

RESEARCH ARTICLE

Open Access



Polarization-based underwater geolocalization with deep learning

Xiaoyang Bai¹, Zuodong Liang¹, Zhongmin Zhu¹, Alexander Schwing^{1,2}, David Forsyth² and Viktor Gruev^{1,3,4,5*} 

Abstract

Water is an essential component of the Earth's climate, but monitoring its properties using autonomous underwater sampling robots remains a significant challenge due to lack of underwater geolocalization capabilities. Current methods for underwater geolocalization rely on tethered systems with limited coverage or daytime imagery data in clear waters, leaving much of the underwater environment unexplored. Geolocalization in turbid waters or at night has been considered unfeasible due to absence of identifiable landmarks. In this paper, we present a novel method for underwater geolocalization using deep neural networks trained on ~10 million polarization-sensitive images acquired globally, along with camera position sensor data. Our approach achieves longitudinal accuracy of ~55 km (~1000 km) during daytime (nighttime) at depths up to ~8 m, regardless of water turbidity. In clear waters, the transfer learning longitudinal accuracy is ~255 km at 50 m depth. By leveraging optical data in conjunction with camera position information, our novel method facilitates underwater geolocalization and offers a valuable tool for untethered underwater navigation.

Keywords Underwater geolocalization, Ocean exploration, Celestial-based navigation, Polarization, Deep learning

1 Main

Earth's water surface is a complex and dynamic environment, encompassing vast oceans, seas, lakes, and rivers. The oceans alone account for over 70% of the Earth's surface area and contain an estimated 97% of the planet's water supply [1, 2]. Despite its importance, in situ monitoring of water properties remains challenging, and less precise satellite imaging is often used to capture water surface temperature, salinity, oxygen/nitrogen

levels, and other parameters [3, 4]. Autonomous underwater sampling robots can provide more accurate *in situ* monitoring, but reliable geolocalization is required for their successful operation [5–9]. As satellite-based global positioning system (GPS) does not work in the underwater environment, alternative methods for underwater localization have been explored with limited success [9–11]. Despite advancements in acoustic navigation, landmark identification, and inertial navigation, underwater geolocalization still has limited area coverage or poor global accuracy [11]. Small underwater vehicles and scuba divers face constraints on size and power for navigation devices, making precise inertial navigation and long-base-line acoustic navigation impractical. Visual-based underwater geolocalization using color and polarization images has demonstrated limited accuracy and is only effective in clear waters and during the day [12, 13]. Therefore, submersible vehicles and scuba divers frequently lack reliable geolocalization, which is crucial for exploratory underwater missions.

*Correspondence:

Viktor Gruev
vgruev@illinois.edu

¹ Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

² Department of Computer Science, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

³ Department of Bioengineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

⁴ Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

⁵ Carle Illinois College of Medicine, University of Illinois at Urbana-Champaign, Urbana, IL 61820, USA

Migratory animals provide examples of precise navigation and geolocalization in both air and water, spanning across the globe [14, 15]. These animals rely on various sensory cues, including polarization-sensitive information from the sky or water [16–18]. Light polarization patterns with structure are ubiquitous in both above- and underwater environments. The scattering of sunlight or moonlight in the upper atmosphere generates distinctive polarization patterns in the sky [19, 20]. Although humans cannot directly perceive light polarization, we may have utilized sky polarization patterns for navigation with appropriate viewing equipment [21].

When viewed from underwater, sky polarization patterns are visible within the Snell window in shallow clear waters and can be leveraged for both geolocalization and navigation [22]. Various marine animals, including the mantis shrimp, rely on these patterns for their navigational needs [17]. However, it is important to note that in environments with low visibility or at greater depths in clear waters, the sky polarization patterns become unobservable, as demonstrated in supplementary videos 1 through 8, rendering them ineffective for navigation purposes. Additionally, it is worth noting that underwater polarization patterns can also be observed outside the Snell window. The formation of these patterns arises from two primary physical phenomena. Firstly, the predominantly unpolarized light emitted by the sun or reflected by the moon undergoes partial linear polarization upon entering the water. Subsequently, this partially polarized light gets scattered by suspended particles, contributing to the observed polarization patterns.

Recordings of underwater polarization patterns date back to the 1950s [23]. As polarization imaging technology has advanced [24–27], better understanding of this hidden world has been gained through *in situ* measurements around the world [9, 28–31]. It was previously thought that underwater light was mainly horizontally polarized [32–35], making it unsuitable for geolocalization. However, Waterman noted that this belief was incorrect [36], likely due to measurement inaccuracies, and suggested that underwater polarization fields could at least provide orientation information and potentially enable navigation. A recent study demonstrated geolocalization accuracy of 1970 km using underwater polarization images in clear water [13]. However, the usefulness of underwater polarization patterns observed in turbid water or at night has not been established. Polarization in turbid water has been dismissed as horizontal [13, 37], and there are no recorded observations of underwater polarization patterns at night.

In open ocean waters or oligotrophic fresh waters with a low scattering coefficient (0.001 m^{-1}), underwater polarization patterns can be accurately represented by

a single scattering model, as depicted in Fig. 1a. Therefore, straightforward inference procedures can be applied to achieve geolocalization in shallow clear water [13]. However, in coastal ocean waters and eutrophic lakes where the scattering coefficient can be as high as 1 m^{-1} , the single scattering model is inadequate for predicting underwater polarization information, as evidenced by the underwater polarization patterns captured with an omnidirectional lens shown in Fig. 1b. The accuracy of predicted underwater polarization patterns can be improved by utilizing multi-scattering models that rely on three-dimensional Monte Carlo techniques. However, integrating these models with geolocalization would necessitate the generation of underwater patterns at numerous locations worldwide, rendering the computational feasibility of such an approach unattainable. Similarly, at night, underwater polarization patterns are influenced by both the moon and night sky contributions, making them challenging to model using the single scattering model, even in clear water at night, as illustrated in Figs. 1c and 2, or at greater depths. This underscores the importance of developing new methods for geolocalization that can handle high-scattering waters and low-light conditions.

Here we show that even though direct inference through predictive models is unmanageable in many underwater situations, polarization patterns produced by daylight in low visibility water and by nightlight in both high and low visibility waters allow accurate geolocalization. First, we collected ~ 10 million images with underwater cameras capable of recording the radial polarization light field from four sites around the globe. We then trained a deep neural network to predict geolocation from underwater angle of polarization (AoP) images collected with an omnidirectional lens, in combination with camera position sensor data (Fig. 1e). We provide systematic comparison for underwater geolocalization accuracy between parametric and data driven model across time, date and different water visibility. We demonstrate that using polarization information instead of intensity-only images results in superior geolocalization accuracy. Additionally, we show, for the first time reported in the literature, geolocalization at night, in low visibility waters, and at a depth of 50 m in clear waters using transfer learning techniques. We provide concluding remarks at the end of the paper.

2 Results

2.1 Underwater data collection and geolocalization methodology

We collected data from four sites with varying visibility and salinity to evaluate our underwater geolocalization method. These included a freshwater lake in Cham-paign, IL, USA with a visibility of around 0.3 m; coastal

