Journal of Aviation/Aerospace Education & Research

Volume 33 | Number 5 Article 1

2024

Real-Time Detection of Sea Turtles Using UAV and Neural Networks on Edge Devices

Jose A. Gonzalez Nunez
Embry-Riddle Aeronautical University, gonzj179@my.erau.edu

Jose G. Gonzalez Nunez
Embry-Riddle Aeronautical University, gonzj180@my.erau.edu

Mustafa I. Akbas
Embry-Riddle Aeronautical University, akbasm@erau.edu

Patrick Currier
Embry-Riddle Aeronautical University, currierp@erau.edu

Nickolas D. Macchiarella
Embry-Riddle Aeronautical University, macchian@erau.edu

Scholarly Commons Citation

Gonzalez Nunez, J. A., Gonzalez Nunez, J. G., Akbas, M. I., Currier, P., & Macchiarella, N. D. (2024). Real-Time Detection of Sea Turtles Using UAV and Neural Networks on Edge Devices. *Journal of Aviation/Aerospace Education & Research, 33*(5). DOI: https://doi.org/10.58940/2329-258X.2060

JAAER

Real-Time Detection of Sea Turtles Using UAV and Neural Networks on Edge Devices

Jose A. Gonzalez Nunez[1a], **Jose G. Gonzalez Nunez**[1b], **Mustafa I. Akbas**[1c], **Patrick Currier**[1d], **Nickolas D. Macchiarella**[1e]
[1]Embry-Riddle Aeronautical University, FL 32124 USA,
[a]gonzj179@my.erau.edu, [b]gonzj180@my.erau.edu, [c]akbasm@erau.edu, [d]currierp@erau.edu, [e]macchian@erau.edu

Abstract

Sea turtle populations continue to diminish around the globe due to various reasons. Therefore, the need for innovative solutions to monitor sea turtles has been increasing. This research paper focuses on an innovative application of artificial intelligence (AI) and machine learning (ML) together with unmanned aerial vehicles (UAV) to improve sea turtle conservation efforts. We outline the design, implementation, and evaluation of a system that deploys UAV equipped with high-resolution cameras, which, coupled with a purpose-built neural network to recognize, and monitor sea turtles. This project thus serves as a platform for understanding the wider applicability and limitations of this technology in the realm of wildlife conservation, while placing particular emphasis on the protection of sea turtles.

Keywords: *Unmanned Aerial Vehicles, Environmental Monitoring, Image Processing, Machine Learning*

Introduction

The increasing urgency to protect our vulnerable marine species has prompted a change towards technology, particularly in the area of wildlife monitoring and conservation. As sea turtle populations continue to diminish due to human activities, poaching, habitat loss, and climate change (Coston-Clements et al., 2009), the need for innovative solutions has never been greater. This research paper focuses on an innovative application of artificial intelligence (AI) and its various sub-disciplines, such as machine learning (ML) and neural networks together with unmanned aerial vehicles (UAVs) to improve sea turtle conservation efforts.

Sea turtles, globally recognized and cherished for their beauty and longevity, are integral to the balance of marine ecosystems. Despite their importance, they are currently under threat due to various factors including pollution, climate change, human activities, and illegal hunting, leading to declining population numbers, and raising concerns among conservationists (Donlan et al., 2010). Accurate and efficient monitoring and estimates of sea turtle populations is pivotal to their conservation (Cohen et al., 2003), but the manual monitoring methods traditionally employed are labor-intensive, unreliable in the long run, and can even pose a danger to the turtle (Anuntachai & Pantuwong, 2019). Leveraging the power of neural networks and image recognition software, we have embarked on a novel journey to automate the process of identifying and tracking sea turtles in their natural habitats. The benefits of such an approach are many, including non-intrusive surveillance, increased surveillance scope, real-time data collection, and potential reduction in human-induced stress to the animals.

In this paper, we outline the design, implementation, and evaluation of a system that deploys UAVs equipped with high-resolution cameras, which, coupled with a purpose-built neural network, will recognize, classify, and monitor sea turtle species in real time. This research will thus serve as a platform for understanding the wider applicability and limitations of this technology in the realm of wildlife conservation, while placing particular emphasis on the protection of sea turtles.

