Aquatic Mammals 2024, *50*(6), 549-568, DOI https://doi.org/10.1578/AM.50.6.2024.549

Automatically Counting Florida Manatees (*Trichechus manatus latirostris*) from Drone Images Using Object-Based Image Analysis

Esteban N. Rodofili[1, 2] and Vincent Lecours[1, 2, 3]

[1]*School of Natural Resources and Environment, University of Florida, Gainesville, FL 32603, USA*
E-mail: erodofili@ufl.edu
[2]*School of Forest, Fisheries and Geomatics Sciences, University of Florida, Gainesville, FL 32603, USA*
[3]*Laboratoire d'expertise et de recherche en géographie appliquée (LERGA), Université du Québec à Chicoutimi, Chicoutimi, Québec, Canada*

Abstract

Florida manatees (*Trichechus manatus latirostris*) require frequent and extensive surveys to inform conservation efforts. Crewed aircraft surveys can be costly, dangerous, and logistically complex. Unoccupied aerial systems (UASs) can assist with these issues. While manual review of UAS imagery can be time- and labor-intensive, automated detection of manatees in aerial survey footage can help. We present an object-based image analysis workflow for the automated detection and count of Florida manatees in *Google Earth Engine*, a free platform for research that allows for scripts and imagery sharing. Training and testing datasets were built from randomly extracted image frames from two stationary, unoccupied aerial system videos over thermal refugia. The workflow captured most manatees (93.98 to 95.62% recall; 4.38 to 6.03% false negative rate), but also counted many objects as manatees incorrectly (4.24 to 14.77% precision; 998.40 to 3,885.54% false positive over the detectable rate). Sun glint, mud plumes, and water close to shore were common causes of false positives. While the automated count was too high, the workflow lays markers over each detection, allowing for quick manual review for more accurate (semi-automated) counts. This study is an early step in automated detection tools for Florida manatees in a cloud-based platform. Future efforts could explore other platforms or may improve this workflow by including new classes for confounding objects.

Key Words: unoccupied aerial systems, object-based image analysis, automated counts, automated detection, Florida manatees, *Trichechus manatus latirostris*

Introduction

The Florida manatee (*Trichechus manatus latirostris*) inhabits coastal and riverine habitats of the United States (Reynolds et al., 2018). Its summer distribution spans from eastern Texas with occasional sightings as far north as Massachusetts, while manatees are found year-round in Florida and southeastern Georgia waters (Reynolds et al., 2018). In Florida, more than 6,000 manatees were counted in statewide aerial surveys from 2015 through 2017 (Reynolds et al., 2018), and the number of manatees in Florida in 2021-2022 was estimated to be 9,790 (with a 95% Bayesian credible interval of 8,350 to 11,730; Gowan et al., 2023). Even though data indicated the population increased during the 21st century (Reynolds et al., 2018), episodic mortality events due to algal blooms, cold stress, and potential starvation associated with seagrass loss have resulted in considerable loss to the manatee population (Bossart et al., 1998; Hardy et al., 2019; Plön et al., 2021). The Florida manatee is also threatened by watercraft collisions and loss of thermal refugia, which are used during the colder months, in addition to other factors (Reynolds et al., 2018).

In this context, abundance surveys can be a useful tool to assess the status and distribution of a population to inform conservation efforts. Additionally, smaller-scale surveys of thermal refugia sites can inform management actions by evaluating changes in manatee site usage following habitat modifications or aiding in the implementation of protection zones to reduce risk of collision (Martin et al., 2012; Udell et al., 2019; Edwards et al., 2021). While crewed aerial surveys have been conducted for Florida manatees (Edwards & Ackerman, 2016; Reynolds et al., 2018), aircraft use can be costly, dangerous, and logistically complex (Sasse, 2003; Martin et al., 2012; Hodgson et al., 2013; Edwards

et al., 2021). Furthermore, crewed aerial surveys rely on observers to accurately identify species and count individuals in a very limited time, where animals visible to observers could potentially be missed (Hodgson et al., 2013). Florida manatees are known to aggregate in large numbers in relatively small thermal refugia, which can make it hard to count individuals from an aircraft and may increase the possibility of observers missing individuals (false negatives) (Edwards et al., 2021) or confusing other objects with manatees (false positives) (Martin et al., 2012; Augustine et al., 2023). Unoccupied aerial systems (UASs) can constitute a complementary tool to crewed aerial surveys worthy of exploration as they can minimize the risk to researchers and costs (Hodgson et al., 2013). UAS imagery also provides a permanent visual record, which would reduce the chances of missing animals (Hodgson et al., 2013) or help clear false positives by allowing researchers to revisit the images and cross-validate their counts with other observers (e.g., Gonçalves et al., 2020). UASs have been used successfully to study manatees and dugongs (*Dugong dugon*). Jones et al. (2006) used a small UAS progressive-scan video for wildlife research and successfully detected Florida manatees visually. Hodgson et al. (2013) successfully analyzed UAS imagery to manually count dugongs at sea, and 95% of their possible dugong sightings were considered certain (as opposed to possible but uncertain or unclear). Turbidity was concluded to be an environmental variable that affected sighting rate, and image overlap was found to be useful to overcome hurdles of image analysis such as sun glint (direct sun reflection over water) and identifying animals initially captured at awkward body angles (Hodgson et al., 2013).

The automation of marine mammal detection in remote sensing imagery (including UAS imagery) constitutes a complementary tool to manual counts worth exploring; it could help save researcher time, effort (e.g., reducing human fatigue from manual analyses), and cost (Rodofili et al., 2022). This is particularly relevant when animals are found over large areas or in large aggregations, like those formed by Florida manatees in thermal refugia during the colder months (Reynolds et al., 2018). However, we found relatively few works on automated detection of sirenians through UAS imagery analysis—mostly in conference papers on dugongs (see Table 1 in Rodofili et al., 2022). For example, Mejias Alvarez et al. (2013) applied an automated workflow to detect dugongs in UAS imagery, obtaining recalls (Equation 1) of 48.57 and 51.4%, and precisions (Equation 2) of 4.01 and 4.97%. Maire et al. (2013) achieved better results with a recall of 69.4% and a precision of 30%, and even better results in calm water conditions (75.4% recall and 87.5% precision). The authors emphasized that dugong detection was difficult as their appearance varied dramatically with sea conditions, and their apparent color changed with depth and water turbidity. Other conference papers used aerial imagery of dugongs and convolutional neural networks (CNNs), a deep learning approach. For instance, Maire et al. (2014) obtained a recall of 59% and a precision of 30%. Also with dugongs, aerial imagery, and deep learning, Maire et al. (2015) obtained recalls of 88.23 and 80.39% with precisions of 3.48 and 15.12% (first and second generation Maxout deep convolutional neural networks [DCNNs], respectively), and recalls of 88.24 and 80.39% with precisions of 2.30 and 27.15% (first and second generation rectilinear DCNNs, respectively).

To see if we could expand on the existing automated detection of sirenians from UAS imagery, we developed and tested an object-based image analysis (OBIA) workflow to count detectable Florida manatees automatically in UAS imagery in the particular setting of thermal refugia. The workflow was developed in *Google Earth Engine* (2022), a cloud-based platform free for research and noncommercial use, to make it more widely accessible to conservation researchers and managers working in conservation.

Methods

Study Area

UAS stationary video data were collected over thermal refugia in Port of the Islands and other nearby thermal sites (Wooten's Pond and Big Cypress National Preserve) in the Ten Thousand Islands region of Collier County, Florida, on 31 January and 2 February 2017, 7 January 2018, and 31 January 2019 (see map location in Edwards et al., 2021). The UAS was piloted at an altitude of 122 m to hover over each canal for 10 min to guarantee complete width coverage and minimize glare and sun glint, with a camera footprint of ~183 m horizontally and ~122 m vertically (Edwards et al., 2021).

Data Collection and Preparation

Study imagery consisted of frames extracted from UAS stationary footage over thermal refugia (e.g., Figures 1 & 2), which allowed for images showing manatees with different postures and levels of submersion (e.g., Figure 3). UAS video data were provided by the Florida Fish and Wildlife Conservation Commission (FWC) and are described in more detail in Edwards et al. (2021). The surveys were conducted with a DJI Phantom 4 Pro UAS version 2 quadcopter.

