System. In ICRA workshop on open source software (Vol. 3, No. 3.2, p. 5).

[17] Hess, W., Kohler, D., Rapp, H. and Andor, D., 2016, May. Real-time loop closure in 2D LIDAR SLAM. In 2016 IEEE international conference on robotics and automation (ICRA)

[18] Besselmann, M.G., Puck, L., Steffen, L., Roennau, A. and Dillmann, R., 2021, August. VDB-Mapping: A High Resolution and Real-Time Capable 3D Mapping Framework for Versatile Mobile Robots. In 2021 IEEE 17th International Conference on Automation Science and Engineering (CASE)

[19] Schnell, T., Plasberg, C., Puck, L., Buettner, T., Eichmann, C., Heppner, G., Rönnau, A. and Dillmann, R., 2020, August. Robot health estimation through unsupervised anomaly detection using Gaussian mixture models. In 2020 IEEE 16th International Conference on Automation Science and Engineering (CASE)

[20] Arjovsky, M., Chintala, S., and Bottou, L. 2017. Wasserstein generative adversarial networks. In ICML.