By developing and refining these tools, we hope to facilitate more informed decision-making, improve conservation policies, and ultimately, contribute to the long-term survival of sea turtles. As we usher in a new era of technology-enabled conservation, it is critical that we continue to push the boundaries of what is possible, leveraging every tool at our disposal to ensure the ongoing vitality of our marine ecosystems.

The main goal of our research is to cultivate and refine an object detection neural network that provides immediate insights regarding the presence of sea turtles within specified geographic zones. We hope to integrate this neural network technology into aerial autonomous vehicles, thus enhancing the efficacy of sea turtle monitoring and eliminating the necessity for ground-based operations. The visual data procured via this methodology will serve to enrich a comprehensive sea turtle database, removing the need for physically capturing them and affixing tracking devices onto them which can sometimes cause health issues for the turtles (Cohen et al., 2003). This innovative approach aspires to create a more sustainable and less intrusive method for monitoring sea turtle populations.

Background and Related Work

The ongoing advancement and broader availability of diverse UAVs has fostered their incorporation into many domains (Neubauer & Akbas, 2022), such as aerial wireless sensor (Akbas & Turgut, 2011), weather observation, urban air mobility (Adkins et al., 2020, 2021) and many others (Muna et al., 2021). A growing body of research in recent years has started to explore the potential of these drones to bolster conservation efforts and facilitate comprehensive surveys of marine fauna (Hensel et al., 2018; Hodgson et al., 2013; Landeo-Yauri et al., 2020). These studies have largely depended upon the UAVs' capability to acquire an abundant number of images, which could be analyzed to discern unique features of marine creatures, thereby enabling their identification.

Photo-identification presents a non-intrusive methodology that allows researchers to conduct seamless monitoring of marine species, eliminating any potential disruption or harm to these creatures (Gope et al., 2005). However, the efficacy of manual identification starts to suffer when dealing with smaller species or those which rarely surface above the water, as their distinct markers are often harder to discern. Moreover, this approach necessitates a labor-intensive process of sifting through a massive array of images to isolate those that contain signs of marine life (Sykora-Bodie et al., 2017).

These impediments have propelled researchers to seek the utilization of neural networks for assistance. In a study by Gray et al. (2019), drone imagery was combined with neural networks to facilitate the identification of sea turtles along the coast of Costa Rica. Their model outperformed manual counts by detecting 8% more turtles, concurrently diminishing the amount of labor required for the process (Gray et al., 2019). Another study by Badawy and Direkoglu (2019) presented a system for sea turtle detection that uses a Faster Region-based Convolutional Neural Network (R-CNN) algorithm that performs detection on a cloud. Their model was able to achieve good results with a precision of 95.7% and a recall of 77.6%, where recall refers to the model's ability to identify all relevant instances (Badawy & Direkoglu, 2019). Both of these studies underscored the viability of using neural networks and drones in tandem to accelerate and enhance the identification process, setting the stage for future endeavors in this direction. These constraints encourage research into the application of neural networks to aid in the process.

Challenges

Nowadays, executing and testing a neural network for object detection is often regarded as a standard procedure, readily accomplished given the volume of research and studies dedicated to this topic. However, the success of these models heavily relies on ample computational resources and large datasets, as they enable the network to learn complex patterns more effectively. Unfortunately, we faced challenges with limited data and computational constraints, which impacted our implementation.

The foremost challenge faced during this research was a lack of suitable data to adequately train the neural network. While photographs of sea turtles are relatively accessible, they predominantly depict lateral views and are captured at close range. The scarcity of aerial imagery of sea turtles, particularly at altitudes suitable for our project (300 feet as required by our flight permit), presents a considerable hurdle. The few available images are typically captured by low-flying drones. However, these images often contain a higher level of detail than would be achievable at the intended flight altitude of our project, leading to initial difficulties in sea turtle identification. While this issue can be mitigated over time by collecting images at the necessary altitude and using these to further train the image recognition software, this presents another challenge. Photographing wild sea turtles is a challenging task due to the seasonal and unpredictable nature of their appearance along coastlines, and planning UAV flights for data collection accordingly is complex and resource intensive. For the neural network's initial training phase, to mitigate the issue of a lack of available images we used a combination of pictures taken from low flying drones obtained online, images used in the previously mentioned studies from Gray et al. (2019), Badawy and Direkoglu (Badawy & Direkoglu, 2019), which are publicly available and some lower quality screenshots taken from public online videos. This allowed us to create a decently sized database of around 500 images that would work during the initial stages of the project while we collect our own images.

