# SMAS: A Smart Alert System for Localization and First Response to Fires on Ro-Ro Vessels

Paschalis Mpeis,
Athina Hadjichristodoulou
University of Cyprus
Nicosia, Cyprus
{pmpeis01,ahadji09}@ucy.ac.cy

Jaime Bleye Vicario
Integral Maritime Safety Centre
Jovellanos
Gijón, Spain
jaimebv@centrojovellanos.es

Demetrios Zeinalipour-Yazti
University of Cyprus
Nicosia, Cyprus
dzeina@ucy.ac.cy

## ABSTRACT

In this paper we present the design and utility of a Smart Alert System (SMAS) for quick first response and effective fighting of fires in their initial stages on roll-on / roll-off (ro-ro) vessels. Given that localization within the steel structures of a vessel is currently an open problem, we develop a ground-breaking infrastructure-free (i.e., "zero" infrastructure) localization architecture to localize on ordinary smartphones in vessel indoor spaces that lack any infrastructure whatsoever (e.g., Wi-Fi, BLE, UWB, RFID, LED). Particularly, we developed a smartphone-based *Computer Vision (CV)* localization technology upon which an innovative nearest neighbor information communication channel for spatio-textual alerting between first responders is constructed. We also develop information sharing channels of multimedia content (e.g., heat scans) but also provide an integrated search and navigation tool for stationary and mobile assets of a vessel and the fire control operation. We have developed a complete functional system of SMAS using a micro-service edge architecture that deploys sharded sqlite micro-databases. We will present SMAS in two modes: (i) Online Mode, where attendees will be able to carry out simulated emergency scenarios; and (ii) Offline Mode, where attendees will be able to observe emergency scenarios recorded on video traces.

## CCS CONCEPTS

• **Information systems** → **Information systems applications**.

**ACM Reference Format:**
Paschalis Mpeis,, Athina Hadjichristodoulou, Jaime Bleye Vicario, and Demetrios Zeinalipour-Yazti. 2022. SMAS: A Smart Alert System for Localization and First Response to Fires on Ro-Ro Vessels. In *Proceedings of The 16th ACM International Conference on Distributed and Event-based Systems (DEBS '22).* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3524860.3543282

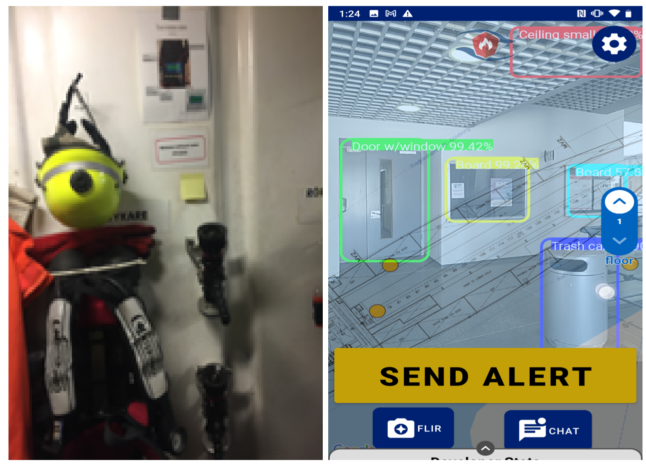


**Figure 1: Smart Alert System (SMAS) provides a mobile application that allows associating physical assets (e.g., deck patterns, bulkhead patterns, hoses, fixed installations, signs, control buttons) in a micro indoor spatial information system. The SMAS application then provides the location of a user using a complete alerting information infrastructure for first responders we have developed.**

## 1 INTRODUCTION

With global trade increasing over the last decade, the maritime industry has grown significantly. Each month, more than 10,000 vessels pass through the Strait of Dover, including cargo ships, tanker ships, roll-on/roll-off ships (Ro/Ro), passenger ships and others. Globalization and the recent advent of low-cost manufacturing facilities has fueled an interest in innovative systems to tackle existing issues to lower the costs, improve safety standards and comply with international and regional regulations.

LASH FIRE[1] is an international EU-funded research project aiming to significantly reduce the risk of fires on board ro-ro ships. A particular focus in the project is the development and validation of smart technical solutions for quick first response and effective fighting of fires in their initial stage. Particularly, the research objective is to develop an innovative geo-positioning technology to allow a more efficient first response for fires on ro-ro vessels. Besides the core geo-positioning technology, the aim is also to provide the building blocks of a novel ship indoor information system.