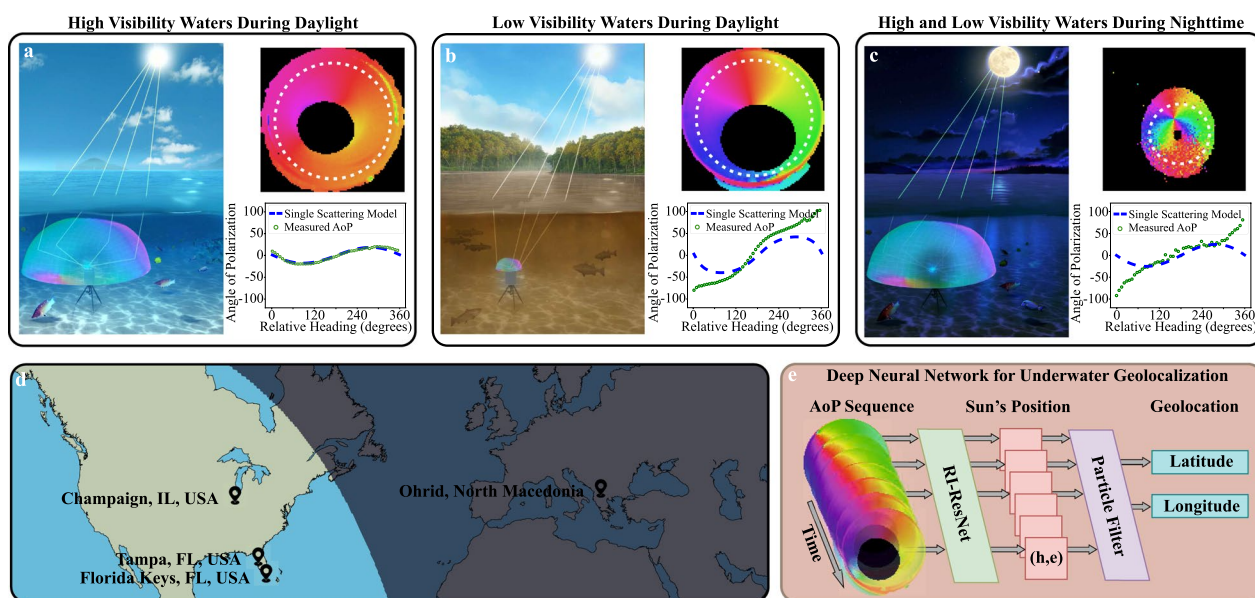


Fig. 1 Deep neural network method for underwater geolocation based on celestial-based underwater polarization information in low and high visibility waters by day and by night. **a–c** We deployed an underwater polarization sensitive imaging system with an omnidirectional lens in high and low visibility waters to collect the required data. False-color images of the measured angle of polarization (AoP) and a graph comparing observed AoP with the parametric model's prediction are displayed next to each drawing. Predictions made by the parametric model are unreliable in low visibility waters and it is ineffective at night. **d** We selected four different sites as indicated on the global map to collect underwater data and to assess the effectiveness of our geolocation method. **e** Our deep neural network, in conjunction with a particle filter, uses sequences of AoP images to estimate the camera's position latitude and longitude

sea waters in Florida Key, FL, USA with variable visibility ranging from 0.5 m to 3 m; sea water in the bay of Tampa, FL, USA with a visibility of around 0.5 m; and a freshwater lake in Ohrid, North Macedonia with visibility exceeding 10 m (Fig. 1d). The imaging instrument was placed on the sea or lake floor at depths of 1 m in Champaign, IL, 2 m in Florida during both winter and summer, and 8 m and 50 m in Ohrid, North Macedonia. Data was collected during the winter in the Florida Keys with a maximum sun elevation of around 40 degrees and during the summer in the bay of Tampa, FL with a maximum sun elevation of approximately 86 degrees (refer to Additional file 1: Video S1, Additional file 2: Video S2, Additional file 3: Video S3, Additional file 4: Video S4, Additional file 5: Video S5, Additional file 6: Video S6, Additional file 7: Video S7, Additional file 8: Video S8 and Fig. 2). Due to their bay-like configuration, all four sites exhibited minimal surface wave activity. This favorable condition allowed us to gather data with minimal surface disturbances.

The data from each site was divided randomly into a training set containing 80% of the data and a testing set containing the remaining 20%. The images in the training and testing data sets were collected on different dates and were spatially and temporally downsampled to 100 by 100 pixels and 1 frame per second, respectively. Frames

in the training data where clouds completely obstructed the sun were manually removed. The purging was performed by personnel with no access to any trained model or test results to avoid introducing bias.

Underwater geolocation was achieved via either parametric or data driven model. In the parametric model, theoretical modeling of single scattering is used to simulate underwater polarization patterns, as illustrated in Fig. 3a. This is achieved by utilizing a Mueller matrix formalism to describe light scattering from particles in water and air-water refraction. To determine the camera's geolocation based on a set of underwater polarization images, the method estimates the sun's heading and elevation angles by minimizing the difference between measured and simulated underwater polarization angles. In our proposed network model, geolocation is predicted using a sequence of angle of polarization images in three stages (Fig. 3c, d). Firstly, a deep network leverages inertial magnetic unit parameters to predict a set of coarse sun locations (azimuth and elevation) for each frame individually. Secondly, temporal information is incorporated by another network to refine these coarse predictions, resulting in fine sun locations (Fig. 3e). Finally, both parametric and data-driven models use a particle filter to estimate geolocation (longitude and latitude) by utilizing a large batch of sun location predictions (see Methods).

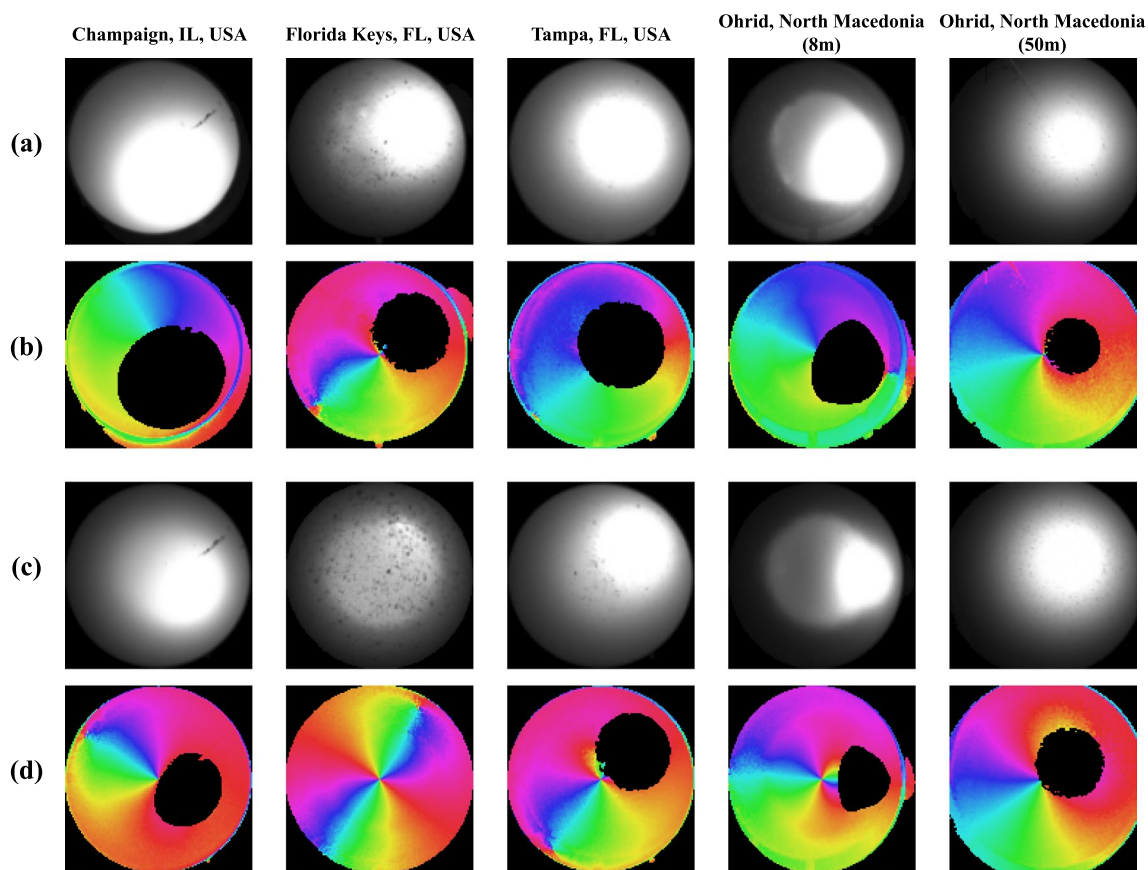


Fig. 2 **a, b** Intensity and AoP images captured at 14:00 local time for different sites and water depths. **c, d** Intensity and AoP images captured at 17:00 local time for different sites and water depths. Note that in the intensity images, the sky is clearly visible in high-visibility waters at a shallow depth (4th column). However, it is not visible in low-visibility waters (1st column) or at greater depths (5th column)

2.2 Daytime underwater geolocation based on polarization images in low and high visibility waters

We compared the accuracy of geolocation using our developed deep neural network approach with that of a parametric method. In clear waters, such as Lake Ohrid, North Macedonia, the difference between the measured underwater angle of polarization and the parametric model is less than 10% when the sun’s elevation is above approximately 30 degrees (Fig. 4a). However, during winter periods and summer sunrise and sunset, the sun’s elevation is below 30 degrees, and the underwater polarization patterns are affected by light from both the sky and the sun (Fig. 4b). These light interactions are not well understood and are not included in the parametric model, resulting in estimated underwater angle of polarization with segments that have errors exceeding 50%. In low-visibility waters, such as those in Florida during the summer and winter periods (Fig. 4c), the estimated angle of polarization

has errors exceeding 50% throughout the entire day due to the lack of multiple scattering effects in the parametric model.