Two 10-min stationary videos were selected (Video 1 from 31 January 2017 and Video 2 from

Figure 1. Frame from Video 1 obtained from UAS overflights of Florida manatees' (*Trichechus manatus latirostris*) thermal refugia in Collier County, Florida (two manatees observable; marked with red arrows)

7 January 2018) from all of the video available due to an optimal viewing angle close to the nadir (see Figure 4 for procedure details). Each video was analyzed to document time intervals in which manatees were visible to the human eye (i.e., detected). These intervals were used to inform both the total number of frames to be extracted for analyses and the interval between subsequent frames. An *R* code (R Core Team, 2022; Supplemental Material 1; the supplemental materials for this article are available on the *Aquatic Mammals* website) was developed in *RStudio* (Posit Team, 2023) to extract a random sample of image frames to use as training data without replacement (Wickham et al., 2016, 2021a, 2021b; Solymos & Zawadzki, 2020; Bivand et al., 2021; Galili, 2021; Garnier & Muschelli, 2021, 2022; Hijmans, 2021; Mouselimis, 2021; Csárdi et al., 2023) for two classes: (1) manatees and (2) water. First, 80 images with one or more manatees and 40 images without manatees were randomly extracted from each video using *R* to avoid capturing any temporal pattern of submersion and emersion that could bias the training process for automated detection. Then, a manual selection from these images was performed to retain 40 images with manatees and five images showing only water. The aim of the manatee images selection was to maximize the number of different postures and behaviors, as well as the differences among images, to increase image heterogeneity. A set separation of at least 11 consecutive image frames was also applied to limit redundancy. We aimed to include only image frames with clearly defined manatees in the training selection; however, this was difficult to fulfill for Video 1 while ensuring the posturing heterogeneity and meeting the time difference between frames requirement. As such, only the manatees we considered clearly defined (the ones recognized as manatees with confidence by the observer) of those image frames were included in the training dataset. Calves in Video 1 were excluded from the training dataset because of their low prevalence. Ultimately, 98 manatees were selected from the 40 training images from Video 1, while the 40 images from Video 2 contained 46 manatees. The selection of five water images targeted variability in water conditions and the presence of sun glint. The selection of water training images from Video 2 was balanced between the first and second halves of the video due to a change in video brightness. Finally, 50 image frames were randomly extracted from each video with an *R* code, independently of intervals with or without manatee presence (Supplemental Material 2), to use as datasets to test the automated detection workflow after training. Each image was investigated manually (by eye) to identify all manatees (if any). Only objects that were unmistakably identified as manatees were considered and counted.

Another *R* code was developed and utilized in conjunction with reference frames geolocated in the *Global Mapper*, Version 21.0 (b100319; 64-bit), with ground control points to convert all the extracted image frames (training and testing) into geolocated red, green, and blue (RGB) .tiff files referenced to the World Geodetic System of 1984 (WGS84; Supplemental Material 3). After this process, image frames from Video 1 covered about 29,590 m^2 with a ground resolution of 10.09 cm per pixel (Figure 1), and those from Video 2 covered about 19,575 m^2 with a ground resolution of 9.44 cm per pixel (Figure 2).

Image Processing and Training Data Preparation

Google Earth Engine (*GEE*) (Gorelick et al., 2017) was chosen to perform the image analyses because of its free access for research, education,

Figure 2. Frame from Video 2 obtained from UAS overflights of Florida manatees' thermal refugia in Collier County, Florida (two manatees observable; marked with red arrows)

Figure 3. Detail of Florida manatees in different postures and submersion levels (upper manatee emerging, manatee in the middle higher above surface but not showing its tail, and lower manatee more submerged and showing its tail)

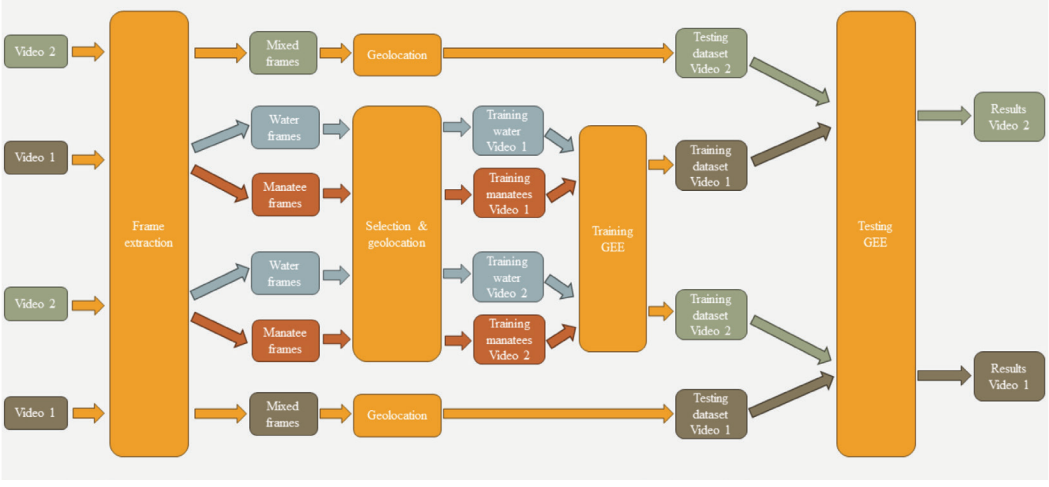

Figure 4. Set of procedures for the object-based image analysis (OBIA) Florida manatee detection workflow

and nonprofit uses, along with its Javascript application programming interface (API) and its ability to process imagery in the cloud regardless of hardware capabilities, to combine user-provided imagery with existing cloud-based imagery, and to share imagery and scripts. The training RGB images were uploaded to *GEE*, and a training script (Supplemental Material 4; see *Google Earth Engine*, 2024, for access to *GEE* information) was put together that integrated new code with Noel Gorelick's EE102 simple non-iterative clustering (SNIC) script (Gorelick, n.d.) as described below.

First, we derived from the RGB bands a water index from Upadhyay et al. (2016) by combining the bands in a mathematical formula (see the formula in point 3 of algorithm B of their study). Then, a low-pass filter, a digital processing function, was applied over each of the original RGB bands with the objective of emphasizing larger, homogenous areas (Canada Centre for Remote Sensing, 2009). In *GEE*, a polygon (or feature in *GEE* terminology) was created to outline the area of analysis. This fixed location was defined to restrain the analyses to the water area. We also added code to enable polygon transformation into a geometry, a particular type of feature in *GEE* that is directly editable with drawing tools, thus enabling the user to quickly move or slightly modify the area of analysis if, for a given frame, an element needed to be excluded from the analysis—for example, an increase in docked vessels along a canal, which would necessitate decreasing the analysis area since docked vessels do not need to be included. This feature was utilized in the analysis of Video 1 due to slight changes in perspective caused by the UAS slightly moving accidentally, which then included some land in the original region of analysis. On other occasions, the region of analysis was modified to include manatees along the canal edges. In general, the area of analysis for Video 1 was > 15,000 m^2 (Figure 5a); and for Video 2, it was > 6,750 m^2 (Figure 5b). The seven bands of the images (red, green, blue, water index, red low-pass filter, green low-pass filter, and blue low-pass filter) were then clipped using the outline features for each area of analysis to remove the terrestrial and anthropogenic elements found along the canals.

Image analysis was based on OBIA principles, which have not been explored extensively for marine mammal detection compared to more traditional pixel-based methods (Rodofili et al., 2022). These two approaches are fundamentally different in their units of analysis (Blaschke et al., 2014). OBIA starts by segmenting the imagery, dividing it into spatially continuous objects where internal heterogeneity is less than the heterogeneity of neighbors (Blaschke et al., 2014). This process yields scale-dependent, potentially more meaningful segments (or clusters) made of many relatively homogeneous pixels (Blaschke et al., 2014). A classification algorithm is then applied to the segments instead of the individual pixels that comprise them. Hence, OBIA is promising as it can provide information about the segments themselves, such as size, shape, or boundary conditions, which can be integrated as variables to inform the classification, something that cannot be done in pixel-based classifications (Blaschke et al., 2014). Before applying the segmentation, the visualization was centered on the region of analysis at a fixed zoom scale, a condition for consistent analysis in *GEE*. The SNIC segmentation algorithm available in *GEE* was used as it

Figure 5. (a) Region of analysis (> 15,000 m^2) overlayed on frame from Video 1 to restrain the analysis to water areas for the OBIA Florida manatee detection workflow; and (b) region of analysis (> 6,750 m^2) overlayed on frame from Video 2 to restrain the analysis to water areas for the OBIA Florida manatee detection workflow.

had been used before in *GEE* OBIA model codes (Gorelick, n.d.). Parameter values were optimized in preliminary testing to capture manatees fully as their own segment (Figure 6; Table 1).

