Proceedings of the Austrian Robotics Workshop 2025

Comparison of neural networks road detection in off-road environments

Jakob Oberpertinger[1], Matthias Eder[1] and Gerald Steinbauer-Wagner[1]

***Abstract*— As unmanned ground vehicles (UGVs) are more frequently deployed in unstructured environments, there is a growing need for robust road and terrain detection systems. The ability to navigate autonomously in challenging terrains depends on the effectiveness of computer vision models.**

Off-road environments encompass rugged terrain, forest roads, agricultural fields, and more, characterized by dynamic changes and unpredictable obstacles. UGVs must discern drivable ground to enable effective navigation while identifying and circumventing obstacles in real-time.

This paper investigates different sensor-based and neural network-driven approaches to address these challenges, focusing on the critical task of identifying forest roads in off-road environments. Using different sensors, we assess their effectiveness in different environmental conditions through a comprehensive comparative analysis of three neural network architectures. Our results highlight the strengths and limitations of different sensor modalities and neural network models. They provide insight into their performance under adverse conditions such as overexposed images, complex shadows, and dense vegetation on forest roads. This research provides valuable insights into developing robust off-road navigation systems essential for advancing autonomous ground vehicle technology.

I. INTRODUCTION

New application areas for unmanned ground vehicles (UGV), such as disaster response or forestry, have led to the need for safe navigation both on and off the road. Effective navigation is essential; it requires not only the ability to identify clear and accessible routes but also the foresight to avoid potential obstacles. Mastering this skill enhances safety and ensures a smoother journey every time. However, research in unstructured environments still lags behind that in structured environments [9]. Off-road environments for anmanned ground vehicles (UGVs) present unique challenges compared to traditional on-road environments. These environments vary widely, including rugged terrain, forest roads, agricultural areas, etc. Off-road environments can experience rapid changes, such as the appearance of lighting and weather conditions, temporary obstacles, or changes in terrain conditions. UGVs must be able to adapt in real-time to meet these dynamic challenges. The absence of clearly defined routes makes it challenging for an unmanned ground vehicle (UGV) to navigate to its destination. For a UGV to find its destination, it must make two important decisions. Firstly, it needs to detect accessible routes that it can safely traverse. Secondly, it must detect obstacles so that it can safely avoid them. There are many different approaches to solving these two challenges, using different sensors or neural network architectures.

Our research focuses on identifying navigable terrain in off-road environments, essential for safe and efficient navigation in unknown terrain. To address this challenge, we are undertaking a comprehensive comparison of three different neural network architectures using a variety of sensors, including RGB and depth images from stereo cameras and point clouds from lidar sensors. By exploring the effectiveness of different sensors, we aim to identify their respective strengths and limitations. This investigation goes beyond pure theoretical analysis, as we are carefully testing the limits of these networks under harsh environmental conditions. These conditions are characterized by significant challenges such as fluctuating sunlight, complex shadow patterns, and dense vegetation. The interplay of sunlight and shadows poses a significant hurdle for camera sensors and neural networks, especially if not adequately trained. In addition, vegetation poses challenges. Forest roads exhibit patches of grass in the center, which complicates the identification of navigable pathways.

The remainder of this paper is structured as follows: Section II discusses current research topics in path detection in off-road environments. Section III presents the three different neural networks evaluated in Section IV. Section V concludes the paper.

II. RELATED RESEARCH

A. Methods

Unmanned Ground Vehicles (UGVs) operating in off-road environments require robust road detection systems for safe and efficient navigation. Recent advancements in neural networks have significantly improved off-road path detection capabilities. However, developing a reliable and stable network for this purpose and selecting the appropriate sensors poses notable challenges. Ilas [8] outlines the key sensor technologies UGVs use to make real-time decisions while monitoring their surroundings. The study explores the various sensors employed across different environments and vehicle prototypes, evaluating the advancements in sensor technology.

Another important technology in off-road road detection is Convolutional Neural Networks (CNNs). CNNs excel in capturing spatial hierarchies of features, making them well-suited for image-based tasks. Researchers have explored various CNN architectures tailored for off-road scenarios. The work of Holder et al. [7] focuses on transfer learning, taking a pre-trained CNN designed for urban road scenes

[1]Jakob Oberpertinger, Matthias Eder, and Gerald Steinbauer-Wagner are with the Institute of Software Technology, Graz University of Technology, Graz, Austria. {jakob.oberpertinger, matthias.eder, steinbauer}@tugraz.at

https://doi.org/10.34749/3061-0710.2025.18

and retraining it to classify off-road scenes. The analysis involves assessing the network performance during various stages of training and exploring different levels of prior training on subsets of off-road data. The study compares the CNN approach with a traditional feature-driven Support Vector Machine (SVM) classifier, demonstrating state-of-the-art results in the challenging problem of off-road scene understanding.

Neural Networks using Lidar data have become a significant advancement in off-road road detection, offering depth information that allows for a more nuanced understanding of the environment. Zhong et al. [14] present a method known as LRTI, designed for identifying drivable areas in challenging off-road scenes. The complexity of this task arises from unstructured class boundaries, irregular features, and noise. By leveraging three-dimensional LiDAR data and a bird's eye view (BEV) perspective, LRTI utilizes texture information derived from LiDAR reflection data. The method incorporates an instance segmentation network to effectively learn this texture information, facilitating the identification of drivable areas. A multi-frame fusion strategy is employed to improve reliability. LRTI successfully achieves real-time processing on unmanned ground vehicles (UGVs).