Significant inaccuracies in the estimated root mean squared errors (RMSEs) of the sun’s heading and elevation angles arise from large modeling errors. For instance, the RMSEs for the sun’s heading and elevation angles are 11.412° and 14.579° in Champaign, IL; 22.093° and 16.874° during the winter in Florida; 60.283° and 35.862° during the summer in Florida; and 14.362° and 11.113° in Lake Ohrid, North Macedonia, respectively (Fig. 5 top row). These errors are substantially reduced with our deep neural network approach, which produces at least one order of magnitude lower RMSEs compared to the parametric model at all sites (Fig. 5 bottom row). The RMSEs for the sun’s heading and elevation angles using our deep neural network approach are 1.135° and 1.623° in Champaign, IL; 2.232° and 1.867° during the winter in Florida; 5.878° and 1.090° during the summer in Florida; and 1.290° and 0.845° at a depth of 8 m in Lake Ohrid,

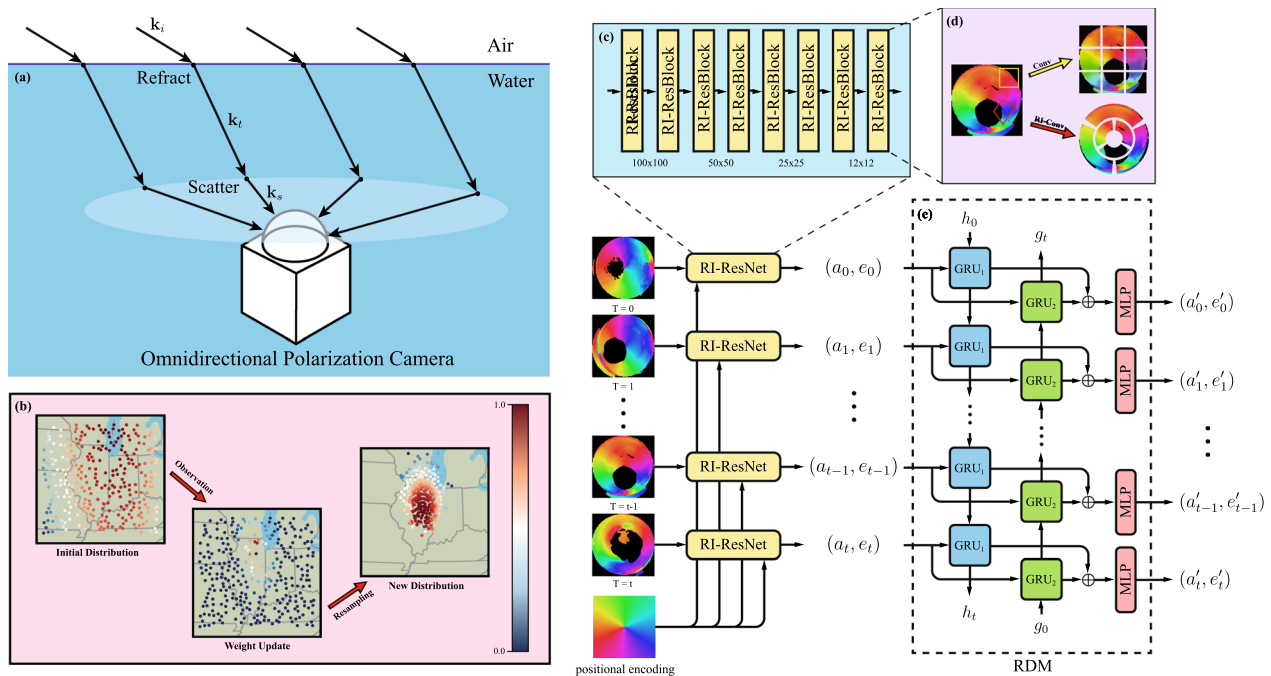


Fig. 3 **a** Underwater polarization patterns mainly result from the refraction of light between air-water interfaces and scattering within the water medium. These patterns can be mathematically modeled using Mueller matrices. **b** The particle filter (PF) pipeline is illustrated with high probability particles shown in red and low probability particles in blue. **c, d** Our proposed network model includes the RI-ResNet architecture, which replaces each convolution layer with its RI-Conv counterpart and accounts for the radial spatial structure in omnidirectional images. **e** The RDM architecture involves a bidirectional recurrent network that models temporal dependencies between images

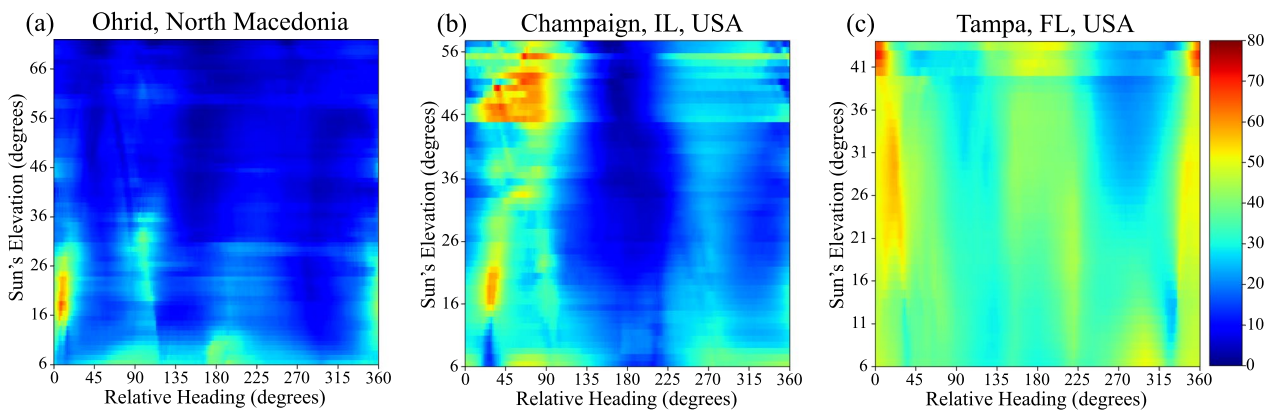


Fig. 4 Relative error between measured and predicted solar elevation and heading angles using parametric model in different sites around the world. **a** Angular prediction errors in Lake Ohrid, North Macedonia are relatively low due to high water visibility. **b, c** Angular error predictions have both high and low errors during different solar elevation due to multiple scattering deficiencies in the parametric model

North Macedonia, respectively. The deep neural network approach consistently exhibits lower RMSEs for the sun's heading and elevation angles compared to the model-based approach throughout the day.

In order to assess the effectiveness of RDM, we compared the RMSEs for solar angle predictions obtained using only the RI-ResNet network module against those

obtained when both network modules (RI-ResNet and RDM) are active. When using RI-ResNet alone, the RMSEs for the sun's elevation and heading were found to be as follows: 3.002° and 1.604° in Champaign, IL; 13.64° and 5.741° during the winter in Florida; 16.43° and 8.591° during the summer in Florida; and 1.474° and 0.713° in Lake Ohrid, North Macedonia, respectively. By

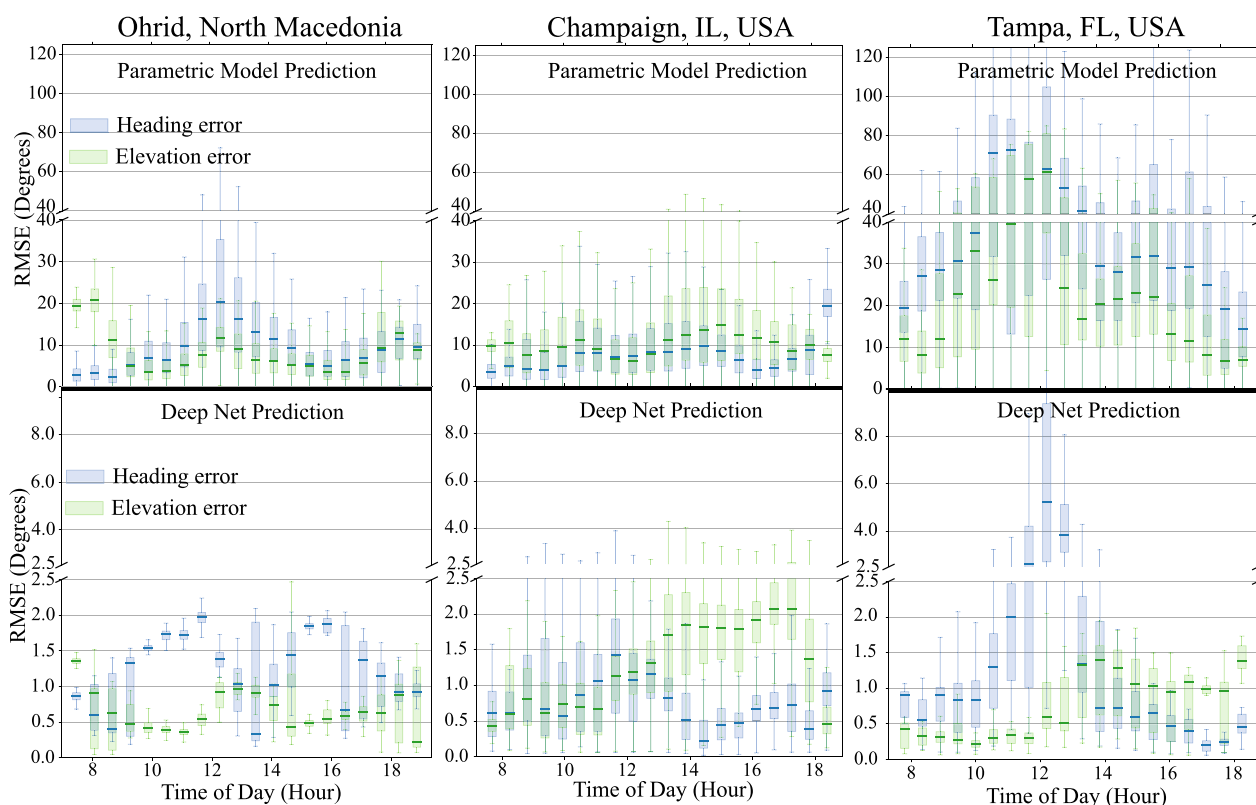


Fig. 5 The top and bottom figures depict the root mean squared error of the estimated solar heading and elevation angles for both the parametric and deep neural network models. The parametric model solely considers single scattering phenomena, leading to greater solar angular errors in low visibility waters (Champaign, IL and Tampa, FL) in comparison to high visibility waters (Lake Ohrid, North Macedonia). Conversely, the deep neural network model learns intrinsic polarization patterns that arise from both single and multiple scattering, which results in similarly low solar angular errors in both low and high visibility waters

