Research Article

International Journal of Advanced Technology and Engineering Exploration, Vol 11(115)
ISSN (Print): 2394-5443 ISSN (Online): 2394-7454
http://dx.doi.org/10.19101/IJATEE.2023.10102512

Efficient dorsal fin-based classification of Risso's and common Bottlenose dolphins using YOLOv7 and YOLOv8 models for real-time applications

Fawaghy Alhashmi[1], Maryam Alhefeiti[2], Shaher Bano Mirza[3*] and Fouad Lamghari Ridouane[4]
Research Engineer, Digital Products and AI Department, Fujairah Research Centre, Sakamkam road, Fujairah, United Arab Emirates[1]
Research Engineer, Marine Department, Fujairah Research Centre, Sakamkam road, Fujairah, United Arab Emirates[2]
Principal Investigator, Academics and Startup Department, Fujairah Research Centre, Sakamkam road, Fujairah, United Arab Emirates[3]
Director FRC, Fujairah Research Centre, Sakamkam Road, Fujairah, United Arab Emirates[4]

Received: 06-November-2023; Revised: 18-June-2024; Accepted: 21-June-2024

Abstract
The existence of whales and dolphins serves as a key sign of the well-being of the marine environment of that area. It is imperative to undertake research and conservation initiatives to safeguard these marine mammals and their ecosystem. This guarantees their persistence for the well-being of future generations. In recent years, marine surveys conducted in Fujairah offshore waters have generated valuable data concerning the distribution of cetacean species. Notably, common Bottlenose dolphins (Tursiops truncatus) and Risso's dolphins (Grampus griseus) have emerged as prevalent species in the region. These data hold significant information that is useful in species identification and its habitat loss mitigation efforts. Computer vision offers an efficient solution for analysing and interpreting vast visual data compared to of the manual detection methods. Therefore, the primary objective of this study is to assess and contrast the efficacy of you only look once version 7 (YOLOv7) and you only look once version 8 (YOLOv8) models in the identification of cetacean species. The findings indicate that both models exhibit strong performance in identifying and categorizing the desired species. Specifically, YOLOv8 demonstrates a slightly superior precision rate of 91.6% compared to YOLOv7. Additionally, YOLOv8 exhibits improved recall (92.5%) and mean average precision (mAP) of 95.9%. The improved performance of YOLOv8 can be attributed to its comprehensive feature map and optimised convolutional network, combined with a novel backbone network.

Keywords
Risso's dolphins, YOLOv7, YOLOv8, Marine, Common bottlenose dolphins.

1.Introduction
The Fujairah offshore water, which is situated within the Gulf of Oman in the United Arab Emirates, is well-known for its rich and diverse marine ecosystem. This area is home to more than 11 distinct species of cetaceans, making it a significant hub of cetacean biodiversity [1]. The presence of whales and dolphins in this region is a significant indicator of a healthy marine ecosystem [2]. Research and conservation efforts are crucial to protect these magnificent creatures and their habitats, ensuring their survival for future generations to appreciate and thrive.

The species Tursiops truncatus, also known as Bottlenose dolphins, is very prevalent in the offshore waters of Fujairah [1].

*Author for correspondence

These dolphins play a crucial role in fostering the richness of marine animals in this location. It has been observed that common Bottlenose dolphins occasionally form social groups that consist of many species, including Risso's dolphins (Grampus griseus) [3–7]. Moreover, this specie is known for its light grey to black dorsally with a light belly [8]. Additionally, their dorsal fins are characterized by a falcate shape and a broad base [8]. The warm temperature and abundant food sources in the waters of Fujairah make it an ideal habitat for these dolphins. Risso's dolphins (Grampus griseus) are frequently observed in the waters of Fujairah, a region known for its rich marine biodiversity. These dolphins are easily recognizable due to their unique physical characteristics, including a robust body shape, a large dorsal fin, and a prominent melon on their forehead, these dolphins may have white or pink patches on their undersides, which are attributed to their interactions with conspecifics [9]. They also exhibit specific behavioral patterns, such as leaping out of the water and riding the bow waves created by boats. The ecological significance of Risso's dolphins lies in their role as top predators in the marine ecosystem. They primarily feed on squid, which helps to regulate the population of these cephalopods and maintain a balanced food web. Additionally, their presence in Fujairah's waters indicates a healthy and thriving marine environment, as these dolphins are known to prefer areas with abundant [1]. Furthermore, their skin often bears numerous scars and scratches, further attesting to their engagement with other dolphins [9]. The exponential increase in anthropogenic marine activities has resulted in oceans becoming increasingly crowded with vessels. This surge poses significant challenges for the management, protection, and conservation of marine mammal biodiversity. Accurate species identification and consistent monitoring of marine life and ecosystems are paramount to addressing these challenges effectively. However, the diverse nature of marine mammals complicates their classification, presenting a formidable obstacle in our efforts to safeguard their populations and habitats. In Fujairah, various measures are being undertaken to monitor and protect the well-being of marine animals. Manual identification of dolphins is a tedious task that requires expert vision and significant labour. Despite the effort, there are still chances of errors or missed identifications due to human limitations. Utilizing computer vision can be an optimal solution, offering robust and accurate identification of dolphins.

Computer vision is a new technology in the artificial intelligence (AI) field that focuses on the identification of objects and the extraction and evaluation of quantitative data from digital images [10]. It has been widely applied across several disciplines for various purposes, including marine ecosystems [11]. Computer vision systems can process large volumes of data quickly, consistently apply identification criteria without fatigue, and learn to improve accuracy over time with machine learning techniques. Marine applications and studies pertaining to computer vision encompass the utilization of visual data analysis techniques for the purpose of animal detection, classification, tracking, and segmentation. These techniques are specifically applied to organisms found in marine environments, including but not limited to plankton, fish, prawns, oysters, marine mammals, and pollution [11]. This technological approach not only enhances efficiency but also minimizes the likelihood of errors, leading to more reliable and scalable dolphin identification. In this research, our attention is directed towards the utilization of computer vision techniques in the classification of images depicting common Bottlenose dolphins and Risso's dolphins. The study focuses on the significance of species identification in the context of mitigating habitat loss. To achieve this objective, computer vision techniques, specifically you only look once version 7 (YOLOv7) and you only look once version 8 (YOLOv8) models, are employed for the purpose of identifying common Bottlenose dolphins and Risso's dolphins through image analysis. Our focus is on how well several versions of you only look once (YOLO) models (YOLOv7 and YOLOv8) perform in terms of classification accuracy when trained on a labelled photo set. Furthermore, in order to evaluate the precision levels of these models and make a comparison of their effectiveness in identifying dolphin species.

