FFI Norwegian Defence Research Establishment

24/01369

FFI-REPORT

Warpath ASV

– situational awareness for autonomous surface vessels

Martin Vonheim Larsen
Ragnar Smestad
Trym Vegard Haavardsholm

Warpath ASV

– situational awareness for autonomous surface vessels

Martin Vonheim Larsen
Ragnar Smestad
Trym Vegard Haavardsholm

Norwegian Defence Research Establishment (FFI) October 28, 2024

Keywords
Autonomi
Situasjonsforståelse
Autonome overflatefarkoster (ASV)
Sensorer
Sensorfusjon

FFI report
24/01369

Project number
1688

Electronic ISBN
978-82-464-3551-0

Approvers
Rikke Amilde Seehuus, *Research Manager*
Håkon Storli Andersen, *Director of Research*

The document is electronically approved and therefore has no handwritten signature.

Summary

In order to operate in a complex and unpredictable environment, autonomous surface vessels (ASVs) must be able to generate their own situational awareness (SA). Here we present *Warpath ASV*, a ASV-tailored specialization of FFI's generic situational awareness framework for autonomous platforms. It implements the situational awareness capabilities deployed on the two ASVs *Odin* and *Frigg*.

With radar and lidar as its primary sensors, *Warpath ASV* can automatically produce occupancy maps, track objects of interest, and conduct geographical self-localization in environments where global navigation satellite system are unavailable (GNSS-denied environments) by fusing measurements across sensors and over time. The resulting situational picture is used for automatic collision avoidance, close formation maneuvers, and general situational awareness for operators.

In 2023, the development of *Warpath ASV* has focused on maturation over adding new features. This report documents current capabilities in *Warpath ASV*, provides an overview of key experiments since 2016, summarizes lessons learned, and points the way forward for future research.

Sammendrag

For å kunne operere i et komplekst og uforutsigbart miljø må autonome sjøfarkoster (ASV-er) være i stand til å skape sin egen situasjonsforståelse. Her presenterer vi *Warpath ASV*, en ASV-tilpasset spesialisering av FFIs generiske rammeverk for situasjonsforståelse på autonome farkoster. Spesialiseringen implementerer situasjonsforståelsesfunksjonaliteten om bord i de to ASV-ene *Odin* og *Frigg*.

Med radar og lidar som hovedsensorer kan *Warpath ASV* automatisk lage hindringskart, følge interessante objekter og gjøre geografisk selvlokalisering i områder uten mulighet for satelittnavigasjon. Dette gjøres ved å fusjonere målinger på tvers av sensorer og over tid. Resultatet er et situasjonsbilde som brukes til automatisk kollisjonsunngåelse, til manøvrering i tett formasjon og til generell operatørstøtte.

I 2023 har utviklingen av *Warpath ASV* handlet om teknisk modning fremfor ny funksjonalitet. Denne rapporten dokumenterer den nåværende tilstanden til *Warpath ASV*, gir en oversikt over viktige eksperimenter siden 2016, oppsummerer viktige lærdommer og peker ut veien videre for fremtidig forskning.

Contents

1 Introduction

This document aims to give a brief overview of the autonomous situational awareness (SA) capabilities of the autonomy package for autonomous surface vessels (ASVs) developed at FFI. In this context, we take “SA” to mean all the steps necessary from acquiring sensor data and processing it, to producing all the high-level actionable information needed by the vessels to perform their autonomous functions. In FFI’s series of “autonomy projects” 1372, 1505, and 1688, we have developed *Warpath*, which is our SA framework for autonomous platforms. The ASVs use a tailored version of Warpath (aptly named *Warpath ASV*), which we in the following will simply refer to as “Warpath” for brevity.

ASVs require sufficient SA to operate safely and efficiently in littoral waters, all-the-while adhering to maritime traffic regulations (COLREGS). In order to be usable for autonomous control, Warpath estimates a (large) set of states that together enable the assessment of risks for grounding, collisions, and COLREGS violations. Specifically, we build a map discerning land and sea, and track objects on the water surface. For each tracked object, we perform motion modelling to predict their future movements. To ensure robustness against large uncertainties and discontinuities in absolute navigation, map and object states are represented relative to the platform. This robustness is critical in GNSS-denied scenarios or with intermittent GNSS reception. Warpath can also conduct geographical self-localization with its own sensors. Together, the localization, land/sea map, and object tracking provides the information needed for the ASVs to safely and efficiently accomplish their mission.

Figure 1.1 shows the overall architecture of Warpath. At the facade, Warpath interfaces several sensors, navigation, and prior map sources. We then process the data from each of these sources in a generic processing graph called Superflow [1]. The resulting measurements are fused in various ways to produce three types of data:

- A map of navigable water vs “not water”, represented as a probabilistic occupancy grid [2].
- A list of discernible object tracks. Each object is at minimum represented as a unique (and consistent) ID, a classification as stationary/moving, and an estimate of their current motion, but other object features may also be included.
- Geographical self-localization and self-odometry.

Figure 1.2 shows a simplified overview of the current processing graph employed on the ASVs.

A visualization of the Warpath SA output is given in fig. 1.3. Here, we see the probabilistic water/land map given as varying shades of gray, where black is certain water and white is certain land. Object tracks are shown as circles with unique IDs. Moving objects are shown in violet and with an arrow indicating their direction of motion, while stationary objects are shown in teal.

The rest of this report is organized as follows. In chapter 2, we describe the sensors currently installed on the ASVs. Next, chapters 3 to 5 give an overview of the mapping, tracking, and localization pipelines. Chapter 6 highlights some key events and experiments with *Warpath ASV* from the beginning in 2016 up until 2024. Finally, in chapter 7, we conclude the work and discuss our current efforts and further plans for 2024.

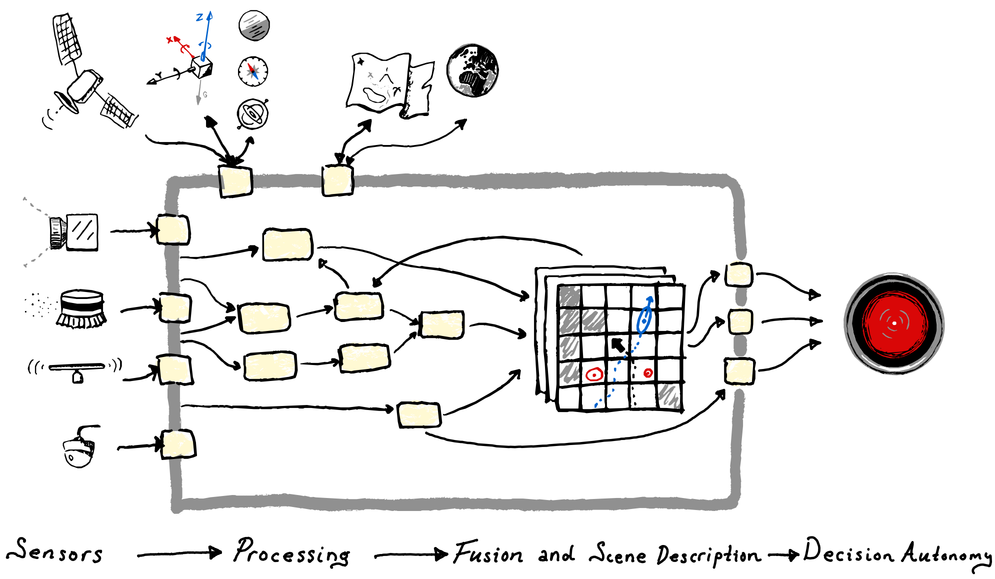


Figure 1.1 *Architectural overview of Warpath. Warpath interfaces several sensors, navigation components, and prior map sources, whose data is fused in a processing graph. The resulting output is self-localization, a land/sea map, and a list of tracked objects, which is handed over to the decision autonomy (HAL).*

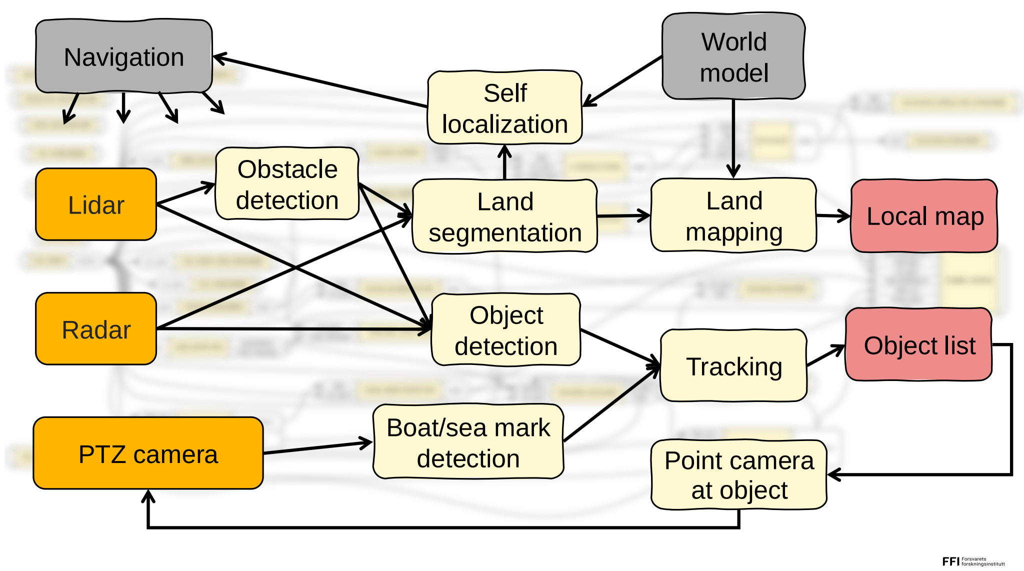


Figure 1.2 *Outline of main features in the processing pipeline of Warpath ASV. Note that navigation is integral to all processes, hence the simplified representation.*