Nate Haddad [5] discusses the challenges of training large deep learning algorithms due to the need for a substantial training dataset and computing power. Transfer learning, a method of transferring knowledge from one domain to another, is introduced as a solution to avoid training from scratch. The focus is on applying transfer learning to large encoder-decoder-style deep neural networks, specifically examining its impact on semantic segmentation tasks. DeepLabv3+, a state-of-the-art architecture from 2018, is highlighted for its efficiency in incorporating techniques from the 2016 Xception model [4].

B. Datasets

Chen Min et al. introduce the first off-road freespace detection dataset, called the ORFD dataset. Recognizing the importance of free space detection in autonomous driving technology, the authors highlight the limitations of existing deep learning methods, which primarily focus on urban road environments. To address this gap, they present the ORFD dataset, comprising 12,198 LiDAR point clouds and RGB image pairs collected in various off-road scenes, weather conditions, and light conditions. The authors propose a novel neural network, OFF-Net, which utilizes a transformer architecture to integrate local and global information, catering to the needs of a large receptive field for free space detection.

Peng et al. [10] address the significance of semantic scene understanding for robust autonomous navigation, particularly in off-road environments. Acknowledging the reliance of recent 3D semantic segmentation advancements on extensive training data, the authors identify a gap in existing datasets, which are either urban-focused or lack multimodal off-road data. The authors introduce RELLIS-3D, a multimodal dataset collected in an off-road setting to bridge this gap. The paper evaluates state-of-the-art deep learning semantic segmentation models on RELLIS-3D, revealing that the dataset introduces challenges distinct from urban environments.

The RUGD dataset [13] provides semantic annotations for unstructured outdoor environments, supporting off-road autonomous navigation. The dataset from a mobile robot platform includes video sequences with dense pixel-wise annotations for terrain classification and obstacle detection. It features 24 semantic categories, including eight terrain types, to enhance path planning and localization in environments lacking structured cues.

III. Evaluated Architectures

In this chapter, we evaluate three previously published neural network architectures, selected for their diverse input modalities and relevance to understanding the off-road scene. Our aim is not to propose new architectures, but to assess how well existing state-of-the-art segmentation methods generalize to off-road environments, particularly in challenging conditions such as forest roads, uneven terrain, and underexposed regions. The motivation behind the selection of these three models is based on their complementary input representations and processing strategies:

- **OFF-Net**: Chosen for using surface normal maps and a transformer-based architecture, offering a high-level representation of terrain structure. It is designed to leverage geometric cues from RGB-D input for improved scene segmentation.
- **DeepLabV3+**: A well-established CNN-based model known for its high segmentation accuracy and strong performance across various domains. It is particularly beneficial when working with limited or domain-specific training data.
- **SalsaNext**: A LiDAR-based semantic segmentation model operating directly on 3D point clouds. Its selection allows us to evaluate how pure LiDAR-based perception compares to image-based methods in unstructured off-road scenes.

This comparative evaluation's significance lies in understanding these architectures' behavior under real-world deployment constraints. By testing on our dataset, comprising RGB imagery, stereo-derived depth, and LiDAR scans collected in diverse environments, we aim to provide practical insight into each network's robustness and adaptability. This evaluation not only identifies the performance boundaries of each modality but also informs future design decisions for autonomous navigation systems in GNSS-denied and visually ambiguous terrain.

In the following section, we will present the concept of the comparison between the three neural networks, which are:

- Off-Road-Freespace-Detection (ORFD)
- DeepLabv3+
- SalsaNext using RELLIS-3D dataset

The three architectures and their design are presented in this chapter in detail.

https://doi.org/10.34749/3061-0710.2025.18

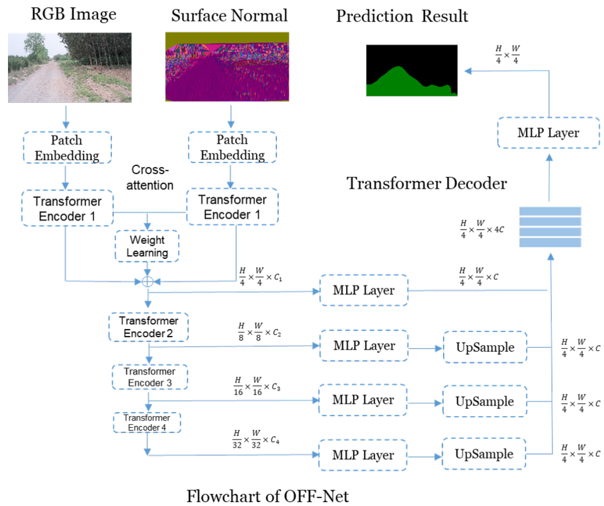


Fig. 1: The architecture of the OFF-Net [12].