Computer vision on dolphin image data introduces many challenges which, if not considered, can affect its application, the challenges include:

1. Many researchers train models using a large number of labelled photo set [12–14] However, training a large photo set is impractical for the majority of smaller-scale research projects like Fujairah dolphins' project.
2. Environmental challenges include varying weather conditions in the photographs, such as heavy rain or harsh lighting/shadows. One possible solution to alleviate these concerns is consider the photo taken during optimal weather conditions to Obtain high-quality datasets, that optimize the YOLOv7 and YOLOv8 modules for accurate classification and addressing real-time processing constraints [15].

The YOLO algorithm is widely acknowledged in the field of computer vision as a highly efficient object detection model. It has garnered significant attention and popularity due to its exceptional performance in accurately identifying objects within an image. The algorithm is particularly renowned for its remarkable inference speed and impressive accuracy [16, 17]. YOLO is available in eight versions, with the most recent being YOLOv8. A new feature in YOLOv7, spatially adaptive denormalization (SPADE) enables the model to modify the normalisation parameters for each feature map, thereby increasing the model's robustness to varying lightening conditions [17, 18]. The YOLOv8 model was specifically developed to incorporate the most efficient characteristics from various real-time object detection algorithms. Its purpose is to seamlessly work with and transition between all iterations of the YOLO framework. This allows researchers engaged in YOLO projects to easily assess and compare the performance of their models [19]. For the following reasons, YOLOv8 is the most extensively used real-time object detector: (a) a lightweight network architecture; (b) efficient feature fusion techniques; and (c) more precise detection outcomes [19].

Based on the importance of species identification in habitat loss mitigation efforts, The study aims to utilise computer vision techniques, specifically YOLOv7 and YOLOv8 models, to identify common Bottlenose dolphins and Risso's dolphins based on image analysis. Moreover, to assess the accuracy levels of these models and compare their performance in dolphin species identification.

The contribution of this study lies in the assessment and comparison of the YOLOv7 and YOLOv8 models for identifying and categorizing common Bottlenose dolphins and Risso's dolphins in the Fujairah offshore waters. The findings demonstrate that both models perform well, with YOLOv8 showing slightly superior precision, recall(R), and mean average precision (mAP). The study highlights the importance of utilizing advanced object detection models like YOLOv8 in marine surveys, specifically for real-time applications, due to its accelerated detection speed without compromising accuracy. The emphasis on practical implementation and the acknowledgment of YOLOv8's suitability for habitat loss mitigation indicate a valuable step forward in leveraging technology for marine species identification. Next, section 2 provides a comprehensive literature review. Then, section 3 explores into the dataset and research methods adopted to carry out the study. Section 4 presents the results of this work, followed by a discussion of the investigated contents in section 5. Finally, the paper draws the main concluding remarks, and future research possibilities.

2. Literature review

Marine mammals exhibit a global distribution, inhabiting both coastal and offshore waters [20, 21]. This taxonomic group encompasses several species such as whales, dolphins, seals, and sea lions [22]. The substantial rise in human marine activities has led to our oceans being densely populated with vessels. These vessels can be detected from distances extending up to 20 kilometres. Previous studies have revealed that this intense presence of vessels has a profound impact on the behaviour of marine animals, with dolphins being especially affected. The constant activity and movement of these ships create an environment where the natural behaviours of these creatures are disrupted, leading to changes in their communication, navigation, and social interactions. The long-term consequences of this pervasive human presence in marine habitats highlight the need for further research and potential measures to mitigate these effects on marine life [23]. In order to effectively manage, protect, and mitigate the decrease in marine mammal biodiversity, it is important to accurately identify the species and consistently monitor the marine life and ecosystem [24].

A recent study off the coast of Fujairah, in the Gulf of Oman, has revealed a rich diversity of cetaceans with 11 or more species now known to occur in the emirate [1]. The diversity of marine mammals makes their classification a challenge [24].

Understanding the impact of environmental and biological factors on species' habitat use is crucial for conservation efforts. In the context of dolphin research, the utilization of machine learning techniques has emerged as a valuable tool for identification and detection tasks. A notable study focusing on the common Bottlenose dolphins (Tursiops truncatus) in the Western Ligurian Sea exemplifies this trend [25]. Despite the frequent presence of these dolphins in the region, their distribution, habitat preferences, and social dynamics remain inadequately understood. This investigation aimed to address these knowledge gaps by examining the spatial distribution and habitat partitioning across different age classes of dolphins. To model the distribution of each group type, researchers employed an ensemble modelling approach known as BIOMOD. This approach underscores the efficacy of machine learning methodologies, particularly ensemble modelling, in providing valuable insights into dolphin behaviour and habitat utilization [25]. The availability of vast image collections using camera traps and crowdsourcing offers unique opportunities for the monitoring and conservation of wildlife. This development underscores the need for automated methods to analyse these images, particularly for the purpose of re-identifying individual animals. Current techniques for re-identification typically rely on two main approaches: manually crafted local features and end-to-end learning of fur pattern similarities. Manually crafted features do not require labelled training data, making them useful in scenarios where such data is scarce. On the other hand, end-to-end learning methods, although requiring a substantial amount of labelled data, tend to deliver superior performance when enough training data is accessible.

The balance between these two approaches depends on the availability of labelled data and the specific requirements of the monitoring and conservation tasks at hand. This highlights the importance of continued research and development in automated image analysis methods to enhance wildlife conservation efforts [26].