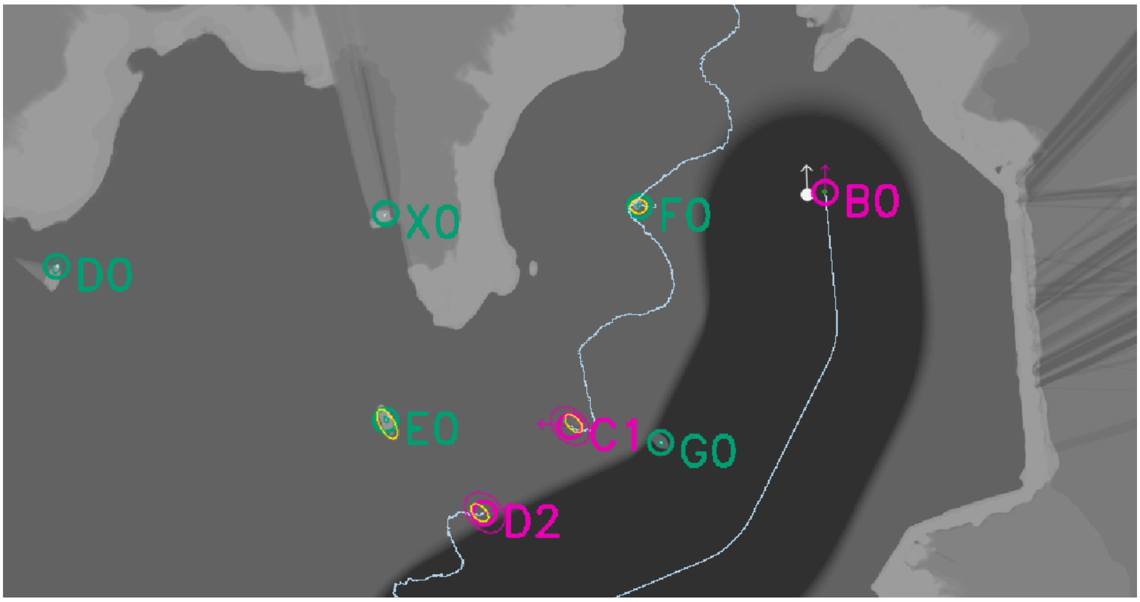


Figure 1.3 *A visualization of the fused situational picture produced by Warpath. Our own vessel (Odin) is shown as the white dot with an arrow. The background shows the "water/not water" map, where the shade of gray corresponds to the certainty level. Here, black represents certain water, white is certain land and 50% gray is "fully uncertain". The colored circles show object tracks, with stationary tracks rendered in teal and moving tracks in magenta. This particular scenario is taken from a formation run, where the other ASV (Frigg) can be seen as track "B0".*

2 Sensors

We are currently using the following sensors on the ASVs:

- **Inertial navigation system (INS) (*NavP*, developed by FFI [3], [4]):**
  The INS is based on a global navigation satellite system (GNSS) receiver, a navigation grade inertial measurement unit (IMU), and a commercial electro magnetical (EM) speed log. The resulting navigation solution contains absolute position, orientation, and velocity of the ASV's body frame in the world, as well as the relative change with respect to the previous time step.
- **Simrad HALO-3 pulse compression radar [5]:**
  A recreational X-band 25 W solid-state pulse compression radar with a 3 ft open array antenna. Depending on configuration, this sensor can deliver short- and long-range performance, from 6 m to 48 NM, and with a 2.4° beam width. The radar also offers a high-speed 48 RPM mode for distances under 2 NM.
- **Ouster OS2-128 [6] and OS1-64 [7] lidars:**
  These are 360° scanning lidars, providing 3D point measurements with a range precision of $\pm 2 - 8$ cm at 10 or 20 Hz. The OS2 is rated for ranges up to 240 m, with 128 lasers distributed over a 22.5° vertical field of view. The OS1 has 120 m range with 64 lasers covering 45° vertically. In addition to range measurements, each data point contains the intensity of the reflected light. The sensors can also return passive measurements of ambient light, which in practice makes them applicable as low-resolution 360° scanning grayscale cameras.
- **Camera:**
  Different cameras have previously been part of the sensor suite for automatic processing, including both panoramic and pan-tilt-zoom (PTZ) cameras. Cameras are not employed in the current state of Warpath, but integration of a new Axis Q6318 PTZ-camera is due in 2024.
- **Time synchronization server:**
  Used for synchronizing clocks around the system and required to fuse information from different sensors correctly.
- **Automatic identification system (AIS):**
  Output from AIS is available for processing, but is currently not utilized in Warpath.
- **Telemetry from cooperating vessels:**
  As an alternative to passively estimating the orientation and velocity of a cooperating vessel, telemetry can be transmitted over the air to aid in tracking and positioning.

With its superior range and robustness, the radar is the natural primary sensor in Warpath. It is used as a source for both mapping, object tracking, and self-localization. Even though we can tune a handful of parameters, it should be noted that since this is a consumer grade navigational radar, we can only receive preprocessed information. Still, it has so far sufficed in our applications. When we operate in the littoral areas around Horten, the range is set to approximately 1 NM. An example of raw radar data is shown in fig. 2.1.

Lidar measurements are also used for mapping and object tracking. When used on surface vessels, we exploit that the water in general will absorb or deflect most of the laser rays (see for instance fig. 2.2), and consequently, most objects detected by the lidar represent potential obstacles. Whitecaps and the

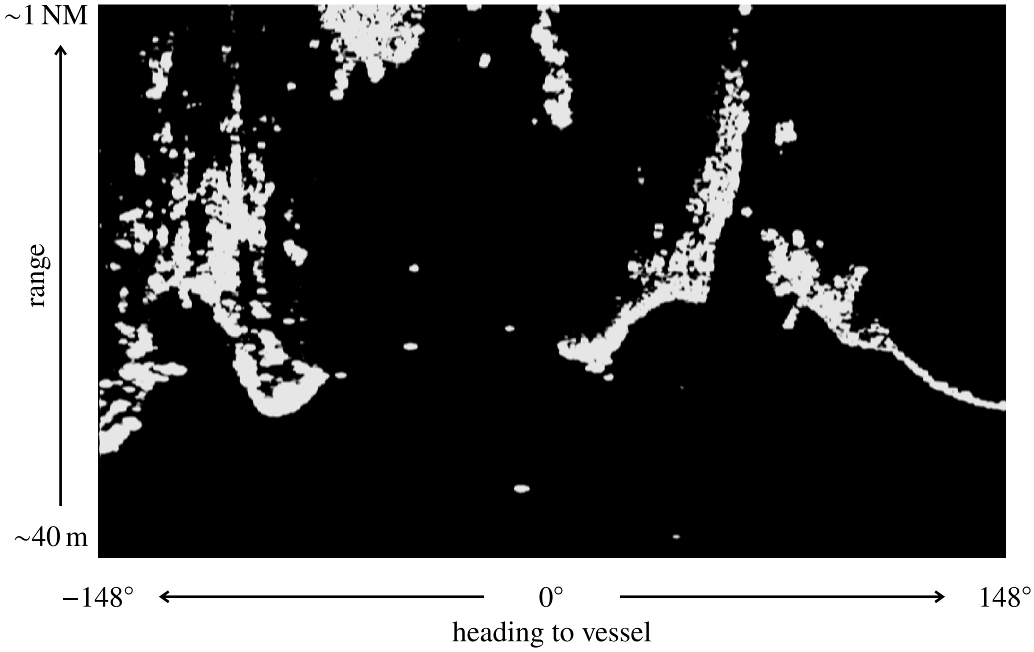


Figure 2.1 Raw polar plot from the HALO-3 Radar. White areas represent strong radar responses, while black represents no response. The large continuous shapes are generally caused by land formations. Smaller free-standing blobs can *indicate smaller objects in the water, but many of the small blobs seen here are just clutter.*

stern wake are known to be false positives, so they must be handled accordingly. To ensure adequate reaction time to obstacles and other vessels, our findings indicate that lidar range measurements should ideally extend to at least 200 m.