A. *Off-Road-Freespace-Detection (ORFD)*

This architecture was presented by Chen Min et al. [12] 2022, which addresses the critical aspect of free space detection in off-road environments for autonomous driving. The paper presents a novel neural network, OFF-Net, which uses a transformer architecture to integrate local and global information, addressing the need for expansive receptive fields in free-space detection tasks, which are critical for accurate detection. Figure 1 shows an overview of the presented OFF-Net. As can be seen in the figure, the network combines two pieces of information: the RGB image and the corresponding surface normal. The paper's authors use LiDAR point cloud information to calculate the surface normal for each image. In our case, we calculate the surface normal from a dense depth image provided by the ZED2 stereo camera[1]. The transformer encoder can extract the features from these two pieces of information, and the transformer decoder predicts the free space. The paper also presents the dataset they have created for off-road freespace detection, called the ORFD dataset. The dataset includes off-road environments such as forests, farmland, and countryside with different weather conditions. The results demonstrate that SNE-RoadSeg, utilizing surface normals instead of depth information, outperforms FuseNet in free space detection. Furthermore, the newly proposed OFF-Net achieves even higher accuracy, surpassing FuseNet by 10.8% in F-score and 16.3% in mIOU. OFF-Net, employing the Transformer framework, efficiently captures local and global information while maintaining real-time processing capabilities, 7 times smaller and 2.7 times faster than SNE-RoadSeg [12].

B. *DeepLabV3+*

The second architecture, DeepLabv3+, is a simple but effective decoder module to improve segmentation results. Chen et. al. [1] describes this architecture as follows: Multiple downsampling of CNN results in a smaller feature map resolution, which leads to lower prediction accuracy and loss of boundary information in semantic segmentation. Similarly, aggregating the context around a feature helps to better segment it, which is achieved with sparse convolutions. DeepLabv3+ helps to solve these problems. The architecture can be seen in Figure 2. To save time and in the absence of a large dataset, we used a pre-trained model from the paper by Nate Haddad [5], who proposes to extend the application of a pre-trained DeepLabv3+ model to the challenging domain of off-road perception. The authors successfully employ transfer learning techniques using the Yamaha-CMU Off-Road Dataset for semantic segmentation of off-road images, showcasing the model's adaptability and effectiveness in a different domain. The Yamaha-CMU Off-Road Dataset [11] consists of 1076 images collected in different environments using three different sensors. It was labeled using eight classes (sky, rough trail, smooth trail, traversable grass, high vegetation, non-traversable low vegetation, and obstacle). The model takes an image as an input parameter, which is provided by the ZED2 stereo camera mounted on the front of the robot.

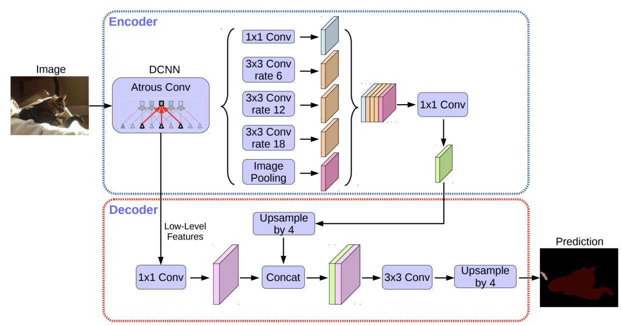


Fig. 2: The architecture of the Deeplabv3+ [1].

C. *SalsaNext*

Last, we used a model using LiDAR data as the input parameter. Peng et al. [10] introduced in their paper SalsaNext, an advanced model designed for real-time uncertainty-aware semantic segmentation of full 3D LiDAR point clouds. The authors made some major improvements to the already existing model SalsaNet. Some improvements are as follows: they replaced the ResNet encoder blocks with a new residual dilated convolution stack with gradually increasing receptive fields and added the pixel-shuffle layer in the decoder. Finally, we implemented a model that utilizes LiDAR data as its input parameter. In their paper, Peng et al. [10] introduced SalsaNext, an advanced framework designed for real-time, uncertainty-aware semantic segmentation of complete 3D LiDAR point clouds. The authors made significant enhancements to the existing SalsaNet model. Notable improvements include the substitution of the ResNet encoder blocks with a novel residual dilated convolution stack that features progressively increasing receptive fields

[1] https://www.stereolabs.com/docs

https://doi.org/10.34749/3061-0710.2025.18

and incorporates a pixel-shuffle layer in the decoder. They also switch from stride convolution to average pooling and apply central dropout treatment. To directly optimize the Jaccard index, they combine the weighted cross-entropy loss with Lovasz-Softmax loss and inject a Bayesian treatment to compute the epistemic and aleatoric uncertainties for each point in the cloud [2]. The improved architecture can be seen in Figure 3. The authors of the paper [10] present the dataset RELLIS-3D, a collection of off-road environments captured at the Rellis Campus of Texas A&M University. The RELLIS-3D dataset comprises a large set of raw sensor data, including color camera images, laser scans, high-precision global positioning measurements, inertial measurements, and depth images from a 3D stereo camera, and is labeled in 20 classes. The results show that SalsaNext achieves a higher mIoU of 43.07% compared to KPConv's 19.07%, which is significantly lower than their performance on the SemanticKITTI dataset, which was 59.5% mIoU and 58.8%, respectively. The imbalance in the point cloud dataset poses a significant challenge for both algorithms, with KPConv showing a more pronounced degradation. Despite attempts to mitigate the imbalance through sampling strategies during training, such efforts only marginally improved the results by 0.6% mIoU [10].