Deep learning models are more suitable for image detection and extraction in challenging environments, as well as the ability to work with a larger amount of data at the same time [16]. Researchers have recently introduced a content-based image retrieval (CBIR) inspired pipeline for identifying individual animals, named aggregated local features for Re-Identification (ALFRE-ID). This approach incorporates interchangeable learned local features, feature aggregation, and feature embeddings to overcome the limitations of existing methods, particularly when dealing with small, labelled datasets [26]. In the field of computer vision, the YOLO model has garnered considerable attention. Researchers have enhanced and added new elements to the architecture, leading to the creation of several traditional models. YOLOv7 is an object detection algorithm, that achieved the highest accuracy among all other real-time object detection models. A study by Kuhlane and the team shows that YOLOv7 outperforms YOLOX, Scaled-YOLOv7, YOLOv5, DEtection TRansformer (DETR), vision transformer adapter (ViT Adapter-B), and many more object detection algorithms in speed and accuracy by 2% higher accuracy at a dramatically increased inference speed (509% faster). Released by Ultralytics in 2023, YOLOv8 marks a significant advancement in this evolution. Compared to earlier versions like YOLOv5 and YOLOv7, YOLOv8 sets the benchmark for detection accuracy. The network architecture of YOLOv8 comprises three main components: the neck, backbone, and head [27].

YOLOv8 incorporates the most efficient characteristics from many real-time object detectors. It is specifically built to be compatible with and seamlessly transition between all versions of YOLO. This allows researchers working on YOLO projects to simply compare their performance [19].

One of the studies introduced a new real-world environmental dataset as a benchmark suit consisting of two large-scale datasets, DeepFish and OzFish. To address the shortcomings of YOLOv3, the paper offers a strong fish detection model that was trained on the combined dataset. However, the module is unable to process the feature information of small objects in the dataset during training and is unable to distinguish fish in the test from a complex sea background. In the study, YOLOv1 and YOLOv2 was able to detect fish in unrestricted real-world marine environments with an average precision of 76.56% and 75.70%, respectively [28].

An underwater fish detection system utilizing deep convolutional neural networks (CNNs) was presented in one of the studies by a group of researchers [29]. YOLOv3 was one of the CNNs used in this system. The study’s primary contribution was the proposal of an automated underwater fish detection system for use in underwater photos and videos. The system was trained and assessed by the authors using a dataset of photos and videos of fish. The study's method for identifying fish in underwater photos and videos demonstrated excellent real-time performance and accuracy. Additionally, YOLOv3 was found to outperform other cutting-edge object detection techniques in terms of both accuracy and computational efficiency in a comparison conducted in the research.

Using underwater videos and the YOLOv2 object detection algorithm, Wang and the team described a real-time fish tracking and detection method [30]. YOLOv2 was pre-trained using a dataset of annotated fish photos, and it was subsequently refined using a dataset of underwater video frames. The model demonstrated a high degree of accuracy in fish detection and tracking when the authors assessed its performance on a test dataset of underwater videos. Because the model could operate in real-time on a standard computer, it could be applied to a variety of underwater tasks, including behaviour research and fish population monitoring [30].

The overall literature review underscores the disruptive impact of increased vessel activity on marine mammals, particularly dolphins, highlighting the need for effective mitigation strategies. Although significant advancements in species identification and ecosystem monitoring have been achieved using machine learning techniques like CBIR, ALFRE-ID, DETR, CNN, and earlier versions of YOLO, there remain gaps in understanding the detailed habitat preferences and social dynamics of dolphins. This study aims to utilize the latest versions of YOLO, specifically YOLOv7 and YOLOv8, to address these gaps. By employing these advanced models, the research seeks to enhance the accuracy and comprehensiveness of dolphin identification, which later may help in providing deeper insights into their behavior and habitat utilization. This approach is expected to significantly improve the identification and consequently monitoring and conservation of dolphin populations, contributing to the broader effort of marine biodiversity protection.

3. Materials and methods

Study area and field surveys

The study area includes whole of Fujairah waters including offshore waters up to the midline of the Gulf of Oman, the survey plan, involving transect lines between inshore (in) start points and offshore (out) endpoints (*Figure 1*). The focus of the survey was on offshore waters, with the start points of transects positioned approximately 20kms from shore in water depths of 80-120m and end points reaching as far as 80kms from shore in water depths of well over 1,000m. During the period spanning from 2017 to 2023, a comprehensive set of 16 boat-based surveys were undertaken to effectively cover the designated systematic transect lines, as seen in *Figure 1*. These transect lines included a combined area of 2219 square kilometers. The duration of each survey varied between four and five days, contingent upon prevailing sea conditions. The survey served the purpose of extensive study based on regional objectives. Within the scope of our research, the task of gathering dolphin photographs was carried out by a team of skilled observers with diverse backgrounds. These observers underwent training to ensure their proficiency in systematically scanning the designated area to locate cetaceans.

Figure 1 Study area location of Fujairah offshore water and illustrating transect lines followed during vessel surveys from 2017 to 2023. The predetermined route is the route followed during vessel surveys, divided by inshore ("IN") and offshore ("OUT) locations at specific start and end points (track waypoint)

Data collection and annotation

A comprehensive collection of 354 images documenting Risso's dolphins (Grampus griseus) and Common Bottlenose dolphins (Tursiops truncates) was conducted using a Canon EOS 5D Mark IV camera paired with a Canon Lens 70-200mm [13, 17]. All the observation and data collection has been done by professional observers. Standardized protocols were established for dolphin sighting and photograph collection to ensure consistency and reproducibility across survey efforts. These protocols outlined specific procedures for locating dolphin groups, approaching them without causing disturbance, and capturing high-quality images while minimizing potential biases. Emphasis was placed on maintaining a safe distance from the animals to avoid disrupting natural behaviors and ensuring the ethical treatment of marine wildlife. Each image was accurately named according to the date, time, and photographer, ensuring thorough documentation of the survey data. These images were carefully stored following standardized procedures on an external hard drive, preserving their integrity and accessibility for future analysis. For instance, images were stored in folders organized by year, with each folder representing a specific survey term. Within these folders, each photo was named in a date-wise manner. This systematic organization facilitated easy retrieval and analysis of images, ensuring that researchers could quickly locate certain record across different surveys and time periods. Further these images are stored on our online server and database. The metadata of each image contained maximum details about the survey, dates, photographer and more. During the data curation process, the image dataset underwent thorough cleaning and manual filtering to ensure high quality. Factors such as resolution and angle of the dorsal fin were carefully examined, resulting in a refined dataset comprising 224 images that met strict quality criteria.

