Capstone Project:

Citizen Science and Artificial Intelligence: A systematic review of the current application on the Research and Conservation of Whale Sharks

Floor H. G. van der Marck (S4338219)

University College Fryslân | Campus Fryslân | University of Groningen

BSc Global Responsibility and Leadership

Capstone Project (CFBGR03610)

Under supervision of Prof Dr. J. A. Beaulieu

June 5, 2023

Leeuwarden, the Netherlands

Acknowledgements

In this research, I express my appreciation to citizen scientists who generously dedicate their time, efforts, and passion to participate in research projects worldwide. Without their invaluable contributions, this research reviewing the potential of citizen science as an integral part of whale shark research and conservation would not have been possible. Often I get the feeling that scientific research has a widely accepted belief that aims to minimise the impact of personal biases and subjective perspectives. This pursuit of objectivity seeks to remove not only personal values and preconceived notions but also tends to overlook the significance of unexpected and chance observations. These points were highlighted by Adler and colleagues (2020) in their work. Citizen science demonstrates the limitations of this approach and tries to embrace a more nuanced perspective. In the context of whale shark research and conservation, the combination of citizen science and AI emerges as a robust framework. I would like to highlight the value of citizen scientists. To sustain the values of justice and respect, their contributions must be transparently and fairly acknowledged.

Abstract

This research explores the current application of citizen science and artificial intelligence (AI) in whale shark research and conservation. The study and protection of whale sharks are of utmost importance in the field of marine conservation due to their vitality to marine ecosystems all over the world. From this perspective, this research aims to investigate the potential of the combination of citizen science AI in enhancing conservation efforts for whale sharks. The objectives include evaluating the advantages and difficulties of combining AI and citizen science for whale shark conservation and their potential to boost public involvement. The methodology employed in this study involved conducting a systematic literature review focused on scientific publications on citizen science and AI relevant to whale sharks. A comprehensive search was performed from the Google Scholar database and WildMe's publication archive, identifying and analysing nine relevant papers that formed the basis of the research's results and discussion. The results reflect that by enhancing understanding of population dynamics, migratory patterns, and habitat utilisation and boosting data collecting, the application of AI and citizen research together can improve whale shark conservation. While citizen science adds to a more extensive range of data and community interaction, AI helps with data processing, validation, and standardisation. This benefits whale shark conservation planning, decision-making, and research. This study offers a unique perspective on how AI and citizen science operate together in whale shark research and conservation by introducing a brand-new conceptual framework that integrates them. In conclusion, this research emphasises the significant contribution that citizen science and artificial intelligence can make to whale shark research and conservation, highlighting the potential for knowledge advancement, informing conservation strategies, involving communities, and opening the door for further interdisciplinary studies.

Keywords: Whale sharks, Marine conservation, Citizen Science, Artificial Intelligence

Table of Contents

Citizen Science and Artificial Intelligence: A systematic review of the current application on the Research and Conservation of Whale Sharks

It is becoming increasingly crucial to do extensive whale shark research and monitoring to preserve the species. The fusion of citizen science and artificial intelligence offers a compelling strategy for furthering initiatives linked to whale shark study and conservation in marine conservation. New possibilities in data collecting, analysis, and conservation tactics can be investigated by combining artificial intelligence systems' computing power with citizen scientists' active participation. In addition to improving our knowledge of whale sharks, this interdisciplinary collaboration also makes it possible to create conservation strategies that are effective. This research begins a scholarly exploration at the nexus of public engagement and technological innovation, aimed to deliver significant contributions to preserving the magnificent marine species known as whale sharks.

Background

Whale Shark Characteristics

The whale shark (Rhincodon typus) is the biggest fish roaming the oceans; it can grow up to 20 metres long and weigh up to 21.5 tons. Some characteristics that distinguish the whale shark from other species are their flattened and wide mouth, which can be up to 1.5 metres wide, and a unique pattern of spots that differs from animal to animal. They are filter feeders, removing plankton, tiny fish, and invertebrates from the water using their five big gill slits (Dove & Pierce, 2021). Whale sharks are known to make long journeys and can be found in tropical and mild-temperature oceans all over the globe (Arzoumanian et al., 2005; Dove & Pierce, 2021). Thought to be lonely beings, they occasionally gather in groups at feeding grounds. Whale sharks are apex predators essential to preserving the harmony of aquatic environments. In addition, they support local communities by being essential for economies and ecotourism (Araujo et al., 2020; 2022). Even though whale sharks are large and considered an emblematic species, a lot is still unclear about them. Their biology, environment, and behaviour, including their social organisation and movement and reproductive biology, are all still being studied in depth by researchers. Before the 1980s, whale shark sightings were uncommon. The IUCN Red List of Threatened Species downlisted whale sharks worldwide to Endangered in 2016 after a more than 60% drop across their range (Pierce & Norman, 2016; Pierce et al., 2021). This raises concerns about what has transpired over the previous decades.

Changing Marine Environment

The marine ecosystem has changed significantly over the past few decades due to human actions such as habitat loss, overfishing, pollution, and climate change (Dove & Pierce, 2021). These actions have decreased marine species, shifted ocean chemistry and temperatures, and changed marine environments. One of the biggest dangers to the aquatic ecosystem is climate change, causing ocean acidification, increasing sea levels, and shifting ocean currents and temperatures (Dove & Pierce, 2021). Whale sharks are among the marine animals whose environments are impacted by these shifts. The spread of plankton, the main food supply for whale sharks, is changing due to rising temperatures. Another effect that alters food supplies further is overfishing. Whale sharks are impacted by overfishing because it makes less of their food available.

Additionally, a severe issue with whale shark bycatch in fishing operations can result in harm or even mortality. The effects of pollution, including plastic contamination, on the marine ecosystem are profound. Ingesting or entangling aquatic life, including whale sharks, pollution can cause damage. Lastly, the loss of habitats, such as coral reefs, significantly affects marine biodiversity. Many aquatic species, including whale sharks, depend on these environments for feeding and nursery grounds (Dove & Pierce, 2021). Overall, due to human activity, marine ecosystems are confronted with severe challenges. Whale sharks must also contend with anthropogenic threats such as bycatch, ship strikes, and targeted fishing.