Cameras have so far been used for detection and classification (based on deep learning), both with stationary- and PTZ-cameras. A FLIR M400 multi-sensor camera [8] was explored in 2018, but was abandoned due to shortcomings in the camera platform. On paper, this camera is quite capable, with both thermal and daylight imaging sensors, good zoom capabilities, and mechanical stabilization. However, the provided APIs for integration are intended for manual operation, and lack the capabilities for automated timestamping and precise control required for computer-based SA systems. In 2024, we are resuming work on automatic operation of a PTZ-camera to aid in object tracking and self-localization. Our current development efforts are now based on the Axis Q6318 PTZ camera, which lacks both thermal sensor and mechanical stabilization, but does enable proper hardware control and clock synchronization. As a result, Warpath has to manage (low-frequency) stabilization, but is able to obtain far more accurate observations thanks to proper timestamping. This functionality is transferable to other low-cost camera systems, as long as proper timing is available.

The INS, time synchronization and proper calibration of sensor poses are all crucial components in a successful SA framework. Some effort has been put into online calibration of radar and lidar, but more work is required to robustify these tasks.

AIS-data is not currently utilized in Warpath, but we have conducted some experiments with it. The primary issues with AIS positional data are the lack of proper timestamps and reporting of positional

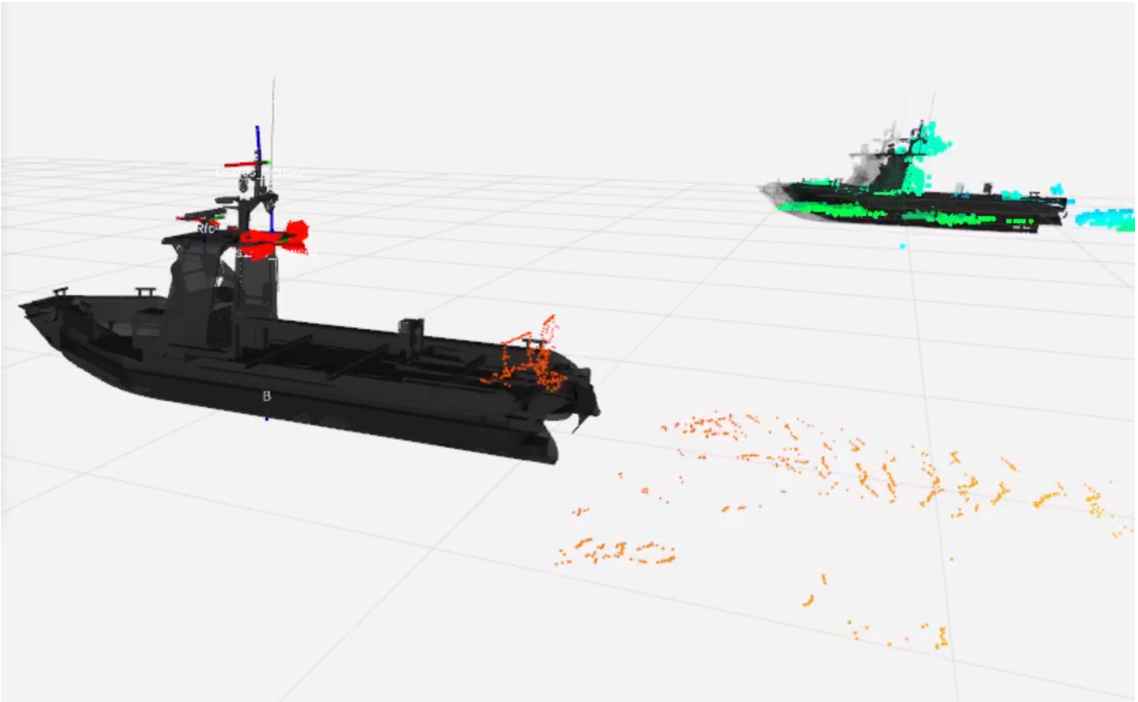


Figure 2.2 Visualization of the Ouster OS-2 lidar pointcloud. The data is captured from the vessel on the left, which observes the cooperating vessel on the right. Lidar returns are shown as the dots color-coded by distance.

accuracy. The timing uncertainty, often as high as 30 s, in turn means that a moving vessel's actual position is highly uncertain at any given moment. Additionally, the accuracy of navigation equipment varies widely from vessel to vessel, forcing us to assume a very poor worst-case positional accuracy. At close range, where precision is crucial, AIS data therefore contributes little value while introducing many opportunities for errors, which is why we disabled it. Nevertheless, we believe AIS can provide useful metadata, such as vessel ID and class. Since AIS reports are easily spoofed, they should always be treated cautiously and not allowed to heavily influence the tracker output.

In the current state, data from radar, lidar, navigation, and telemetry are serialized and transmitted via ROS-topics and can thus be recorded into ROS-bags. Even though this introduces extra latency, the benefit is that we can easily play back the data and do sensor processing offline while expecting the processing pipeline to behave in the same way as when processing live data on the vessel.

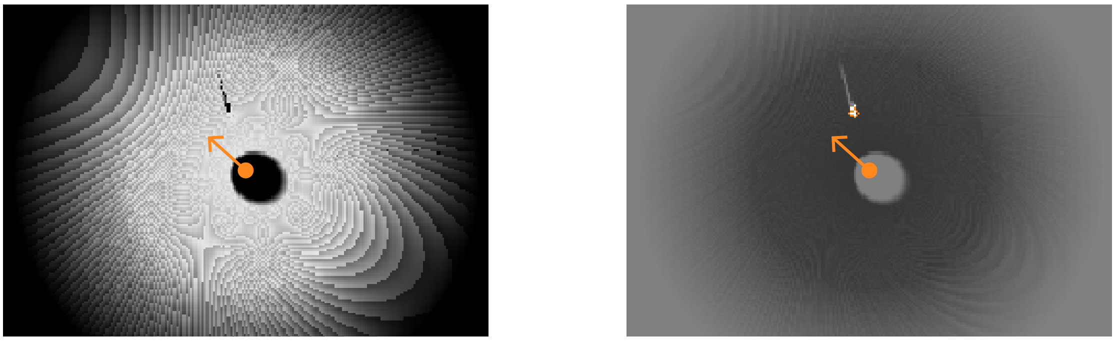
3 Overview of the mapping pipeline

We use three sources of information for building the water/land map: Pre-existing sea charts and measurements from the radar and lidar sensors. The resulting map is a north-oriented raster grid centered approximately (but with a known offset) at the vessel origin. Each cell represents the logarithm of the land-to-water odds [2]. Large positive or negative values indicate land or water areas with high certainty, while uncertain cells have values near zero. This representation has the nice properties that both class label (land or water) and uncertainty is encoded in a single scalar, and that measurement updates can be done with simple summation [2].

When mapping with the lidar, we exploit that the lidar rays are largely not reflected back from the water, which means that any lidar return is an indication of structure (or “not water”) in the corresponding grid cell. The higher the points are above sea level, the more likely it is that there is an obstacle. Similarly, a lidar ray that passes unobstructed through the volume above a grid cell is an indication that the current area does not contain structure. Starting from the top ray, the further down we can follow a series of unobstructed rays, the less likely the grid cell is to contain an obstacle. Depending on the sea state, we use these measures to put a confidence-level on the positive/negative observation for each cell, as illustrated in fig. 3.1. The computations needed here fit well inside a standard graphics pipeline, enabling low-latency processing at full rate with minimal impact on system load. Thanks to the high rate of the lidar, the resulting temporally fused map is surprisingly robust to inconsistent false detections close to the sea plane, such as foaming waves.

In the radar mapping pipeline, we generally consider all returns to be land or structure (or “not water”), but ignore smaller blobs that are already represented as tracks. To account for deficiencies in the onboard processing of the radar, we only trust negative observations (“water”) along each ray until the first return. Upon the first echo, we trust immediately following returns as positive observations (“not water”), but with gradually decaying confidence. Any observations along a ray beyond the first contiguous blob of positive observations are simply disregarded (log odds 0). The effects of this scheme can be seen in fig. 3.2, where landforms gradually become fainter before casting shadows of gray uncertainty in the radial direction.

As discussed, the log odds representation allows us to fuse observations across sensors and over time through simple summation. To account for unmodelled effects, we also add a small process noise (effectively a slow decay to 0 log odds) to each cell. In our probabilistic modelling, we place much more trust in what we see with our sensors than in the prior sea charts. As a result, the log odds values from the prior map lie close to 0, which can be seen as the faintest landforms in fig. 1.3. In principle, the lidar and radar are trusted equally, but due to higher resolution and measurement rate, the lidar can reach far higher certainty levels. The effect of this higher certainty is evident in fig. 1.3, where cells covered by the lidar swath reach darker blacks and brighter whites.