To facilitate species annotation, robust image labeling was conducted using the advanced Roboflow software (https://roboflow.com/) [19]. This powerful tool enabled manual labeling of dorsal fins, a crucial step in species identification. The dorsal fins were accurately categorized into two classes: Grampus griseus and Tursiops truncates, ensuring precise classification and annotation of each image [1]. Following annotation, the dataset underwent further processing steps using the roboflow platform. This included data augmentation techniques to enhance model robustness and generalizability, such as rotation, flipping, and scaling. Data augmentation is a crucial step in enhancing the robustness and generalizability of machine learning models for image analysis tasks such as species identification. This process involves generating new, altered versions of the original images by applying a variety of transformations and modifications. These transformations can include rotation, flipping, scaling, translation, shearing, and changes in brightness and contrast. By introducing these variations, data augmentation effectively increases the diversity of the training dataset, allowing the model to learn from a wider range of scenarios and conditions. For example, rotation involves rotating the image by a certain angle, which helps the model become invariant to the orientation of the object being detected. Flipping horizontally or vertically creates mirrored versions of the images, further expanding the dataset without changing the object's characteristics. Scaling alters the size of the objects within the image, simulating different distances or perspectives, while translation shifts the object's position within the frame, mimicking variations in camera angles or focal points. Shearing distorts the shape of the object, simulating different viewpoints or perspectives. Additionally, changes in brightness and contrast simulate variations in lighting conditions, ensuring that the model can accurately detect objects under different lighting environments. This process ultimately improves the model's performance and reliability. Additionally, advanced preprocessing methods were employed to standardize image dimensions and optimize feature extraction, further improving the accuracy of subsequent analysis. *Table 1* shows a partial image of the dataset for the most sighted dolphins in the surveys. Sample images of the dataset are shown in *Figure 2*.

Figure 2 Training batch sample with annotation (0) presents the Risso's dolphins' class and (1) present the common Bottlenose dolphin

Data Enhancement

Training a deep learning model requires a significant amount of data for training and validation. This is crucial for the model to efficiently extract features and gain knowledge. Using various techniques, efficient to expand the dataset by applying producer enhancement, pre-processing with Auto-Orient features, and Augmentation to generate new training examples. This method enables the model to gain knowledge from enhanced versions of every image in the training set. Modifying the images includes adjusting their saturation, exposure, and introducing noise within a range of -25% to +25%. Consequently, the size of the dataset has tripled (*Figure 3*). After labelling 224 images, the augmentation method increased the number of training set images to 468. The validation set contained 23 images, and the testing set comprised 45 images. A split rate of 87%, 4%, and 8% is employed for the purposes of training, validation, and testing, correspondingly. *Figure 2* illustrates enhanced images accompanied by their corresponding transform parameters. The annotation files were exported as text files, following the precise format needed for the YOLO algorithm.

Table 1 shows the most sighted dolphins in the surveys from 2016- 2022

Type of dolphin	Scientific name	Dorsal fin categorization	Data set	Sample from data
Risso's Dolphin	*Grampus griseus*	Long, and thin Heavily scarred	104	
Common Bottlenose dolphin	*Tursipos truncates.*	Short and Broad Falcate and board at base Pointed pectoral fins	113	

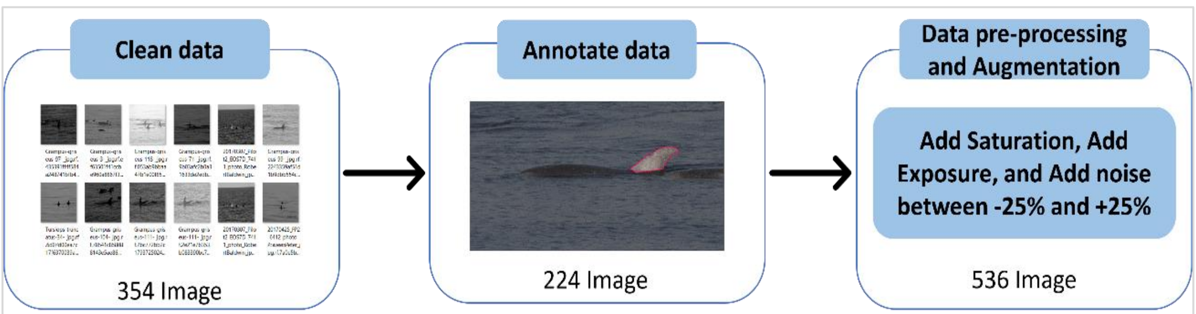


Figure 3 The process of cleaning, annotating data and applying augmentation

Experimental environment and parameter adjustment

The experimental operating system used in this study is the Colab environment, with PyTorch serving as the framework for the deep learning models created. The impressive computing power of the Intel Xeon CPU @2.20 GHz and NVIDIA® Tesla T4 have been utilized. The model underwent training for 30 epochs using a batch size of 16. It utilised an initial learning rate of 0.01, a momentum factor of 0.973, and a weight decay of 0.0005[11].

Model Evaluation Indicators This study utilized precision (P), recall (R), and mAP as accuracy evaluation indicators. Precision-Recall Curve (P-R Curve) is a curve with recall as the x-axis and precision as the y-axis Equation 1 and Equation 2 are used to calculate the values of R and P. In these equations, true positive (TP) represents a correctly predicted positive class, false positive (FP) represents an incorrectly predicted positive class, and false negative (FN) represents an incorrectly predicted negative class, as shown in Equations 1, 2, and 3.