As the classes did not include forest roads, we selected a subset of the 20 available classes, focusing only on those relevant to detecting passable ground. This subset includes dirt, grass, puddles, asphalt, and mud.

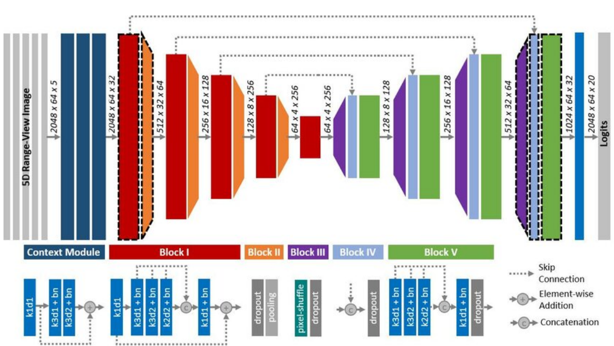


Fig. 3: The architecture of the SalsaNext [2].

After implementing, we conducted rigorous testing for various environments and sensors. The following chapter describes the results and evaluations of these tests in detail.

IV. EVALUATION

Autonomous navigation in off-road scenarios presents unique challenges that demand robust and accurate perception systems.

A. Data Generation

To evaluate the three networks and generate test data, we are utilizing the robots, Mercator [6], developed by Graz University of Technology, and Husky[2], developed by Clearpath. Mercator is a universal off-road platform developed for autonomous navigation in disaster response scenarios. It is a four-wheeled mobile platform with double Ackermann steering, an onboard computer, and a mounting frame for various sensor setups. Husky is a medium-sized robotic development platform with a large payload capacity. It is a customizable robot with the ability to add multiple sensors. The assessment spanned diverse environments, ranging from optimal visibility forest roads to challenging off-road terrains covered with grass.

To collect and record data for analysis, we equipped the two unmanned robots mentioned above with a ZED2 stereo camera[3] and 3D LiDAR scanners. Our data collection spanned a variety of environments and locations, including mountainous areas, rural landscapes, and forest roads in Styria, Austria, capturing different weather and terrain conditions. We selected challenging scenarios from the collected data for network testing, including varying light conditions, narrow forest roads, off-road paths with grass tracks, and grass-covered terrain, as shown in Figure 4. The ground truth annotation of the data was conducted manually.

B. Network Performance Metrics:

To evaluate the three different models, we have used the widely used mean intersection-over-union (mIOU) metric [3], which is given by

$$mIOU = \frac{1}{C}\sum_{c=1}^{C}\frac{TP_c}{TP_c + FP_c + FN_c} \tag{1}$$

where C is the number of classes, and TP (=true positive), FP (=false positive) and FN (=false negative) are the predictions for class c. The analysis focused on two classes: traversable and non-traversable areas.

C. Quantitative Results

We chose 250 images from our generated data for a quantitative analysis, described in IV-A. This ensures a well-distributed selection that captures key challenges such as lighting conditions, vegetation, and shadows.

Table I outlines the mIOU rates for each network for each image shown in Figure 4 and the mIOU (=mean IOU). Notably, on the mIOU, the DeepLabV3+ model outperformed OFF-Net by 9.64%, despite OFF-Net utilizing the surface normal as additional information. The SalsaNext network achieved a mIOU rate of 27.86%, emphasizing its ability to distinguish between traversable and non-traversable areas.

	Reference	Green Strip	Shadow	Underexposed	**mIOU**
DeepLabV3+	95.11%	51.55%	53.29%	1.53%	**76.91%**
OFF-Net	73.60%	29.25%	50.60%	0.12%	**67.27%**
SalsaNext	30.82%	28.28%	35.53%	19.87	**27.86%**

TABLE I: mIOU of the three neural networks.

[2]https://clearpathrobotics.com/husky-unmanned-ground-vehicle-robot/

[3]https://www.stereolabs.com/docs

https://doi.org/10.34749/3061-0710.2025.18

D. Comparative Analysis:

Despite OFF-Net incorporating additional information, the DeepLabV3+ model outperformed it. This raises questions about the effectiveness of the extra data and underscores the importance of careful feature selection and integration. Figure 4 shows the difference between the two networks using camera information, in which the second column shows the ground truth in light green, the third column shows the prediction of the DeepLabV3+ model in blue, and finally, the last column shows the prediction of the OFF-Net network in dark green.

The four different scenarios visualize the main problems and limitations of the two networks. The first scenario (reference) shows a well-visible, clear, and wide forest road, which both networks can predict quite well, with both mIOU values higher than 70%, as shown in Table I. The next scenario (green strip) shows an off-road divided by a grass strip. Here, both networks have difficulty accurately delineating the entire road and only manage to identify segments without grass. Again, the DeepLabV3+ scores a higher mIOU value compared to the OFF-Net.