In this paper, we expose the design and utility of a *Smart Alert System (SMAS)* for quick first response and effective fighting of fires on roll-on / roll-off vessels in their initial stages. Even though in the incident of a fire all crew members may act as first responders, there are some of them (e.g., *fire patrol members, able seamen, personnel form the engine control room*) that are more likely to act as first responders due to their normal access to restricted cargo decks

[1]EU Project LASH FIRE. https://lashfire.eu/


*DEBS '22, June 27–30, 2022, Copenhagen, Denmark*

ACM ISBN 978-1-4503-9308-9/22/06…$15.00
https://doi.org/10.1145/3524860.3543282

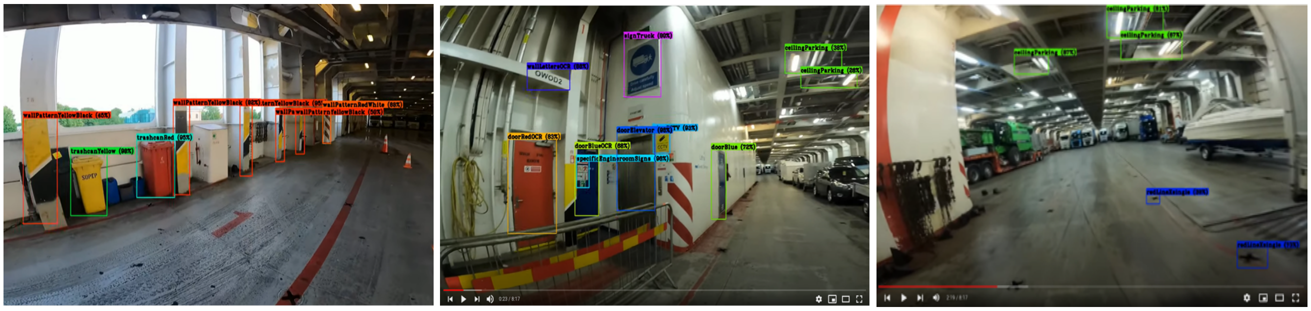

**Figure 2: Computer Vision Localization: Training machine learning models on footage we received from a real ro-ro vessel has shown that static non-movable objects can be captured under a variety of lighting conditions with great accuracy, yielding a robust referencing system for localization in environments that lack infrastructure (e.g., Wi-Fi, BLE, UWB, RFID, LED).**

spaces. Fire patrol members are designated first responders and are the target personnel for the technology we develop in the scope of SMAS.

By equipping first responders with powerful mobile computing devices will allow them to increase their *cyber-physical senses* (i.e., multiple sensing devices like heat scanner in a tiny device), be *connected* (with the bridge and other personnel, discarding possibly outdated communication gear), be *informed* (e.g., carrying bulky manuals and maps in digital form), and *location-aware* (i.e., localization, navigation and tracking of mobile and static assets). These are all dimensions that will increase fire safety by the means of state-of-the-art information technology that has proven itself in everyday life scenarios and that is for the same reason also unobtrusive, with a low learning curve, adaptable through software, and economically viable for massive deployment.

The SMAS architecture builds upon many years of experience with Anyplace, which is a Wi-Fi localization, navigation, crowdsourcing and indoor modeling platform developed over the years at the University of Cyprus. Given that Wi-Fi technology is not widely available on ro-ro vessels, we developed a "zero" infrastructure localization system using Computer Vision (CV) localization.

Particularly, the developed localization technology requires no infrastructure whatsoever (e.g., Wi-Fi, BLE, UWB, RFID, LED) and it can be deployed at scale with minimal cost on newer or older vessels. Providing a network channel, through sparse Wi-Fi or Mobile 4G/5G antennas can provide the full potential of the developed alerting application and provide a communication channel between first responders. This, however, is only complimentary to the primitive fingerprinting localization technology we develop in this work.