$$Recall = \frac{Tp}{Tp+FN} \quad (1)$$

$$Percision = \frac{Tp}{Tp+Fp} \quad (2)$$

The F-score is a commonly used metric to evaluate the model accuracy, providing a balanced measure of performance, F1 calculated according to Equation 3

$$F_1 = \frac{2 \times Percision \times Recall}{Percision + Recall} \quad (3)$$

4. Result

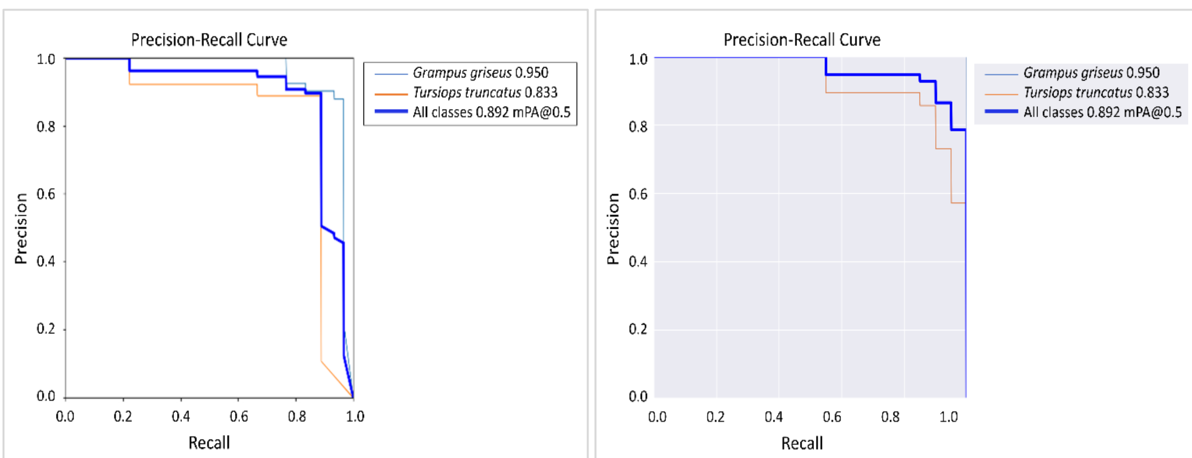
Among 338 sightings of cetacean groups, 139 common Bottlenose dolphin groups were recorded. Common Bottlenose dolphins were normally sighted among groups in which group sizes ranged from 12 to 130 individuals. Most records were in the southeastern part of the survey area in waters ranging from 92 to 1,218 m in depth and at distances of 24 to 38 nm (numerical mile) offshore. The sightings records occurred between February 2021 and May 2023.

Among the 338 observations of cetacean groups, a total of 31 individuals belonging to the Risso's dolphin species were documented. Risso's dolphins were often observed in aggregations characterised by group sizes ranging from 2 to 50 individuals. Most records were in the southeastern region of the survey area, namely in seas with depths ranging from 142 to 1,173 metres. These records were found at distances of 24 to 38 nautical miles offshore. The 31 instances of sightings were documented over the time frame spanning from April 2017 to June 2022.

Following the training part both models YOLOv7 and YOLOv8 shows excellent results numerically and visualization, both models can classify the dolphin species in both images and videos format. It can be observed from *Table 2* the performance of YOLOv7 instances segmentation and yolov8 standard instances segmentation. Its show that the average precision is the standard YOLOv8 (91.6%) with impeccable visualization results the best mAP of the models where YOLOv7 with closely precision of (89.0%). Comparing the precision of the two models, the results from YOLOv8 standard for segmentation show the best precision on the P-R curve. *Figure 4(a)* displays the results of YOLOv7, which shows lower precision across all classes at 89.2%. On the other hand, *Figure 4(b)* indicates that YOLOv8 achieves higher precision of 95.8% across all classes.

Table 2 Result of segmentation models considering precision, recall and mAP

Model	Precision	Recall	mAP_0.5	mAP_0.5:0.95
YOLOv7	0.89002	0.91111	0.89176	0.71819
OLOv8	**0.916**	**0.925**	**0.95906**	**0.788**



(a) (b)

Figure 4 The precision-confidence curve of (a) YOLOv7 and (b) YOLOv8

In *Figure 5*, a sample prediction from the test set is illustrated. *Figure 5* depicts the visual detection and segmentation outcomes of the model on the testing dataset. The species type can be deduced from the dorsal fin images by the model. Instances segmentation has the capability to identify the type based on the fin shape, which possesses various distinguishing characteristics. The trained model accurately identifies the Grampus griseus dolphin based on its fluke shape and classify it as such, as shown in *Figure 5a*. *Figure 5b* illustrates the process of identifying Tursiops turncatus. Additionally, *Figure 6* illustrates the confusion matrix, which provides insights into the trained model's classification performance.

(a) (b)

Figure 5 Testing examples of dolphin classification (a) trained model accurately identifies the Grampus griseus dolphin based on its fluke shape (b) identifying Tursiops turncatus

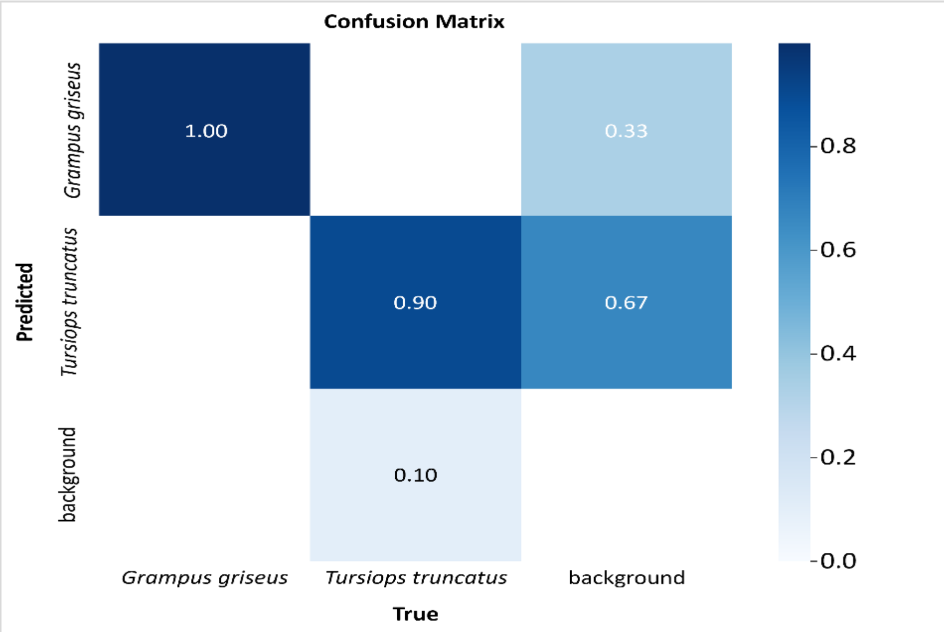


Figure 6 Normalised confusion matrix

5. Discussion

The findings of this study provide valuable insights into various important aspects of the distribution of cetaceans and the effectiveness of machine learning models in classifying dolphin species. The results indicate a notable presence of common Bottlenose dolphins, with group sizes varying from 12 to 130 individuals. These dolphins are mainly found in around the Fujairah coast. This indicates a significant persistence in occurrence of this species in particular habitats throughout the Fujairah coastal waters over the duration of the study. however, extended studies are required to understand the residence behavior of these dolphins. In the investigated region, Risso's dolphins were often seen in groups, mostly in the southeastern area although at slightly greater depths. This suggests that these dolphins may have specific habitat preferences or ecological niches within this region.