The third scenario (shadow) shows a narrow forest path in a partially shaded wooded area. The OFF-Net has difficulty distinguishing between shaded and sunlit areas. However, DeepLabV3+ shows superior performance in this respect, suggesting that the OFF-Net model could be improved by refining the training dataset. DeepLabV3+ detects areas at the side of the path, which can lead to difficult or impassable paths. If we look at the mIOU values from Table I, we can see that DeepLabV3+ has a slightly higher mIOU value, but if we look at the images, OFF-Net is more accurate on the path. Last but not least, a road is completely covered with grass, which neither network can predict. Both networks have an mIOU value lower than 2%. It shows the networks are not trained for this type of off-road.

E. Insights into SalsaNext Network:

While SalsaNext demonstrated its ability to distinguish between drivable surfaces such as grass, dirt, and bush, ... its limitation lies in its lack of specificity in identifying true off-road. As a result, it is not a good choice for off-road detection and, therefore, scores the worst mIOU values. Future improvements could focus on refining the training data to include a wider range of off-road surfaces, thereby improving its ability to make nuanced distinctions. Figure 5 shows the predicted point cloud for different environments. The first environment is a wide forest road; the second is a narrow forest path.

F. Challenges and Solutions for OFF-Net:

OFF-Net faced challenges related to sun reflection and shadows, impacting its predictions. Bright reflections and rapid changes in brightness, especially transitioning from shadows to sunlight, were identified as major concerns. Moreover, the network can be improved by adding more difficult scenarios to the training data, such as underexposure, forest roads divided by grass strips, or fully covered roads with grass.

V. CONCLUSION

This paper evaluated three neural networks—DeepLabV3+, OFF-Net, and SalsaNext—for autonomous navigation in off-road environments using the Mercator robot. Tests covered forest paths, narrow trails, and grass-covered terrain, highlighting each model's strengths and limitations.

Future work should improve SalsaNext's training data and improve OFF-Net through adaptive mechanisms or filtering. These insights support further optimization of network robustness for real-world off-road navigation.

ACKNOWLEDGMENT

This work was funded by the Austrian Research Funding Association (FFG) under the scope of the THINK.WOOD.INNOVATION program.

REFERENCES

[1] L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, "Encoder-decoder with atrous separable convolution for semantic image segmentation," arXiv, 2018.

[2] T. Cortinhal, G. Tzelepis, and E. E. Aksoy, "Salsanext: Fast, uncertainty-aware semantic segmentation of lidar point clouds for autonomous driving," arXiv, 2020.

[3] M. Everingham, S. Eslami, L. Van Gool, C. Williams, J. Winn, and A. Zisserman, "The pascal visual object classes challenge: A retrospective," *International Journal of Computer Vision*, vol. 111, pp. 98–136, Jan. 2015.

[4] R. Garg, A. Kumar, N. Bansal, M. Prateek, and S. Singh, "Semantic segmentation of polsar image data using advanced deep learning model," *Scientific Reports*, vol. 11, pp. 15 365:1–18, 07 2021.

[5] N. Haddad, "Semantic segmentation of off-road images using transfer learning and deeplabv3+," *https://github.com/nmhaddad/semantic-segmentation/tree/master*, 2022.

[6] R. Halatschek, K. Ramanna, W. Url, and G. Steinbauer-Wagner, "Universal offroad robot platform for disaster response," in *2020 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR)*, 11 2020, p. 6.

[7] C. J. Holder, T. P. Breckon, and X. Wei, "From on-road to off: Transfer learning within a deep convolutional neural network for segmentation and classification of off-road scenes," in *Computer Vision – ECCV 2016 Workshops*, G. Hua and H. Jégou, Eds. Cham: Springer International Publishing, 2016, pp. 149–162.

[8] C. Ilas, "Electronic sensing technologies for autonomous ground vehicles: A review," in *2013 8TH INTERNATIONAL SYMPOSIUM ON ADVANCED TOPICS IN ELECTRICAL ENGINEERING (ATEE)*, 2013, pp. 1–6.

[9] F. Islam, M. M. Nabi, and J. E. Ball, "Off-road detection analysis for autonomous ground vehicles: A review," *Sensors*, vol. 22, no. 21, 2022. [Online]. Available: https://www.mdpi.com/1424-8220/22/21/8463

[10] P. Jiang, P. R. Osteen, M. B. Wigness, and S. Saripalli, "RELLIS-3D dataset: Data, benchmarks and analysis," *CoRR*, vol. abs/2011.12954, 2020. [Online]. Available: https://arxiv.org/abs/2011.12954