The segmentation process produced a band with a unique identification number for each segment created and, for each of the seven original bands, one band that recorded the mean value of each segment for that band. In *GEE*, the segmentation process adapts to the scale of observation (i.e., the zoom level) of the interface and is therefore recomputed when the user interacts with the interface. To allow for a consistent, fixed segmentation product, we exported the segmentation products as

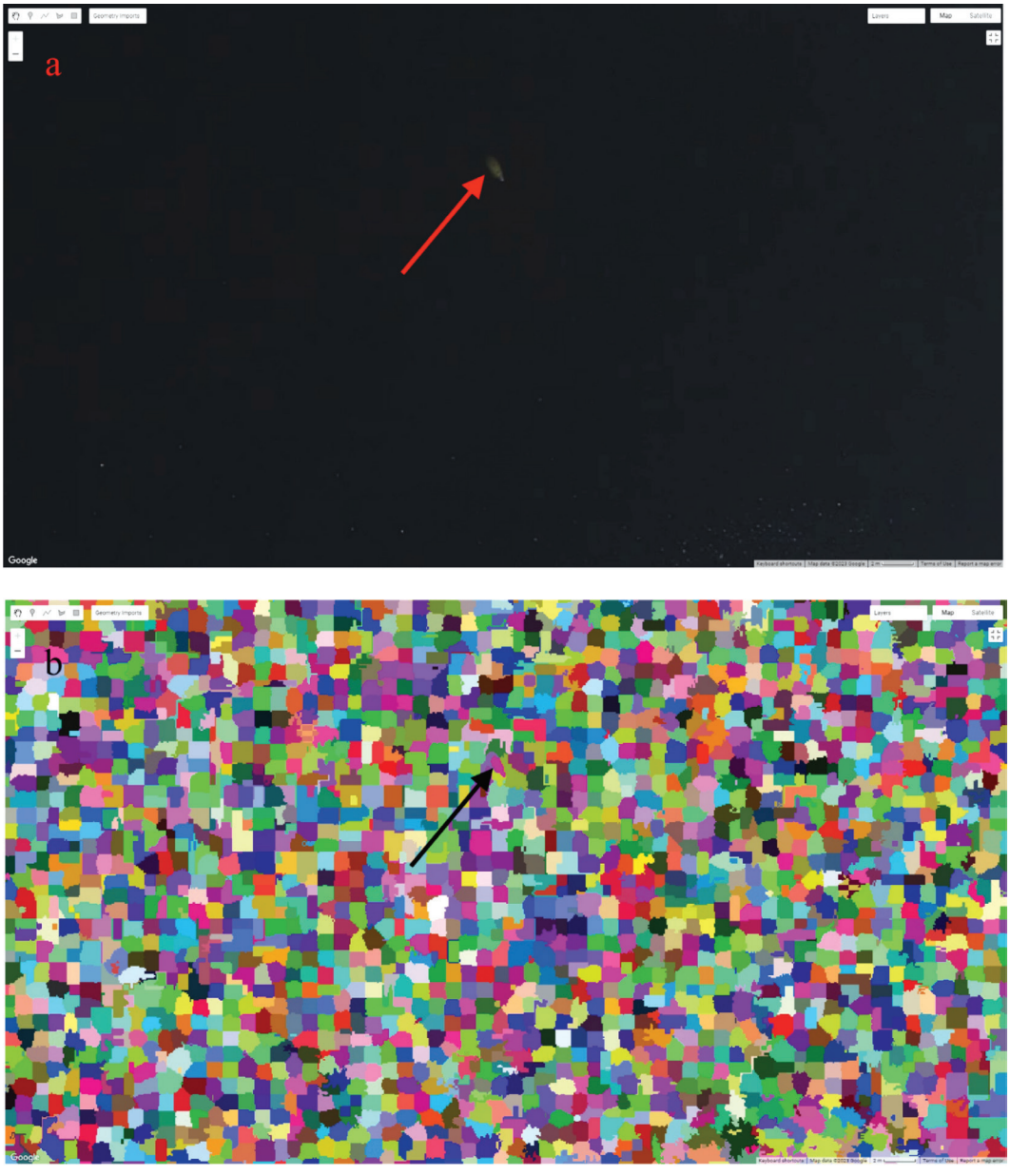

Figure 6. (a) Original frame used in the OBIA Florida manatee detection workflow implemented in *Google Earth Engine* (*GEE*) with red arrow pointing to a manatee; and (b) frame segmentation with black arrow pointing to manatee segmentation.

GEE assets at a fixed scale corresponding to each video's original resolution. The exported segmentations were then used to compute the standard deviation of each segment for each band, as well as the area, perimeter, half-width, and half-height of each segment. These variables were combined with the mean of each band for each segment in a new image, corresponding to the inputs for the classification. The manatees within the image frames retained for training of the manatee class were then delineated manually by tracing a polygon to create a *GEE* geometry (Figure 7a & b). Each of these polygons was used to extract the values of the descriptive statistics (i.e., mean, standard deviation, area, perimeter, half-width, and half-height) of the corresponding segments

Table 1. Segmentation parameters used in the object-based image analysis (OBIA) workflow implemented in *Google Earth Engine* (*GEE*) for Florida manatee (*Trichechus manatus latirostris*) detection

Parameter	Definition	Value
Seed points	Starting points for segment building process	12
Compactness factor	Causes segments to be more or less square	5
Connectivity	Number of neighboring pixels considered	8
Neighborhood size	Reduces tile boundary effects	24

or sections of segments inside the polygon. All polygons corresponding to manatees were then combined into one single *GEE* asset. The same was done for the frames retained for training the water class, although polygons were drawn to include the entire area of analysis, encompassing variable water conditions within the canals. While Video 1 showed a boat stationed in the middle of the canal in all its frames (Figure 1), Video 2 only showed boats docked by the shore, already left out by the area of analysis, extended as needed to include manatees in the water along the edge of the canals, as explained earlier. As a result, a boat class was not included in the training and, therefore, polygons for water training were drawn to avoid the boat in Video 1. Lastly, the training datasets for water and manatees were finally merged (Supplemental Material 4) to obtain the training from each video.

Testing of Workflow in Manatee Detection

We implemented a cross-validation approach under which the training dataset from Video 1 was used to detect manatees in the testing dataset from Video 2 (50 images), and vice versa. The testing script (Supplemental Material 5) followed the same processes described above for filter creation; segmentation; and the calculation of bands' standard deviation, area, perimeter, half-width, and half-height per segment. It differed from the training script in that, because of a longer extension for additional steps, the script called on functions defined in separate scripts to perform particular steps such as the segmentation (Figure 8). The main script also defined inputs for particular functions and retrieved outputs that served as input for the following functions (Figure 8). This prevented the testing script from being excessively long and confusing for future users.

A function that performed the segmentation was followed by a second function that calculated the bands' standard deviation, area, perimeter, half-width, and half-height per cluster, and also performed the image classification (see Supplemental Material 5). A random forest algorithm (a machine learning classifier consisting of an ensemble of decision trees in which each one performs a classification and the result of the majority of decision trees is the final classification result; Jensen, 2016) was used for the classification per se within this second function. This classifier was trained with the training data from the opposite video (e.g., classifier to test in Video 2 trained with data from Video 1) and a determined number of trees that was the same for both trainings. The number of trees used was determined by the computational limits of *GEE* in preliminary testing. The maximum allowable number of trees for the training from Video 1 was 800. While *GEE* allowed more trees to be used for the training from Video 2, methodology was kept consistent, utilizing 800 trees throughout the analysis. The relative importance of variables per cluster in the classification was computed by adapting a SERVIR-Mekong (2019) approach. The image compiling all the variables per segment (mean and standard deviation of each band, area, perimeter, half-width, and half-height per cluster) was classified using the trained classifier. In the final step within the second function, the classification was also exported as a *GEE* asset to fix it at the ground resolution scale and original coordinate reference system.

The exported classification was then used as input into a third function (pre-count function; see Supplemental Material 5). The first step of this function retained only the pixels of the manatee class, therefore keeping only segments of the manatee class. This function then took the manatee pixels left and identified close pixels through a 3-pixel kernel, assigning a unique ID. The size of this kernel was established in preliminary testing to have it be sufficiently small enough to allow different manatees within an aggregation to have their pixels considered far enough to be given different object IDs to have them counted separately by the next function. However, a compromise had to be made so that the kernel size did not consider parts of the same manatee as different objects, which would result

Figure 7. (a) Manatee from Video 2 as seen in frame used in the OBIA Florida manatee detection workflow implemented in *GEE*; and (b) geometry (red) delineating said manatee to extract variable information for the manatees' training dataset.

in them being assigned a different object ID and being counted separately by the next function. As a result of this function, the pixels within a segment that was classified as manatee were given a unique ID as a whole, and also various segments recognized as manatees were integrated (if considered close enough by the kernel) into objects whose pixels shared that unique ID. The assignment of the unique ID step required a maximum size for the objects. If objects were bigger than that size, they would be considered background and masked. To limit manatee aggregations from being identified as background, we used the maximum size accepted by *GEE* (1,024 pixels). However, this is a limitation as there could be a situation in which a manatee aggregation larger than that size would be masked as part of the background. The product image with the object

IDs was exported as a *GEE* asset with the original ground resolution scale.

A fourth function (count function; see Supplemental Material 5) was called from the main script as a last step. This function was given the image of the unique ID manatee objects as input to reduce each of them to a vector or feature (as opposed to an image or raster). This allowed *GEE* to count them and retrieve the automated count value in *GEE*'s console. The fourth function also obtained the centroids from each of the features and converted them to a geometry (as explained in the prior section) to better visualize such centroids over the image. While centroids could be visualized over the image as points, the conversion to geometry allowed for more conspicuous markers representing each centroid that did not decrease in size when zooming out. The last step in the fourth function script was to export and save the centroids outside of *GEE* to enable the user to open them in a geographic information system.