The assessment of the effectiveness of YOLOv7 and YOLOv8 machine learning models in classifying dolphin species showed promising results. Both models had excellent precision rates, with YOLOv8 exhibiting slightly improved performance compared to YOLOv7 across multiple evaluation criteria. YOLOv8 had superior performance in terms of average precision and consistently outperformed other methods in precision-recall curves. This highlights its robustness and effectiveness in accurately identifying dolphin species from photos and videos in the Fujairah coastal area. The results indicate the potential of advanced machine learning approaches in improving our capacity to precisely categories and monitor marine animal populations, a vital aspect of successful conservation endeavors.

The study's visual detection and segmentation results provide further confirmation of the effectiveness of the trained models in identifying dolphin species based on specific morphological characteristics, such as the form of the dorsal fin. The models' capacity to effectively distinguish between Risso's dolphins and common Bottlenose dolphins, based on fin traits, highlights their potential for species-specific identification in field surveys and monitoring programs throughout the Fujairah coast. Furthermore, the confusion matrix offers valuable insights into the categorization performance of the models, providing a quantitative assessment of their accuracy and identifying potential areas for development.

This study enhances our comprehension of the distribution patterns of cetaceans along the Fujairah coast and emphasizes the effectiveness of machine learning methods in identifying species and promoting conservation in this area. Future research ought to focus on the improvement and streamlining of machine learning models for the purpose of monitoring and surveilling marine mammal populations in the coastal waters of Fujairah. This will ultimately assist in the implementation of specific conservation measures aimed at protecting these important species and their habitats.

Limitation and suggestions

The study's limitation regarding the relatively small sample size of cetacean sightings, particularly for Risso's dolphins, could have impacted the accuracy and representativeness of the findings. To mitigate this, extending monitoring periods could increase the sample size, providing a more comprehensive understanding of cetacean distribution patterns along the Fujairah coast. Additionally, while machine learning models such as YOLOv7 and YOLOv8 showed promising results in species classification, there is room for improvement in accuracy and efficiency. Fine-tuning these models, as suggested in one of the studies, could enhance their performance, especially in distinguishing between closely related species or challenging environmental conditions [29]. Furthermore, incorporating additional environmental variables like oceanographic parameters or prey availability could enhance species distribution predictions and deepen our understanding of ecological drivers shaping cetacean habitat use in the area. Moreover, the study's focus on visual detection and segmentation based on dorsal fin characteristics may overlook other indicators of species' identity, such as vocalizations or behavior patterns. Integrating acoustic monitoring techniques or behavioral observations, as suggested in previous research, could complement visual data and enrich our understanding of cetacean communities in the Fujairah coast. Combining various approaches, such as the CBIR-motivated pipeline ALFRE-ID and ensemble modeling approach BIOMOD, as demonstrated in different studies, could address these limitations stepwise, enhancing the study's methodology and outcomes [23,25,26].

Since Fujairah is the fifth largest emirate in the United Arab Emirates, boasts diverse terrains and marine environments, making it a notable tourist destination regionally and internationally. It is home to various natural resources, among which coastal waters stand out as particularly vital, contributing significantly to Fujairah's sustenance while remaining highly vulnerable to environmental impacts. As the second largest bunkering port globally, Fujairah's strategically advantageous location poses significant environmental challenges, primarily due to increased maritime traffic, elevating the risk of oil spill incidents. Additionally, the coastal ecosystem faces various threats, including the direct discharge of brines from desalination facilities, the proliferation of harmful algae blooms, and industrial outfalls. These environmental concerns make Fujairah's coast a compelling area for research, as preserving these coastal resources is crucial for Fujairah's ecological health and economic growth. Without concerted conservation efforts, Fujairah's status and future aspirations may face jeopardy [31]. Considering the dynamic nature of marine ecosystems and the potential impacts of anthropogenic activities, longitudinal studies and regular monitoring efforts are crucial. These efforts would track changes in distribution patterns and assess the effectiveness of conservation measures over time, ensuring the continued protection of cetacean populations along the Fujairah coast. By carefully integrating these suggestions and methodologies from relevant studies, future research can overcome limitations and contribute to a more comprehensive understanding of cetacean ecology and conservation in the region. The result of all class is shown in *Appendix I*. A complete list of abbreviations is listed in *Appendix II.*

6. Conclusion and future work

In this study, the YOLOv7 and YOLOv8 models were assessed for their effectiveness in classifying dolphin species (common Bottlenose dolphins and Risso's dolphins), with promising results. Both models demonstrated excellent precision rates of 89.2%, with YOLOv7 and 95.8% with YOLOv8 at all classes, with YOLOv8 exhibiting superior performance across multiple evaluation criteria. This highlights YOLOv8's robustness and effectiveness in accurately identifying dolphin species from photos and videos in the Fujairah coastal area. The study's visual detection and segmentation results further confirmed the models' effectiveness in identifying dolphin species based on specific morphological characteristics, such as the dorsal fin shape. The models' ability to distinguish between Risso's dolphins and common Bottlenose dolphins based on these traits underscores their potential for species-specific identification in field surveys and monitoring programs throughout the Fujairah coast.

Moreover, the confusion matrix provided valuable insights into the models' categorization performance, offering a quantitative assessment of their accuracy and highlighting potential areas for improvement. The results indicate the significant potential of advanced machine learning approaches, particularly YOLOv8, in enhancing our capacity to precisely categorize and monitor marine animal populations, a vital aspect of successful conservation endeavors.