Our proposed method relies on three stages: (i) *Training:* vessel owners supply to the software team video recordings of the vessel's interior spaces with a particular camera lens focus on static objects. The video recordings are analyzed on a deep learning data center to produce a YOLO Machine Learning model [3] then transformed to a Tensorflow Lite runtime. This process is carried out once per vessel type (or vessel family - in case multiple vessels have similar internal objects); (ii) *Logging:* Subsequently, the model is loaded to a smartphone app provided to the vessel owners, which are asked to associate surrounding objects with their location by clicking on a map, yielding a Fingerprint database. Logging is carried out once per unique vessel; and (iii) *Localization:* The first responders utilize a smartphone application using the fingerprint database that shows both to them and any nearby user (who is connected to a telecommunication network) the location of the responders using the app (e.g., also on the bridge).

The contributions of this work are summarized as follows:

- We present the developments of a ground-breaking localization technology using computer vision.
- We present the developments of a location-oriented smart alert system relying on computer vision localization, providing quick solutions to a variety of spatial operators.

## 2 SMAS CV LOCALIZATION

In this section we describe the three stages of our innovative CV Localization approach we prototyped as part of SMAS, namely: *training*, *logging*, and *localization*.

### 2.1 Training

In this sub-section we describe an extensive annotation of videos we carried out with the CVAT video on our dedicated deep learning datacenter. Computer Vision Annotation Tool (CVAT) - CVAT.org - is a free, open source, web-based image and video annotation tool which is used for labeling data for computer vision algorithms. Originally developed by Intel, CVAT is designed for use by a professional data annotation team, with a user interface optimized for computer vision annotation tasks.

Particularly, we obtained 15 lengthy and shorter videos from a real ro-ro vessel as part of the LASH-FIRE project (captured by using a chess mounted GoPro camera). The videos were annotated by a team of 5 persons over a period of 4 weeks with several meetings iterations, to refine the quality of the constructed neural network model we built. We have additionally processed video capture at the University of Cyprus (UCY) to develop additional video footage for a second CV model, designated for our testing environment.

To execute CVAT in our environment we have setup a dedicated rackable server featuring a state-of-the-art NVIDIA Tesla V100 GPU card. This card reduces training time down to a few hours from several days or weeks we required initially on Google's free Colab

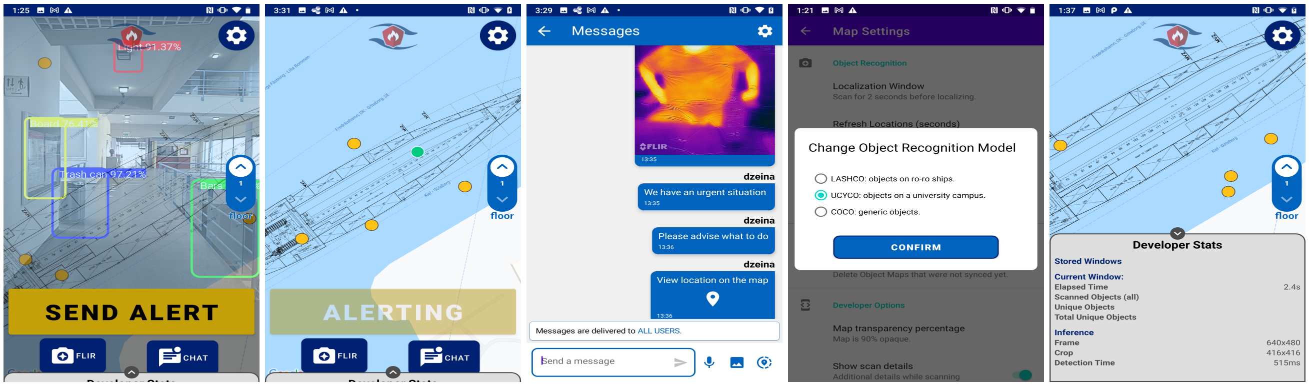


**Figure 3: SMAS Mobile App: (i) *CV Localization:* happens in the background and can be overlayed on the vessel map shown (see background); (ii) *Alert:* can be send out to other first responders with a central button or voice command at variety of spatial granularities (i.e., KNN, Bounding Box, Deck, All); (iii) *Chat:* enables the communication between first responders with text messages, images, or transcribed text at variety of spatial granularities; (iv) *Model:* one can change the model from LASHCO, COCO or UCYCO, depending on the target environment. Additional CV models are structured in the backend and can easily be integrated; (v) *Develop:* various advanced functionalities are available through the menu.**

environment [2]. Additionally, by processing our data locally allows us to improve I/O performance, as handling large video traces over a slow network can become the bottleneck.