The second major challenge revolves around hardware limitations of the UAVs. Given the relatively small payload capacity of most drones, the addition of any substantial device could adversely impact the drone's flight time. Therefore, hardware selection was highly constrained. Our selection was the Nvidia Jetson Xavier from Leopard Imaging, a compact device capable of capturing images and running the neural network without significantly affecting the UAV's weight or size. However, this device's limited internal storage necessitated the creation of lightweight image recognition software, further complicating the implementation process.

Methodology

The combination of Unmanned Aerial Vehicle (UAV) technology and object detection methodology is poised to become an instrumental facet in conservation and rescue endeavors across a diverse array of habitats. Thus, the primary objective of this paper is the implementation and evaluation of a neural network dedicated to object recognition, designed to enable the real-time detection and categorization of sea turtles. This technological approach is pursued with the goal of bolstering the preservation and sustainability of these marine species. The following section shows our overarching research goal and elaborates upon the intricate challenges associated with the concurrent deployment of object detection technology and UAVs. Additionally, it provides an account of the current initiatives in this realm and delineates prospective objectives.

Development Process

Hardware

The device supporting the software must adhere to the weight and size constraints associated with UAV payloads. In light of this requirement, our choice has been the Nvidia Jetson device series, known for their potent yet compact configurations designed specifically for running AI software. More precisely, we utilized an Nvidia Jetson Xavier from Leopard Imaging, model LI-XNX-CB-6CAM, equipped with the capability to operate a neural network and simultaneously support up to six MIPI CSI cameras. This device, weighing 68 grams and measuring 100 mm x 79 mm, comfortably aligns with drone payload specifications. Nonetheless, as previously acknowledged, the Jetson device is constrained by a relatively small internal storage capacity, which dictated the maximum size of the image recognition software.

Software

Figure 1

Overview of Object Detection Model Training

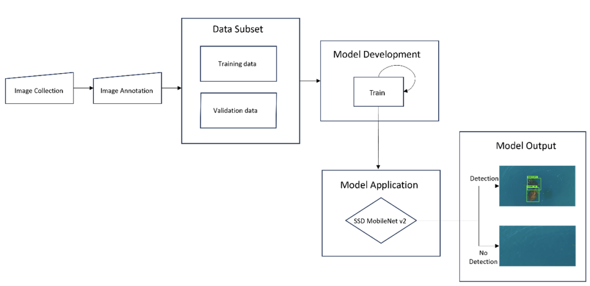


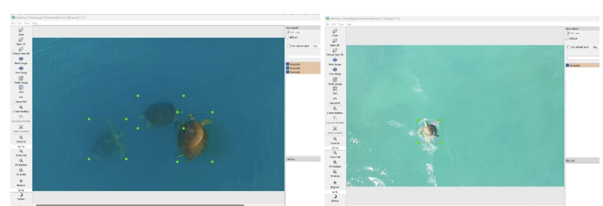
Image Processing

In the context of this paper, our main focus is image processing and the training of the neural network to enable swift and accurate detection of sea turtles from images taken at an altitude of 300 feet, the flight altitude allowed by our flight permit. Image processing encapsulates the alteration of digital images via computer algorithms, typically executed to enhance image quality or to extract specific information. For the purpose of this study, this procedure was employed to delineate and annotate the portions of images where sea turtles were discernible. This procedure is carried out in order to provide data about the object shown in the image to the computer vision model. Labeling the image greatly improves the training process of the model as well as its overall accuracy and precision by removing unnecessary information from the image, allowing the model to focus on relevant features and patterns of the main object.