While this study employed the basic version of YOLOv8, future investigations can explore alternative frame sizes, such as YOLOv8n (nano) and larger iterations of YOLOv8, to cater to a diverse range of conservation needs. These developments underscore the versatility of artificial intelligence (AI)-driven models in ecological research and conservation, opening doors for more sophisticated ecological analysis, real-time applications, and the handling of larger datasets. Future work could focus on increasing the dataset size to further improve accuracy and developing real-time applications for differentiating closely related marine species. The ongoing surveys and data collection efforts demonstrate the value of continued model improvement, showing significant potential for further enhancing the capabilities of these models with new data.

Acknowledgment

The authors acknowledge that Five Oceans Environmental Services (5OES) and Balázs Buzás are a partner to run the whale and dolphin project in the field survey.

Conflicts of interest

The authors have no conflicts of interest to declare.

Data availability

The data utilized in this research paper is available upon request from the corresponding author. However, due to the sensitive nature of the data, it cannot be made publicly accessible. Please contact the corresponding author directly for access to the dataset used in this study.

Author’s contribution statement

Fawaghy Alhashmi: Data curation, analysis, and interpretation of results, investigation, methodology, software, writing – original draft. **Maryam Alhefeiti:** Data collection, Investigation, methodology, writing – original draft. **Shaher Bano Mirza:** Conceptualization, validation, writing – review & editing. **Fouad Lamghari Ridouane:** Conceptualization, project administration, resources, supervision, validation.

References

[1] Baldwin R, Willson A, Looker E, Buzás B. Growing knowledge of cetacean fauna in the Emirate of Fujairah, UAE. Tribulus. 2018; 26:32-42.

[2] Herrera I, Carrillo M, Cosme DEM, Haroun R. Distribution of cetaceans in the Canary Islands (Northeast Atlantic Ocean): implications for the natura 2000 network and future conservation measures. Frontiers in Marine Science. 2021; 8:669790.

[3] Scott MD, Chivers SJ. Distribution and herd structure of bottlenose dolphins in the eastern tropical Pacific Ocean. The Bottlenose Dolphin. 1990:387-402.

[4] Bearzi M. Dolphin sympatric ecology. Marine Biology Research. 2005; 1(3):165-75.

[5] Maze-foley K, Mullin KD. Cetaceans of the oceanic northern Gulf of Mexico: distributions, group sizes and interspecific associations. Journal of Cetacean Research and Management. 2006; 8(2):203-13.

[6] Weir CR. Distribution and seasonality of cetaceans in tropical waters between Angola and the Gulf of Guinea. African Journal of Marine Science. 2011; 33(1):1-5.

[7] Bacon CE, Smultea MA, Fertl D, Würsig B, Burgess EA, Hawks-johnson S. Mixed species associations of marine mammals in the Southern California Bight, with emphasis on Risso’s dolphins (Grampus Griseus). Aquatic Mammals. 2017; 43(2):177-84.

[8] Wells RS, Scott MD. Bottlenose dolphin, Tursiops truncatus, common bottlenose dolphin. In Encyclopedia of Marine Mammals 2018 (pp. 118-25). Academic Press.

[9] Perrin WF, Würsig B, Thewissen JG. Encyclopedia of marine mammals. Academic Press; 2009.

[10] Hartman KL. Risso’s dolphin: Grampus Griseus. In encyclopedia of marine mammals 2018 (pp. 824-7). Academic Press.

[11] Hong HHH, Yang XYX, You ZYZ, Cheng FCF. Visual quality detection of aquatic products using machine vision. Aquacultural Engineering. 2014; 63:62-71.

[12] Gümüş B, Balaban MÖ, Ünlüsayin M. Machine vision applications to aquatic foods: a review. Turkish Journal of Fisheries and Aquatic Sciences. 2011; 11(1).

[13] Norouzzadeh MS, Nguyen A, Kosmala M, Swanson A, Palmer MS, Packer C, et al. Automatically identifying, counting, and describing wild animals in camera-trap images with deep learning. Proceedings of the National Academy of Sciences. 2018 ;115(25): E5716-25.

[14] Willi M, Pitman RT, Cardoso AW, Locke C, Swanson A, Boyer A, et al. Identifying animal species in camera trap images using deep learning and citizen science. Methods in Ecology and Evolution. 2019; 10(1):80-91.

[15] Tabak MA, Norouzzadeh MS, Wolfson DW, Sweeney SJ, Vercauteren KC, Snow NP, et al. Machine learning to classify animal species in camera trap images: applications in ecology. Methods in Ecology and Evolution. 2019; 10(4):585-90.

[16] Kuhlane L, Brown D, Brown A. Exploring the incremental improvements of YOLOv7 on bull sharks in Mozambique. Master Thesis, Rhodes University, Grahamstown, South Africa. 2023.

[17] Upulie HD, Kuganandamurthy L. Real-time object detection using YOLO: a review. real-time object detection using YOLO: a review. Academia.edu 2021 (pp. 1-6).

[18] Wang CY, Bochkovskiy A, Liao HY. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. In proceedings of the IEEE/CVF conference on computer vision and pattern recognition 2023 (pp. 7464-75). IEEE.

[19] Lou H, Duan X, Guo J, Liu H, Gu J, Bi L, et al. DC-YOLOv8: small-size object detection algorithm based on camera sensor. Electronics. 2023; 12(10):1-14.