(a) (b)

Figure 3.1 *Feature images used in the lidar detection and mapping pipelines. Our own vessel has position and heading as given by the orange dot with arrow. Fig. (a): Confidence plot. The confidence is (mostly) a function of the height of the lowest passing lidar ray above the sea plane, resulting in the conspicuous "onion rings". Fig. (b): Log odds plot of "not water" vs "water" from the lidar mapping stage. Black indicates "water", white indicates "not water", and 50% gray indicates that we have no information. Notice the small white blob, which is also detected as an object (the tiny orange marker).*

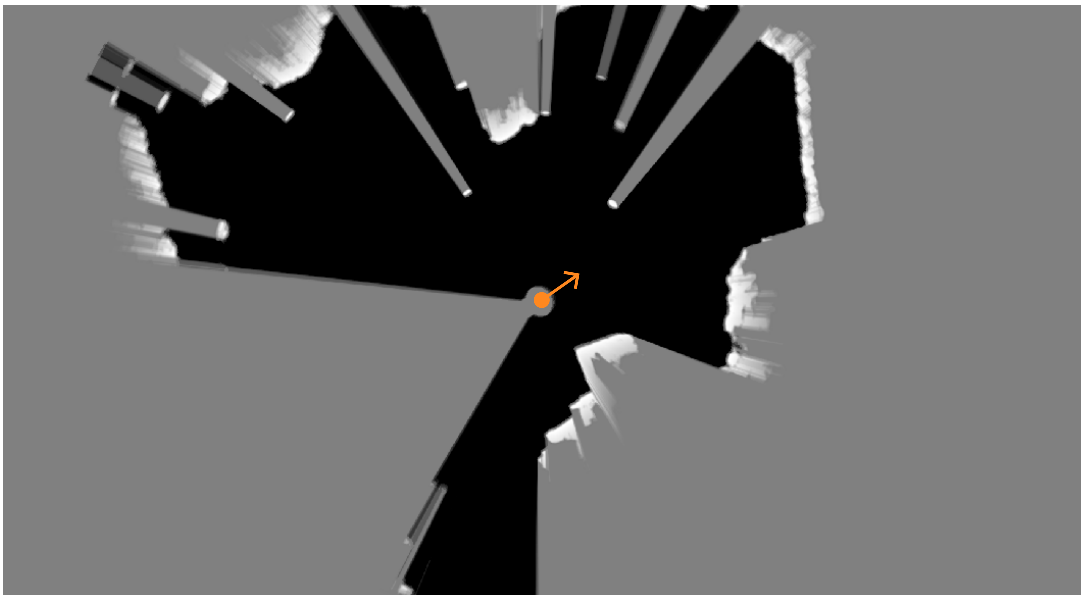


Figure 3.2 *Log odds plot of "not water" vs "water" from the radar mapping stage. Our vessel has position and heading as given by the orange dot with arrow. Black indicates "water", white indicates "not water", and 50% gray indicates that we have no information. As can be seen from the gray "shadows" we have no confidence in radar returns that are obstructed by nearer returns. The gray triangle behind our vessel is due to the* ~ 64° *rearward blind-zone.*

4 Overview of the target tracking pipeline

The Warpath target tracking architecture is built around a central tracker that fuses detections from each of the sensors directly. To succeed with centralized tracking, it is essential to have accurate time synchronization, sensor calibration, and platform navigation. In this section, we will give a brief overview of the considerations that must be made in the tracker setup to properly exploit the information made available by all the sensors.

In the tracking setup, we mostly employ typical techniques from the tracking-by-detection paradigm [9], but enforce probabilistic track management [9], [10]. Using probabilistic modelling in the track management enables us to easily handle a mix of sensors with different fields of view and non-uniform detection performance across the scene. By also formulating measurement-to-track comparison in terms of probability, we can seamlessly incorporate any measurements and estimation we can model probabilistically. This generalization enables us to use nontraditional object features such as visual appearance, classification, size or external ID (for instance MMSI) as first-class citizens in the tracking — not just position, orientation, and motion. We refer to this concept as *multi-feature* tracking. The key motivation behind this multi-feature tracking is to be able to combine well performing radar tracking with the state-of-the-art techniques from visual tracking, which we shall discuss further in chapter 7.

To relieve the requirements for accurate absolute navigation, we perform all positional estimation in a relative frame. Specifically, we represent track motion state in a frame centred at the vessel and aligned with the horizontal sea plane and the forward axis of the vessel's body frame. Compared to using a standard earth-fixed tracking frame, this setup is more cumbersome during tracker prediction. The benefit, however, is that the relative frame enables detection-to-track matching and measurement update independently of navigation, only requiring vessel-internal calibration. As a result, the overall tracking performance is limited by the short-term accuracy of the inertial navigation system instead of the absolute navigation accuracy. The relative tracker therefore achieves better performance in general, but is also especially well-suited for GNSS-denied scenarios.

As input to the tracker, we use detections from the radar and lidar, and telemetry from the cooperating vessels. In coastal areas, we operate the radar with settings that give a detection range from about 40 m out to roughly 1 NM. The theoretical detection range for the lidar is 240 m, but beyond 50 m the probability of detection drops drastically depending on the object in question. Therefore, the radar constitutes most of the SA for general traffic, (small) static obstacles, and sea markers, while the lidar fills the gap at very close range. Figure 4.1 shows a scenario where this combination is needed to track an object first appearing at long range before approaching closely. Meanwhile, when performing team sweeping with two ASVs in formation (typically roughly 40 m apart), the lidar is the primary sensor for localizing the cooperating vessel, used in addition to received cooperative telemetry.

With the radar, we detect objects by first finding (small) blobs in raw polar data, followed by processing to determine confidence level for each detection. This process is illustrated in fig. 4.2. Because the preprocessing performed by the HALO-3 radar tends to cause both a lot of false negatives and false positives behind the first returns, we mostly disregard detections without unobstructed line-of-sight. The key here is that we still leverage the obstructed data in the track

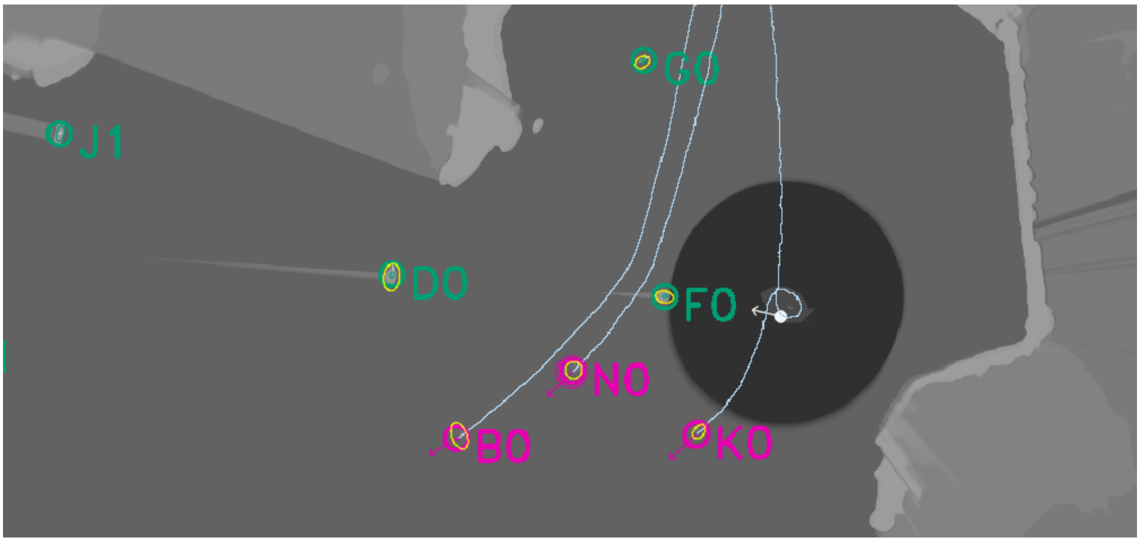


Figure 4.1 The fused situational picture from a scenario where our own vessel (Odin, white dot with arrow) has been approached at close range by another vessel (here shown as track "K0"). All three moving objects ("B0", "N0", and "S0") are first detected by the radar at range. As "K0" approaches, it will at some point appear in the lidar before moving inside the minimum range of the radar. The tracker here fuses radar and lidar detections into a single track moving from radar-only range into lidar-only range, and then back out to radar-only range.

management, but with a suitable confidence modelling. This way we get the best of both worlds: Few issues with false tracks in the shade behind land, but can still maintain well-established tracks.

For general object detection with the lidar, we detect (small) blobs in the land-vs-water log odds plots from fig. 3.1b. To improve tracking performance on the cooperating vessel in formation runs, we have added a retroreflector in the mast, and apply a special detector pipeline for this use case. The major benefit of the reflector is that it represents a well-defined point-of-measurement on the other vessel, in addition to making detection and association trivial, as shown in fig. 4.3.

In addition to the lidar reflector detections, we also leverage telemetry (position and velocity) from the cooperating vessel's internal navigation system in formation tracking. The position-component of the telemetry is mostly useful for asserting that the reflector we are tracking with the lidar is indeed our cooperating vessel. Since the lidar already measures the relative position of the cooperating vessel with centimeter-level accuracy, the contribution of the absolute position (with meter-level accuracy) from telemetry is negligible. Meanwhile, the velocity-component of the telemetry contributes greatly in improving tracking performance, since the lidar does not measure velocity directly. When equipped with a navigation grade IMU and either a EM-log or other sensor-based relative navigation (see for instance chapter 5), the two cooperating vessels can obtain and communicate absolute velocity even in GNSS-denied scenarios. Because the telemetry position relies on GNSS and only negligibly contributes to tracking performance, we only actively use the velocity-component of the telemetry in formation tracking.