Most of these changes come into play over the single life span of many mature whale sharks present-day (Dove & Pierce, 2021). Outperforming the limitations of biological evolution, these species need more time to adapt to their changing marine environment.

Definitions

The term “citizen science” describes the participation of the general public in scientific studies and data collection. Through public participation, citizen science fosters cooperation, democratises scientific research, and increases public awareness of and interest in scientific activities. Citizen scientists assist scientific research in various domains by collecting, analysing, and interpreting data. Citizen science in marine conservation is applied through reporting sightings, monitoring behaviour, and collecting environmental data.

The creation and use of computer systems and algorithms that display intelligent behaviours and capacities generally associated with human intellect is referred to as “artificial intelligence” (AI). AI allows computers to sense and comprehend their surroundings, reason, gain knowledge through experience, and act independently or with little human involvement. AI is used in many different contexts, such as data analysis, automation, decision-making, and problem-solving, and it has the potential to improve accuracy, precision, and innovation across a range of areas.

Role of Citizen Science and Artificial Intelligence

Gaining knowledge of the change in the marine environment and modelling these shifts can benefit from both citizen science and AI. Presently, citizen science is a valuable resource for studying global change (Chandler et al., 2016; Theobald et al., 2015). Participation in citizen science projects from local communities, tourists and the passionate general public can help gather information on aquatic animals, ecosystems and environmental factors (Norman et al., 2016). Often, this can be over a large and broad scale, be cost-efficient and facilitate the involvement of the local community (Araujo et al., 2019; 2022; Norman et al., 2016; Theobald et al., 2015). Although citizen science has yet to achieve its full potential, as initiatives expand in extent and the degree of public participation, there will be a more significant need for innovative tools in database administration, scientific analysis, and educational instruments. (Bonney et al., 2009). This is where AI may come into play and advance the utilisation of citizen science.

Large databases produced by citizen science initiatives and other sources can be analysed using AI (Arzoumanian et al., 2005). For instance, machine learning methods can find trends in abundance, distribution, behaviour, and reaction to external factors. This knowledge can then be used to better forecasting models on how species and ecosystems may transform under future climate scenarios. Citizen research has greatly benefited our knowledge of whale sharks' distribution, population growth, and migration habits (Araujo et al., 2022). Predictive models can be created using data gathered by citizen scientists to help pinpoint regions crucial for whale shark protection. A further application of AI is analysing satellite monitoring data to locate whale shark concentrations and comprehend how they use various oceanographic characteristics.

Citizen science plays a crucial part in improving AI by utilising the intelligence of a collective to improve several aspects of the technology. As mentioned above, citizen science enables data collection on a large and broad scale because citizen scientists may help compile enormous amounts of data from various contexts and disciplines. It can also aid in the time-consuming process of labelling and annotating data, which is essential for training AI models. In order to ensure the correctness and dependability of the technology, citizen scientists can also help to control quality by checking and verifying AI-generated outputs. This leads to testing and improving algorithms; they can offer insightful feedback that results in improvements and optimisations. Finally, citizen science efforts encourage public participation and boost awareness of AI capabilities by incorporating the community and establishing ownership and knowledge of the technology's potential and limits.

Research Question, Aim and Objectives

The article integrates citizen science and AI potential for endangered species conservation, explicitly focusing on the application of whale sharks. In light of the changing marine environment, this research aims to determine how well citizen science and AI can help our understanding of whale shark ecology and behaviour and benefit conservation efforts. Within this context, the research aims to evaluate the effectiveness of the combination of this approach. Furthermore, it aspires to assess the prospective future of citizen science and AI. The research question "How can the combination of citizen science and artificial intelligence help conservation efforts of whale sharks?" serves as a guiding force to facilitate the aim of this study.

The objective of this research is to examine the benefits and drawbacks of using citizen science and AI for the conservation of whale sharks, discuss the challenges associated with the implementation, and assess how well these methods might encourage community participation. Furthermore, the study will highlight the research gaps associated with citizen science and AI and recommend areas for the future sustainability of the field.

Relevance

Effective marine conservation projects depend on knowledge about the behaviour and ecology of whale sharks. Understanding the whale shark's behaviour and ecology under global climate change is of utmost significance, given the growing dangers presented by anthropogenic activities. This research seeks to enhance conservation efforts by examining the integration of citizen science and AI in tracking whale sharks. The knowledge from this research is essential for conservation activities, including facilitating handlebars and helping to create effective legislation and conservation plans for the whale shark species.

Outline

Following, the methodology of literature collection and data extraction will be explained. Results will highlight the findings of the systematic literature review and will present advantages, disadvantages and implications of the application of citizen science and AI on whale shark research and conservation. The discussion section summarises the findings, compares approaches, identifies strengths and limitations, and addresses the research gaps found. The conclusion summarises vital findings and emphasises the relevance of citizen science and AI for advancing whale shark research and conservation.

Methodology

This research reviews the literature on citizen science and AI in the context of whale sharks to understand how the combination can help us benefit in designing and managing conservation efforts. The peer-reviewed papers examining the topics and structures of initiatives that have been academically published are the main emphasis of the study. The literature will be collected through a systematic analysis methodology. This section will elaborate on the data collection and extraction methods.

The collection of literature has been conducted employing two methods. Google Scholar has been searched between January and March 2023 for articles reporting ongoing or past citizen science, and AI projects relevant to whale sharks between 2005 and the present. In 2005, a novel technique for identifying individual whale sharks through numerical pattern-matching of their natural surface ‘spot’ colourations was presented (Arzoumanian et al., 2005). After this onset, the quality and quantity of studies on the use of AI in whale shark monitoring have significantly improved, producing a substantive body of information that, with careful analysis, can be used to identify new trends and assess particular methods. The main search items “citizen science”, “artificial intelligence” and their synonyms and “whale shark” and their scientific name, every time combined with the logical operators OR and AND, were used to extract articles: (“citizen science” OR “crowdsourcing” OR “volunteer monitoring”) AND (“AI analysis” OR “machine learning” OR “artificial intelligence”) AND (“whale sharks” OR “Rhincodon typus”). The search yielded 132 results.