[20] Richards VP, Greig TW, Fair PA, Mcculloch SD, Politz C, Natoli A, et al. Patterns of population structure for

inshore bottlenose dolphins along the eastern United States. Journal of Heredity. 2013; 104(6):765-78.
[21] Wells RS, Rhinehart HL, Hansen LJ, Sweeney JC, Townsend FI, Stone R, et al. Bottlenose dolphins as marine ecosystem sentinels: developing a health monitoring system. EcoHealth. 2004; 1:246-54.
[22] Natoli A, Al HS. Marine mammals of the emirates: whales, dolphins, porpoises and dugongs. In a natural history of the emirates 2023 (pp. 557-84). Cham: Springer Nature Switzerland.
[23] Diamant R, Testolin A, Shachar I, Galili O, Scheinin A. Observational study on the non-linear response of dolphins to the presence of vessels. Scientific Reports. 2024; 14(1):6062.
[24] Berta A. Return to the sea: the life and evolutionary times of marine mammals. University of California Press; 2020.
[25] Elena F, Ascheri D, Giulia BC, Francesca S, Mcginty N. Bottlenose dolphin (Tursiops truncatus) habitat partitioning in relation to age classes in the northwest Mediterranean Sea. Marine Biology. 2024; 171(7):137.
[26] Nepovinnykh E, Chelak I, Eerola T, Immonen V, Kälviäinen H, Kholiavchenko M, et al. Species-agnostic patterned animal re-identification by aggregating deep local features. International Journal of Computer Vision. 2024:1-6.
[27] Chetoui M, Akhloufi MA. Fire and smoke detection using fine-tuned YOLOv8 and YOLOv7 deep models. Fire. 2024; 7(4):135.
[28] Al MA, Hasan F, Emon MF, Haque MR, Anwary AR, Shatabda S. YOLO-fish: a robust fish detection model to detect fish in realistic underwater environment. Ecological Informatics. 2022; 72:101847.
[29] Asyraf MS, Isa IS, Marzuki MI, Sulaiman SN, Hung CC. CNN-based YOLOv3 comparison for underwater object detection. Journal of Electrical and Electronic Systems Research. 2021; 18:30-7.
[30] Wang M, Liu M, Zhang F, Lei G, Guo J, Wang L. Fast classification and detection of fish images with YOLOv2. In OCEANS-MTS/IEEE kobe techno-oceans (OTO) 2018 (pp. 1-4). IEEE.
[31] Alhmoudi SA, Aldhanhani HR, Ridouane FL, Ateeg M, Al MA, Mirza SB. Study the impact of the anthropogenic activities on the marine environment of Fujairah offshore waters of UAE based on baseline surveys and buoy data. Journal of Marine Sciences. 2024; 2024(1):1998158.

Fawaghy Alhashmi received her B.Sc degree in computer science from United Arab Emirates University computer engineer at the Fujairah Research Center and a master's student at Mohamed bin Zayed Smart University. Has mainly focused on computer vision in environmental research. Within this capacity, she carried out research on integration technology within cetaceans and wildlife projects. She wins second place in the Senior Graduation Project award at the College of IT, United Arab Emirates University. Moreover, NLP project where she used several models to digitize the Arabic language for a competition and got first place, the competition was organized by the AI ministry and Cisco Systems.
Email: fawaghyalhashmi@gmail.com

Maryam Alhefeiti received her B.Sc degree in Biochemistry from United Arab Emirates University. Currently, she is a marine researcher at the FRC (Fujairah Research Center). Has mainly focused on marine ecosystem restoration, biodiversity assessment, genetic diversity, and environmental monitoring. Within this capacity, she carried out research on oyster reefs, cetaceans, and turtles. She participated in the Fujairah 1st fisheries assessment survey at the most developed scientific research vessel in the region. She started the 1st genetic diversity assessment of Fujairah dolphins. And now project leader for eDNA and oyster reef.
Email: Maryam.alhefeiti@gmail.com

Dr. Shaher Bano Mirza received her Ph.D. degree in biomathematics, bioinformatics, and computational biology. Currently holds the position of Principal Investigator and Head of Academics and Outreach at Fujairah Research Centre, United Arab Emirates. She specializes in data science and environmental protection, with a particular focus on sustainability and the utilization of native plants. In the past, she worked at Massachusetts Institute of Technology (MIT, USA) in the field of computational biophysics before joining FRC. Furthermore, she is the Project Manager for water resource mapping at FRC, which is a collaborative project with NWEC-UAEU.
Email: shaher.bano@frc.ae

Dr. Fouad Lamghari Ridouane received his Ph.D. degree in chemistry from Aix-Marseille University, in 1990. Currently, he is the director of Research at the Fujairah Research Centre, adviser of the Fujairah Authority, and leading applied research programs related to Weather forecasting, Artificial intelligence and deep learning, Air quality, and Air pollution monitoring and modeling, Underground water and hydrogeological mapping, Marine ecosystem monitoring, and Oil Spill early detection and monitoring. He is actively engaging in startup creation, technology scale-up, and interacting with global manufacturers, partners, and technology providers for data analysis and knowledge transfer according to international standards. Fouad Lamghari is a renowned Laboratory expert who has successfully participated in Aramco's billion-dollar projects.

He commissioned some of the world's most iconic research centers, namely King Abdullah University of Science and Technology (KAUST) and The Francis Crick Institute in London. Before joining the Fujairah Research Centre, Fouad Lamghari was the R&D Manager of the DEWA R&D Centre at the Mohammed bin Rashid Al Maktoum Solar Park, where he oversaw PV testing, Advanced Materials, 3D Printing, Power Systems, Smart Grids, Water Analysis, Experimental Desalination, Robotics and Drones, Building Performance, and HVAC development.
Email: Fouad.lamghari@frc.ae

Appendix I

Class	Precision	Recall	mAP_0.5	mAP_0.5:0.95
All	0.916	0.925	0.958	0.788
Grampus griseus	0.94	1	0.995	0.825
Tursiops truncatus	0.891	0.85	0.921	0.891

Appendix II

S. No.	**Abbreviation**	**Description**
1	5OES	Five Oceans Environmental Services
2	AI	Artificial Intelligence
3	ALFRE-ID	Aggregated Local Features for Re-Identification
4	CBIR	Content-Based Image Retrieval
5	CNNs	Convolutional Neural Networks
6	DETR	DEtection TRansformer
7	FP	False Positive
8	FN	False Negative
9	mAP	Mean Average Precision
10	n	nano
11	nm	Numerical Mile
12	P	Precision
13	P-R Curve	Precision-Recall Curve
14	R	Recall
15	SPADE	Spatially Adaptive Denormalization
16	TP	True Positive
17	ViT Adapter-B	Vision Transformer Adapter
18	VOLOv3	You Only Look Once Version 3
19	YOLO	You Only Look Once
20	YO2LOv1	You Only Look Once Version 1
21	YOLOv2	You Only Look Once Version 2
22	YOLOv5	You Only Look Once Version 5
23	YOL4Ov7	You Only Look Once Version 7
24	YOLOv8	You Only Look Once Version 8