We combine all radar, lidar, and telemetry detections in a single multi-object tracker. The reflex detector and the received telemetry produce "special" detections in the sense that they can only

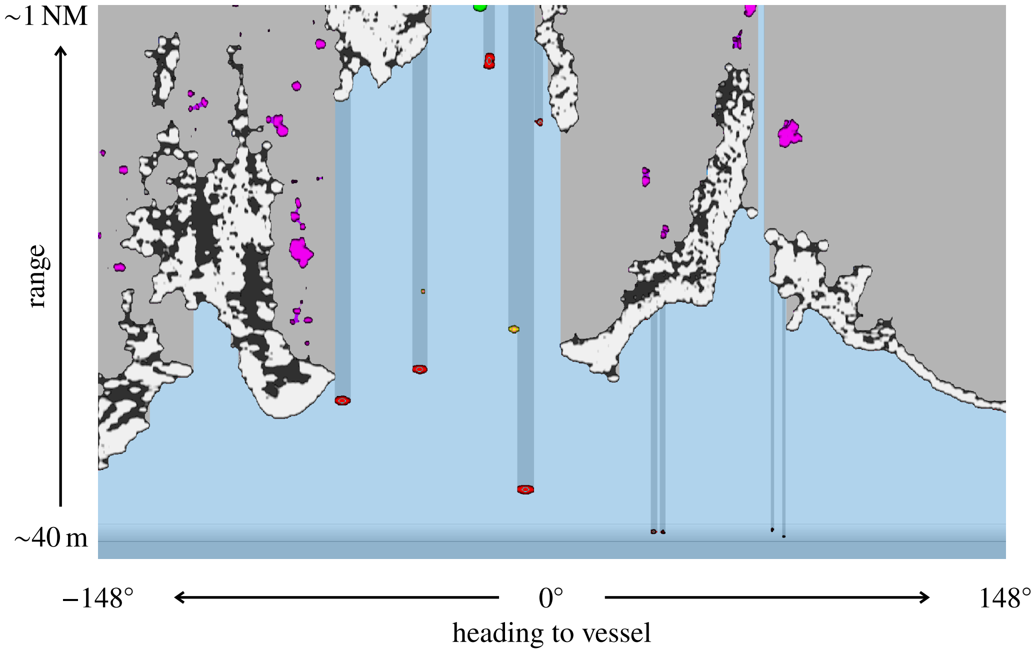


Figure 4.2 *Visualization of the radar object detection stage. The plot is shown in polar coordinates, with heading on the x-axis (0° is straight forward) and range on the y-axis. We detect (small) blobs, but assume different confidence level in different sections of the plot. Large blobs, here rendered in black/white, are assumed to be landforms and do not result in detections. Unobstructed blobs (red) result in detections with maximum confidence. Partially obstructed blobs (yellow) are given lower confidence. Blobs behind landforms (magenta) or tangent to the maximum range (green) are given zero confidence.*

Figure 4.3 *Ambient near-IR image from the Ouster OS2 lidar observing the cooperating vessel. Specular reflections are marked in red, while diffuse and ambient light are rendered as grayscale. The actual retroflector lies underneath the strongest return (inside the green circle), while the weaker red dots result from blooming around the reflector. We use the strongest return as the detection.*

detect the cooperating vessel. To help control association for these special detections, we leverage the multi-feature paradigm to include an “external ID” (which may be unknown) as part of each track’s estimated state. In formation scenarios, we can then tag telemetry “detections” with high confidence to have, for instance, external ID “ODIN”. The tracker will then ensure to associate these measurements with the “ODIN”-track, if it exists. Similarly, we add “ODIN”-tags to reflex measurements, albeit with slightly lower confidence. To fuse positional information, we rely on good calibration to transform observations from each sensor frame to the tracker frame (which is centred on the vessel and aligned with the horizontal sea plane and the forward axis of the vessel). Thanks to the probabilistic multi-feature formulation, the tracker can then associate and fuse measurements from all sensors seamlessly. In formation scenarios, the resulting tracker solution will contain a single track with ID “ODIN”, while all other tracks have unknown external ID.

5 Geographical localization by coastline matching

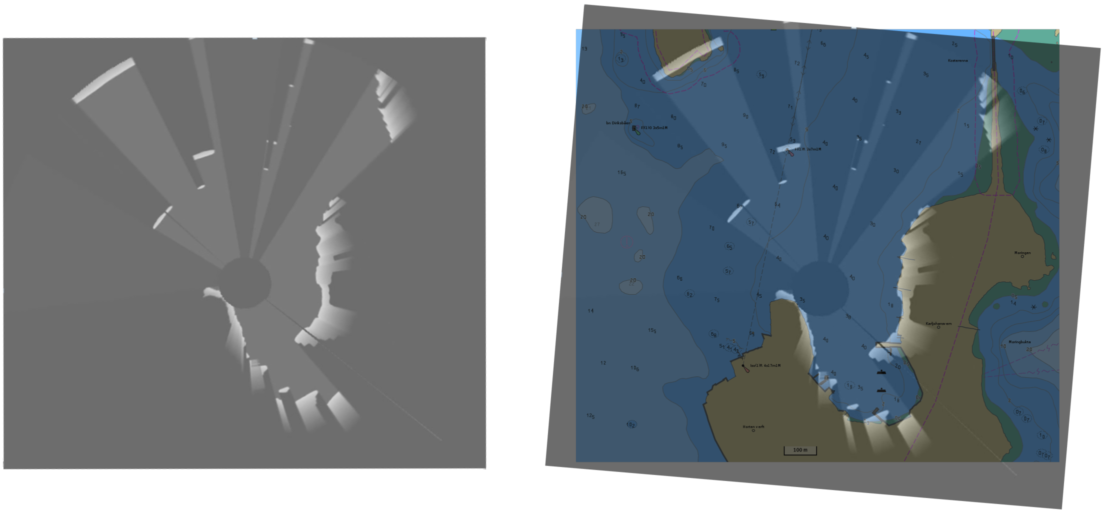
A very common sensor aboard many different types of sea vessels is the marine radar. It is capable of detecting land and objects at a range of several kilometers, even when there is low visibility due to bad weather or at night. Given an appropriate georeferenced sea chart, or a map of recognizable landmarks, it is possible to use on-board sensors to estimate the geographical position and heading of the vessel by matching sensor measurements with the georeferenced data. As part of a radar-aided INS demonstrator system [4], we have developed a method that estimates the vessel's geographical 2D-position by matching radar measurements with a georeferenced sea chart. The process is illustrated in fig. 5.1.

For each raw polar radar sweep image, we first construct a local Cartesian map measurement of the coastline. Next, we perform the actual localization by matching the local map against a pre-stored georeferenced coastline map. Based on direct image alignment techniques [11], the matching process finds a position estimate with uncertainty, together with a confidence score and a quality score. The confidence score is a measure of how much we trust the reported position, given the visible surrounding landscape. If we have few to none objects detected by the radar, we must put lower confidence in the result from the alignment algorithm. The quality reflects how much we trust the reported position based on the consistency of the radar measurement compared to the reference chart, as illustrated in fig. 5.2.

The method is expected to work with reduced performance in situations where terrain formations in our surroundings are ambiguous. In such situations, the method may report high confidence and high quality, but still the reported position may be inaccurate. Figure 5.3 illustrates some relevant scenarios.

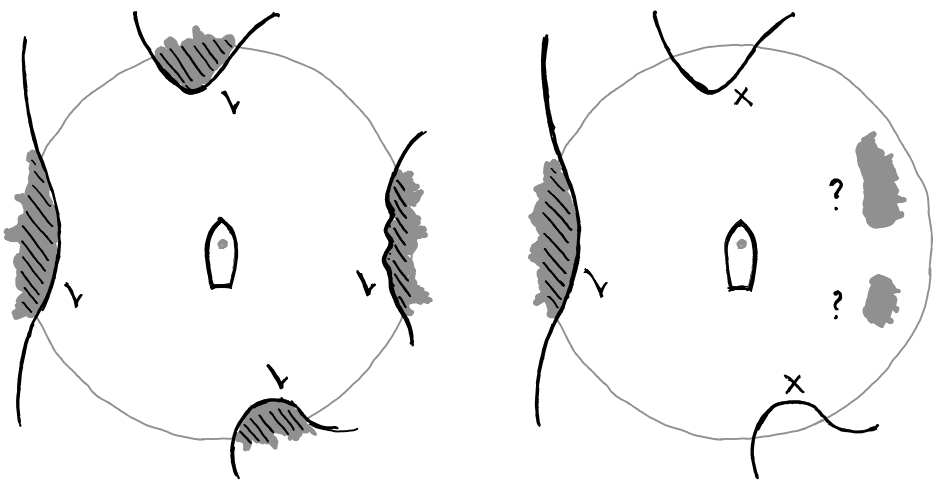
When the radar loses sight of land, there is no matching to perform, and the output of the method will be labeled invalid. However, the matching process is susceptible to "snapping" onto land in situations where the coastline is just beyond reach, but some object shows up close to the outer edge in the direction of land. The output will then be valid, but report low quality. Figure 5.4 illustrates this.