Although the search algorithms of Google Scholar for ranking and classifying publications are ambiguous, its use is primarily driven by the fact that it helps people comprehend academics’ diverse and global perspectives (Meho & Yang, 2007). This applies to heterogeneous areas that comprise a wide range of stakeholders, viewpoints, disciplinary boundaries, and publishing formats, of which citizen science is a part. For this review, it should be emphasised that citizen science projects frequently do not appear in peer-reviewed academic publications (Theobald et al., 2015). Additionally, when published, they either fail to give involved volunteers enough recognition or fail to use language that is suitable for citizen science (Cooper et al., 2014). Therefore, this thesis will focus on the academic results of citizen science community efforts that are relevant to AI and whale sharks.

Secondly, WildMe's publication archive has been searched. WildMe (https://www.wildme.org/) is a US-based non-profit organisation that develops and deploys open-source software and technology for researching and conserving various species. The organisation primarily focuses on using computer vision and AI to identify individual animals from photographs or video footage. WildMe offers several tools and platforms for researchers and citizen scientists to upload and analyse their wildlife data. WildMe has developed a community-driven platform called Wildbook for Whale Sharks, more recently known as Sharkbook (https://www.sharkbook.ai/). No global database was available to record whale shark sightings until 2003 when the Wild Book became the primary database (Arzoumanian et al., 2005).

In March 2023, articles reporting ongoing or past citizen science and, or AI projects relevant to whale sharks (and other species) between 2005 and the present have been searched in WildMe's publication archive. Including other species in this search has been chosen as they may serve as relevant examples of similar applications in research on other species. The main search items, "citizen science" and "machine learning" and their synonyms, were used in a keyword search (ctrl F) to extract articles. This search yielded 16 results.

Finally, during the analysis of the papers from both Google Scholar and WildMe search, I have conducted backwards snowballing on the selected articles for further possible literature on the topic using the reference list and a keyword search (ctrl F) of the main search items "citizen science", "artificial intelligence" and their synonyms and "whale shark" and their scientific name. The final result yielded 101 peer-reviewed articles of relevance merged from both Google Scholar and WildMe's archive search efforts. The papers in English have been organised in alphabetical order of the first author's name, according to the journal type as peer-reviewed or not, based on the year of release. The abstracts and results of these papers have been analysed to eliminate any whose topic emphasis does not relate to whale sharks and citizen science or AI. From this point, articles have been excluded or included based on their relevance to citizen science, AI and whale sharks through the data extraction form, elaborated below. Figure 1 shows the flowchart of inclusion and exclusion. After exclusion, the final number of extracted papers was nine relevant to citizen science and, or AI, concerning whale sharks.

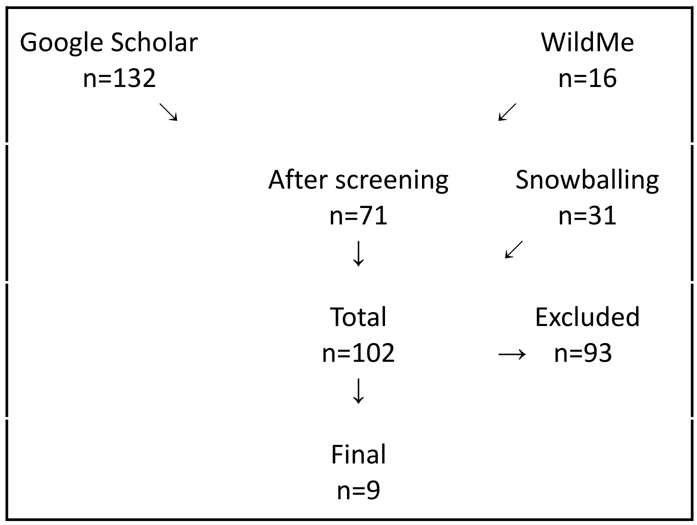

Figure 1: Flowchart of exclusion and inclusion

Appendix I shows the classification of these articles according to data extraction form. Characteristics key to the practice of citizen science, AI and whale sharks have been included. The standardised data extraction form has been used to extract relevant information from each article through a thorough analysis. The data extraction form has been developed based on the research question and inclusion criteria established for the review. The form includes fields for study context, participants, the extent of citizen science and AI involvement, the aim of the research and research outcomes. Before the complete data extraction, the data extraction form was trialled on articles selected from the WildMe archive. By testing the method on a smaller subset, an assessment of the form's effectiveness has been achieved to make adjustments where necessary. However, changes were deemed unnecessary as the form provided the means for sufficient analysis. Eventually, the data extraction form was utilised for the complete set of selected articles.

Deviating from a traditional systematic literature review, which usually assesses each paper individually, in the results section of this paper, collective findings, themes and trends across the selected papers have been analysed that are relevant to the main research question and objectives. In doing so, the analysis of the relevant literature is informed by the systematic selection and review process, as depicted above. This choice was made due to the interconnectedness of citizen science and AI in whale shark research and conservation.

Results

Citizen Science and Whale Sharks

Gaining knowledge of the global ecology of whale sharks, including their exposure to the dangers of the changing marine environment and anthropogenic threats, especially over a large spatial and temporal scale, is essential in the conservation of the whale shark (Araujo et al., 2017). Precisely, the study of aquatic animals with great mobility such as the whale shark, poses particular difficulties that call for multidisciplinary collaboration and multi-methodological approaches (Araujo et al., 2022; Norman et al., 2017). It requires large-scale biogeographic research projects which exceed the capabilities and capacities of a single research team (Norman et al., 2017). To overcome these difficulties, citizen science has gained popularity as a useful methodology. Involving the general public in scientific programmes through citizen science is a useful strategy, by utilising the power of observations from a huge number of people, assisted by growing public awareness and easily available digital technology (Araujo et al., 2020; Norman et al., 2017). In this section, the application of citizen science on whale sharks will be discussed along the lines of common themes found in a selection of 7 papers. These themes include data collection, method of analysis, exemplary findings, advantages and disadvantages, implications for conservation.

Data Collection in Literature

The data collection in citizen science projects as described in the papers is executed by the general public (i.e. Araujo et al., 2017; 2020; 2019; 2022). In some papers, researchers also directly contributed to the data collection process (e.g. Araujo et al., 2017). Participants were found to range from tourists, tour operators, diving and snorkelling communities, local fishermen, to anyone with an interest in science and willingness to contribute to whale shark research. By Magson and colleagues (2022), to entice visitors and operators to contribute images and video through one of the platforms, local presentations, educational handouts and posters, and awareness campaigns were run. Whale shark tourism has grown tremendously over the past two decades (Norman et al., 2017), which has resulted in an increase in citizen science contribution from the tourism sector. The predominant data collection method used in whale shark research involving citizen scientists is photographic data collection.