After the testing script was run on each frame, the automated count was recorded. True and false positives and false negatives were calculated by comparing the manual and automated counts. Following Rodofili et al. (2022), the false negative rate (FNR) (Equation 3), the false positive over detectable rate (FPDR) (Equation 4), and the automated count deviation (ACD) (Equation 5) were calculated as accuracy metrics for each of the 50 test images of both Video 1 and Video 2 individually (per-frame basis) given that each frame yielded an automated count value. Finally, mean FNR, FPDR, and ACD were calculated for each testing dataset, and a combined mean was calculated to obtain a final global accuracy measure of our workflow. Moreover, we calculated the FNR, FPDR, and ACD on a whole-video basis to replicate what would happen in a context where images would be mosaicked together by combining the false positives and negatives from all frames from each video. We then also averaged the results from the two videos. We calculated the average recall (Equation 1) and precision (Equation 2) for each testing dataset (Videos 1 & 2) on both bases (per-frame and whole-video) and their total average within a basis to compare our results with those of the other automated detection studies on Sirenia:

$$(1)\quad \text{Recall} \times 100 = \frac{TP}{TP+FN} \times 100$$

Recall = ratio between true positives and the sum of true positives and false negatives, presented as a percentage; TP = true positives; and FN = false negatives.

$$(2)\quad \text{Precision} \times 100 = \frac{TP}{TP+FP} \times 100$$

Precision = ratio between true positives and the number of positive detections by an automated method, presented as a percentage; TP = true positives; and FP = false positives.

(3)

$$(1\text{-recall}) \times 100 = \frac{FN}{TP+FN} \times 100 = \frac{FN}{\text{Detectable animals}} \times 100$$

False negative rate = ratio between the number of animals missed by an automated method (numerator) and the number of animals detectable in the image (the animals counted manually) (denominator), presented as a percentage; FN = false negatives; and TP = true positives.

$$(4)\quad \frac{FP}{TP+FN} \times 100 = \frac{FP}{\text{Detectable animals}} \times 100$$

False positive over detectable rate = ratio between the number of objects wrongly classified as animals by an automated method (numerator) and the number of animals detectable in the image (the animals counted manually) (denominator), presented as a percentage; FP = false positives; FN = false negatives; and TP = true positives.

$$(5)\quad \frac{FN+FP}{TP+FN} \times 100 = \frac{FN+FP}{\text{Detectable animals}} \times 100$$

Automated count deviation = ratio between the sum of animals missed and the objects wrongly classified as animals by an automated method (numerator) and the animals detectable in the image (the animals counted manually) (denominator), presented as a percentage; FN = false negatives; FP = false positives; and TP = true positives.

When analyzing the discrepancy between manual and automated counts of each frame, different sources of false negatives and false positives had to be considered in our accuracy metrics. Aggregated manatees could be detected as a single object, leaving remaining manatees unaccounted for and, therefore, yielding false negatives. Moreover, at times, the tip of a particular segment (small extensions from the main body of a segment) could be counted as a separate object by the workflow, generating a false positive in the automated count (Figure 9). For these reasons, false negatives and positives were grouped under diverse criteria: those that were more directly related to errors in the classification and those that, while not completely unrelated to the rest of the processes in the workflow, were more directly related to errors in the transition from image to vector and counting process of the

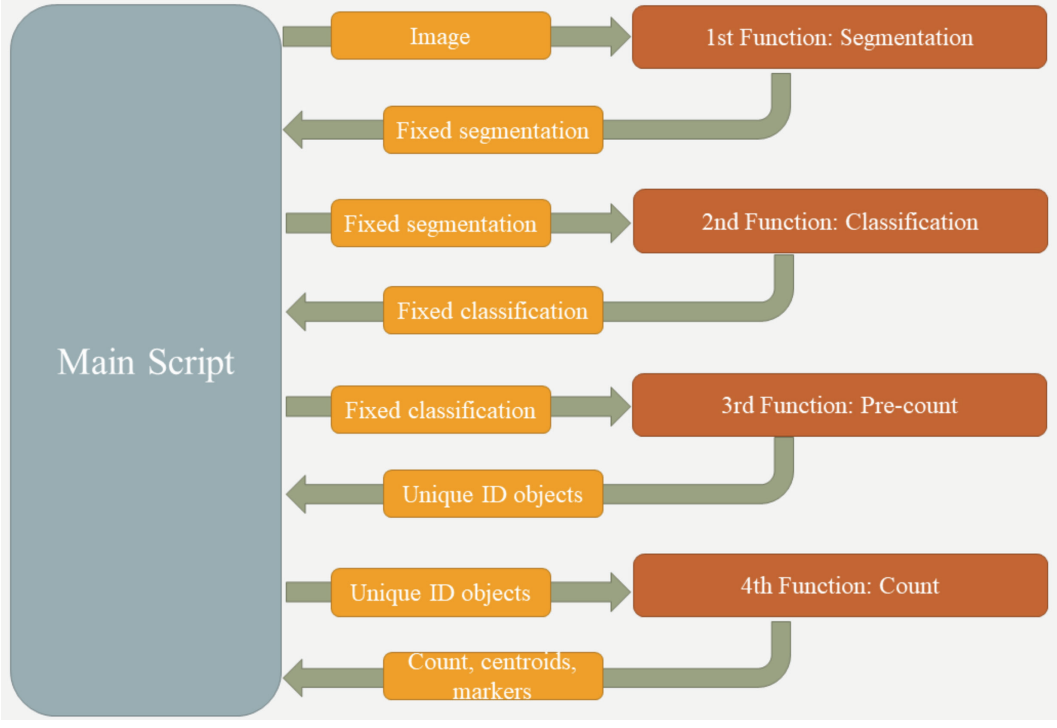

Figure 8. Main script of the OBIA Florida manatee detection workflow implemented in *GEE* for testing, along with its four functions and their main inputs and outputs

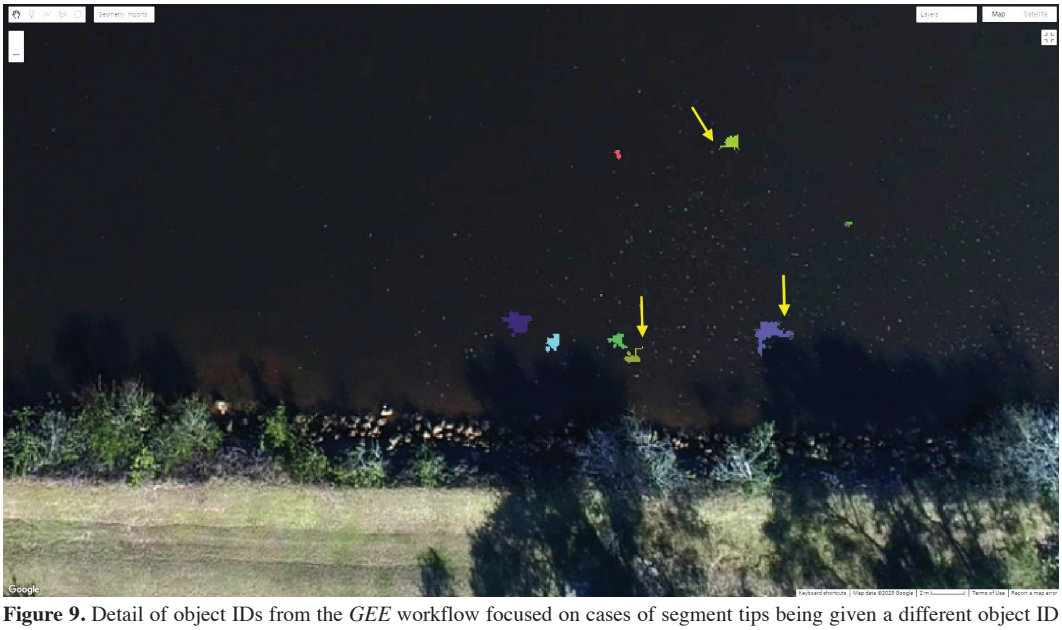

Figure 9. Detail of object IDs from the *GEE* workflow focused on cases of segment tips being given a different object ID (different color) than the main section of the segment. Three cases are indicated with yellow arrows, with upper case with green main section and violet tip, lower left case with darker green main section and light violet tip, and lower right case with violet main section and red tip.

workflow. For false negatives, we differentiated those that resulted from the classification of a manatee as water (misclassification false negatives) and those that resulted from aggregated manatees being counted as a single object (merging false negatives). In the case of false positives, we differentiated those that resulted from classification of, for example, water as a manatee (misclassification false positives) and those that resulted from tips of segments detected as different objects (tips false positives). If there was an actual manatee recognized as such that had a tip recognized as a separate manatee object, one tip false positive was added. If water was recognized as a manatee object and one tip was recognized as a separate feature, the misclassification false positives were increased by one, and the tips false positives were also increased by one. Cases in which multiple parts of a water segment were recognized as different features (concatenated tips or a main segment with tips to the side) were considered as one misclassification false positive and the rest as tips false positives. In the case where two manatees were merged as one object but a tip (not an entire manatee object itself or an object that equated to a considerable part of a manatee) was counted as a separate object, the merging false negatives were increased by one and the tips false positives were also increased by one. Finally, if only a section of a manatee or just its surrounding waters were detected as such, the manatee was counted as a false negative, and the object detected as a manatee was counted as a false positive.