The results from the estimation are transmitted back to the navigation system. Experiments show that the method can greatly improve absolute geographical positioning in GNSS-denied environments.



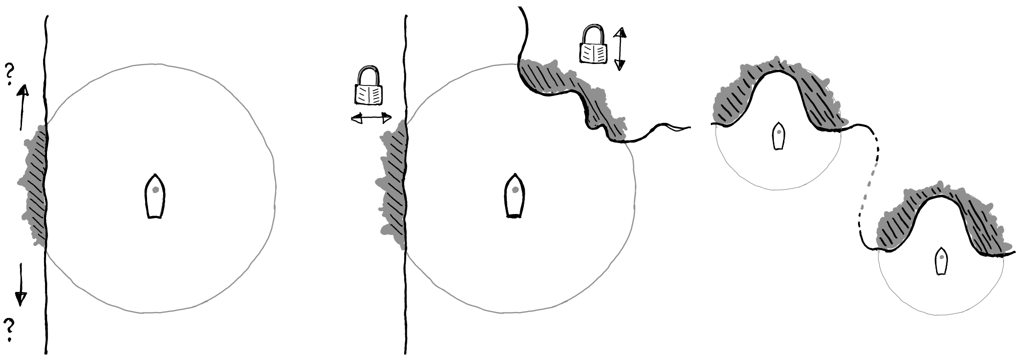
(a) (b)

Figure 5.1 Absolute geographical localization by coastline matching. Fig. (a): The result from coastline detection in a single radar sweep. Fig. (b): The same plot overlaid a georeferenced sea chart. Even though we have noise and detections of sea markers in the water, the method is able to align the two images.



(a) (b)

Figure 5.2 Quality scoring in radar localization. The gray circle around the vessel illustrates the range of the radar. Shaded gray areas are radar detections, and hatched areas are where the detections overlap coastline in the map. Fig. (a): High quality. All radar detections recognized as coastline matches actual coastline in the map. Fig. (b): Low quality. We have radar detections that do not match any coastline in the map (right), and given the estimated position, we would expect the radar to detect coastline at the top and bottom.



(a) *(b)* *(c)*

Figure 5.3 Land features that may influence the result. Fig. (a): With land formations like this, there will be a potentially large uncertainty in position along the direction of travel, as any position would yield a good match. Fig. (b): Detections in multiple directions and land formations with more unique features will give better results. Fig. (c): Similar looking land formations can give a good match locally, but result in an incorrect position globally.

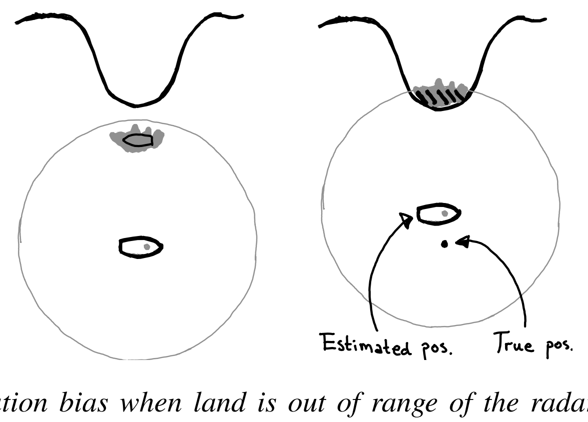


Figure 5.4 Localization bias when land is out of range of the radar. The method may incorrectly recognize the detection of the oncoming vessel as coastline and thus match the detection with the headland above. The low amount of radar returns will yield low confidence.

6 Key experiments 2016 - 2024

The development of *Warpath* in general, the components it is built from, and *Warpath ASV* in particular, has occurred gradually and incrementally over a period of approximately eight years. Along the way, experiments and demonstrations have been the main way of marking progress and improved functionality. Key demonstrations and experiments are illustrated chronologically along the timeline in fig. 6.1, and photos from a selection of events are shown in fig. 6.3.

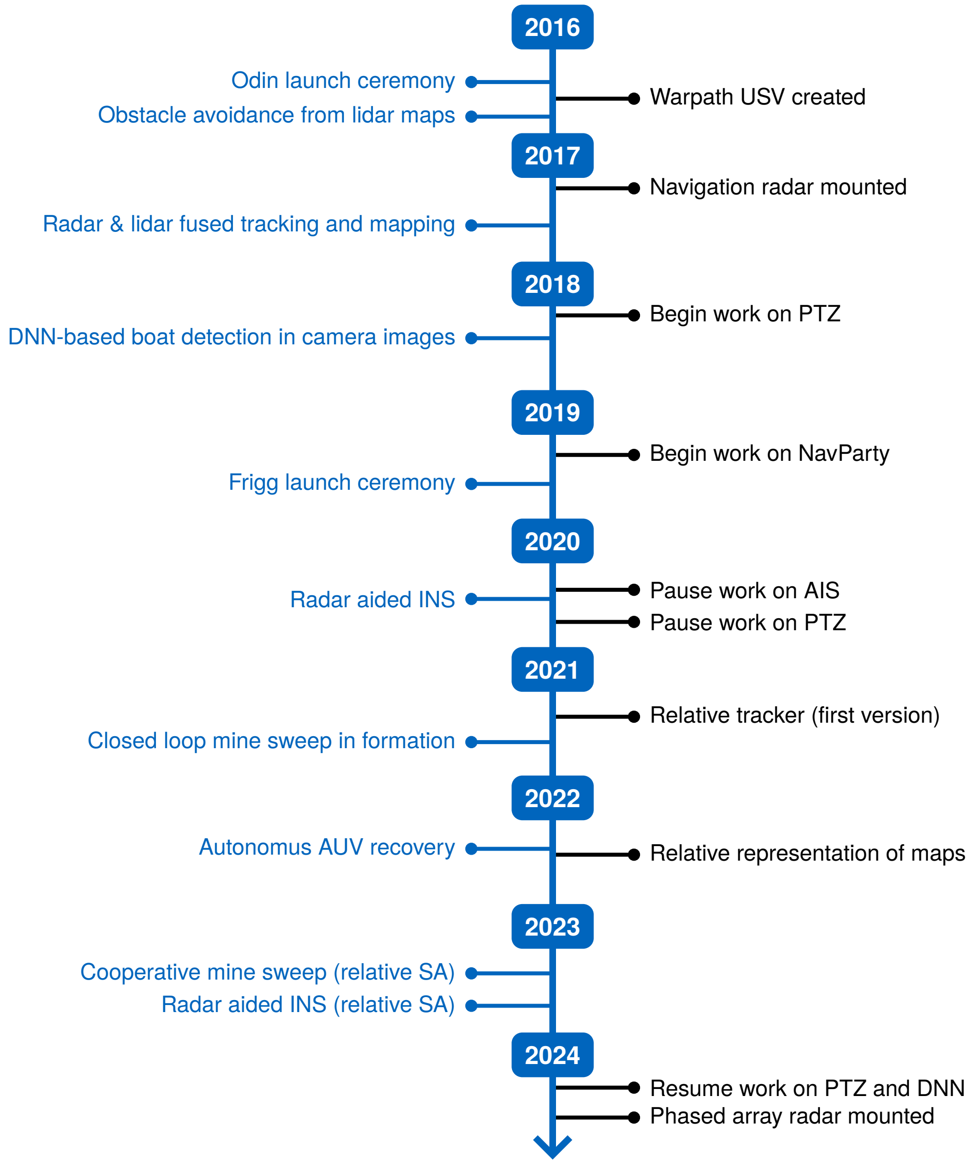


Figure 6.1 Key demonstrations and experiments (blue), and important events and development milestones (black) for Warpath ASV between 2016 – 2024.

Figure 6.2 August 2016: Launch ceremony for Odin, starring Minister of Defence Ine Eriksen Søreide and FFI Research Manager Morten Nakjem.

A lot of the ground work, the foundation of which *Warpath ASV* still stands on today, was developed in the period from 2016 to approximately mid 2019 when *Frigg* was acquired. The initial focus was to demonstrate safe navigation through obstacle detection, mapping, sensor fusion, and tracking. The modular nature of Warpath and its processing framework allows us to show off basic functionality at an early stage, and then improve the individual capacities independently later, piece by piece.

In the period 2018 – 2020, work was spent improving the internals of Warpath. A new high level interface to navigation and coordinate frames, *NavParty*, aided greatly in the later transition to a relative representation of tracks and local maps. Meanwhile, we also demonstrated object detection in camera images, and automatic control of PTZ-camera by the decision autonomy (HAL). The effort was eventually halted due to challenges with the camera platform and that other tasks were prioritized.