The labeling process was done by using an open-source program called LabelImg. This program is one of the most popular annotation tools, it was a great option to quickly label objects in multiple images and minimize the time required for this process. LabelImg is no longer being developed; it's now part of an improved version called "Label Studio" which contains multiple new features and allows the user to annotate text, audio, and videos. Figure 2 illustrates an example of the labeling process. The data procured from this exercise was then used in the training of our neural network, allowing the object detection model to independently identify the defined data features.

Figure 2

Examples of Annotation Process



Object Detection Model

For the creation and training of the machine learning module, we utilized TensorFlow (Abadi et al., 2016), a complimentary and open-source software library developed by Google. The primary aim of this library is to facilitate the development and application of machine learning models in a streamlined and efficient manner, extending its functionality to tasks such as image and speech recognition, and processing of language, numbers, and sound. TensorFlow is structured on data flow graphs, where mathematical operations are represented as nodes, facilitating the flow of data, or tensors. The potency of TensorFlow lies in its capacity to effectively distribute computations across multiple CPUs or GPUs, rendering it an optimal tool for the management of extensive neural networks.

Employing TensorFlow's Detection Model Zoo, we were able to expedite the model's training phase. The Detection Model Zoo offers an assortment of pre-trained models, each with unique detection speeds and mean average precisions, providing users with a wide spectrum of options. The model Single Shot Detector (SSD) MobileNet v2 FPNLite 320x320 was selected for this project due to its balance between speed and precision, achieving 22 ms per detection with a precision of 22.2 mAP. This makes it well-suited for devices with limited computational resources. The architectural model, SSD MobileNet v2, is a composite model, encompassing a MobileNet v2 network and an SSD layer, characterized by linear bottlenecks and shortcut connections. The SSD layer detects the object of interest using features derived from the MobileNet base network.

MobileNetV2 is a convolutional neural network architecture designed for mobile and embedded vision applications, characterized by its efficiency and lightweight design. The foundation of MobileNets is a streamlined architecture that creates lightweight deep neural networks using depth-wise separable convolutions. To cut down on computation time and parameters, MobileNet employs a Depthwise separable convolution rather than the standard convolution. It operates by first applying convolution to each channel of the image rather than as a block n times, and then obtaining n filters via 1x1 convolution.

Finetuning a pre-trained model like SSD MobileNet v2 offers numerous advantages compared to training a model from scratch. Firstly, it drastically reduces the demand for extensive datasets and training time, saving valuable resources. Secondly, it enhances model performance by leveraging the knowledge gained during pretraining.

After the training phase of the object detection model, it was integrated with the Jetson device using TensorFlow Lite. TensorFlow Lite enables on-device machine learning, offering a toolkit that assists developers in deploying their models on mobile, embedded, and edge devices. It aids users in converting and optimizing machine learning models to ensure efficient performance on limited-resource platforms, such as the Jetson device. As stated in the Challenges section, one of the issues faced during the integration of the object detection model with the Jetson device was the limited memory capacity. However, TensorFlow Lite allowed us to curtail the memory requirement of our model without compromising on performance or accuracy.