In concordance with the situations described above, the FNR was calculated separately for misclassification and all false negatives together. The same applied to the misclassification false positives and all false positives in the case of the FPDR. The ACD was calculated for misclassification false positives and false negatives, and separately again for all types of false positives and negatives. However, under the per-frame basis, frames that did not have any manatees in the manual count did not have any equations calculated for them as this would entail a division by zero in the equations. Furthermore, these frames would never have false negatives. Nevertheless, in these frames, records of misclassification and merging false positive counts were documented.

Results

In the Video 1 testing dataset, 39 frames of 50 had manatees manually counted. The Video 2 testing dataset had less manatee presence, with only six frames of 50 having manually counted manatees. After segmentation, for each testing frame, we obtained the classification, the manatees recognized with unique IDs, the manatee features, and the markers over each object identified as a manatee (Figure 10). The process to export the classification in the second function in the testing of Video 1 image frames took between 6 h for the first of the testing image frames after the classifier tree testing to 3 min. For Video 2, processing time ranged from 11 h for the first of the testing image frames to 3 min.

For the per-frame results (raw data are available upon request to the corresponding author), in both datasets, the FNR was considerably lower compared to the FPDR, regardless of the metric being calculated over misclassification cases or all cases (Table 2). These metrics measure error and, therefore, lower results are better. As a consequence of the higher FPDR, the ACD values responded to the FPDR, while the FNR contribution was considerably lower (Table 2). Only Video 1 had false negatives, and a comparison between misclassification and all false negatives revealed a lower contribution from plainly missed manatees compared to manatees missed in the count by merging with other manatees in a larger object still classified under the manatee class. Two frames in which this happened corresponded to the close proximity of a mother and calf (one case in which both were initially part of the same segment and another in which they were different segments but merged into one feature later), with the calf not being a class with specific training in the workflow. In both datasets, the misclassification false positives were most of the false positives, surpassing tips false positives, and yielding a comparatively high misclassification FPDR compared to the FPDR for all false positives. Looking at the frames without manually counted manatees (left out from the metrics calculations), for the 11 frames from the Video 1 dataset, the average number of false positives was 21.09 per frame (range: 3 to 60); while for the 44 frames from the Video 2 dataset, the false positive average was 10.82 per frame (range: 2 to 28). This is congruent with the high false positive number reflected in the FPDR for the frames with manatees of each dataset. Similar trends were also reflected in the results on a whole-video basis (Table 3; raw data are available upon request to the corresponding author).

Results per frame indicated that FPDR metrics for the Video 2 dataset were lower than those of the Video 1 dataset; and while it is worth taking into account the low number of frames in Video 2 FPDR calculations, the false positives of the frames excluded from metrics calculations show

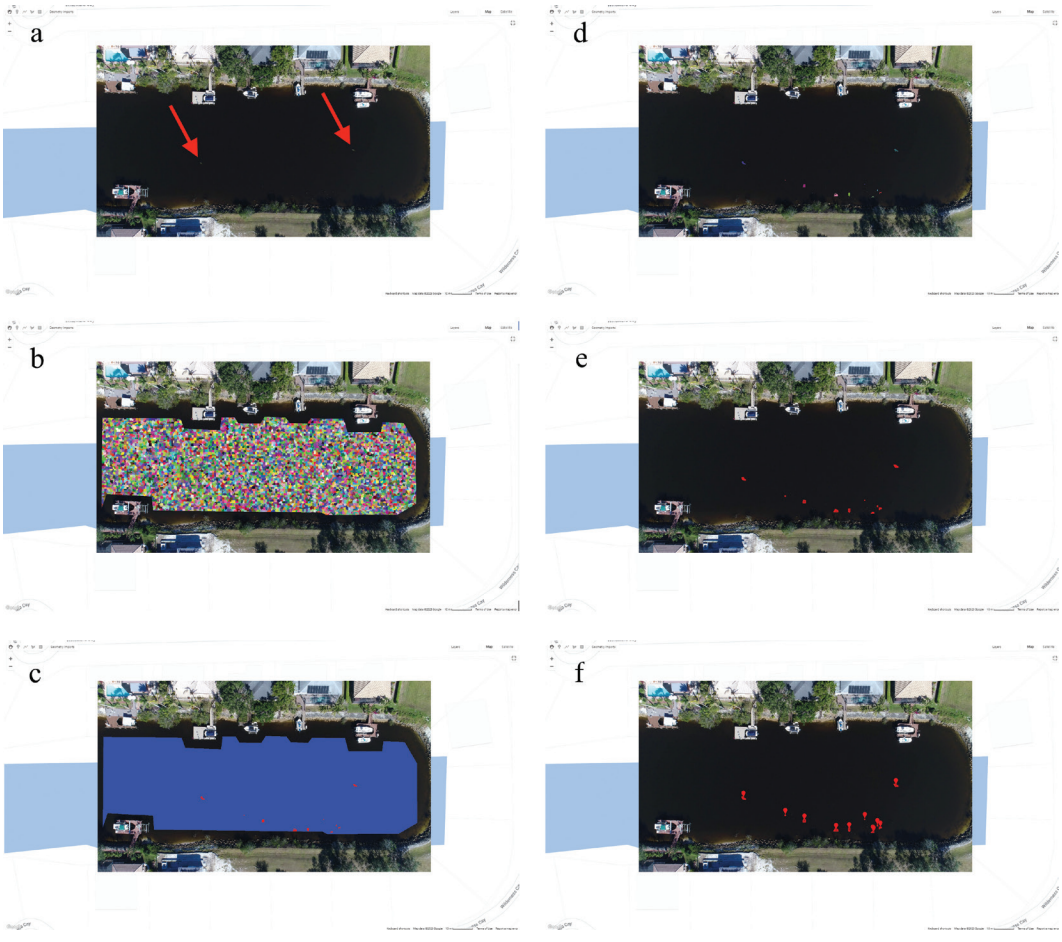

Figure 10. Steps of the OBIA Florida manatee detection workflow implemented in *GEE*: (a) original frame with manatees marked by red arrows; (b) segmentation; (c) classification with segments recognized as manatees in red (including false positives), and those recognized as water in blue; (d) objects recognized as manatees with unique object IDs (different colors); (e) objects recognized as manatees converted to features; and (f) markers over each feature recognized.

Table 2. Summary of the false negative rate (FNR), false positive over detectable rate (FPDR), and automated count deviation (ACD) results calculated on a per-frame basis for the OBIA Florida manatee detection workflow. The table includes the number of frames with manatees manually counted per testing dataset, and the FNR, FPDR, and ACD averages for each dataset among frames with manatees for misclassification and all false negatives and false positives. The last row shows the average between both datasets. **Note:** Misclassification is abbreviated as "Misc."

Testing dataset	Frames with manatees	FNR (Misc.) (%)	FNR (All) (%)	FPDR (Misc.) (%)	FPDR (All) (%)	ACD (Misc.) (%)	ACD (All) (%)
Video 1	39	2.56	8.76	1,072.44	1,263.46	1,075.00	1,272.22
Video 2	6	0.00	0.00	625.00	733.33	625.00	733.33
Average	--	1.28	4.38	848.72	998.40	850.00	1,002.78

Table 3. Summary of the false negative rate (FNR), false positive over detectable rate (FPDR), and automated count deviation (ACD) results calculated on a whole-video basis for the OBIA Florida manatee detection workflow. The table includes the FNR, FPDR, and ACD for each dataset for misclassification, and all false negatives and false positives. The last row shows the average between both datasets. **Note:** Misclassification is abbreviated as "Misc."

Testing dataset	FNR (Misc.) (%)	FNR (All) (%)	FPDR (Misc.) (%)	FPDR (All) (%)	ACD (Misc.) (%)	ACD (All) (%)
Video 1	2.41	12.05	989.16	1,171.08	991.57	1,183.13
Video 2	0.00	0.00	5,887.50	6,600.00	5,887.50	6,600.00
Average	1.21	6.03	3,438.33	3,885.54	3,439.54	3,891.57

a similar situation. Sun glint was identified as a source of false positives (see Figure S1a & b in Supplemental Material 6). Mud plumes (see Figure S1c & d in Supplemental Material 6) and water close to shore (see Figure S1e & f in Supplemental Material 6) also contributed to generating false positives. The boat presence in the middle of the canal in Video 1 yielded some false positives, mostly not due to the boat itself but to surrounding waters (see Figure S2 in Supplemental Material 6). The FPDR calculated for boat-related false positives was 232.05% on a per-frame basis (and 196.39% on a whole-video basis). Moreover, between the two videos, three image frames showed a bird's presence, which in two frames had water segments surrounding it recognized as manatees, with a section of the bird included as the positive object in one of those frames. Segmentation yielded segments that included manatees as well as surrounding water (see Figure S3 in Supplemental Material 6), and segments of water surrounding manatee segments could also be classified as positive and integrated into a positive feature that included the manatee (see Figure S4 in Supplemental Material 6). Examining results calculated on a whole-video basis (Table 3), the FPDR metrics were many times greater for Video 2 compared to Video 1, possibly a result of low numbers of actual manatees in Video 2, while the FPDR values of Video 1 varied little compared to results on a per-frame basis. These changes were reflected in the ACD values, and, as a result, the main contributor to the total averages of FPDR and ACD was Video 2 (Table 3).