From 2020, the main focus has been cooperative operation of platforms, mainly two ASVs in formation, but also ASV and AUV (autonomous underwater vehicle) together for automatic recovery. We have completed the transition to exclusively relative representation of states within Warpath, and in line with parallel development on other platforms, we have had constant improvements to the tracker. Demonstration of radar-aided INS was done in this period, both before and after the transition to relative states. In 2023, *Warpath ASV* went through a relatively stable period, focusing on maturation over adding new features. A significant milestone was reached in September 2023 during a demonstration for the Ministry of Defence, where all our prior efforts culminated in a successful showcase of autonomous, cooperative mine sweeping using ASVs. In the coming time, we will shift focus back to development of new capabilities over the improvement of existing features.

(a) Aug. 2017: An uncrewed Odin, seconds after autonomously performing an evasive manoeuvre against a fused radar+lidar track during a demonstration near Kongsberg Maritime, Horten.

(b) June 2018: Initial tests with convolutional neural networks (CNN)-based boat detection in images from the FLIR M400 PTZ camera. Even stock models achieved surprisingly good results.

(c) June 2018: Odin's wake from autonomously performing a COLREGS-compliant maneouvre after detecting a head-on-head situation with the small boat "Småen" in the top-right corner of the photo.

(d) August 2018: Odin, about to demonstrate evasive manoeuvres against "Småen" for the Minister of Transport. As can be deduced from fig. (c), the manoeuvres were successful.

(e) September 2021: Odin and Frigg in tight formation based on fused lidar tracking and telemetry while towing a closed loop influence sweep [12].

(f) May 2022: Frigg autonomously moving into position for recovery of Hugin AUV, based on fused lidar tracking and telemetry [13].

Figure 6.3 Photos from key demonstrations with Warpath ASV between 2016 – 2024.

7 Conclusion

We have presented an overview of *Warpath ASV*, a specialized adaptation of FFI's generic SA framework for autonomous platforms, which implements the SA capabilities deployed on the two ASVs *Odin* and *Frigg*. *Warpath ASV* integrates data from lidar, radar, and camera sensors, along with navigation components and prior map sources, all of which are processed and fused within a processing graph. The resulting output is self-localization, a land/sea map, and a list of tracked objects, which is handed over to the decision autonomy (HAL) for safe navigation and motion planning. The development of *Warpath* in general, the components it is built from, and *Warpath ASV* in particular, has evolved gradually and incrementally over the last eight years.

Lessons learned

Looking back on our work with Warpath, not just for ASV, but for platforms in all domains, we have tested a large selection of sensors and worked with a variety of hardware. These are some of the key lessons we have learned.

Choose sensors with open interfaces. Many sensors, if not most, do not fully expose controls or the observed data through their provided interfaces. For example, the HALO-3 radar only offers heavily preprocessed intensity plots. Similarly, the FLIR M400 PTZ camera does not allow proper control when performing stabilization, and its interface becomes overwhelmed by the frequent interactions required for autonomy. These restrictions increase development time and prevent the sensors from being used to their full potential. Unfortunately, the peculiarities of the sensor interface are not detailed in datasheets and requires hands-on experimentation to uncover.

Choose sensors that can be time synchronized. On a continuously moving platform, such as a ASV, it is crucial to know where the sensor was pointing when the data was acquired for the data to be useful. This is only possible if the timestamps of the sensor data can be aligned with the navigation system. For autonomous sensor processing, this time synchronization of navigation and sensor data is an ever-present challenge, so it's wise to choose sensors that simplify this process as much as possible.

For ASVs, radar is the primary sensor, but must be paired with other short-range sensors. While the HALO-3 radar does not provide all the data we would like, it has consistently shown that a navigation radar delivers most of the situational awareness needed to navigate around land and other vessels, from long distances down to about 100 meters. At shorter distances, we have demonstrated the use of lidar measurements for maintaining tight formation between vessels and reactive collision avoidance. For positioning in formation, simpler solutions might be more effective, such as using radio beacons, range finders, or a combination of cameras and markers. For general short-range SA, both lidar and cameras are well-suited. Lidar excels at mapping and provides robust measurements of void. Although cameras are more susceptible to poor lighting conditions, they are particularly effective for object detection and provide superior angular precision at a lower cost. If possible, combining both radars, lidars, and cameras can therefore be advantageous.

Relative representations improve precision and seamlessly handle unreliable GNSS. When working with time-synchronized high-end INSs, it might seem logical to perform tracking and mapping using absolute coordinates. However, for this type of autonomous platforms, our experience is that using relative coordinates is a clearly better choice. Although it may be more cumbersome, relative representations inherently provides robustness against unreliable GNSS and is far more flexible when integrating alternative sensor-based sources of absolute navigation. In most cases, a relative representation also has better precision, since we defer the uncertainty cost that comes with converting to absolute coordinates as long as possible.

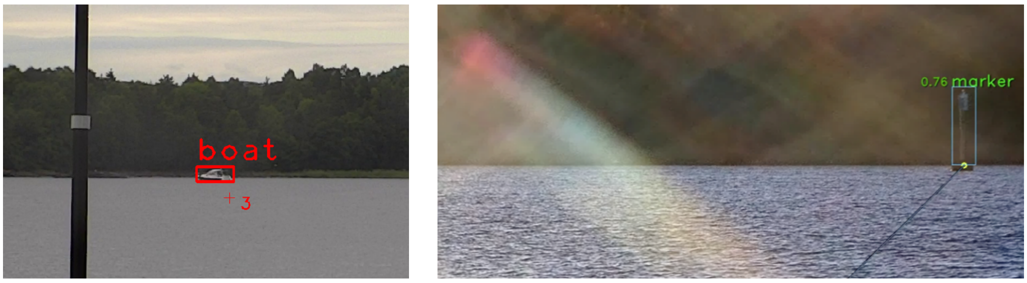
AIS should not be integrated via measurement-level fusion. Our initial tests in 2019 integrated AIS through *measurement-level* fusion, where AIS reports are treated like any other sensor input. However, due to the low precision of AIS positional data, this approach can potentially have worse tracking performance compared to only using data from our sensors. *Track-level* fusion, where AIS and sensor tracks are mostly kept separate and gradually combined, is a better approach. With this type of tracker architecture, we could achieve the best of both worlds: The performance of the sensor-based tracker seamlessly combined with the metadata provided by AIS. However, since AIS reports can be spoofed, it's crucial to ensure that this data cannot compromise the tracker output.

Investing in robust hardware that "just works" pays off. It's easy to underestimate the physical stress that uncrewed military vehicles are actually exposed to during normal operation. In addition to challenging climates, shock and vibration with significant damage potential will occur, and this must be taken into account when purchasing equipment and designing your system. Thoroughness and attention to detail are keys to success in all aspects of hardware. Proper mounting, cooling, shielding of cables, sufficient power, and robust data links are all examples of measures that will reduce the risk of occasional hardware related errors, which can be hard to debug. Tracking down errors in software that turns out to be caused by hardware will always feel like a waste of time. While opting for robust hardware often ensures reliability, this choice might also mean sacrificing the latest technological advancements.

A modular software architecture allows us to be flexible on where to focus development. We have prioritized demonstrating a complete and functional system. The modularity of our architecture enables us to quickly develop a full-stack solution, placing greater emphasis on the integration between components rather than optimizing each individual part. By starting with simple building blocks, we can quickly identify and resolve major issues in the overall system, and then address the limitations of individual components directly and independently.

Ongoing development 2024

Our research efforts for 2024 are mainly focused toward supporting FFI project 1635, "New military applications for autonomous uncrewed surface vessels (ASV)". We will continue to explore various aspects of collaboration and cooperation between uncrewed systems, mainly within the topic of shared SA in multi robot systems (MRS). For collaborating autonomous units to converge on common solutions to their tasks, they need to have aligned situational awareness. Here, we distinguish between *close* and *distant collaboration* based on whether or not the platforms observe much of the same scene simultaneously. In distant collaboration, our primary focus is on communicating target data, where one platform must be able to produce a description that another platform can later use to re-identify the same target. Close collaboration, on the other hand, offers far more opportunities for enhancing the shared SA. We can use joint observations to improve



(a) (b)

Figure 7.1 Detection of a boat (a) and a sea marker (b) in daylight images. The detector in (a) [15] was trained on the stock MS COCO dataset [19], while the detector in (b) [16] was also trained on a private dataset focusing on Norwegian sea markers [20], [21]. Both figures are from work carried out in 2017 – 2018.

Figure 7.2 Phased array radar from Kongsberg Seatex mounted on Frigg.

both absolute and relative navigation between platforms, fill in blind zones, improve resolution in mapping, and improve tracking performance. If high-bandwidth communications are available, one platform can function as a loosely connected sensor rig for the other. Alternatively, some processing can be done on both platforms, with refined states shared and merged into a joint situational picture.

The outlined scenario for project 1635 is for the ASVs, within a delimited area, to be able to find and identify a specific target based on a prior description of said target. Specifically, we use PTZ cameras to both classify objects and recognize the special target based on visual signature, all integrated in the existing tracking pipeline. For object detection and classification in images, we use mostly standard CNN- and transformer-based detector architectures [14]–[18] trained on public and private datasets, as illustrated in fig. 7.1. Previously, we have been working in land-based applications with methods for visual re-identification of objects and aim to build on these experiences for the ASV application. Once again, we use the multi-feature paradigm to incorporate both classification and visual re-identification directly in the tracking.