The recurring platform for data collection in the literature is Wildbook for Whale Sharks (i.e. Araujo et al., 2017; 2022; Hoffmayer et al., 2008; Magson et al., 2022; Norman et al., 2017). The process of data collection via this platform is described by Norman and colleagues (2017). Photographs of whale sharks are gathered by researchers or citizen scientists who capture the distinctive spot pattern of a whale shark, which serves as a means for identification. The online database receives the gathered photographs and uses them as a platform to store and analyse the data. Participants who submit an entry receive an automated response email with a web link where they may access it. Additionally, the submitter is automatically informed if the same shark is seen again. Other pertinent data related to the sighting, such as the whale shark's estimated total length, location, and possible sex may also be uploaded by participants. Repeated sightings of identified individuals boost confidence in confirming each whale shark's characteristics. The uploaded photographs are processed by researchers utilising computer-assisted pattern-matching technologies, which will be elaborated on further in the results.

As the Wildbook for Whale Sharks platform grew, its expansion was constrained by the varying growth rates of data for each whale shark hotspot. This allowed for training and outreach programmes to be expanded with a focus on particular whale shark aggregation locations, which led to acceptance by researchers and other important stakeholders. Moreover, this helped to guarantee that a solid data set was available for the current evaluation of the biology and ecology of the whale shark. This expansion of knowledge on whale shark coincidentally led to a relatively recent expansion of citizen science monitoring (Norman et al., 2017).

Analysis Method in Literature

A trustworthy method used in citizen science initiatives to analyse the gathered information on species with distinctive identification marks is photo-identification (Araujo et al., 2020) and has been used in seven papers (i.e. Araujo et al., 2017; 2019; 2020; ; Arzoumanian et al., 2005; Holmberg et al., 2008; Magson et al., 2022; Norman et al., 2017). Underwater cameras may be used to gather information on the whale shark's distinctive and consistent spot patterns without harming the animal. Since the mid-2000s, photo-identification has grown in popularity with the help of programmes like the Wildbook for Whale Sharks (Araujo et al., 2022). The Wildbook for Whale Sharks database received a total of 28,776 reports of whale shark encounters between 1992 and 2014, leading to the identification of 6091 distinct whale sharks from 54 different countries (Norman et al., 2017). Present day, the database counts 10,2316 approved encounters (May 2023, https://www.sharkbook.ai/encounters/searchResults.jsp?state=approved).

Due to their regular aggregations and the accessibility of affordable technology, whale sharks are especially well suited for photographic-identification. Social media and the availability of underwater camera equipment have made it easier for the general public to participate in citizen science initiatives to collect whale shark photo-identification data (Araujo et al., 2022). The ease of image sharing, driven by the pursuit of capturing the best whale shark photo, has caused a significant increase over the last decade (Magson et al., 2022). With the use of these data, it is possible to monitor trends, migrations, seasonality, habitat usage, behaviour, and to determine events such as vessel strikes, pollution, and targeted fishing (Araujo et al., 2022). Researchers have improved their knowledge of whale sharks in a number of areas across the world, by enhancing existing databases or depending exclusively on the basis of photo-identification data from the public (Araujo et al., 2020).

Exemplary Findings

With over 500,000 visitors each year, Oslob in the Philippines is the world's most popular place to see whale sharks (Araujo et al., 2020). Whale sharks are provisioned at Oslob to enable encounters between humans and sharks, and since 2012, researchers have been performing daily in-water photo-identification via Wildbook for Whale Sharks. A juvenile female whale shark was identified in Oslob, Philippines, in December 2017. It was frequently seen in Oslob but not at other observation points. The same individual was discovered in Pulau Sipadan, Malaysia, by a team of divers in October 2019, demonstrating the species' capacity for long-distance travel (Araujo et al., 2020). This pioneering photo-identification that tracked a whale shark's migration from the Philippines to Malaysia shows the potential of data from citizen scientists and is exemplified in other studies as well (i.e. Norman et al., 2017).

Another example can be found in Sabah, northern Borneo, in Malaysia. As described by Araujo and colleagues (2022), whale shark information is scarce in Malaysia, with the majority of the information coming from reports in grey literature, citizen science sightings reported in Wildbook for Whale Sharks, and other local researchers. Therefore, citizen science is essential in completing conservation efforts. Sabah's "Malaysia Whale Sharks" initiative makes use of social media sites like Facebook, Instagram, YouTube, and Flickr to compile even more information (Araujo et al., 2022). Furthermore, the paper provides two instances of how social media exposes Malaysia's unlawful whale shark hunting. Researchers were able to locate interactions with whale sharks, match them to records already in the database, or create new identifications by mining videos and images posted on social media sites like Facebook and WhatsApp. These illustrations help us comprehend the continual risks to the species, especially in places where there are already laws and protections in place (Araujo et al., 2022).

On the other side of the world, a study carried out by Hoffmayer and colleagues (2021), displays the seasonal occurrence, horizontal movement and habitat use patterns of whale sharks in the Gulf of Mexico. Various data sources were employed in the study, including opportunistic sightings reported by boat captains and local fishermen and photo-identification information provided by citizen scientists through the Wildbook for Whale Sharks database. The study discovered evidence of seasonal trends in the appearance and migration of whale sharks, with summer being the season of greatest abundance. Through the use of photo-identification data, citizen science participants contributed to the study's geographic coverage and gave useful information on the movements of certain whale sharks.

Furthermore, the study of Magson and colleagues (2022) explored the occurrence and seasonality of whale sharks in the waters of Thailand, using a citizen science-based approach to supplement the limited quantitative data available. This study heavily utilises citizen science as a main method for gathering information on whale shark encounters. Sighting reports were given by divers, other users of marine resources, and members of the general public. The researchers disseminated infographics and a Code of Conduct to encourage participation from photographers, visitors, and local tour operators. Out of 249 observations between 2004 and 2019, the research identified a total of 178 individual whale sharks. The majority of these sightings happened between 2015 and 2019, most likely as a result of increased reporting made possible by social media, direct marketing and the development of the underwater photography community in Thailand. Despite limits in the accuracy of size and sex estimate, the study emphasises the relevance of citizen scientific engagement in data-poor places like the Gulf of Thailand. According to the study, citizen science can produce outcomes that are comparable with those of more established techniques when it comes to observing whale shark aggregations.