As previously mentioned, we calculated the average precision and recall for each dataset and their total average to compare our results with those of these other automated detection studies on sirenians on a per-frame and whole-video basis (Tables 4 & 5; raw data are available upon request to the corresponding author).

Discussion

The results of our workflow show FNRs among the lowest compared to other marine mammal automated detection studies (see Table 1 in Rodofili et al., 2022). Comparing our results with those of studies using UAS imagery of dugongs, on a per-frame and whole-video basis, our recall was higher than that in both Maire et al. (2013) and Mejias Alvarez et al. (2013). This was while facing some common challenges, such as sun glint, animals in different positions, and turbidity, although we must note our imagery comes from a different environment (the particular thermal refugia habitat used by Florida manatees). Our recalls also surpassed those of works using dugong aerial imagery and deep learning (Maire et al., 2014, 2015); however, our results show a considerable number of false positives, with an FPDR revealing false positives from approximately 10 to more than 30 times the number of visible animals, depending on the basis for results calculation. On a per-frame basis, in terms of precision, we could see an improvement over the results of Mejias Alvarez et al. (2013), but our precision was lower than that of Maire et al. (2013, 2014). And, while higher than the first generation DCNNs in Maire et al. (2015), our precision was lower than both of their second generation DCNNs. Our precision on a whole-video basis was between the results of Mejias Alvarez et al. (2013)—4.01 and 4.97%. However, our precision on a whole-video basis was again lower than that of Maire et al. (2013, 2014); and while slightly greater than the first generation DCNNs in Maire et al. (2015), our precision was lower than both their second generation DCNNs. Our results underscore the limitations of our workflow based on OBIA and Random Forest given the large number of false positives, at least in the context of a limited dataset. Future studies could explore deep learning as these methods have resulted in improved precision (Maire et al., 2014, 2015) while still working on improving recalls. Future studies may also employ

Table 4. Summary of recall and precision results calculated on a per-frame basis for the OBIA Florida manatee detection workflow. The table includes the number of frames with manatees manually counted per testing dataset, and the recall and precision averages in their percentual form for each dataset among frames with manatees based on all false negatives and all false positives. The last row shows the average between both datasets.

Testing dataset	Frames with manatees	Recall (%)	Precision (%)
Video 1	39	91.24	12.26
Video 2	6	100.00	17.28
Average	--	95.62	14.77

Table 5. Summary of recall and precision results calculated on a whole-video basis for the OBIA Florida manatee detection workflow. The table includes the recall and precision in their percentual form for each dataset based on all false negatives and all false positives. The last row shows the average between both datasets.

Testing dataset	Recall (%)	Precision (%)
Video 1	87.95	6.99
Video 2	100.00	1.49
Average	93.98	4.24

a pixel-based approach. In summary, our results illustrate the need for additional studies to improve the methodology and to undertake other workflows and platforms, especially when it comes to decreasing confusion with sun glint, mud plumes, and water close to shore.

The steps taken after classification for an automated count of manatees (beyond just their detection) also played a key role in contributing to false positives and negatives. The kernel was selected in preliminary testing to assign aggregated manatees different object IDs so that they could be counted separately by the next function, and to avoid the same scenario for tips from segments. Nevertheless, this was not always possible. The fourth function further complicated the delineation of objects to be counted. Although the function used the object ID, it was noticed the "reduce to vectors" tool employed in said function (see Supplemental Material 5) only put pixels that were adjacent or that were connected diagonally (not sharing a side but a corner) into a feature; elsewhere, they went into separate features. As a result, separate but close segments that had been given the same object ID by the third function would be converted into different features by the fourth function and counted separately. Conversely, while segment tips were part of the same segment, if the kernel from the third function failed to give tips the same object ID as the main body of the segment, they were converted into separate features and counted separately. These aspects of the *GEE* workflow, for which some processes involved an underlying pixel-basis, added difficulty in the overall transition from segment to a counted object and in predicting which setting changes in the workflow might yield a more accurate count. We suggest further investigation into new workflow alternatives in *GEE* and other platforms that count with a more uniform OBIA-underlying logic in all workflow processes.

The segmentation yielded clusters corresponding to manatees that also included water, which may have impacted workflow accuracy. This could occur in the segmentation step during testing and training, affecting the manatee class in the resulting training dataset. These observations underscore the added complexity of OBIA use in the sense of having the additional segmentation process beyond the classification. Sometimes real-life objects in the image do not match the segments obtained (Ye et al., 2018). This can make it difficult to assess classification accuracy, and it also highlights that it may be recommendable for future studies to assess segmentation accuracy (Ye et al., 2018) and to explore different segmentation methods such as those discussed by Maire et al. (2015). One possible explanation for false positives in our workflow could be imperfect segmentation during training. As a result, values of variables used in the classification from segments that

combined manatee and water would have been included in the manatee class training, leading to recognition of water as manatee objects during testing. The cases of segmentation of manatees into multiple segments instead of a single one could also have affected the workflow's performance by confusing the variables related to shape and size (e.g., area, perimeter), with sections of a manatee being classified in testing or used in training instead of a whole manatee. In relation to the last two points on segmentation and the steps related to obtaining individual object IDs, it is important to emphasize that even within the structure of the testing script (Figure 8), the errors from one function are passed on to the next function. As such, it could be beneficial to incorporate steps to control for or correct errors between functions in future workflows (OBIA or otherwise).

In particular, if we compare the accuracy metrics of the two datasets, we notice higher FNR and FPDR in the Video 1 testing dataset on a per-frame basis, and the opposite on a whole-video basis for the FPDR, possibly as a result of the low number of actual manatees in Video 2. While the metrics on a per-frame basis could have also been affected by the scarce number of image frames that had manatees counted manually and were part of the Video 2 averages, the false positives in frames without manatees (and overall) of Video 2 still reflect fewer false detections than in Video 1. Therefore, we can at least take into account the comparison between the false positives of both testing datasets and examine possible causes. The difference in ground resolutions and area recorded, together with the general layout of each canal and the land area covered in each video, led to a difference in area of analysis, with that of Video 1 being more than twice as large. As the water class training was standardized to five frames per training dataset, this led to the training from Video 1 applied on the testing dataset from Video 2 having more area of water training, offering a possible explanation for less false positives on the Video 2 testing dataset. The number of image frames with manatee presence used for the manatee class training was also standardized for both training datasets (40 frames) and, given the higher density of manatees in Video 1, the training derived from that video and applied to the testing dataset from Video 2 had 98 manatees compared to its counterpart (46 manatees). This suggests the training from Video 1 being more extensive in manatees as well, offering an additional plausible explanation for fewer false negatives and false positives in the Video 2 testing dataset. These two explanations can be complemented by the fact that the testing in Video 2 involved less water area in which to have false positives in the first place, even though the water training dataset was bigger than that of manatees for both videos as the entire area of analysis was incorporated for training in each of the five water training frames. In the future, creating training datasets from multiple sources may yield more homogenous and better results across testing imagery datasets.

While the automated count from our workflow was not satisfactory, our workflow produced markers over each detection that maintain their size as zooming out, allowing the user to locate and examine each detection to remove false positives. Therefore, a semi-automated use of the workflow by having automated detection and a manual revision of the positives assisted by the markers is a viable alternative, especially since most workflow errors were false positives with a few missed animals (for the misclassification fraction of these, the markers would not be able to help). Semi-automation has been useful in reducing analysis time in the past (e.g., Thums et al., 2018). Furthermore, false negatives mostly occurred due to merging of various manatees into one detection as opposed to misclassification false negatives. This would also support the semi-automation approach in the sense that merging false negatives can be spotted during review because they still correspond to a detection and have a marker assigned to them. The advantages of this semi-automated use of the workflow would be more evident in the context of a satellite image or a UAS orthomosaic for which the markers allow for a quick review in particular areas without having to inspect the large extent of such images manually. The initial manual review of all imagery to obtain a count against which to measure the workflow accuracy was done by one person, with zooming options in *GEE* because of the limited extent of each image and the limited number of images in our study. However, we still consider this a limitation of our study. Moreover, in the context of a UAS orthomosaic and satellite images, human error could be more considerable, and it would be worth having more researchers to achieve a consensus count (e.g., Gonçalves et al., 2020) or to get multiple manual counts to obtain an estimate of an error rate in manual counts that may be incorporated into the rest of the analysis to carry through the uncertainty. This approach could be used in future workflows for the full manual revision of imagery to test their accuracy or in the manual counts of a semi-automated use of such workflows.

