

# Horizon Based Cross-View Geolocalization in Maritime Environment

Extended Abstract

Alexandre Foucher

Lab-STICC, CNRS UMR 6285, Université Bretagne Sud  
Lorient, France  
alexandre.foucher@univ-ubs.fr

Dominique Heller

Lab-STICC, CNRS UMR 6285, Université Bretagne Sud  
Lorient, France  
dominique.heller@univ-ubs.fr

Cédric Seguin

Lab-STICC, CNRS UMR 6285, Université Bretagne Sud  
Lorient, France  
cedric.seguin@univ-ubs.fr

Johann Laurent

Lab-STICC, CNRS UMR 6285, Université Bretagne Sud  
Lorient, France  
johann.laurent@univ-ubs.fr

## ABSTRACT

Geolocation is essential for autonomous navigation of surface maritime systems. Most current methods rely on GNSS or radar systems, which are not suitable for hostile environments. Visual Geolocalization (VG) offers a passive alternative, but poses a significant challenge, particularly with a limited camera field of view. This paper introduces a novel VG approach utilizing a 50° camera for horizon-based cross-view correlation. We correlate ground images with simulated horizons rendered from a Digital Elevation Model (DEM) using a siamese network and compact visual descriptors. Our results demonstrate the feasibility of GNSS-free navigation for maritime systems, achieving a mean localization accuracy of 38 m and heading accuracy up to 8° over a 25 km<sup>2</sup> area, starting from an unknown position. This opens new possibilities for robust and autonomous maritime operations in unfamiliar environments with limited ground imagery.

## KEYWORDS

Computer-vision; GNSS-free; USV; Embedded System

### ACM Reference Format:

Alexandre Foucher, Cédric Seguin, Dominique Heller, and Johann Laurent. 2026. Horizon Based Cross-View Geolocalization in Maritime Environment: Extended Abstract. In *Proc. of the 25th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2026)*, Paphos, Cyprus, May 25 – 29, 2026, IFAAMAS, 3 pages. <https://doi.org/10.65109/YIUE1134>

## 1 INTRODUCTION

One critical aspect of achieving autonomy is geolocation, this enables autonomous navigation planning by the system without human intervention. Usually, Global Navigation Satellite System (GNSS) is a reliable and accurate method. However, the accuracy can be compromised due to dense vegetation, structure obstruction or also malicious attacks such as spoofing and jamming.

In this context, it is essential to develop fail-safe localization method for Unmanned Surface Vehicle (USV). The scientific literature offers various alternatives to GNSS in scenarios where positioning is no longer reliable. Most of them rely on active sensors such as radar [6, 11, 12], lidar [14] or sonar [5, 9]. However, in a hostile environment, discretion can be a crucial element of success, limiting the use of these emitting methods.

Otherwise, dead-reckoning is a passive well known alternative for navigation [2, 10]. Nevertheless, this method remains highly dependent on the quality of the sensor and inevitably suffers from drift over time.

In this paper, we propose an topographic cross-view visual geolocalization approach using only DEM as reference in coastal environment. Our method aims to enable real-time geo-localization using only limited field of view camera. The geo-localization is based on the correlation between the rendered horizon from known DEMs and ground images captured by the system’s camera. This approach could allow system to operate on terrain where the only preliminary informations available are elevation map.

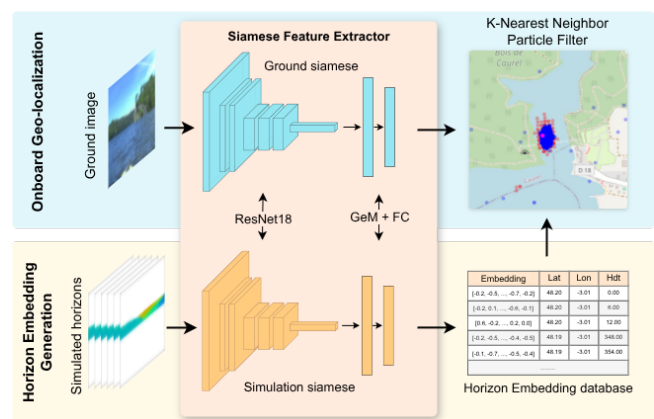
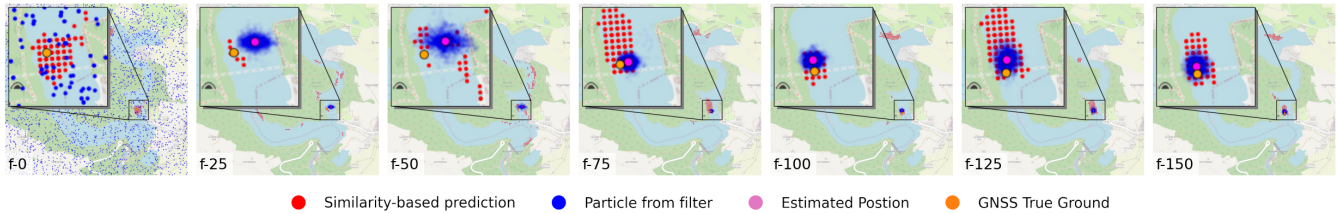