As a final example; in the study of Araujo and colleagues (2017) the Wildbook for Whale Sharks database, and direct contributions resulted in the collection of 276 images. However, due to problems like low image quality or inappropriate framing, only (or 58% could be used for identification. 31% of the acceptable images featured whale sharks that were seen at the research location, including 14 newly identified animals. The information gathered was used to create residency models and provide light on how whale sharks used the research location. Insights on the travel of the species were also provided by the citizen science data, with matches appearing in a variety of places, including feeding grounds. Notably, one whale shark who had been released in Taiwan was subsequently recognised on Panaon Island, marking the first instance in Southeast Asia where two photos from different countries were matched. The report also highlights the distances travelled within particular durations by mentioning some individual whale sharks' fastest-known travels.

Advantages

The following list presents the sum of advantages citizen science applied in whale shark research and conservation, as found in the literature.

(1) Expanded Spatial and Temporal Coverage: By making it possible to gather a sizable amount of data from many places and time periods, citizen science in whale shark study and protection offers a huge advantage. The involvement of citizen scientists allows for more regional and temporal coverage of whale shark research than is possible with traditional studies alone. Through this partnership, considerable data gathering is possible in areas that may otherwise be difficult for researchers to reach (Araujo et al., 2017; 2019; Norman et al., 2017).

(2) Cost-Effective and Long-Term Monitoring: Initiatives in citizen science have the added benefit of being practical, inexpensive, and able to conduct long-term monitoring. The logistical and budgetary costs associated with conventional research techniques can be greatly minimised by involving citizen scientists. The excitement and involvement of volunteers enable cost reductions in data collection, fieldwork, and monitoring tasks (Araujo et al., 2019). Additionally, citizen

science also makes it easier to continuously gather data and monitor the environment (Norman et al., 2017).

(3) Community Engagement and Education: Initiatives in support of citizen science offer a chance to involve and inform the general people about whale sharks and the need for their conservation. Citizen science encourages a sense of ownership and responsibility among participants and equips them to support conservation initiatives by incorporating volunteers in data collecting. This involvement may result in more understanding, backing, and sustainable methods for the long-term preservation of whale sharks (Araujo et al., 2017; 2019; Magson et al., 2022).

(4) Rapid Response and Early Warning Systems: In order to identify risks to whale sharks early on and take appropriate action, citizen scientists can be quite helpful. Citizen scientists can operate as an early warning system by reporting any odd events, such as entanglements, illegal fishing, or habitat destruction, thanks to their presence in local communities and regular contact with the maritime environment. This quick reaction time improves conservation efforts and helps to reduce possible threats (Araujo et al., 2022; Norman et al., 2017).

Disadvantages

The following list presents the sum of disadvantages citizen science applied in whale shark research and conservation, as found in the literature.

(1) Data Quality, Consistency and Reliability: Concerns about the dependability and quality of the data produced arise when citizen science is used in whale shark research. Peak tourist seasons may affect temporal and geographical efforts, which might lead to biases and errors in the statistics. Inconsistencies in data quality and completeness might also result from discrepancies in citizen scientists' experience and methodologies. Moreover, lack of oversight during data collecting may result in observer biases, inaccuracies, or discrepancies in the data that is

gathered. Due to these limitations, studies and comparisons are less reliable, which might lead to incorrect interpretations of the dynamics of the whale shark population (Araujo et al., 2017; 2019; Magson et al., 2022; Norman et al., 2017).

Conservation Implications of Citizen Science

The following list presents the sum of implications of citizen science applied in whale shark conservation, as found in the literature.

(1) Knowledge Advancement: Through citizen scientist involvement, the knowledge of the species and its habitat advances by providing clear information about the ecology of whale sharks. Effective conservation strategies and decision-making depend on this information (Araujo et al., 2017; 2019; 2020; Magson et al., 2022; Norman et al., 2017).

(2) Tourism Management and Whale Shark Wellbeing: The regular monitoring of whale sharks, particularly during ecotourism operations, furthermore offers insights into the effects of tourism on appearance and frequency of return to hotspot locations. The management of tourism-related activities can be guided by this information in order to avoid disturbances and safeguard the well-being of whale sharks (Magson et al., 2022; Norman et al., 2017).

(3) Increased Collection Methodologies: The public can participate in data collection thanks to the expansion of the marine wildlife tourism industry, as well as the availability of smartphones and underwater cameras. By utilising citizen scientists' abilities, a sizable number of observations can be produced, providing important data that would otherwise be difficult to gather through traditional research techniques. The greater data accessibility aids in management of conservation activities and advances our knowledge of whale sharks (Araujo et al., 2020; 2022; Magson et al., 2022).

(4) Identification of Crucial Habitats and Connectivity: Citizen science initiatives can assist in identifying places that are crucial to the survival of whale sharks. Additionally, citizen research can uncover patterns of population connectedness and mobility by tracking certain whale sharks throughout several areas, indicating the necessity for collaborative conservation and management approaches across regions and countries (Araujo et al., 2020; Norman et al., 2017).

(5) Population Assessment: Mark-recapture analysis can be applied with citizen science data, to estimate population numbers and mixing. This knowledge is essential for figuring out the dynamics of the whale shark population and creating conservation plans based on population assessments (Norman et al., 2017).

(6) Climate Change Impacts: When paired with environmental factors, data from citizen science can shed light on the long-term effects of climate change on the movements of whale sharks. This information assists the development of adaptation plans and aids in pinpointing geographic regions where the effects of climate change may be most noticeable (Norman et al., 2017).

(7) Citizen Science Limitations: Directed research programmes can be created to fill up gaps left by the limitations of citizen science initiatives in terms of sampling outside of peak tourist seasons, or to increase data quality. In doing so, more thorough and reliable data can be gathered, improving our understanding of whale shark populations, by devoting resources to independent sampling outside of typical tourist activity and in information on how to collect valid photographs (Magson et al., 2022; Norman et al., 2017).