We also acknowledge that our *GEE* workflow would need improvements in its code or in *GEE*'s processing speed to become a more effective, time-saving tool. The classification of different image frames showed a reduction in time as more image frames were processed, with those following the initial testing image frame generally taking

minutes. However, the first image frames took hours for a small area, and it was the case that some image frames after the initial one could take hours again. Including training data from image frames could also take hours, and the testing of different numbers of trees for the Random Forest classifier took several hours as well. It is possible our workflow could yield a better time-to-analyzed-area performance if used in UAS orthomosaics and satellite images, which is arguably a target that automated detection of marine mammals pursues in the end but that remains to be tested. *GEE* also has limitations in time and memory when the user runs a particular workflow (*Google Earth Engin*e, 2022), which could indeed occur when running our workflow over larger areas. However, code improvements could be implemented to avoid reaching time and memory limits (*Google Earth Engine*, 2022); and *GEE* itself, in the future, may have improvements to address these situations.

This study aimed to utilize stationary video footage for training and testing with manatees in different postures and levels of submersion for automatic detection. Our results on Florida manatees contribute to the field of marine mammal automated detection in remote sensing imagery—first, by evaluating a species that has had less attention and is in the challenging environment of turbid waters; second, by using an approach that is less explored in marine mammal detection: OBIA (Rodofili et al., 2022); and lastly, by exploring the capabilities of a free platform instead of paid software for image analysis (e.g., *ArcGIS*, *eCognition*). In the future, we recommend Florida manatee researchers and conservation managers improve this workflow by incorporating further training or modifying the code and testing it on images of aggregated manatees, or on UAS orthomosaics that cover larger areas where manatees are more widely distributed. As mentioned in Rodofili et al. (2022), these workflows are concerned with the automated detection of marine mammals in imagery to obtain animal counts as a first step. Given our results, a semi-automated approach could aid in clearing the counts from false positives from the workflow as a second step. Moreover, counts could potentially be corrected for false negatives from more FNR or recall estimates from other workflow evaluation studies such as this one. Further steps are needed after counts to calculate abundance or density estimates over extended areas such as making adjustments for submerged animals not shown in the imagery—animals available in the study area but not at the surface (i.e., not detectable) as has been implemented for whale densities calculated from satellite imagery (e.g., Bamford et al., 2020). These adjustments could involve additional information obtained through tracking devices such as surfacing and submersion times (e.g., Bamford et al., 2020). In their abundance estimation of Florida manatees in 2015-2016 from crewed aerial surveys, Hostetler et al. (2018) used a method described by Martin et al. (2015) that involved the detectability of a manatee replica at 0.5-m increments of submersion and visibility categories for different sites. Martin et al. (2015) also incorporated information on manatee diving behavior from telemetry and time-depth data to account for submerged animals.

This study also used imagery derived from stationary video for workflow training and testing with manatees in different postures and levels of submersion. However, if imagery over the same position was aimed for use in abundance or density estimates, accounting for detection of the same manatee would be important (e.g., Edwards et al., 2021). In consonance with an application of automated detection workflows to larger areas of Florida manatee habitat, we recommend the inclusion of object classes that acted as confounding objects in our study, such as birds, boats, mud plumes, water by the shoreline, sun glint, and manatee calves. The inclusion of boats, together with the capability to recognize land, would aid in the application to larger areas. Exploring other free platforms for OBIA, like *R*, and investigating if accuracy and processing time can be improved to make OBIA an effective and accessible tool for Florida manatee researchers and conservation managers is recommended. Furthermore, deep learning has shown promise for marine mammal automated detection (Rodofili et al., 2022) and, as we have seen from Maire et al. (2014, 2015), continuing to explore free platforms for deep learning workflows is encouraged. In conclusion, as UASs are used more in marine mammal surveys for their cost and logistical advantages, detection and count workflows have the potential to become a useful time-saving tool in marine mammal research and conservation programs.

Note: The supplemental materials for this article are available in the "Supplemental Material" section of the *Aquatic Mammal*s website: https://www.aquaticmammalsjournal.org/supplemental-material.

Acknowledgments

Thanks are due to the Florida Fish and Wildlife Conservation Commission (FWC) and Holly Edwards (FWC), Julien Martin (U.S. Geological Survey), Timothy Gowan (FWC), Jeffrey Hostetler (FWC), and Brad Smith (independent researcher) for sharing the UAS videos for the training and testing datasets and for their assistance and feedback throughout this work. Special thanks to Noel

Gorelick (Google) for his assistance with the code. Thanks are due as well to Michelle LaRue (University of Canterbury), and to Alina Zare, Debra Murie, Edward Camp, Amr Abd-Elrahman, Reza Khatami, Youssef Omar Kadoura, and Michael Espriella from the University of Florida for their assistance and feedback throughout this work. Thanks are also due to the reviewers and handling editor for their constructive feedback. This work was supported by funds allocated to Vincent Lecours by the University of Florida Senior Vice President for Agriculture and Natural Resources, and by an Early-Career Research Fellowship to Vincent Lecours from the Gulf Research Program of the National Academies of Sciences, Engineering, and Medicine. Funding was also provided by the Natural Sciences and Engineering Research Council of Canada through a Discovery Development Grant (#DDG-2024-00018) awarded to Vincent Lecours. Esteban N. Rodofili was funded jointly by the School of Natural Resources and Environment and the School of Forest, Fisheries and Geomatics Sciences (College of Agricultural and Life Sciences) of the University of Florida. The authors declare there are no competing interests. Random extraction *R* codes for training and testing frames and frame geolocation, and *GEE* links to example training scripts, merging scripts, testing main scripts, and functions' scripts are available in the supplemental materials. Image frames; geolocation reference frames, their *R* extraction code, and ground control points; GEE water and manatee training geometries; regions of analysis (original or modified); and training sets are available from the authors upon reasonable request and with permission of Edwards et al. (2021). For datasets from Edwards et al. (2021), please see Edwards et al. (2021).

Literature Cited

Augustine, B. C., Koneff, B. A., Pickens, B. A., & Royle, J. A. (2023). Towards estimating marine wildlife abundance using aerial surveys and deep learning with hierarchical classifications subject to error. *bioRxiv*. https://doi.org/10.1101/2023.02.20.529272

Bamford, C. C. G., Kelly, N., Dalla Rosa, L., Cade, D. E., Fretwell, P. T., Trathan, P. N., Cubaynes, H. C., Mesquita, A. F. C., Gerrish, L., Friedlaender, A. S., & Jackson, J. A. (2020). A comparison of baleen whale density estimates derived from overlapping satellite imagery and a shipborne survey. *Scientific Reports*, *10*(1), 12985. https://doi.org/10.1038/s41598-020-69887-y

Bivand, R., Keitt, T., & Rowlingson, B. (2021). *rgdal: Bindings for the "Geospatial" Data Abstraction Library*. https://CRAN.R-project.org/package=rgdal

Blaschke, T., Hay, G. J., Kelly, M., Lang, S., Hofmann, P., Addink, E., Feitosa, R. Q., van der Meer, F., van der Werff, H., van Coillie, F., & Tierde, D. (2014). Geographic object-based image analysis – Towards a new paradigm. *ISPRS Journal of Photogrammetry and Remote Sensing*, *87*, 180-191. https://doi.org/10.1016/j.isprsjprs.2013.09.014

Bossart, G. D., Baden, D. G., Ewing, R. Y., Roberts, B., & Wright, S. D. (1998). Brevetoxicosis in manatees (*Trichechus manatus latirostris*) from the 1996 epizootic: Gross, histologic, and immunohistochemical features. *Toxicologic Pathology*, *26*(2), 276-282. https://doi.org/10.1177/019262339802600214

Canada Centre for Remote Sensing. (2009). *Fundamentals of remote sensing*. https://natural-resources.canada.ca/sites/nrcan/files/earthsciences/pdf/resource/tutor/fundam/pdf/fundamentals_e.pdf

Csárdi, G., Hester, J., Wickham, H., Chang, W., Morgan, M., & Tenenbaum, D. (2023). *remotes:* R *package installation from remote repositories, including "GitHub."* https://remotes.r-lib.org

Edwards, H. H., & Ackerman, B. B. (Eds.). (2016). *Aerial surveys of manatee distribution in Florida, 1984–2004*. Florida Fish and Wildlife Conservation Commission. https://www.researchgate.net/publication/305278000_Edwards_Holly_H_and_Bruce_B_Ackerman_eds_2016_Aerial_surveys_of_manatee_distribution_in_Florida_1984-2004_Fish_and_Wildlife_Research_Institute_Technical_Report_TR-19_iv_273_pp#fullTextFileContent

Edwards, H. H., Hostetler, J. A., Stith, B. M., & Martin, J. (2021). Monitoring abundance of aggregated animals (Florida manatees) using an unmanned aerial system (UAS). *Scientific Reports*, *11*(1), 12920. https://doi.org/10.1038/s41598-021-92437-z

Galili, T. (2021). *installr: Using* R *to install stuff on Windows OS (such as:* R*, "Rtools," "RStudio," "Git," and more!)*. https://CRAN.R-project.org/package=installr

Garnier, S., & Muschelli, J. (2021). *{Rvision} – A computer vision library for* R. https://swarm-lab.github.io/Rvision