comparing these values with the results obtained when RDM is included, it can be concluded that the incorporation of RDM significantly improves the output accuracy of RI-ResNet.

Figure 6 displays the mean and one standard deviation for the particle filter covariance at the end of the day for both the parametric model (dashed line) and the deep neural network model (solid line) for the four locations around the globe. The parametric model-based geolocation predictions at the end of the day have a median error of 738 km and 1416 km in the East–West and North–South directions in Champaign, IL; 1519 km and 567 km in Florida during the winter; 3947 km and 2275 km in Florida during the summer; and 1034 km and 629 km in Lake Ohrid, North Macedonia, respectively. By contrast, the deep neural network model yields more accurate initial angular estimates, leading to geolocation errors of 55 km and 156 km in the East–West and North–South directions in Champaign, IL; 128 km and 78 km in Florida during the winter; 56 km and 64 km in Florida during the summer; and 50 km and 160 km at 8 m depth in Lake Ohrid, North Macedonia, respectively. The deep

neural network-based geolocation results are at least one order of magnitude more accurate than those of the parametric models.

We conducted a geolocation accuracy evaluation throughout the day for both parametric and deep neural network methods. Figure 7 displays the results for two sites with high mid-day solar elevations: one with high visibility waters (Ohrid, North Macedonia) and one with low visibility waters (Tampa, Florida during the summer). For the parametric model estimation, geolocation accuracy is highest at mid-day and decreases towards the end of the day due to the lack of skylight contributions in the model. In low visibility waters, geolocation error is uniformly high throughout the day, but the standard deviation decreases towards the end of the day due to particle filter noise reduction. The deep neural network model exhibits relatively constant but low errors throughout the entire day in both clear and low visibility waters, with slightly higher errors around mid-day in low visibility waters. This is likely due to the network having only observed a small number of images with high solar elevations (i.e., above 75°). An intriguing observation was

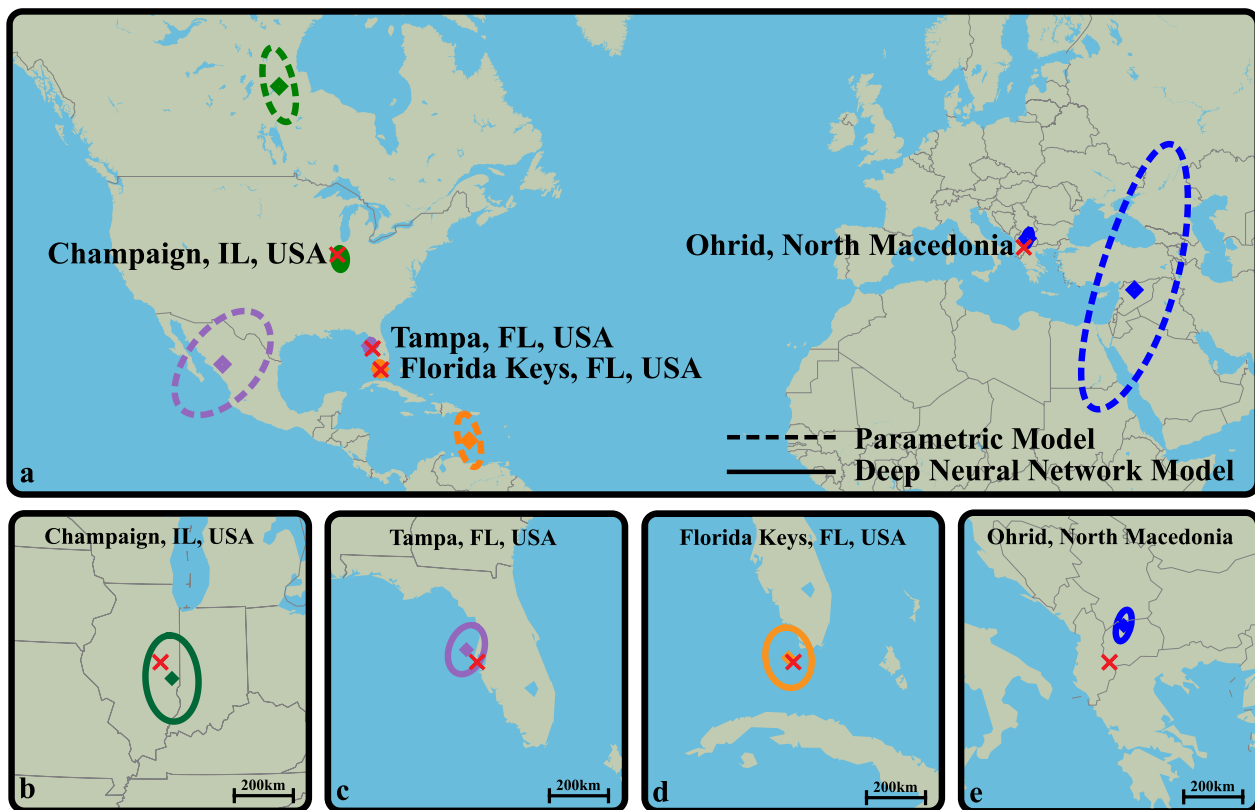


Fig. 6 **a** The accuracy of underwater geolocalization predictions across the globe is significantly improved using a deep neural network (shown as a solid line) compared to a parametric model (shown as a dashed line). The global map illustrates the mean (shown as a diamond) and first standard deviation (shown as either a solid or dashed line) of the particle filter estimate for geolocation at the end of a day. The large errors observed in the mean and standard deviation of the estimated geolocation using the parametric approach are primarily due to a lack of understanding of the various physical phenomena that contribute to underwater polarization. **b–e** The close-up maps display the errors in the network model at a scale that allows the resolution of the covariance

made that the variation in tides had no impact on the prediction of the sun's angle and, consequently, on the accuracy of geolocalization. Despite the camera's depth fluctuating by more than 50% at the Florida sites due to different tidal cycles, the underwater polarization field remained relatively consistent.

2.3 Transfer learning for underwater geolocalization at depth

To assess the accuracy of underwater geolocalization at greater depths, we collected polarization data at 50 m depth in Lake Ohrid, North Macedonia and evaluated the transfer learning capability of our approach. Due to logistical constraints, we were only able to collect continuous data for several hours over 2 days. We used data from 8 m depth to train our neural network, and data from 50 m depth to test geolocalization accuracy. The RMSEs for the sun's heading and elevation angle predictions were 6.605° and 4.236° , respectively. However, the geolocalization error in the East–West

and North–South directions increased to 473 km and 255 km, respectively, as shown in Fig. 8 and Additional file 5: Video S5.

The lower accuracy of geolocalization at greater depths is attributed to two factors. First, the interactions between light and water change as depth increases. Light undergoes multiple scattering and absorption events as it travels through deeper water. Although angle of polarization images at 8 m and 50 m depth appear visually similar, images at 50 m have lower degree of linear polarization and intensity than those at 8 m. The maximum degree of linear polarization recorded at 50 m is approximately 15%, compared to 35% at 8 m. Second, since the neural network is trained on angle of polarization images with higher degrees of linear polarization and intensity, it does not perform as well on images with lower degrees of polarization and intensity. As a result, the differences in noise profiles between the training and test data sets limit the accuracy of geolocalization predictions.