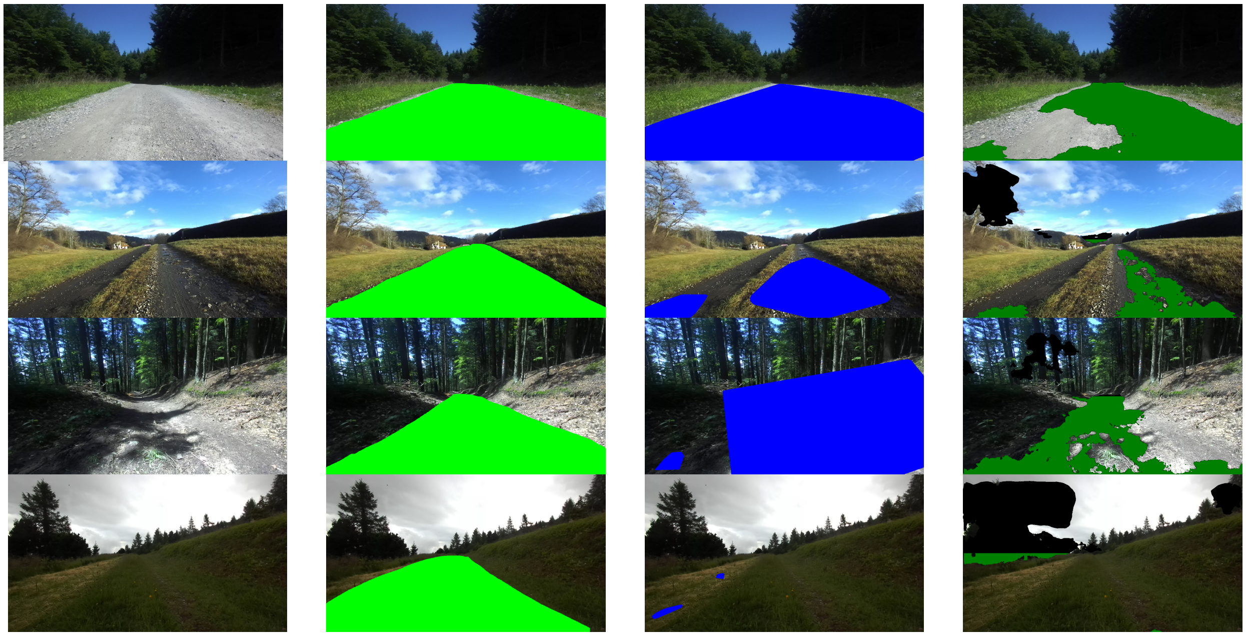
[11] D. Maturana, P.-W. Chou, M. Uenoyama, and S. Scherer, "Real-time semantic mapping for autonomous off-road navigation," in *Field and Service Robotics*. Springer, 2018, pp. 335–350.

[12] C. Min, W. Jiang, D. Zhao, J. Xu, L. Xiao, Y. Nie, and B. Dai, "Orfd: A dataset and benchmark for off-road freespace detection," 2022.

[13] M. Wigness, S. Eum, J. G. Rogers, D. Han, and H. Kwon, "A rugd dataset for autonomous navigation and visual perception in unstructured outdoor environments," in *International Conference on Intelligent Robots and Systems (IROS)*, 2019.

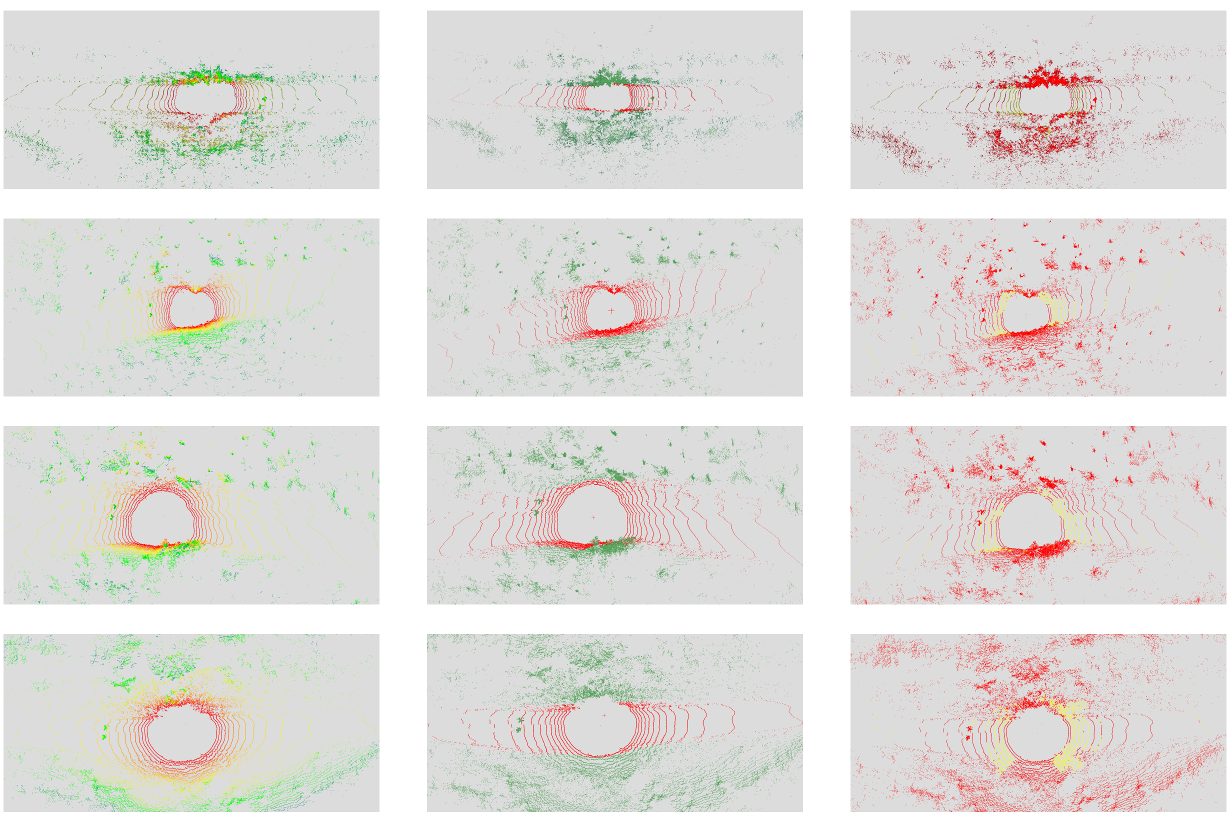
[14] C. Zhong, B. Li, and T. Wu, "Off-road drivable area detection: A learning-based approach exploiting lidar reflection texture information," *Remote Sensing*, vol. 15, no. 1, 2023. [Online]. Available: https://www.mdpi.com/2072-4292/15/1/27