Results and Discussion

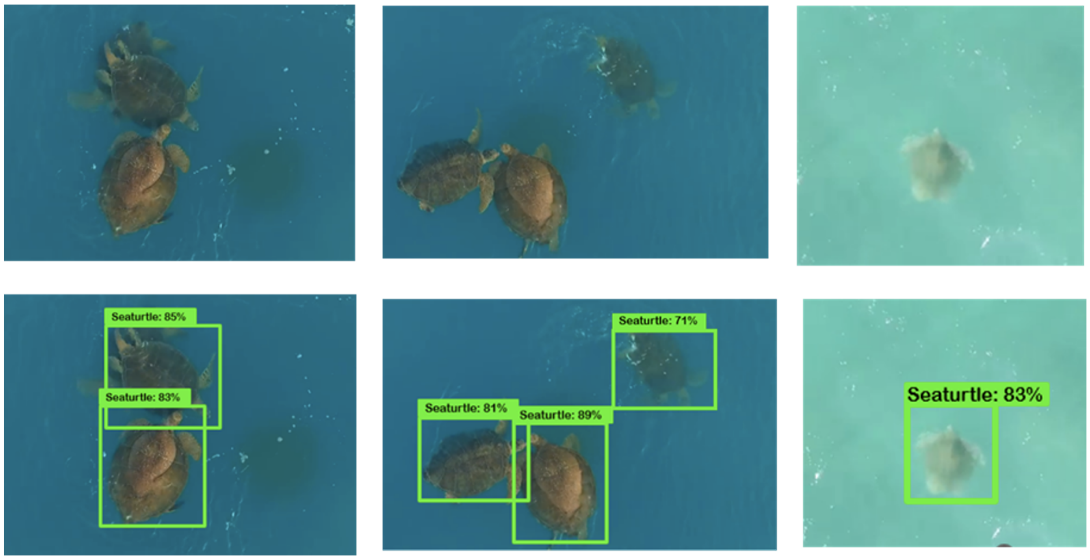
This section discusses the results acquired from the trained object detection model. Our initial dataset encompassed 300 images, all captured from a UAV's perspective. The code created for the training of this model separates the images into two sets: a training set that comprised 85% of the images, and a testing set that encompassed the remaining 15% of the images. While the most common practice when training a model is to follow the 70/30 or 80/20 rule, where 70% or 80% of the images are used for training and the remaining 30% or 20% of the images for testing, given the small size of our dataset we opted for an 85/15 split to obtain the best results. As delineated in the Development Process section, the object detection model was trained using TensorFlow and subsequently integrated and tested on a Jetson Nano. To ensure no turtle sighting would be missed during real-time processing, we set the minimum confidence threshold for our model at 50%. This value represents the lowest certainty level the model allows in its predictions, with any detection below this threshold being disregarded. This approach was essential for promptly identifying areas where turtles could potentially be found. When the neural network was run outside of real-time conditions, the threshold was increased to 75

Figure 3 presents three instances of the results derived after operating the object detection model. The images in Figure 3 (a) demonstrate the model's capability to detect turtles on the water's surface and in various positions. Figure 3 (b) showcases the model's ability to identify turtles at greater depths, while Figure 3 (c) depicts a similar detection scenario in a different habitat with an alternate color scheme.

Out of the 60 images set aside for testing, 49 were accurately identified, leading to an overall accuracy of approximately 81.67%, a F1-score of 81.77%, (which balances precision and recall), and a recall of 82.03% for our object detection model. Figure 4 exhibits a few examples of images that were incorrectly interpreted. Figure 4 (a) and (b) display images where turtles were entirely overlooked, while Figure 4 (c) presents a shark that was erroneously identified as a turtle.

Figure 3

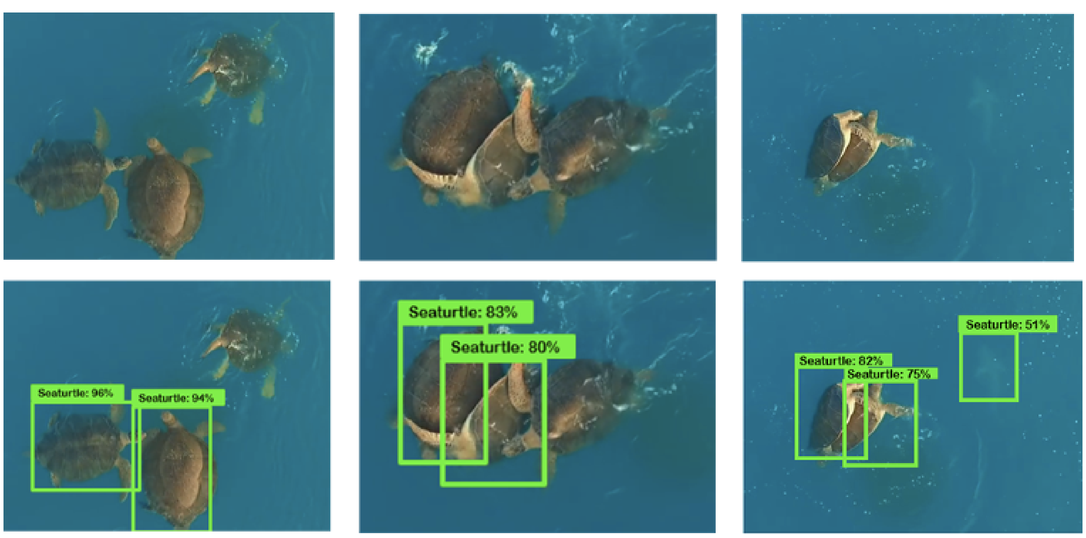
Object Detection Results: Correct Detections