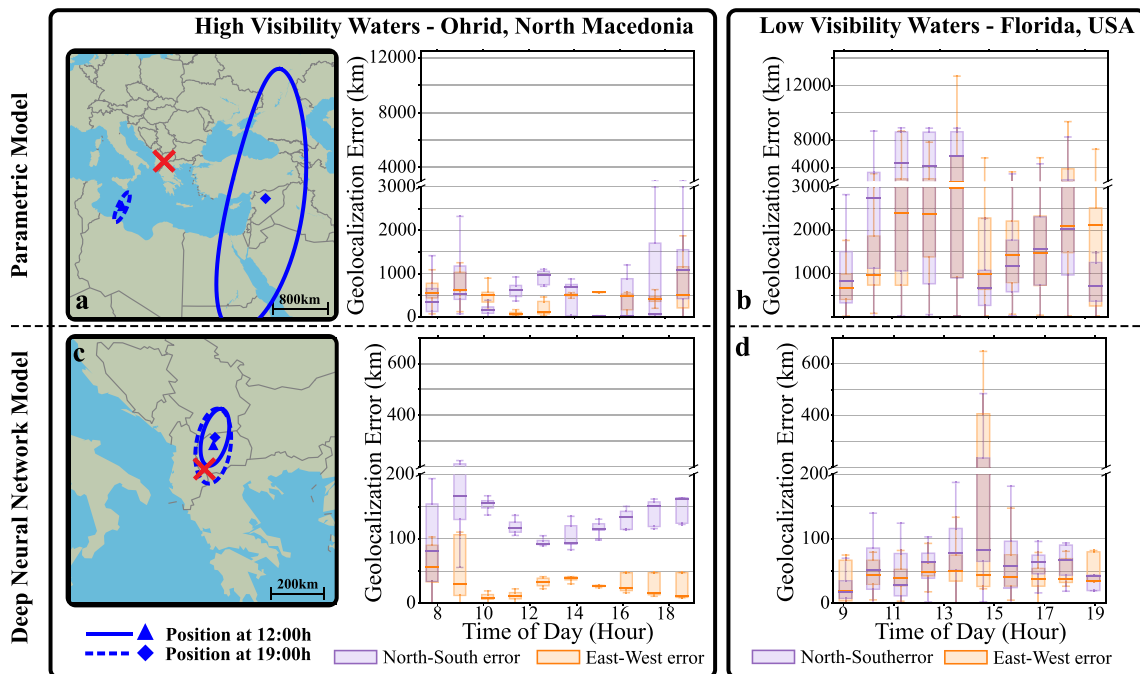


Fig. 7 Geolocalization throughout the day in low and high visibility waters. The top row (**a** and **b**) and bottom row (**c** and **d**) show geolocalization accuracy throughout the day using the parametric model and the deep neural network model, respectively, in both high (left) and low (right) visibility waters. The parametric-based underwater geolocalization has moderate to low accuracy in low visibility waters due to model deficiencies in incorporating all physical phenomena that contribute to underwater polarization. In contrast, the deep neural network geolocalization performs uniformly well throughout the day in both high and low visibility waters. The individual maps display the mean (triangle and diamond) and first standard deviation (solid and dashed line) of the covariance of the particle filter estimate of geolocation at noon and at the end of the day, respectively. The box plots represent the median and upper/lower quartiles for the North–South (purple) and East–West (orange) geolocalization prediction errors

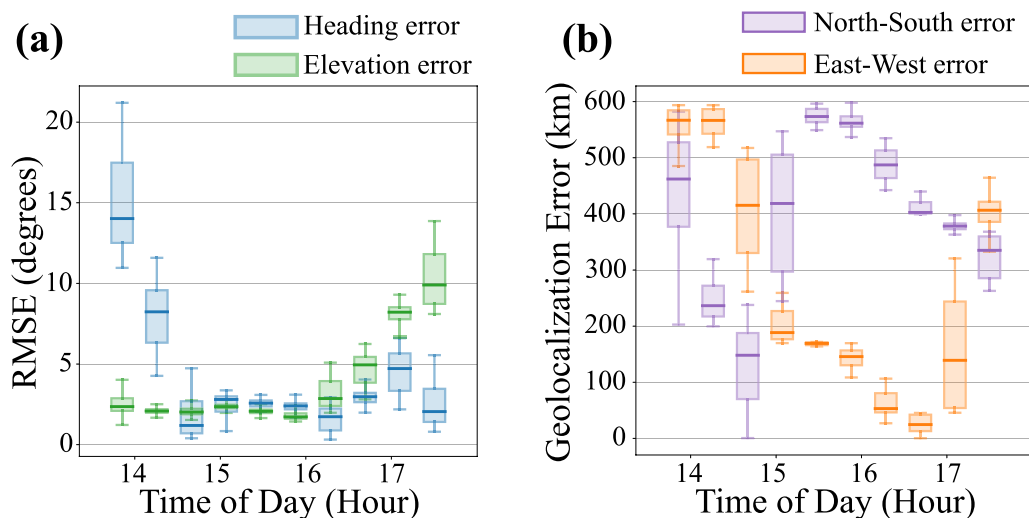


Fig. 8 Underwater geolocalization data at 50 m depth in Lake Ohrid, North Macedonia. **a** Solar angular error and **b** geolocalization error across several hours at 50 m depth

2.4 Nighttime underwater geolocalization

To assess the geolocalization accuracy at different moon phases, we collected underwater polarization data at night across all four sites (Fig. 9 and Additional file 6: Video S6, Additional file 7: Video S7, Additional file 8: Video S8). Because the underwater light intensity is much weaker at night than during the day, the camera exposure was set to 1 s for a moon cycle between full and gibbous, and to 10 s when the moon cycle was between crescent and quarter. However, due to the short recordings of less than two hours for crescent moon, the number of nighttime images used to train the deep neural network was limited. Despite this constraint, the RMSEs for the moon’s heading and elevation were 19.404° and 7.160° in Champaign, IL; 43.019° and 10.056° in Florida during the winter; 37.392° and 15.297° in Florida during the summer; 12.947° and 3.552° in Lake Ohrid, North Macedonia, respectively. The final output from the particle filter provided nighttime geolocalization with East–West and North–South errors of 32 km and 357 km in Champaign, IL; 786 km and 1307 km in Florida during the winter; 2131 km and 1650 km in Florida during the summer; 1020 km and 285 km in Lake Ohrid, North Macedonia, respectively. Notably, the geolocalization accuracy was independent of the moon cycle (Fig. 9b).

2.5 Daytime underwater geolocalization based on intensity images

Two physical phenomena, refraction and in-water scattering, generate a radial intensity profile that depends on the sun’s position, and it is possible to predict the solar angular position using intensity images from an omnidirectional gray scale camera. To test this hypothesis, we used the same data set and deep neural network architecture as described earlier, but trained it with intensity-only images. We added the intensity data from four super pixels to generate an intensity image. Interestingly, for data collected in Champaign, IL and Lake Ohrid, North

Macedonia, the solar angle predictions based on intensity images were similar to those based on polarization images. However, in Tampa, FL and the Florida Keys, the total RMSEs for solar angle predictions based on intensity images were 6.760° and 16.071°, respectively, compared to polarization-based RMSEs of 2.758° and 2.134°, respectively.

The water visibility in Lake Ohrid and in Champaign is different, but it remained relatively constant during the data collection period. The local lake in Champaign is small and not affected by wind conditions or rain, while the visibility and temperature of Lake Ohrid remain constant during summer periods. Despite the high turbidity and multiple light scattering events in the local lake in Champaign and the few scattering events due to the water clarity in Lake Ohrid, the network can recapitulate the underwater intensity image’s dependence on solar angles in both cases due to the constant water environment. However, in both Florida sites, water visibility varied throughout the day and between different days due to currents, tides, and other environmental factors. These small changes in visibility introduce enough noise into the training set that prevented accurate solar prediction based on intensity images. However, angle of polarization remains robust to scattering perturbations in the water environment and helps improve solar angle predictions by the neural network [38].

3 Conclusion

This paper presents a learning-based method for highly accurate underwater geolocalization using an omnidirectional polarization camera. The approach is effective across multiple sites worldwide and improves geolocalization accuracy compared to the traditional parametric model by an order of magnitude [13]. In addition, for the first time reported in the literature, polarization-based geolocalization in low visibility waters, high visibility waters at 50 m depth and at night is demonstrated.



Fig. 9 Geolocalization accuracy during nighttime under different moon phases. **a** The global map displays the mean (represented by a diamond) and first standard deviation (represented by a solid line) of the particle filter estimate of geolocation at the four sites (crosses) during the new moon and full moon phases. **b** The box plots indicate the median and upper/lower quartiles for the North–South (purple) and East–West (orange) geolocalization prediction errors

Underwater imaging in turbid waters or at night poses a significant challenge due to the limited amount of ambient light available to capture images. In order to maintain a high signal-to-noise ratio in the images, the exposure time is adjusted individually, ensuring that the average intensity of the image falls within the middle range of the imager's dynamic range. During midday when water visibility is high, the typical image exposure time is approximately 1 millisecond. However, for low visibility water, the exposure time falls within the range of 50 to 500 milliseconds. Conversely, during sunrise and nighttime conditions, the exposure times can vary from 0.5 to 10 s.