Garnier, S., & Muschelli, J. (2022). *{ROpenCVLite} – Helper package for installing OpenCV with* R. https://swarm-lab.github.io/ROpenCVLite

Gonçalves, B. C., Spitzbart, B., & Lynch, H. J. (2020). SealNet: A fully-automated pack-ice seal detection pipeline for sub-meter satellite imagery. *Remote Sensing of Environment*, *239*, 111617. https://doi.org/10.1016/j.rse.2019.111617

Google Earth Engine. (2022). *Debugging guide*. https://developers.google.com/earth-engine/guides/debugging

Google Earth Engine. (2024). *Earth engine access*. https://developers.google.com/earth-engine/guides/access

Gorelick, N. (n.d.). *SNIC*. https://code.earthengine.google.com/?accept_repo=users/gorelick/EE102

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). *Google Earth Engine*: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, *202*, 18-27. https://doi.org/10.1016/j.rse.2017.06.031

Gowan, T. A., Edwards, H. A., Krzystan, A. M., Martin, J., & Hostetler, J. A. (2023). *2021-2022 statewide abundance estimates for the Florida manatee* (Fish and Wildlife

Research Institute Technical Report No. 27). Florida Fish and Wildlife Conservation Commission. https://f50006a.eos-intl.net/ELIBSQL12_F50006A_Documents/Final_1.5_TR27_231027.pdf

Hardy, S. K., Deutsch, C. J., Cross, T. A., de Wit, M., & Hostetler, J. A. (2019). Cold-related Florida manatee mortality in relation to air and water temperatures. *PLOS ONE*, *14*(11), e0225048. https://doi.org/10.1371/journal.pone.0225048

Hijmans, R. J. (2021). *raster: Geographic data analysis and modeling*. https://CRAN.R-project.org/package=raster

Hodgson, A., Kelly, N., & Peel, D. (2013). Unmanned aerial vehicles (UAVs) for surveying marine fauna: A dugong case study. *PLOS ONE*, *8*(11), e79556. https://doi.org/10.1371/journal.pone.0079556

Hostetler, J. A., Edwards, H. H., Martin, J., & Schueller, P. (2018). *Updated statewide abundance estimates for the Florida manatee* (Fish and Wildlife Technical Report No. 23). Florida Fish and Wildlife Conservation Commission. https://f50006a.eos-intl.net/F50006A/OPAC/Details/Record.aspx?BibCode=1864664

Jensen, J. R. (2016). *Introductory digital image processing: A remote sensing perspective* (4th ed.). Pearson Education, Inc.

Jones, G. P., Pearlstine, L. G., & Percival, H. F. (2006). An assessment of small unmanned aerial vehicles for wildlife research. *Wildlife Society Bulletin*, *34*(3), 750-758. https://doi.org/10.2193/0091-7648(2006)34[750:aaosua]2.0.co;2

Maire, F., Mejias Alvarez, L., & Hodgson, A. (2014). A convolutional neural network for automatic analysis of aerial imagery. In W. Li, L. Wang, S. L. Phung, A. Bouzerdoum, & P. Bouzerdoum (Eds.), *Proceedings of the 2014 International Conference on Digital Image Computing: Techniques and Applications (DICTA)* (pp. 1-8). Institute of Electrical and Electronics Engineers Inc. https://eprints.qut.edu.au/77510/1/dicta14_dugong_cnn.pdf; https://doi.org/10.1109/DICTA.2014.7008084

Maire, F., Mejias Alvarez, L., & Hodgson, A. (2015). Automating marine mammal detection in aerial images captured during wildlife surveys: A deep learning approach. In B. Pfahringer & J. Renz (Eds.), *AI 2015: Advances in artificial intelligence* (*AI 2015. Lecture Notes in Computer Science*, Vol. 9457). Springer, Cham, Switzerland. https://doi.org/10.1007/978-3-319-26350-2_33

Maire, F., Mejias Alvarez, L., Hodgson, A., & Duclos, G. (2013). Detection of dugongs from unmanned aerial vehicles. In *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems* (pp. 2750-2756). IEEE. https://doi.org/10.1109/IROS.2013.6696745

Martin, J., Edwards, H. H., Fonnesbeck, C. J., Koslovsky, S. M., Harmak, C. W., & Dane, T. M. (2015). Combining information for monitoring at large spatial scales: First statewide abundance estimate of the Florida manatee. *Biological Conservation*, *186*, 44-51. https://doi.org/10.1016/j.biocon.2015.02.029

Martin, J., Edwards, H. H., Burgess, M. A., Percival, H. F., Fagan, D. E., Gardner, B. E., Ortega-Ortiz, J. G., Ifju, P. G., Evers, B. S., & Rambo, T. J. (2012). Estimating distribution of hidden objects with drones: From tennis balls to manatees. *PLOS ONE*, *7*(6), e38882. https://doi.org/10.1371/journal.pone.0038882

Mejias Alvarez, L., Duclos, G., Hodgson, A., & Maire, F. (2013). Automated marine mammal detection from aerial imagery. In R. Wernli, Sr. (Ed.), *Proceedings of the OCEANS '13 Conference* (pp. 1-5). Institute of Electrical and Electronics Engineers Inc. https://eprints.qut.edu.au/61589/1/PID2867597.pdf

Mouselimis, L. (2021). *OpenImageR: An image processing toolkit*. https://CRAN.R-project.org/package=OpenImageR

Plön, S., Roussouw, N., & Cockcroft, V. (2021). Unprecedented mortality of Florida manatees along the Atlantic coast. *Sirenews* (No. 73), 24-26. https://mission.cmaquarium.org/app/uploads/2021/05/Sirenews-73-April2021-update.pdf

Posit Team. (2023). RStudio*: Integrated development environment for* R. Posit Software, PBC.

R Core Team. (2022). R*: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.r-project.org

Reynolds III, J. E., Powell, J. A., Keith Diagne, L. W., Barton, S. L., & Scolardi, K. M. (2018). Manatees: *Trichechus manatus*, *T. senegalensis*, and *T. inunguis*. In B. Würsig, J. G. M. Thewissen, & K. M. Kovacs (Eds.), *Encyclopedia of marine mammals* (3rd ed., pp. 558-566). Elsevier. https://doi.org/10.1016/B978-0-12-804327-1.00165-5

Rodofili, E. N., Lecours, V., & LaRue, M. (2022). Remote sensing techniques for automated marine mammals detection: A review of methods and current challenges. *PeerJ*, *10*, e13540. https://doi.org/10.7717/peerj.13540

Sasse, D. B. (2003). Job-related mortality of wildlife workers in the United States, 1937-2000. *Wildlife Society Bulletin*, *31*(4), 1015-1020. https://www.jstor.org/stable/3784446

SERVIR-Mekong. (2019). *Script for variable importance in classification*. SERVIR-Mekong.

Solymos, P., & Zawadzki, Z. (2020). *pbapply: Adding progress bar to "*apply" functions*. https://CRAN.R-project.org/package=pbapply

Thums, M., Jenner, C., Waples, K., Salgado Kent, C., & Meekan, M. (2018). *Humpback whale use of the Kimberley: Understanding and monitoring spatial distribution* (Report of Project 1.2.1). Prepared for the Kimberley Marine Research Program, Western Australian Marine Science Institution, Perth, Western Australia, Australia. https://wamsi.org.au/wp-content/uploads/bsk-pdf-manager/2019/07/Final-Report-WAMSI-KMRP-Whales-Humpback-Whale-Use-of-the-Kimberley-Thums-et-al.pdf

Udell, B. J., Martin, J., Fletcher, R. J., Jr., Bonneau, M., Edwards, H. H., Gowan, T. A., Hardy, S. K., Gurarie, E., Calleson, C. S., & Deutsch, C. J. (2019). Integrating encounter theory with decision analysis to evaluate collision risk and determine optimal protection zones for wildlife. *Journal of Applied Ecology*, *56*(5), 1050-1062. https://doi.org/10.1111/1365-2664.13290

Upadhyay, P., Mahadik, S., & Kamble, A. (2016). Image classification using visible RGB bands. In *2016 3rd*

International Conference on Computing for Sustainable Global Development (INDIACom) (pp. 2660-2663). https://ieeexplore.ieee.org/document/7724745

Wickham, H., François, R., Henry, L., & Müller, K. (2021a). *dplyr: A grammar of data manipulation*. https://CRAN.R-project.org/package=dplyr

Wickham, H., Hester, J., Chang, W., & Bryan, J. (2021b). *devtools: Tools to make developing* R *packages easier.* https://CRAN.R-project.org/package=devtools

Wickham, H., Chang, W., Henry, L., Lin Pedersen, T., Takahashi, K., Wilke, C., Woo, K., Yutani, H., Dunnington, D., & van den Brand, T. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag. https://ggplot2.tidyverse.org

Ye, S., Pontius, R. G., & Rakshit, R. (2018). A review of accuracy assessment for object-based image analysis: From per-pixel to per-polygon approaches. *ISPRS Journal of Photogrammetry and Remote Sensing*, *141*, 137-147. https://doi.org/10.1016/j.isprsjprs.2018.04.002