Figure 1: Diagram of the geo-localization process, which consists of two primary stages: offline horizon simulation dataset generation, and onboard geo-localization using nearest horizon descriptor search and particle filtering for geolocation estimation.

This work is licensed under a Creative Commons Attribution International 4.0 License.

Proc. of the 25th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2026), C. Amato, L. Dennis, V. Mascardi, J. Thangarajah (eds.), May 25 – 29, 2026, Paphos, Cyprus. © 2026 International Foundation for Autonomous Agents and Multiagent Systems ([www.ifaaamas.org](http://www.ifaaamas.org)). <https://doi.org/10.65109/YIUE1134>



**Figure 2: Demonstration of the complete geolocalization process on test scenario 0. The top image sequence shows the evolution of geolocalization over time based on specific log frame noted f.**

## 2 PROPOSED APPROACH

To obtain accurate geolocation from an image despite the limited field of view, we propose a pipeline based on deep-learning as shown in figure 1. This pipeline illustrates our cross-view geolocalization method, highlighting in one hand the pre-processing part using elevation model simulation and in the other hand the onboard geolocation estimation. Horizon correlation is based on image retrieval task where the highest correlation score is inversely proportional to descriptor distance. Descriptors also known as embedding or feature vector depending of the field of research are computed via feature extraction made using Convolutional Neural Network (CNN). Those descriptors are compared using euclidean based nearest neighbor search to retrieve highest correlation hypotheses. Then, hypotheses are filtered using Monte Carlo localization particle filter to reduce outlier and increase accuracy [3, 4].

As our approach is based on DEM, we assume that we know approximately the location where the system operates to limit the searching area. The principle is to generate a sufficiently exhaustive dataset of digital horizons rendered from topographic data. Subsequently, after training a model to recognize similarities and differences in the horizon, a vectorial database is created to store descriptors from the simulation dataset. It is important to notice that all computationally expensive processes are performed beforehand. Finally, only the database of descriptors and the trained model are stored onboard the USV, limiting memory usage and required computing power. For localization, the drone will then need to perform only an inference and a search for the most similar horizons. As mentioned, we distinguish between the dataset, which represents image sets for deep-learning training, and the database, which represents descriptor storage.

Following the observation by Berton et al. [1], recent research tends to increase the complexity of model architecture. Resulting in heavier and slower architectures that may not be suitable for long-term use on embedded systems with limited resources, such as battery capacity and computing power. For this reason, we explore several compact descriptors with fewer than 256 dimensions using the lightweight network ResNet18 [7] combined with GeM [13] aggregation layer.

Each network is trained on an identical dataset for 10 epochs, using the same hyperparameters, with only the descriptor size differing. We use the weighted soft margin triplet loss [8] for training by exploiting triplets based on a relevance score. Indeed, we identify similarities based on the difference in coordinates: the closer the elements are physically, the higher the score.

## 3 RESULTS

In order to evaluate our approach, we initially extracted four complete and unique logs, ordered by difficulty according to occlusion, viewing distance, and weather conditions. Table 1 indicates the method accuracy for the position and the heading when the particle filter has converged. In the best case, we obtain a mean error of 38m for test set 1 in a 25km<sup>2</sup> area where we don't know the initial localization. We also retrieve a mean heading accuracy up to 8° in the best scenario.