https://doi.org/10.34749/3061-0710.2025.18



(a) RAW Image (b) Ground Truth (c) DeepLabV3+ (d) OFF-Net

Fig. 4: Predictions of the image based networks DeepLabV3+ and OFF-Net.



(a) RAW Pointcloud (b) Ground Truth (c) SalsaNext

Fig. 5: SalsaNext Network prediction

Multi Robot Route Planning for ROS2

Matthias Reicher[1] and Markus Bader[1]

***Abstract*— This work presents the implementation of a multi robot route planner based on the prioritized planning approach as well as its integration into ROS2 and the well-known Nav2 stack. Further, a method to increase the resilience towards uncertainty and unpredictability in timing during the execution of found routes is introduced. These so-called routing preconditions are shown to be effective on a subset of routing scenarios and offer significant opportunity for further exploration.**

***Index Terms*— multi robot system, path planning, ROS2, Nav2**

I. INTRODUCTION

To leverage the advantages of a multi-robot system (MRS), large fleets of mobile robots must be able to effectively compute routes from one point in the environment to another without risking collision. This makes multi-robot-route-planning a fundamental problem for MRS, as it lays the groundwork for more complex behavior [3]. Many approaches to solving this problem have been discussed in the literature, with so-called "prioritized planning" appearing in a significant number of publications [2]. However, up to current knowledge, no publicly available ROS2-compatible software packages provides an easy integration of such functionality. This work aims to close the identified gap, similar to the previous work of [1] on ROS, but by taking advantage of the advanced capabilities offered by the well-known Nav2 stack. Results are presented by using a simulated environment as shown in Fig. 1.

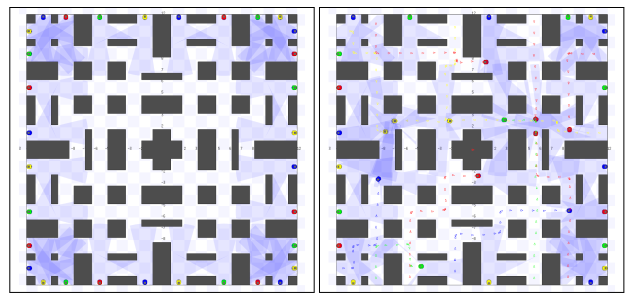
II. PRIORITIZED PLANNING

Prioritized Planning refers to the practice of decomposing the multi-robot-route-planning problem into a series of single-robot-route-planning (SRRP) problems. Each of the SRRP-problems concerns itself with finding a collision-free route for an individual robot and must take static obstacles as well as robots for which a route has already been found into consideration. Since routes are planned in descending order according to some priority metric, higher-priority robots represent dynamic obstacles in the planning space of low priority robots.

III. IMPLEMENTED PLANNING ALGORITHM

To realize this specification of a Prioritized Planner, some considerations need to be made: First, a planning algorithm which is able to handle dynamic obstacles is required to solve the individual SRRP-problems. Second, the routes generated by the prioritized planner need to be suited for execution by a real MRS.

[1]The authors are with Faculty Informatics at TU Wien, Vienna, Austria. `firstname.lastname@tuwien.ac.at`



(a) Initial position (b) During navigation with Nav2

Fig. 1: Stage-simulation of a 32-robot MRS.

A. Sequential Planner

The chosen planning algorithm can be described as a variant of the spatio-temporal A*-Algorithm introduced in [4] operating on a graph-based abstraction of the environment. This abstraction is able to emulate 4/8-connected grid maps, as well as higher level concepts such as voronoi graphs with multi-edges. The key difference to the well-known A*-Algorithm is given by additional occupancy checks whenever a graph vertex is explored and added to the frontier: should it be occupied by another robot at the point in time in which the planning robot expects to enter, time must be spent waiting earlier along the currently considered route. If it is impossible to insert this waiting time at some point along the path without risking collisions, the proposed node is not marked for further exploration. These iterative planning processes result in a detailed record describing at which points in time any particular graph vertex is expected to be occupied by a robot if no unexpected delays occur.