## 2.2 Logging

In this sub-section we start out with a brief description of the design methodology behind our CV Logger. Particularly, this is a mobile application that can be used by the people setting up the localization system on the vessel to collect reference points on the vessel with a smartphone. The localization setup personnel (e.g., crew members or first respondents) scan the vessel by following some simple guidelines provided on the first tab of the tool. We also experimented with logging in front of video capturing played on a computer screen and preliminary results are quite encouraging, meaning that in the future this step might be voided. The logger tool is equipped with a variety of machine learning models. Particularly, the logger allows selecting between the following three machine learning corpuses:

**Common Objects in Context (COCO)** [3]: this is a large-scale object detection, segmentation, and captioning dataset (24MB). COCO has several features, namely: Object segmentation, Recognition in context, Superpixel stuff segmentation, 330K images (200K labeled), 1.5 million object instances, 80 object categories, 91 stuff categories, 5 captions per image, 250,000 people with keypoints. COCO was used for development and for testing the platform for localization accuracy in ordinary spaces but will also be used for the demonstration at the conference venue.

**University of Cyprus Objects in Context (UCYCO)**: this is a small-scale object detection, segmentation, and captioning dataset specifically for the University of Cyprus buildings (23MB). This model was particularly useful for the development stage and was constructed with the assistance of Google Colab (i.e., GPU resources in the cloud).

**LASH Objects in Context (LASHCO)**: this is a mid-scale object detection, segmentation, and captioning dataset specifically for a vessel as part of LASH FIRE project (24MB). LASHCO has approximately 100 classes. The given object detection, segmentation, and captioning dataset will be used for testing the platform and localization accuracy in vessel spaces as part of the LASH FIRE project and has been constructed on a dedicated deep learning server, namely an HP DL380 Gen10 with 80 logical processors and a powerful NVIDIA v100 card.

## 2.3 Localization

Our system uses CV Radiomap-based indoor localization, which stores computer vision signals from cameras in a database at a high density. The localization subsystem utilizes the following high-level routine: in an offline phase, a logging application records the so called CV fingerprints, which comprise the various identified objects bounding rectangles at certain locations (x,y) pin-pointed on a building floor map (e.g., every few meters). Subsequently, in a second offline phase, the CV fingerprints are joint into a NxM matrix, coined the CV RadioMap, where N is the number of unique (x,y) fingerprints and M the total number of CV objects. Finally, a user can compare its currently observed CV fingerprint against the RadioMap to find the best match, using known algorithms such as KNN or WKNN. A similar methodology can be applied to other types of signals, for instance, our earlier Anyplace system was using Wi-Fi and we aim to integrate all measurements under the same roof in the future.

# 3 THE SMAS APPLICATION

In this section we describe the two basic elements of the SMAS application, namely: (i) the SMAS mobile application used by first

[2] Google Colab. https://colab.research.google.com/
[3] COCO dataset. https://cocodataset.org/

responders; and (ii) the SMAS Backend, comprising of an indoor GIS and a CV Localization method, both developed in-house.

The SMAS mobile application allows a user to interact through a very simple interface that is designed on the premise that the user won't always look at the smartphone screen and that usage of the on-screen keyboard might not be accessible either (as such we have voice interface with voice-to-text transcription). The app is written in Android X and Kotlin. They follow the MVVM paradigm and asynchronously handles communication and computation using the *coroutine* invocation mechanism. This is particularly important for involved operations, such as *object recognition* or *inference*, which are quite heavy computationally for the smartphone. There are numerous tunning features that we will present during the demonstration to better explain both the computer vision localization subsystem and the emergency first response application.

Regarding the backend, anything related to Smart Alert was implemented in the separate SMAS API service to allow the service to run on the edge to minimize latency. We utilize a pure relational embedded implementation to provide high performance localization on the edge. We also opted for a micro-database architecture, which uses different SQLite instances that serve as separate queues for user messages or alerts (using shared and exclusive locks) to provide a primitive operational functionality on basic hardware that can be something as little as a Raspberry PI. All localization algorithms are implemented in SQL with respective Views, Indexes and UDFs. The provision of localization algorithms using a relational SQL implementation made it particularly simple to fine tune the performance of the algorithms before deployment. Particularly, for data-driven (i.e., fingerprint-based) localization algorithms this SQL paradigm has exposed particular benefits that will be formalized further in the future. Finally, caching is carried out on the handheld on a another embedded sqlite instance.