The light scattering caused by suspended particles in the water diminishes the degree of linear polarization even further. In highly turbid waters like Tampa, Florida, and Champaign, IL, the maximum degree of linear polarization was approximately 15%. On the other hand, in the clear waters of Lake Ohrid, the maximum degree of linear polarization exceeded 35%. Due to the relatively high degree of linear polarization and the high signal-to-noise ratio of the intensity images, there is no need for any denoising techniques before estimating the geolocalization information from the angle of polarization images.

The physical properties of water can significantly impact the underwater polarization patterns, leading to distinctive changes in polarization at different locations. For instance, variations in particle density, oxygenation, pollution, and depth can all contribute to alterations in polarization patterns. Moreover, complex relief in the water can cause blocked scattered light, further influencing polarization. In this context, creating maps of local water properties may enable more precise geolocation, and localization at a fine scale may be feasible even in the presence of relief [39].

Our research presents a novel method for high-accuracy geolocalization using polarization in clear and turbid waters, day or night, and at greater depths. This method offers a potential new way for aquatic creatures to navigate, even in low-visibility conditions. By using underwater background polarization information, they may be able to find their way around and reach their destination with greater accuracy. This could have significant implications for marine life, as well as for human activities such as underwater exploration and search and rescue missions.

4 Methods

4.1 Underwater imaging instrument

The use of two distinct underwater housings, each with dome ports, allowed for the collection of the underwater data. The first housing was created by retrofitting a Blue Fin housing, while the second housing was custom-designed using Autocad and manufactured by PCBWay

Incorporated. Both housings housed a polarization imaging sensor (FLIR Blackfly Polarization Monochrome Camera) which was equipped with a fisheye lens (Fujinon FE185C057HA-1) and an inertial magnetic unit (TCM-XB, PNI Sensor Corporation). Communication between the inertial magnetic unit (IMU) and polarization camera was achieved via an I2C protocol. A 100 m underwater Ethernet cable was used to connect the polarization camera to a computer located on the shore. This cable provided power to the camera and IMU while simultaneously transmitting data to the computer. Data acquisition software, which was developed in Python, was used to record all video data in h5 format with IMU information. The camera could transmit up to 20 frames per second, and the information was stored in a 64 TB network area storage where the data was compressed every night for efficient storage.

4.2 Underwater data collection and preprocessing

The underwater camera system was mounted on an extruded aluminum platform, which was able to rotate freely for calibration data collection of the IMU. The calibration data was processed in Python using the *imucal* package. The imaging platform was then placed on the sea or lake floor at various depths, such as 1 m in Champaign, IL, 2 m in Florida during both winter and summer, and 8 m and 50 m in Ohrid, North Macedonia. In some sites, the entire platform was randomly rotated every day to collect a more diverse training set. During the day, the exposure time was set between 0.2 and 2 ms, and the frame rate was set to 20. At night, the camera exposure was set to 1 s for a Moon cycle between full and gibbous (i.e. 1 frame per second) and 10 sec when the moon cycle was between crescent and quarter (i.e. 0.1 frame per second).

Prior to conducting experiments, we preprocessed the raw angle of polarization (AoP) image data. Firstly, we averaged over a 15-frame temporal window to obtain images at 0.66 Hz (daytime) or 0.1 Hz (nighttime) to reduce stochastic noise and data redundancy. Next, we cropped out background rows and columns from each frame, rescaled it to 100×100 , and performed a calibration algorithm based on a previously published method [40]. Finally, a human agent who had no access to the geolocalization models or their results identified noisy frames where the sun was either occluded by thick clouds or below the horizon.

4.3 Polarization-based underwater geolocalization with parametric model

The underwater geolocalization parametric method employs a theoretical model of single scattering to simulate underwater polarization patterns (Fig. 3a). To determine

the camera’s geolocation based on a set of underwater polarization images, the method estimates the sun’s heading and elevation angles by minimizing the difference between measured and simulated underwater polarization angles. Subsequently, a particle filter is employed to determine the geolocation \mathbf{g} (i.e., longitude and latitude) using a sequence of sun angle predictions.

4.3.1 Parametric model for underwater polarization patterns

To model underwater polarization patterns, a Mueller matrix formalism is utilized to describe light scattering from particles in the water (\mathbf{M}_S) and air-water refraction (\mathbf{M}_R). Rotational matrices ($\mathbf{M}_{S \rightarrow D}$ and $\mathbf{M}_{R \rightarrow S}$) are also included to account for any offsets between the different coordinate systems. The process begins with unpolarized sunlight, which is represented by a Stokes vector \mathbf{S}_i , and undergoes transmission from air to water before scattering from the particles suspended in the water. The following equation describes this process and yields the Stokes vector \mathbf{S}_d , which corresponds to the underwater light as detected by the polarization-sensitive camera:

$$\mathbf{S}_d = \mathbf{M}_{S \rightarrow D} \mathbf{M}_S \mathbf{M}_{R \rightarrow S} \mathbf{M}_R \mathbf{S}_i. \tag{1}$$

The Mueller matrix for air-water refraction (\mathbf{M}_R) can be calculated as follows:

$$\mathbf{M}_R = \begin{bmatrix} \alpha + \beta & \alpha - \beta & 0 & 0 \\ \alpha - \beta & \alpha + \beta & 0 & 0 \\ 0 & 0 & \gamma & 0 \\ 0 & 0 & 0 & \gamma \end{bmatrix} \tag{2}$$

where α , β and γ are represented by the following three equations:

$$\alpha = \frac{1}{2} \left[\frac{2 \sin \theta_t \cos \theta_i}{\sin(\theta_i + \theta_t) \cos(\theta_i - \theta_t)} \right]^2, \tag{3}$$

$$\beta = \frac{1}{2} \left[\frac{2 \sin \theta_t \cos \theta_i}{\sin(\theta_i + \theta_t)} \right]^2, \tag{4}$$

$$\gamma = \frac{4 \sin^2 \theta_t \cos^2 \theta_i}{\sin^2(\theta_i + \theta_t) \cos^2(\theta_i - \theta_t)}. \tag{5}$$

In the given equations, θ_i and θ_t represent the incident and transmitted angles, respectively, and are determined by Snell’s law using the refractive index of water relative to air, denoted by n :

$$\sin \theta_i = n \sin \theta_t. \tag{6}$$

The final step is summarized by the following equation, which utilizes the Mueller matrix for Rayleigh scattering in the water medium:

$$\mathbf{M}_S = \frac{1}{2} \begin{bmatrix} 1 + \cos^2 \theta & \cos^2 \theta - 1 & 0 & 0 \\ \cos^2 \theta - 1 & 1 + \cos^2 \theta & 0 & 0 \\ 0 & 0 & 2 * \cos \theta & 0 \\ 0 & 0 & 0 & 2 * \cos \theta \end{bmatrix}. \tag{7}$$

It is worth noting that the rotation matrix $\mathbf{M}_{R \rightarrow S}$ and $\mathbf{M}_{S \rightarrow D}$ are applied to rotate the coordinate system from the incident light beam to the transmitted beam and from the transmitted beam plane to the scattering plane, respectively. The rotational matrix can be expressed as follows:

$$\mathbf{M}_{R \rightarrow S, S \rightarrow D} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(2\varphi) & \sin(2\varphi) & 0 \\ 0 & -\sin(2\varphi) & \cos(2\varphi) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{8}$$

where φ is the angle of rotation. Our underwater imaging system captures the radial underwater polarization field in a single snapshot. The angle of polarization is computed using standard parametric equations from the raw data recorded by our polarization camera. To estimate the heading and elevation angles of the sun (\mathbf{h}), a linear regression is performed by comparing the measured angle of polarization to the angle predicted by the single scattering model for different solar angles.

4.3.2 Geolocalization via particle filter

We convert sun position (\mathbf{h}) to geolocation estimates (\mathbf{g}) (longitude and latitude) in the last stage. We further improve the accuracy of geolocation estimates by collecting a sequence of sun’s heading and elevation observations \mathbf{h}_t over a period of time $t \in 1, \dots, T$ since the recording camera is stationary. To achieve this, we utilize a particle filter to describe the posterior probability $P(\mathbf{g} \mid \mathbf{h}_1, \dots, \mathbf{h}_T)$ [41].