In the course of 2024, we have also installed a new phased array radar from Kongsberg Seatex on *Frigg* (fig. 7.2). The radar is ideally suited for short range drone detection and counter-UAS (CUAS), but we intend to also use it for mapping and safe navigation in tandem with our existing sensors.

Acronyms

AIS automatic identification system
A radio-based system for marine vessels to broadcast their position and related metadata.

ASV autonomous surface vessel

AUV autonomous underwater vehicle

CNN convolutional neural networks

COLREGS convention on the international regulations for preventing collisions at sea

CUAS counter-UAS

EM electro magnetical

GNSS global navigation satellite system

IMU inertial measurement unit

INS inertial navigation system

MMSI maritime mobile service identity
9-digit code that uniquely identifies maritime radio communication equipment, including AIS transceivers.

MRS multi robot systems

NM nautical mile

PTZ pan-tilt-zoom

SA situational awareness

References

[1] R. Smestad, M. V. Larsen, and T. V. Haavardsholm, “Superflow – An efficient processing framework for modern C++,” Forsvarets Forskningsinstitutt (FFI), Kjeller, Norge, FFI-Rapport 24/00556, 2024.

[2] S. Thrun, “Probabilistic robotics,” *Communications of the ACM*, vol. 45, no. 3, pp. 52–57, 2002.

[3] B. Jalving, K. Gade, O. K. Hagen, *et al.*, “A toolbox of aiding techniques for the hugin auv integrated inertial navigation system,” in *Oceans 2003. Celebrating the Past... Teaming Toward the Future (IEEE Cat. No. 03CH37492)*, IEEE, vol. 2, 2003, pp. 1146–1153.

[4] O. K. Hagen, R. Smestad, E. B. Nilssen, *et al.*, “GNSS-independent maritime navigation – High speed vessels,” Forsvarets Forskningsinstitutt (FFI), Kjeller, Norge, FFI-Rapport 21/00390 (BEGRENSET), 2021.

[5] “Simrad halo-3 pulse compression radar.” (2024), [Online]. Available: https://www.simrad-yachting.com/simrad/series/halo-radar/halo-3-radar-w3-arrayri1220m-cbl/ (visited on 04/25/2024).

[6] “Ouster os2 lidar.” (2024), [Online]. Available: https://ouster.com/products/hardware/os2-lidar-sensor (visited on 04/25/2024).

[7] “Ouster os1 lidar.” (2024), [Online]. Available: https://ouster.com/products/hardware/os1-lidar-sensor (visited on 04/25/2024).

[8] “Flir m400 multi-sensor marine camera.” (2024), [Online]. Available: https://www.flir.eu/products/m400/?vertical=maritime&segment=solutions (visited on 04/25/2024).

[9] S. S. Blackman and R. Popoli, *Design and analysis of modern tracking systems*. Artech House Publishers, 1999.

[10] M. Larsen., S. Rolfsjord., D. Gusland., *et al.*, “Base: Probably a better approach to visual multi-object tracking,” in *Proceedings of the 19th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications - Volume 4: VISAPP*, INSTICC, SciTePress, 2024, pp. 110–121, ISBN: 978-989-758-679-8. DOI: 10.5220/0012386600003660.

[11] R. Szeliski, “Image alignment and stitching: A tutorial,” *Foundations and Trends® in Computer Graphics and Vision*, vol. 2, no. 1, pp. 1–104, 2007, ISSN: 1572-2740. DOI: 10.1561/0600000009. [Online]. Available: http://dx.doi.org/10.1561/0600000009.

[12] M. S. Rundhovde, E.-L. M. Ruud, F. Hermansen, *et al.*, “Unmanned Surface Vessel Autonomy for Cooperative Mine Sweeping Capabilities,” Forsvarets Forskningsinstitutt (FFI), Kjeller, Norge, FFI-Rapport 22/01517, 2022.

[13] G. I. Bitar, E.-L. M. Ruud, J. Sandrib, *et al.*, “Automatic underway auv recovery using a usv,” Forsvarets Forskningsinstitutt (FFI), Kjeller, Norge, FFI-Rapport 22/01519, 2022.

[14] S. Ren, K. He, R. Girshick, *et al.*, *Faster r-cnn: Towards real-time object detection with region proposal networks*, 2016. arXiv: 1506.01497 [cs.CV].

[15] W. Liu, D. Anguelov, D. Erhan, *et al.*, "Ssd: Single shot multibox detector," in *Lecture Notes in Computer Science*. Springer International Publishing, 2016, pp. 21–37, ISBN: 9783319464480. DOI: 10.1007/978-3-319-46448-0_2. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-46448-0_2.

[16] J. Redmon and A. Farhadi, "Yolov3: An incremental improvement," *arXiv*, 2018.

[17] Z. Ge, S. Liu, F. Wang, *et al.*, "YOLOX: Exceeding YOLO Series in 2021," *arXiv*, Jul. 2021. [Online]. Available: http://arxiv.org/abs/2107.08430.

[18] Q. Wei, B. Zeng, J. Liu, *et al.*, *Litetrack: Layer pruning with asynchronous feature extraction for lightweight and efficient visual tracking*, 2023. arXiv: 2309.09249 `[cs.CV]`.

[19] T.-Y. Lin, M. Maire, S. Belongie, *et al.*, "Microsoft coco: Common objects in context," in *Computer Vision – ECCV 2014*, Cham: Springer International Publishing, 2014, pp. 740–755, ISBN: 978-3-319-10602-1.

[20] J.-K. Mathisen and J. Lowzow, "Hvordan benytte maskinlæring i navigasjon," *Sjøkrigsskolen Bacheloroppgave*, Dec. 2017.

[21] M. Engvik and E. Omland, "Maskinlæring som klassifiseringsverktøy – den nye marinen?" *Sjøkrigsskolen Bacheloroppgave*, Dec. 2018.

About FFI

The Norwegian Defence Research Establishment (FFI) was founded 11th of April 1946. It is organised as an administrative agency subordinate to the Ministry of Defence.

FFI's mission

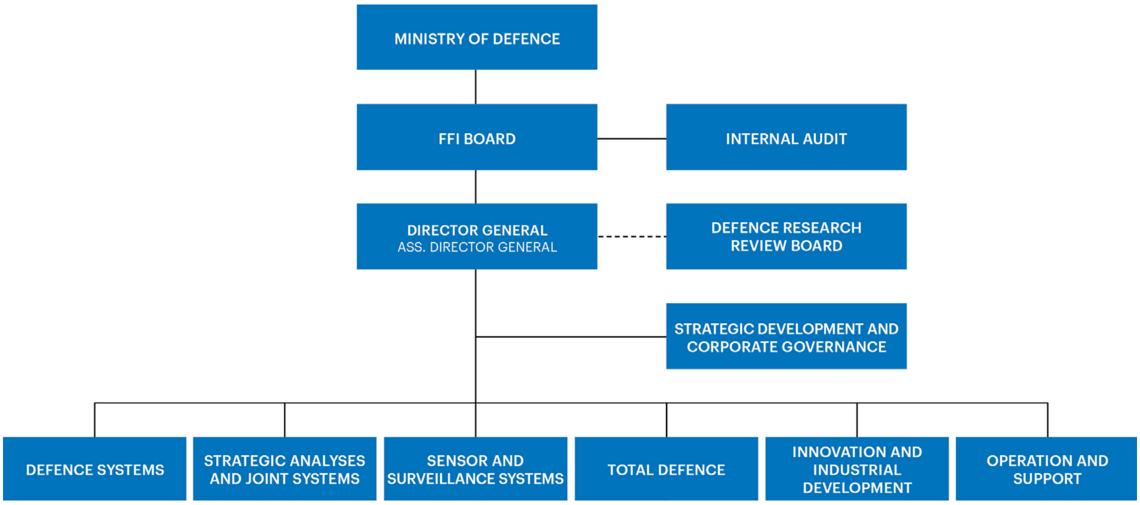
FFI is the prime institution responsible for defence related research in Norway. Its principal mission is to carry out research and development to meet the requirements of the Armed Forces. FFI has the role of chief adviser to the political and military leadership. In particular, the institute shall focus on aspects of the development in science and technology that can influence our security policy or defence planning.

FFI's vision

FFI turns knowledge and ideas into an efficient defence.

FFI's characteristics

Creative, daring, broad-minded and responsible.



Forsvarets forskningsinstitutt (FFI)
Postboks 25
2027 Kjeller

Besøksadresse:
Kjeller: Instituttveien 20, Kjeller
Horten: Nedre vei 16, Karljohansvern, Horten

Telefon: 91 50 30 03
E-post: post@ffi.no
ffi.no

Norwegian Defence Research Establishment (FFI)
PO box 25
NO-2027 Kjeller
NORWAY

Visitor address:
Kjeller: Instituttveien 20, Kjeller
Horten: Nedre vei 16, Karljohansvern, Horten

Telephone: +47 91 50 30 03
E-mail: post@ffi.no
ffi.no/en