**Table 1: Analysis of the geolocalization accuracy made of 200 simulations using 5000 particles on all tests using best model according previous results.**

Test set	Desc. Dim.	Position Acc.		Heading Acc.	
		Min	Mean	Min	Mean
0	256	0.76m	57.78m	0.02°	<b>7.38°</b>
1	128	0.78m	<b>37.39m</b>	0.02°	18.17°
2	32	0.56m	62.54m	0.01°	14.24°
3	256	1.70m	89.14m	0.01°	16.05°

This approach allows a system to be located within a restricted area without divergence over time, unlike dead reckoning or visual odometry. As shown in Figure 2, the particle filter in blue rapidly converges toward the ground truth in orange in 3 iterations. Particles also continues to follow the ground truth even in an occluded environment, from frame 25 to frame 50.

## 4 CONCLUSION AND PERSPECTIVE

In this study, we demonstrate a visual geo-localization approach in a maritime environment without GNSS, using only a camera with a limited field of view. We use the horizon simulated from a digital elevation model as a reference for a method based on cross-view and particle filter. This work highlights the feasibility and potential of autonomous navigation in hostile areas where the ground image is limited and GNSS is no longer reliable.

Our method achieves geolocalization with an average accuracy of 38m in best case and a heading difference of less than 8° from the actual ground truth in an area of 25km<sup>2</sup> with an unknown starting point and using only a 50° field of view camera.

## ACKNOWLEDGMENTS

This study is being conducted with the funding support of Agence de l'Innovation de Défense - Ministère des Armées - France.

## REFERENCES

- [1] Gabriele Berton, Carlo Masone, and Barbara Caputo. 2022. Rethinking visual geolocalization for large-scale applications. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 4878–4888.
- [2] Nadav Cohen and Itzik Klein. 2024. Inertial navigation meets deep learning: A survey of current trends and future directions. *Results in Engineering* (2024), 103565.
- [3] Frank Dellaert, Dieter Fox, Wolfram Burgard, and Sebastian Thrun. 1999. Monte carlo localization for mobile robots. In *Proceedings 1999 IEEE international conference on robotics and automation (Cat. No. 99CH36288C)*, Vol. 2. IEEE, 1322–1328.
- [4] Lena M Downes, Ted J Steiner, Rebecca L Russell, and Jonathan P How. 2023. Wide-area geolocalization with a limited field of view camera in challenging urban environments. *arXiv preprint arXiv:2308.07432* (2023).
- [5] Ove Kent Hagen and Kjetil Bergh Anonsen. 2014. Using terrain navigation to improve marine vessel navigation systems. *Marine Technology Society Journal* 48, 2 (2014), 45–58.
- [6] Jungwook Han, Yonghoon Cho, and Jinwhan Kim. 2019. Coastal SLAM with marine radar for USV operation in GPS-restricted situations. *IEEE Journal of Oceanic Engineering* 44, 2 (2019), 300–309.
- [7] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2016. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 770–778.
- [8] Sixing Hu and Gim Hee Lee. 2020. Image-based geo-localization using satellite imagery. *International Journal of Computer Vision* 128, 5 (2020), 1205–1219.
- [9] Rohan Kapoor, Alessandro Gardi, and Roberto Sabatini. 2020. Acoustic positioning and navigation system for gnss denied/challenged environments. In *2020 IEEE/ION Position, Location and Navigation Symposium (PLANS)*. 1280–1285.
- [10] Wenwen Liu, Yuanchang Liu, and Richard Bucknall. 2019. A robust localization method for unmanned surface vehicle (USV) navigation using fuzzy adaptive Kalman filtering. *Ieee Access* 7 (2019), 46071–46083.
- [11] Hongjie Ma, Edward Smart, Adeel Ahmed, and David Brown. 2017. Radar image-based positioning for USV under GPS denial environment. *IEEE transactions on intelligent transportation systems* 19, 1 (2017), 72–80.
- [12] Jonatan Olofsson, Gustaf Hendeby, Fredrik Gustafsson, Deran Maas, and Stefano Marano. 2020. GNSS-free maritime navigation using radar and digital elevation models. In *2020 IEEE 23rd International Conference on Information Fusion (FUSION)*. IEEE, 1–8.
- [13] Filip Radenovic, Giorgos Tolias, and Ondrej Chum. 2018. Fine-tuning CNN image retrieval with no human annotation. *IEEE transactions on pattern analysis and machine intelligence* 41, 7 (2018), 1655–1668.
- [14] Qing Zha and Yanwei Huang. 2021. Research on positioning and navigation of USV based on Lidar. In *Proceedings of 2021 Chinese Intelligent Systems Conference: Volume I*. Springer, 700–709.