# 4 RELATED WORK

The research behind a smart, location-aware, alert system that is tailored for restricted environments, such as ro-ro ships, spans across two different domains. These are indoor localization and real-time chat applications.

## 4.1 Indoor Localization

Indoor Localization systems [2] usually consist of indoor models, like floormaps and *Points-of-Interest (POIs)*, along with wireless signals used to localize users. Anyplace[4] is a fingerprint Wi-Fi localization system that provides superior accuracy but requires unfortunately Wi-Fi access points for the localization task. The localization subsystem of SMAS on the other hand relies on Computer Vision that has made enormous progress over recent years [3]. As such, it can readily digest high-volumes of video sources and generate machine learning models that understand the environment. These models can be loaded to a diverse hardware range.

## 4.2 Messaging Systems

Messaging systems have drawn significant attention, both by the research and industrial communities [1]. Leading industrial applications in terms of user base, such as Whatsapp and Telegram, focus on end-to-end encryption. SnapChat became the first widely-used service that focused around ephemeral social interactions. Completely anonymous, location-aware platforms have also emerged. Yik Yak and Rayzit allows posting to nearby users that are within a 5km radius. While anonymity is compelling, it is not needed in the context of a smart alert system. In fact, a user's location and identity is required to carry out an efficient first response. Additionally, these systems are quite rigid when it comes to tackling emergencies due to the lack of spatial awareness.

[4]Anyplace. `https://anyplace.cs.ucy.ac.cy/`

# 5 DEMONSTRATION SCENARIO

During the demonstration, the attendees will be able to appreciate the functionality, the visualization abstraction and the performance of SMAS in a real setting.

**Demo Artifact:** We will deploy the latest version of SMAS that includes: i) the SMAS Logger; ii) the SMAS Object Recognition and Localization Engine; iii) the SMAS Chat App; and iv) the SMAS Search and Navigation Engine. Our system has a low setup time because it relies on existing infrastructure available at the conference venue (i.e., COCO model) or the University campus (UCYCO model). We will not require anything beyond a Wi-Fi network to achieve communication between first responders (Wi-Fi is not necessary for the localization part). Particularly, our aim is to both demonstrate the live capabilities of SMAS but also the capabilities of SMAS over pre-recorded footage running on the screen screen of the demonstrator's PC. Our demonstration is fully compatible for both a physical in-person presentation and quite engaging even for a remote / hybrid presentation through video traces.

**Demo Plan:** We will carry our demonstration out in 2 modes: (i) Online Mode, where we will ask conference attendees to collect CV fingerprint data using the Logger and upload them to the SMAS backend. We will provide respective Android devices to the attendees. After uploading the fingerprints, attendees will be asked to observe and interact with the collected data to improve decision making (e.g., where to collect more fingerprints for better localization accuracy). Finally, the audience will be asked to use the SMAS Navigator to take advantage of the updated indoor models; and (ii) Offline Mode, where we show to attendees how we localize in various indoor spaced with the SMAS system.

## Acknowledgments

The Project LASH FIRE has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 814975.

# REFERENCES

[1] Georgios Chatzimilioudis, Constantinos Costa, Demetrios Zeinalipour-Yazti, Wang-Chien Lee and Evaggelia Pitoura, "Distributed In-Memory Processing of All k Nearest Neighbor Queries," *IEEE Transactions on Knowledge and Data Engineering*, vol. 28, no. 4, pp. 925-938, doi: 10.1109/TKDE.2015.2503768, April 2016.
[2] Paschalis Mpeis, Thierry Roussel, Manish Kumar, Constantinos Costa, Christos Laoudias, Denis Capot-Ray Demetrios Zeinalipour-Yazti, "The Anyplace 4.0 IoT Localization Architecture," *21st IEEE International Conference on Mobile Data Management (MDM)*, pp. 218-225, doi: 10.1109/MDM48529.2020.00045, 2020.
[3] YOLOv4: Optimal Speed and Accuracy of Object Detection Alexey Bochkovskiy, Chien-Yao Wang, Hong-Yuan Mark Liao, "Scaled-YOLOv4: Scaling Cross Stage Partial Network," *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 13024-13033, doi: 10.1109/CVPR46437.2021.01283, 2021.