We start by initializing a set of N particles within a rectangular area of size 1000 km by 1000 km with uniform weight $1/N$. Each particle represents a possible location of the camera, and its weight represents the probability of the particle being the true location. When a new measurement \mathbf{h}_t is received, the weight of the j -th particle is updated as follows:

$$P'(\mathbf{g}_j) = P(\mathbf{g}_j) \cdot P(\mathbf{h}_t \mid \mathbf{g}_j), \tag{9}$$

and after all particles have been updated, we normalize their weights to sum up to 1:

$$P(\mathbf{g}_j) = \frac{P'(\mathbf{g}_j)}{\sum_{k=1}^N P'(\mathbf{g}_k)}. \tag{10}$$

To obtain the conditional probability $P(\mathbf{h}_t | \mathbf{g}_j)$, we first observe from geolocation \mathbf{g}_j at the same time as observation \mathbf{h}_t to determine the ground truth sun location, represented by \mathbf{h}'_t . Then, we use a radial basis function (RBF) kernel to compute the probability value, ranging from 0 to 1, as the similarity between \mathbf{h}'_t and \mathbf{h}_t . Formally:

$$P(\mathbf{h}_t | \mathbf{g}_j) = \exp\left(-\frac{\|\mathbf{h}'_t - \mathbf{h}_t\|^2}{2\sigma^2}\right). \tag{11}$$

By comparing the computed sun locations of particles at each time-point with the prediction from the estimation model, we can assign high posterior probabilities to particles that closely match the model's prediction across all \mathbf{h}'_t 's on the list. Conversely, particles that deviate from the model's prediction will receive low posterior probabilities. Ultimately, we arrive at a distribution that indicates each particle's likelihood of being the camera's true geolocation.

Initially, we distribute N particles over a large $1000 \times 1000 \text{ km}^2$ rectangle, which results in a sparse distribution where even the most accurate particle can be tens of kilometers away from the ground truth location. To overcome this issue, we employ a resampling strategy, where we resample particles based on their weights every M observations (M is a hyper-parameter chosen empirically). The new particles are perturbed with Gaussian noise to prevent overlap. This resampling procedure concentrates the particles closer to the current posterior mean and increases the resolution of geolocation. Our particle filter's architecture is shown in Fig. 3b. We obtain the mean geolocation by computing the weighted mean of the particle locations, and the covariance is also calculated from the particles. We apply the same particle filter design to the solar predictions from our deep neural network model, which will be explained in the following section.

4.4 Polarization-based underwater geolocalization with neural network model

Our proposed network model utilizes a sequence of angle of polarization images, denoted as $(\mathbf{x}_0, \dots, \mathbf{x}_t)$, to predict geolocation in three steps. Firstly, a deep network predicts a set of coarse sun locations (azimuth and elevation), denoted as $\mathbf{y}_i = (a_i, e_i)$, for each frame \mathbf{x}_i separately by incorporating its IMU parameters. Secondly, another network refines these coarse predictions using temporal information, resulting in fine sun locations denoted as $(\mathbf{y}'_0, \dots, \mathbf{y}'_t)$. Finally, a particle filter

estimates the geolocation \mathbf{g} (longitude and latitude) using a large batch of fine sun location predictions.

4.4.1 Coarse sun location prediction: RI-ResNet

We based our neural network on the ResNet-18 architecture [42]. However, this architecture has two drawbacks when applied to omnidirectional polarization images. Firstly, its convolution kernels are aligned to the sides of the frame, and hence they do not account for the true spatial relationship between pixels in the omnidirectional image. Secondly, the architecture is not rotation-invariant by design, meaning it cannot provide consistent predictions for frames that have similar sun locations but different camera orientations.

We introduce a solution to these issues with a rotation-invariant ResNet (RI-ResNet), which replaces standard convolution layers with deformable convolutions [43]. These convolutions have kernels oriented towards the center of the frame (Fig. 3d). Additionally, we add a positional encoding map to each frame that is calculated from IMU data. This map helps to recover the true orientation of each pixel, irrespective of the camera's heading. The positional encoding map is defined as:

$$\phi_{ij} = \text{atan2}(i - H/2, j - W/2) + \theta, \tag{12}$$

$$\mathbf{p}_{ij} = \begin{bmatrix} \cos(\phi_{ij}) \\ \sin(\phi_{ij}) \end{bmatrix}. \tag{13}$$

The variable θ represents the yaw component of the IMU vector. As a result, $\mathbf{p}_{i,j}$ contains the absolute heading of pixel (i, j) . By incorporating this positional encoding map, the RI-ResNet architecture produces more reliable sun location predictions and enables rotation-based data augmentation. Specifically, the calculation of the approximate sun location is expressed as:

$$\mathbf{y} = f_\varphi(\mathbf{x} \oplus \mathbf{p}). \tag{14}$$

In the equation above, f_φ denotes the RI-ResNet, parametrized by φ . Note that the input angle of polarization image \mathbf{x} and the positional encoding map \mathbf{p} are joined together by pixel-wise concatenation (denoted by *oplus*). We train the RI-ResNet using a mean-squared-error (MSE) loss which compares predicted sun location to ground-truth. Formally:

$$\mathcal{L}_{MSE} = \|\mathbf{y} - \mathbf{y}_{GT}\|_2^2. \tag{15}$$

Here \mathbf{y}_{GT} is the ground truth sun location (azimuth and elevation). The overall architecture of RI-ResNet is visualized in Fig. 3c.

4.4.2 Fine sun location prediction: recurrent denoising module

Next, we use the orderly path of the sun across the sky to refine the per-frame sun location estimation. This is achieved by training a BiGRU network, referred to as the recurrent denoising module (RDM), which smooths out the raw per-frame estimates. The RDM considers the entire list of RI-ResNet outputs $(\mathbf{y}_0, \dots, \mathbf{y}_t)$ and reduces the overall zigzaggedness of the curve they form. Since the refinement of the sun location at a specific timestep depends on all previous and subsequent estimations, bidirectional recurrent neural networks like BiGRU are preferred. The architecture of the RDM is illustrated in Fig. 3e. Mathematically,

$$h_{k+1}, o_{k+1}^{(1)} = GRU_1(\mathbf{y}_k, h_k), \tag{16}$$

$$g_{k+1}, o_{k+1}^{(2)} = GRU_2(\mathbf{y}_{t-k}, g_k), \tag{17}$$

$$\mathbf{y}'_k = MLP(o_{k+1}^{(1)} \oplus o_{t-k}^{(2)}). \tag{18}$$

Rather than utilizing the RI-ResNet output directly to train the RDM, we generate the noisy input by adding Gaussian noise to the ground truth. This approach is more effective and resilient. Only during the evaluation phase do we combine the RI-ResNet and RDM. To optimize the RDM, we use the same MSE loss, as shown in Eq. 12.

$$\mathcal{L}_{MSE} = \|RDM(\mathbf{y}_{GT} + \epsilon) - \mathbf{y}_{GT}\|_2^2, \epsilon \sim \mathcal{N}(0, \sigma). \tag{19}$$

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s43593-023-00050-6>.

Additional file 1: Video S1. Time lapse video of both intensity and angle of polarization recorded during the day at 8 m depth in high visibility waters in Lake Ohrid, North Macedonia.

Additional file 2: Video S2. Time lapse video of both intensity and angle of polarization recorded during the day in Tampa, Florida, USA during the summer of 2021.

Additional file 3: Video S3. Time lapse video of both intensity and angle of polarization recorded during the day in Florida Keys, Florida, USA during the winter of 2020/2021.

Additional file 4: Video S4. Time lapse video of both intensity and angle of polarization recorded during the day in Champaign, IL, USA during the summer/fall of 2020.

Additional file 5: Video S5. Time lapse video of both intensity and angle of polarization recorded during the day at 50 m depth in high visibility waters in Lake Ohrid, North Macedonia.

Additional file 6: Video S6. Time lapse video of both intensity and angle of polarization recorded during the night at 8 m depth in high visibility waters in Lake Ohrid, North Macedonia.

Additional file 7: Video S7. Time lapse video of both intensity and angle of polarization recorded during the night in Florida Keys, Florida, USA during the winter of 2020/2021.

Additional file 8: Video S8. Time lapse video of both intensity and angle of polarization recorded during the night in Champaign, IL, USA during the summer/fall of 2020.

Author contributions

XB, AS, DF and VG designed the study. ZL, ZZ and VG designed the underwater instrument and collected underwater data. XB, AS, DF and VG designed the learning and parametric based approach. XB, AS, DF and VG analyzed the data and wrote the manuscript with input from all other authors.

Funding

This work was funded by grants from the Office of Naval Research (N00014-19-1-2400 and N00014-21-1-2177) and U.S. Air Force Office of Scientific Research (FA9550-18-1-0278).

Declarations

Competing interests

The authors declare no competing interests.