(a) (b) (c)

Figure 4

Object Detection Results: Sea Turtles Not Detected



(a) (b) (c)

Despite the successful implementation and testing of our neural network with existing images, a significant limitation was the absence of sufficient original image data collected by our team. This was due to the inherent difficulties in locating and photographing sea turtles in their natural habitats. Despite numerous UAV flights aimed at collecting these images, we only managed to spot sea turtles on one occasion. Further intensifying this challenge was our inability to capture high-quality images due to constraints related to the camera equipment and other external factors. Consequently, the testing of the neural network was constrained mostly to images that we did not capture ourselves, which placed certain limitations on our ability to fine-tune and validate the model with in-the-field data.

As we continue our efforts and begin to amass a more substantial collection of original image data, we anticipate that the effectiveness and precision of our object detection model will be considerably improved. Looking forward, our intentions extend beyond just detecting the presence of sea turtles. We aspire to evolve our model to differentiate between sea turtle species and, optimistically, even identify individual turtles. This objective will further advance our understanding of sea turtle behaviors and distributions, contributing invaluable knowledge to the broader field of marine conservation.

Conclusion

Our research was able to uncover both significant opportunities and notable constraints. As highlighted in our discussions, the limited dataset that we used in this research significantly influenced the performance of our model. A comprehensive representation of the diverse habitats of sea turtles and the variations in their appearance due to lighting, shadows, and water reflections, unfortunately, fell short in our current dataset. This resulted in challenges regarding the model's consistency and reliability in turtle detection under varying conditions. It is important to recognize that these obstacles are not insurmountable. The nature of machine learning and neural networks is such that the accumulation of more images captured in various conditions and settings will substantially augment the training of our object detection model. This continued learning, driven by increasing volumes of high-quality data, will improve the model's robustness and predictive accuracy, thereby overcoming the limitations observed in the current study.

Looking ahead, our roadmap for the development of this project is to perform targeted enhancements and an expanded application scope. Foremost among these is the intent to further refine and optimize our detection model on the Jetson Xavier platform. By improving the efficiency of our algorithms and optimizing their interaction with the hardware, we aim to significantly reduce the overall processing time. This step is pivotal in increasing the model's operational feasibility in real-time applications, ultimately accelerating the pace at which conservation efforts can be informed and actioned.

Beyond this technical improvement, our ambitions also encompass broadening the horizons of our model's functionality. Our long-term goal is to expand the model's detection capability to include a diverse array of other aquatic and terrestrial species. By doing so, we aim to evolve our model from a focused tool for sea turtle conservation into a versatile asset for global wildlife conservation efforts. This extension would dramatically amplify the positive impact of our work, contributing invaluable insights to the protection and preservation of biodiversity across multiple ecosystems.

In conclusion, our research has laid the groundwork for a potentially transformative tool in wildlife conservation. While challenges remain, the possibilities for improvement and extension of this work are significant. As we continue to refine our model and widen its applications, we believe that our research will play a pivotal role in shaping innovative and effective conservation strategies for the future.

References

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., & Zheng, X. (2016). Tensorflow: Large-scale machine learning on heterogeneous distributed systems. *arXiv [Cs.DC]*. http://arxiv.org/abs/1603.04467

Adkins, K. A., Akbas, M., & Compere, M. (2020). Real-time urban weather observations for urban air mobility. *International Journal of Aviation, Aeronautics, and Aerospace*, *7*(4). https://doi.org/10.15394/ijaaa.2020.1540

Adkins, K. A., Akbas, M., & Compere, M. (2021). Real-time urban weather observations for aviation. *AIAA Aviation 2021 Forum*. https://doi.org/10.2514/6.2021-2359