(8) Specialised Monitoring for Accurate Demographics: Implementing specialised monitoring programmes at locations with scarce data can help provide more accurate population demographics. The construction of a global evaluation of whale sharks across their range is made possible by the collective analysis of data from many sites, giving a thorough grasp of their conservation status (Magson et al., 2022; Norman et al., 2017).

(9) Centralised Data Management: For the purpose of creating future national and international management plans, it is advantageous to have all citizen science monitoring data safely kept in a single area with shared access (i.e. Wildbook for Whale Sharks). Access to information facilitates the planning of long-term conservation efforts for whale sharks and other species worldwide (Araujo et al., 2020; Holmberg et al., 2008; Norman et al., 2017).

(10) Standardised Methodologies: To facilitate research and conservation efforts, it is recommended to establish a minimum code of conduct for interacting with whale sharks and to implement information and education campaigns targeting stakeholders such as fishing communities and local governments. Furthermore, standardised methodologies are essential for citizen scientific initiatives involving whale shark research because they foster uniformity, dependability, and cooperation. The quality of identification data may be increased, enabling precise comparisons and analysis, by creating stringent processes for data collection, including rules for identifying images, documenting dates and locations, and minimising disturbance to the animals. Collectively, these actions improve the scientific rigour and efficacy of whale shark research and conservation (Araujo et al., 2017; 2020; Magson et al., 2022; Holmberg et al., 2008).

(11) Social Media: By enabling the efficient distribution of information and encouraging public involvement, social media significantly contributes to the study and conservation of whale sharks. Sharing updates on research results, conservation actions, and educational campaigns enables scientists and conservation groups to increase public awareness of the need to conserve whale sharks and build support for conservation efforts (Araujo et al., 2022; Magson et al., 2022).

(12) Local Community Engagement: The involvement of citizen scientists can help raise awareness about whale sharks among the local community and beyond, fostering a sense of stewardship and promoting conservation actions from a local level. As they frequently have valuable

information and experience with the species and its environment, local communities must be involved for conservation programmes to be effective. According to the authors, local communities' involvement in data gathering can result in data that are more accurate and of higher quality (Araujo et al., 2017; 2022).

Artificial Intelligence and Whale Sharks

It may be possible to handle low quantities of images and identify individuals by eye. However, this approach becomes infeasible and unreliable when data are gathered from numerous animals spotted in several locations throughout the globe. Large data sets are now readily available, making human photo-identification impractical. This has spurred the development of computer-assisted methods for reliably and quickly scanning photographic libraries. Mid-2000's, a novel automated algorithm was developed by Arzoumanian and colleagues, that enabled pattern matching and identification of individual whale sharks. Whale sharks are born with a distinctive pattern of lines and spots that does not appear to change significantly through time and may be used to identify certain sharks.

The majority of recent papers discussing whale shark research cite Arzoumanian and colleagues (2005) when mentioning the use of photographic identification techniques. Simultaneously, when mentioning the use of this technique, the same papers mention the use of for example the Wildbook for Whale Sharks platform as a means of executing the methodology. Accordingly, the technique as introduced by Arzoumaninan and colleagues is presently integrated in Wildbook for Whale Sharks, possibly with technological advancements over time. From the WildMe website (https://www.wildme.org/codex-and-wildbook.html); wildbook uses computer vision and machine learning to analyse photographs, find animals, classify them according to their species, and propose matched people from the database. The application of AI on whale sharks will be discussed along the lines of common themes found in a selection of 2 main papers that discuss the application of AI specifically, where one paper serves as a full description of the method and the other as an example of application. These themes include method emergence, exemplary findings, advantages and disadvantages, and implications for conservation.

Method Emergence

When conventional tagging of whale sharks has limited success, Arzoumanian and colleagues (2005) introduced the first numerical method for identifying individual whale sharks based on their spot patterns (Figure 2). This technique is based on an automated algorithm that is used in astronomy to identify star patterns. It was considered that this method would help to improve global knowledge of the life cycles, migratory patterns, and demography of whale sharks and, eventually, guide conservation and management activities (Arzoumanian et al., 2005). The Wildbook for Whale Sharks was published online 2003, incorporating a database of whale shark photos that was originally known as the ECOCEAN Whale Shark Photo-Identification Library, which was created in 1995 (Araujo et al., 2017; Norman et al., 2017). The newly developed technique was incorporated into the photo-identification library (Arzoumanian et al., 2005).

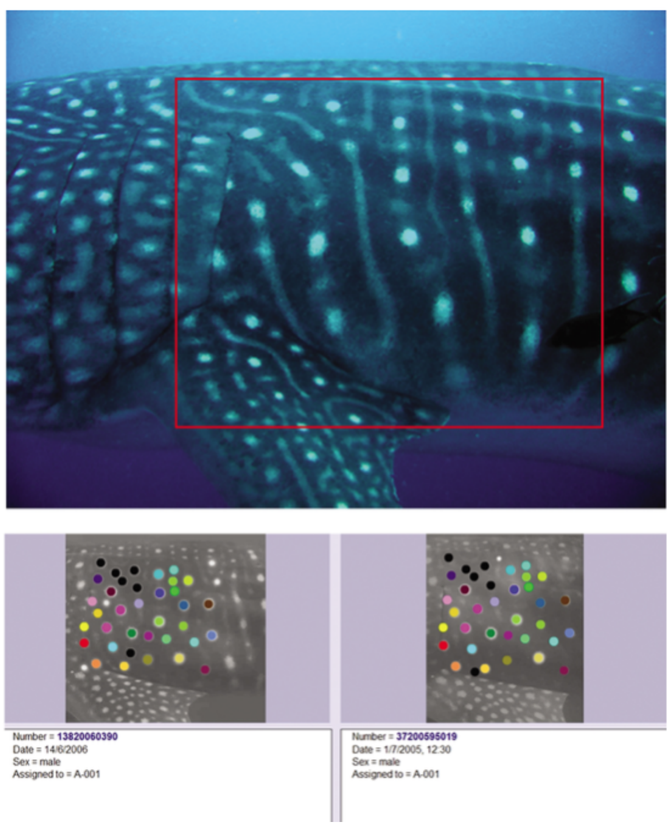

Figure 2: Visualisation of automated photo-identification techniques (Norman et al., 2017).