B. Route Representation

After planning an ideal path for a robot in the system, post-processing is done to create a route suited for execution by a real MRS. Routes consist of a series of indexed route segments, each describing a move from one vertex of the graph to one of its neighbors. In addition to the timestamps during which this move is expected to take place, a set of preconditions for the segment is generated by considering all other robots scheduled to pass the destination of the move before it occurs. A precondition is considered to be satisfied as soon as the robot it is referencing has completed the noted segment of its own route (i.e. it has passed through the vertex at which both routes cross). This creates clear precedence relations, which serve to improve the systems resilience towards neglected or unexpected delays during navigation.

https://doi.org/10.34749/3061-0710.2025.19

IV. ROS2 Integration

The ROS2 integration of the implemented planner is split between multiple communicating system components, pictured in Fig. 2.

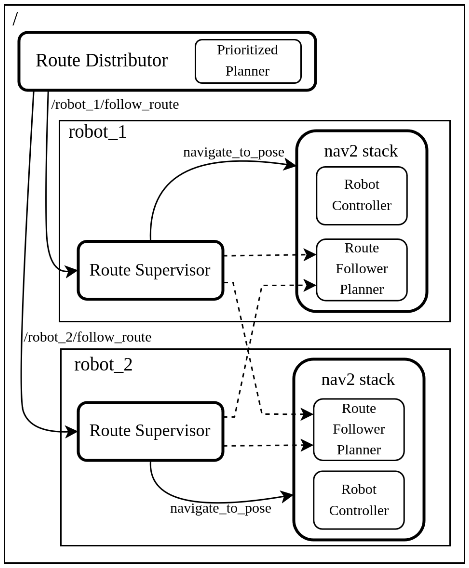


Fig. 2: Architecture of the ROS2 integration.

A. Route Distributor

The Route Distributor node acts as the central coordinator of the MRS. It is responsible for initializing navigation by generating each robots route using the implemented prioritized planning algorithm and distributing them among the MRS using ROS actions. During route execution, it monitors the received feedback and aborts navigation should unexpected issues arise.

B. Route Supervisor

The communication between robots and the Route Distributor is handled by an individual Route Supervisor node for every robot. Each of these nodes also monitors the robots progress along its own route and publishes this information for consumption by all the Route Followers in the system. This enables robots to wait on unsatisfied preconditions to in order to avoid situations not considered during planning.

C. Route Follower

To enable the use of the wide variety of localization strategies, local planners and other software components available within Nav2, the system integrates with a Nav2-planner-plugin known as the Route Follower.

V. Evaluation

The implemented planning algorithm was tested on randomly generated routing problems featuring 8-32 robots concurrently attempting to find a route through a heavily restricted warehouse-like environment. Through varying the order in which routes are planned, a solution to each of these routing problems was found. The systems capability of executing these found routes was then evaluated by simulating navigation using the Stage simulator.

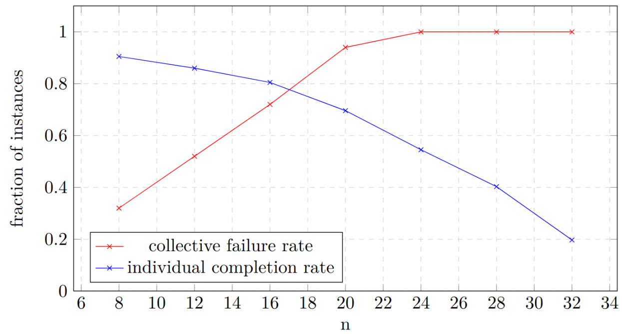


Fig. 3: Routing success in a highly constrained environment.

Fig. 3 depicts the ratio of individual robots which were able to reach their goals as well as the chance of any robot failing to finish its route due to an emergency stop, a collision or similar reasons. Both metrics behave in a roughly linear fashion, resulting in sharply degrading reliability as more concurrently navigating robots are added to the system.

Two central causes for these failures were identified:

1) Off-the-shelf Nav2 local planner solutions navigating based on a generic path representation deviating from the strictly defined pre-planned routes.
2) Endless waiting on an unsatisfied precondition referring to a stuck robot causing cascading failure in the system.

VI. Summary and Outlook

Collision-free routes for members of a multi-robot systems can be found by the implemented algorithm, but it is evident that this does not guarantee that these routes can be executed without issue in realistic conditions. While routing preconditions were introduced to counteract timing-related failures, they have proven insufficient to avoid them entirely without addressing flaws in the systems architecture and implementation. Introducing additional mechanisms to increase robustness such as on-line re-planning in case of a detected deadlock represents another avenue for future work.

VII. Acknowledgment

This research is supported by the Austrian Science Fund (FWF) under project No. 923138, GreenFDT.

References

[1] B. Binder, F. Beck, F. König, and M. Bader, "Multi Robot Route Planning (MRRP):
Extended Spatial-Temporal Prioritized Planning," in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Nov 2019, pp. 4133–4139.

[2] J. Heselden and G. Das, "Heuristics and rescheduling in prioritised multi-robot path planning: A literature review," *Machines*, vol. 11, no. 11, p. 1033, 2023.

[3] G. Kyprianou, L. Doitsidis, and S. A. Chatzichristofis, "Towards the achievement of path planning with multi-robot systems in dynamic environments," *J. Intell. Robot. Syst.*, vol. 104, no. 1, 2022.

[4] W. Wang and W.-B. Goh, "Multi-robot path planning with the spatio-temporal A* algorithm and its variants," in *Advanced Agent Technology*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 313–329.

https://doi.org/10.34749/3061-0710.2025.19