Received: 7 April 2023 Revised: 19 May 2023 Accepted: 1 June 2023
Published online: 10 July 2023

References

1. T.L. Frölicher, E.M. Fischer, N. Gruber, Marine heatwaves under global warming. *Nature* **560**(7718), 360–364 (2018)
2. L. Yu, S.A. Josey, F.M. Bingham, T. Lee, Intensification of the global water cycle and evidence from ocean salinity: a synthesis review. *Ann. N. Y. Acad. Sci.* **1472**(1), 76–94 (2020)
3. S. Jasechko, Z.D. Sharp, J.J. Gibson, S.J. Birks, Y. Yi, P.J. Fawcett, Terrestrial water fluxes dominated by transpiration. *Nature* **496**(7445), 347–350 (2013)
4. T. Sohail, J.D. Zika, D.B. Irving, J.A. Church, Observed poleward freshwater transport since 1970. *Nature* **602**(7898), 617–622 (2022)
5. J.J. Leonard, A. Bahr, Autonomous underwater vehicle navigation. in *Springer Handbook of Ocean Engineering*. Springer Handbooks, ed. by M.R. Dhanak, N.I. Xiros (Springer, Cham, 2016)
6. B.H. Robison, K.R. Reisenbichler, R.E. Sherlock, The coevolution of midwater research and roV technology at mbari. *Oceanography* **30**(4), 26–37 (2017)
7. D.R. Yoerger, A.F. Govindarajan, J.C. Howland, J.K. Llopiz, P.H. Wiebe, M. Curran, J. Fujii, D. Gomez-Ibanez, K. Katija, B.H. Robison et al., A hybrid underwater robot for multidisciplinary investigation of the ocean twilight zone. *Sci. Robot.* **6**(55), 1901 (2021)
8. S.S. Afzal, W. Akbar, O. Rodriguez, M. Doumet, U. Ha, R. Ghaffarivardavagh, F. Adib, Battery-free wireless imaging of underwater environments. *Nat. Commun.* **13**(1), 5546 (2022)
9. J.S. Jaffe, Underwater optical imaging: the past, the present, and the prospects. *IEEE J. Ocean. Eng.* **40**(3), 683–700 (2014)
10. L. Paull, S. Saeedi, M. Seto, H. Li, Auv navigation and localization: a review. *IEEE J. Ocean. Eng.* **39**(1), 131–149 (2013)
11. Y. Yang, Y. Xiao, T. Li, A survey of autonomous underwater vehicle formation: performance, formation control, and communication capability. *IEEE Commun. Surv. Tutor.* **23**(2), 815–841 (2021)
12. Y. Wu, X. Ta, R. Xiao, Y. Wei, D. An, D. Li, Survey of underwater robot positioning navigation. *Appl. Ocean Res.* **90**, 101845 (2019)
13. S.B. Powell, R. Garnett, J. Marshall, C. Rizk, V. Grucev, Bioinspired polarization vision enables underwater geolocalization. *Sci. Adv.* **4**(4), 6841 (2018)
14. R. Muheim, J.B. Phillips, S. Akesson, Polarized light cues underlie compass calibration in migratory songbirds. *Science* **313**(5788), 837–839 (2006)
15. C. Buehlmann, M. Mangan, P. Graham, Multimodal interactions in insect navigation. *Anim. Cogn.* **23**, 1129–1141 (2020)

16. R. Wehner, M. Müller, The significance of direct sunlight and polarized skylight in the ant's celestial system of navigation. *Proc. Natl. Acad. Sci.* **103**(33), 12575–12579 (2006)
17. R.N. Patel, T.W. Cronin, Mantis shrimp navigate home using celestial and idiothetic path integration. *Curr. Biol.* **30**(11), 1981–1987 (2020)
18. M. Dacke, D.-E. Nilsson, C.H. Scholtz, M. Byrne, E.J. Warrant, Insect orientation to polarized moonlight. *Nature* **424**(6944), 33 (2003)
19. N.J. Pust, A.R. Dahlberg, M.J. Thomas, J.A. Shaw, Comparison of full-sky polarization and radiance observations to radiative transfer simulations which employ aeronet products. *Opt. Express* **19**(19), 18602–18613 (2011)
20. I. Pomozi, G. Horváth, R. Wehner, How the clear-sky angle of polarization pattern continues underneath clouds: full-sky measurements and implications for animal orientation. *J. Exp. Biol.* **204**(17), 2933–2942 (2001)
21. R. Hegedüs, S. Åkesson, R. Wehner, G. Horváth, Could vikings have navigated under foggy and cloudy conditions by skylight polarization? On the atmospheric optical prerequisites of polarimetric viking navigation under foggy and cloudy skies. *Proc. R. Soc. A Math. Phys. Eng. Sci.* **463**(2080), 1081–1095 (2007)
22. S. Sabbah, A. Barta, J. Gál, G. Horváth, N. Shashar, Experimental and theoretical study of skylight polarization transmitted through Snell's window of a flat water surface. *JOSA A* **23**(8), 1978–1988 (2006)
23. T.H. Waterman, Polarization patterns in submarine illumination. *Science* **120**(3127), 927–932 (1954)
24. V. Gruev, R. Perkins, T. York, CCD polarization imaging sensor with aluminum nanowire optical filters. *Opt. Express* **18**(18), 19087–19094 (2010)
25. A. Tonizzo, J. Zhou, A. Gilerson, M.S. Twardowski, D.J. Gray, R.A. Arnone, B.M. Gross, F. Moshary, S.A. Ahmed, Polarized light in coastal waters: hyperspectral and multiangular analysis. *Opt. Express* **17**(7), 5666–5683 (2009)
26. Y. Qian, Y. Zhao, Q.-L. Wu, Y. Yang, Review of salinity measurement technology based on optical fiber sensor. *Sens. Actuators, B Chem.* **260**, 86–105 (2018)
27. X. Zhang, L. Hu, M.-X. He, Scattering by pure seawater: effect of salinity. *Opt. Express* **17**(7), 5698–5710 (2009)
28. Y. You, A. Tonizzo, A.A. Gilerson, M.E. Cummings, P. Brady, J.M. Sullivan, M.S. Twardowski, H.M. Dierssen, S.A. Ahmed, G.W. Kattawar, Measurements and simulations of polarization states of underwater light in clear oceanic waters. *Appl. Opt.* **50**(24), 4873–4893 (2011)
29. P.C. Brady, A.A. Gilerson, G.W. Kattawar, J.M. Sullivan, M.S. Twardowski, H.M. Dierssen, M. Gao, K. Travis, R.I. Etheredge, A. Tonizzo et al., Open-ocean fish reveal an omnidirectional solution to camouflage in polarized environments. *Science* **350**(6263), 965–969 (2015)
30. D. Stramski, E. Boss, D. Bogucki, K.J. Voss, The role of seawater constituents in light backscattering in the ocean. *Prog. Oceanogr.* **61**(1), 27–56 (2004)
31. N. Shashar, S. Johnsen, A. Lerner, S. Sabbah, C.-C. Chiao, L.M. Mäthger, R.T. Hanlon, Underwater linear polarization: physical limitations to biological functions. *Phil. Trans. R. Soc. B Biol. Sci.* **366**(1565), 649–654 (2011)
32. R. Wehner, Polarization vision—a uniform sensory capacity? *J. Exp. Biol.* **204**(14), 2589–2596 (2001)
33. Y.Y. Schechner, N. Karpel, Recovery of underwater visibility and structure by polarization analysis. *IEEE J. Ocean. Eng.* **30**(3), 570–587 (2005)
34. T.W. Cronin, N. Shashar, R.L. Caldwell, J. Marshall, A.G. Cheroske, T.-H. Chiou, Polarization vision and its role in biological signaling. *Integr. Comp. Biol.* **43**(4), 549–558 (2003)
35. N. Shashar, S. Sabbah, T.W. Cronin, Transmission of linearly polarized light in seawater: implications for polarization signaling. *J. Exp. Biol.* **207**(20), 3619–3628 (2004)
36. T.H. Waterman, Reviving a neglected celestial underwater polarization compass for aquatic animals. *Biol. Rev.* **81**, 111–115 (2006)
37. T.W. Cronin, J. Marshall, Patterns and properties of polarized light in air and water. *Phil. Trans. R. Soc. B Biol. Sci.* **366**(1565), 619–626 (2011)
38. J.S. Tyo, M. Rowe, E. Pugh, N. Engheta, Target detection in optically scattering media by polarization-difference imaging. *Appl. Opt.* **35**(11), 1855–1870 (1996)
39. J. Hays, A.A. Efros, im2gps: estimating geographic information from a single image. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2008)
40. X. Bai, Z. Zhu, A. Schwing, D. Forsyth, V. Gruev, Angle of polarization calibration for omnidirectional polarization cameras. *Opt. Express* **31**(4), 6759–6769 (2023)
41. P. Del Moral, Non linear filtering: interacting particle solution. *Markov Process. Related Fields* **2**(4), 555–580 (1996)
42. K. He, X. Zhang, S. Ren, L. Sun, Deep residual learning for image recognition. In: *Proc. CVPR*, pp. 770–778 (2016)
43. J. Dai, H. Qi, Y. Xiong, Y. Li, G. Zhang, H. Hu, Y. Wei, Deformable Convolutional Networks. In: *Proc. ICCV* (2017)