The pattern-matching algorithm utilised in this study was created and tested on the ECOCEAN library's testing ground. While one of the researchers (i.e. Norman) acquired the majority of the raw data used to test the algorithm, contributions to the ECOCEAN Library have come from researchers, and ecotourists. A precise area, the measurement region, was chosen behind each whale shark's gill slits in order to identify spot patterns. In order to capture the needed area, photographers were urged to place their cameras exactly over the centre of the measurement zone. Additionally, in order to prove the shark's identity further, it was also suggested to photograph secondary identifying markers such as fin or body scars (Arzoumanian et al., 2005).

A computer-driven pattern-matching system has to be able to distinguish between important aspects of a picture (Arzoumanian et al., 2005). Blob extraction, a method for measuring the positions and dimensions of groupings of pixels of a single colour, is ideally suited to the contrast of white whale shark markings on their darker skin. Individual identification for each shark is based on the spatial interactions between these groups, which are represented by a set of derived (x, y) coordinates. Due to limiting conditions including limited visibility and intense surface sunshine that may wash away spots, the unpredictability of the underwater environment presents a barrier to feature detection. It is also possible that pictures are taken of whale sharks where the image forms an unfavourable angle of the whale shark. Using various image-processing techniques, these unwanted circumstances can be corrected. Subsequently, blob extraction is used to identify spots after the photo has been reduced to a cropped, rotation-corrected, and contrast-enhanced grayscale image. In 2005, the extraction procedure took around 10 minutes in total for a skilled operator.

Advancements in computer vision and AI have allowed further automation of this process and can expedite it. The pattern-matching and identification processes have been greatly accelerated by the continued use of AI, which has decreased the need for human operators. The advanced algorithms can identify patterns, extract features, and adjust their behaviour based on the input data they receive. They can improve their performance over time by iteratively learning from examples or feedback. Algorithms as described by Arzoumanian and colleagues (2005) operate based on predefined rules and procedures.

Arzoumanian and colleagues (2005) suggested that work on method improvements and an evaluation of the method's potential would continue. Their application requires a skilled operator to extract spot coordinates from provided photographs and review the outcomes of the automated scan across the image collection. Although the latter is a desired check on how well the approach scores picture comparisons, the former process can be further automated to increase productivity and reduce the chance of operator error.

In present practice, this technology assists in determining if the whale shark being seen is a newly occurred individual or a reoccurrence of a shark that has already been registered in the database. Whale sharks that have been successfully identified are given a prefix depending on the location of their initial confirmed observation. Each recently discovered whale shark is given a unique number associated with the sighting site. Some encounters may not be matched to a single shark since some may not have identification photos of high enough quality to verify the identity of the whale shark. Through the global online database, the information gathered is disseminated to interested parties and open to the general public. This makes it possible to match movements between places on an international scale. The Wildbook database has search capabilities that enable the assessment of each location's encounter rates, sizes, and sex ratios for individually recognised whale sharks (Norman et al., 2017).

Exemplary Findings

From the papers described in the previous section, a few mention the use of AI methodology through photo-identification as described by Arzoumanian and colleagues (2005). These include; Araujo et al. (2017; 2019; 2022), Holmberg et al. (2017), Magson et al. (2022), Norman et al. (2017), and . Each of these papers focus more on the citizen science contribution to the research results rather than the application of AI. It is not stated how these algorithms are utilised and what their implications are. In this section, a light is shed on the specific application of AI on whale sharks. If not explicitly mentioned, they suggest making use of AI through Wildbook for Whale Sharks or briefly mention the usage of AI algorithms for automated photo-identification (i.e. Hoffmayer et al., 2021).

The study of Holmberg and colleagues in 2008 focuses on collaborative photo monitoring of whale sharks for reliable and comparable population characteristics. The utilisation of AI can be attributed to the use of Arzoumanian and colleagues' (2005) photographic identification technique. The aim of this study was to implement and evaluate the novel technique for identifying individual whale sharks. The researchers used diving boats and spotter planes to gather data from 1995 to 2006 from a variety of sources, including researchers, visitors, the ecotourism industry, and government authorities. Photographs with, encounter dates, an estimate of the whale sharks' length, and information on their gender were all recorded. Individual animals were accurately identified using computer-assisted photo-identification for mark-recapture analyses. The study was deemed successful; over 90% of the time in testing, utilising a collection of previously recognised shark photos, the approach correctly matched pairings showing the same pattern, along with variables such as scarring and other markings. More than 100 additional matches were generated from a bigger database of previously unidentified photos using the technique. However, performance was found to be negatively impacted by elements like image quality and photographic perspective (Holmberg et al., 2008).

Advantages

The following list presents the sum of advantages of artificial intelligence applied in whale shark research and conservation, as found in the literature.

(1) Efficient Data Processing: AI systems can handle vast amounts of picture data quickly, making it possible to analyse extensive datasets that would otherwise take a long time to examine

manually. This effectiveness makes it possible to extract useful information from a large number of photos, expanding the size and reach of the whale shark study (Arzoumanian et al., 2005).

(2) Accurate and Objective Identification and Behaviour Analysis: Computer vision approaches based on AI may successfully recognise and categorise whale sharks in photographs, enabling effective individual recognition and tracking. With the help of this capacity, population dynamics over time may be better understood, and migrations and behavioural patterns can be tracked (Holmberg et al., 2008).

(3) Enhanced Conservation Efforts: AI tools allow for the precise identification and tracking of individual whale sharks, facilitating long-term population monitoring and advancing conservation efforts. Computer vision powered by AI can offer useful information on population numbers, movement patterns, habitat use, and migration routes. This data aids in the formulation of sensible conservation plans and the evaluation of the potency of adopted tactics (Araujo et al., 2019; Arzoumanian et al., 2005).

(4) Noninvasive: Photo-identification monitoring enhanced with AI in whale shark research has provided a noninvasive manner to tag and identify individuals (Araujo et al., 2019; 2022; Norman et al., 2017).

Disadvantages

The following list presents the sum of disadvantages of artificial intelligence applied in whale shark research and conservation, as found in the literature.