Akbas, M. I., & Turgut, D. (2011). Apawsan: Actor positioning for aerial wireless sensor and actor networks. In T. Pfeifer, A. Jayasumana, & N. Aschenbruck (Eds.), *2011 ieee 36th conference on local computer networks* (pp. 563–570). https://doi.org/10.1109/LCN.2011.6115518

Anuntachai, A., & Pantuwong, N. (2019). An image-based sea turtle identification using postorbital facial feature points matching technique [[Paper presentation]]. *2019 19th International Conference on Control, Automation and Systems (ICCAS)*, 1058–1063. https://doi.org/10.23919/ICCAS47443.2019.8971623

Badawy, M., & Direkoglu, C. (2019). Sea turtle detection using faster r-cnn for conservation purpose. In R. Aliev, J. Kacprzyk, W. Pedrycz, M. Jamshidi, M. Babanli, & F. Sadikoglu (Eds.), *10th international conference on theory and application of soft computing, computing with words and perceptions - icsccw-2019 - advances in intelligent systems and computing* (pp. 535–541). Springer. https://doi.org/10.1007/978-3-030-35249-3_68

Cohen, J. E., Jonsson, T., & Carpenter, S. R. (2003). Ecological community description using the food web, species abundance, and body size. *Proceedings of the National Academy of Sciences of the United States of America*, *100*, 1781–1786. https://doi.org/10.1073/pnas.232715699

Coston-Clements, L., Waggett, R. J., & Tester, P. A. (2009). Chaetognaths of the united states south atlantic bight: Distribution, abundance and potential interactions with newly spawned larval fish. *Journal of Experimental Marine Biology and Ecology*, *373*(2), 111–123. https://doi.org/10.1016/j.jembe.2009.03.008

Donlan, C. J., Wingfield, D. K., Crowder, L. B., & Wilcox, C. (2010). Using expert opinion surveys to rank threats to endangered species: A case study with sea turtles. *Conservation Biology*, *24*, 1586–1595. https://doi.org/10.1111/j.1523-1739.2010.01541.x

Gope, C., Kehtarnavaz, N., Hillman, G., & Würsig, B. (2005). An affine invariant curve matching method for photo-identification of marine mammals. *Pattern Recognition*, *38*(1), 125–132. https://doi.org/10.1016/j.patcog.2004.06.005

Gray, P. C., Fleishman, A. B., Klein, D. J., McKown, M. W., Bézy, V. S., Lohmann, K. J., Johnston, D. W., & Gaggiotti, O. (2019). A convolutional neural network for detecting sea turtles in drone imagery. *Methods in Ecology and Evolution*, *10*(3), 345–355. https://doi.org/10.1111/2041-210X.13132

Hensel, E., Wenclawski, S., & Layman, C. A. (2018). Using a small, consumer grade drone to identify and count marine megafauna in shallow habitats. *Latin American Journal of Aquatic Research*, *46*(5), 1025–1033. https://doi.org/10.3856/vol46-issue5-fulltext-15

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

Landeo-Yauri, S., Ramos, E., Castelblanco-Martínez, D., Torres, C., & Searle, L. (2020). Using small drones to photo-identify antillean manatees: A novel method for monitoring an endangered marine mammal in the caribbean sea. *Endangered Species Research*, *41*, 79–90. https://doi.org/10.3354/esr01007

Muna, S. I., Mukherjee, S., Namuduri, K., Compere, M., Akbas, M. I., Molnár, P., & Subramanian, R. (2021). Air corridors: Concept, design, simulation, and rules of engagement. *Sensors*, *21*(22). https://doi.org/10.3390/s21227536

Neubauer, T., & Akbas, M. I. (2022). Sosuas: Stability optimized swarming for unmanned aerial systems [[Paper presentation]]. *2022 IEEE/AIAA 41st Digital Avionics Systems Conference (DASC)*, 1–6. https://doi.org/10.1109/dasc55683.2022.9925853

Sykora-Bodie, S. T., Bezy, V., Johnston, D. W., Newton, E., & Lohmann, K. J. (2017). Quantifying nearshore sea turtle densities: Applications of unmanned aerial systems for population assessments. *Scientific Reports*, *7*(1). https://doi.org/10.1038/s41598-017-17719-x