(1) Images must be annotated by humans in order to train AI models, and this procedure may unintentionally add biases or inaccuracies. In order to do an accurate study on whale sharks, annotations that identify particular individuals, sex, or behaviours must be accurate. Annotations

that are inaccurate or biassed might result in incorrect interpretations of the data and incorrect conclusions (Arzoumanina et al., 2005).

(2) Environmental aspects (e.g. water clarity, lighting and photo quality) can have an impact on how well computer vision algorithms work.These variables can differ greatly in the case of whale shark photography underwater, making it difficult to reliably identify individuals or get useful information from the data. The usefulness of AI-based computer vision systems in different, real-world environmental situations is constrained by their reliance on clear, high-quality pictures (Arzoumanian et al., 2005; Holmberg et al., 2008).

Conservation Implications of Artificial Intelligence

The following list presents the sum of implications of artificial intelligence applied in whale shark conservation, as found in the literature.

(1) Enhanced Population Monitoring: Whale shark monitoring and efficient individual identification are made possible by computer vision techniques based on AI. Conservation managers may evaluate population size, travel patterns, and habitat utilisation with the help of this information, which aids in long-term population monitoring activities. The creation and execution of focused conservation initiatives rely heavily on these findings (Arzoumanian et al., 2005; Holmberg et al., 2008).

(2) Conservation Planning and Decision-Making: Making educated decisions on conservation can be facilitated by the availability of precise and complete data gathered using AI-based computer vision algorithms. This information may be used by conservation managers to prioritise conservation efforts, spend funds wisely, and create management plans that are suited to the unique requirements and vulnerabilities of whale shark populations (Arzoumanian et al., 2005).

Discussion

Commonalities, contradictions, patterns and relationships between the advantages, disadvantages and implications of citizen science and AI in whale shark research and conservation can be derived from the results of this research. This discussion will display the results and try to find explanations for them.

The biggest commonality is found in the expanded data collection that citizen science and AI both provide. A larger and more thorough understanding of whale sharks is made possible by the compilation of large datasets, which is facilitated by both citizen science and AI. Data from different places and times can be collected through citizen science, at the same time as AI can process this photographic data efficiently. With this expanded information, both approaches contribute to the advancement of conservation efforts. By offering useful data on whale shark population dynamics, migratory patterns, and habitat use, conservation efforts can be better targeted and optimised to the needs of the whale shark species.

Commonalities between the disadvantages of citizen science and AI are also found in the literature. Data reliability, consistency, and quality are issues that both citizen science and AI must contend with. Concerns about possible biases, mistakes, and discrepancies in data gathering procedures, especially during tourist season, occur in citizen science. Likewise, precise annotations are essential to AI, and errors or biases in the annotation process can have an impact on the validity of the data and subsequent interpretations. Additionally, environmental variables can have an impact on data gathering and analysis in both citizen science and AI. The quality of the data obtained from citizen science projects may be impacted by factors like underwater lighting and water clarity. Similarly, since AI-based computer vision systems depend on clear, high-quality photos, changes in environmental factors may reduce their ability to accurately identify individuals and extract meaningful data.

Certain contradictions are evident when comparing citizen science and AI side by side. While citizen science can provide a sizable quantity of data, data quality may be an issue owing to restrictions on sampling outside of the busiest tourist seasons or a lack of standardised procedures. As a contradiction, AI-based computer vision algorithms can deliver accurate and comprehensive data but may fall short in terms of contextual data and insights. Moreover, both approaches require resources that may not align with one another. Citizen research depends on public engagement and may need large resources to teach and engage people, assuring data quality. On the other hand, the implementation and upkeep of AI demands technical expertise, infrastructure, and computational resources.

Citizen science advantages can significantly help in overcoming the drawbacks of AI. While AI has limits in terms of being able to appropriately annotate data and handle a variety of environmental factors, citizen scientists can make up for these shortfalls. Citizen science's wider scope can aid in offsetting any biases or mistakes that may be introduced through the AI annotation process. Their observations can be a useful source of information and validation for AI systems, resulting in more accurate and thorough data processing. Additionally, citizen science provides long-term monitoring at a low cost, which helps overcome the difficulties presented by few resources and financial limitations in AI-based research. Moreover, knowledgeable and involved citizen scientists are more likely to provide accurate data, reducing biases and mistakes introduced during the AI training process. Therefore, citizen science can help increase public knowledge and support for conservation programmes through community engagement and awareness.

Equally, the advantages of AI can help mitigate the disadvantages of citizen science. The reliability and correctness of the datasets may be increased by automated data validation performed by AI, which also helps to minimise biases and errors introduced during data collecting by citizen scientists. AI systems may detect any flaws or inconsistencies in the data gathered by citizen scientists by analysing large volumes of image data efficiently. In doing so, the objective nature of AI-based identification and behaviour analysis eliminates the possibility of observer biases and contradictions in the data collected by citizen scientists. Moreover, AI aids in the development of standardised methodologies for data collection and analysis. Citizen scientists may obtain direction and feedback on their data gathering methods by using AI algorithms, maintaining consistency and increasing data quality.

In spite of complementary factors, some assumptions about citizen science and AI can be made when comparing the two approaches side by side. It is possible to make the assumption that because of differences in methodology, observer experience, or temporal and geographic coverage, citizen science data may be less accurate or subject to biases. On the other hand, there may be presumptions that data analysis based on AI is always reliable and impartial. Nevertheless, it is crucial to understand that AI models depend on the quality of annotated data used for training and may be subject to biases or mistakes if the training data is faulty or limited. Furthermore, citizen science emphasises community participation, enlisting volunteers and local groups in data gathering initiatives. One would assume that participation in citizen science fosters a higher feeling of ownership and accountability among participants and improves local awareness of whale sharks and their conservation. In contrast, there may be presumptions that AI lacks the capacity to engage communities or draw on local expertise because it is a technology-driven approach. However, it is crucial to understand that citizen scientists may contribute to AI by adding their skills and local knowledge to the processes of data gathering, annotation, and validation. In both citizen science and AI techniques, recognising the importance of community involvement and local expertise promotes a more inclusive and cooperative research environment.

Strengths and Limitations

This research has introduced a brand-new conceptual framework that unifies AI and citizen science, offering a distinctive viewpoint on how they function together in whale shark research and conservation. Future research in the area will benefit greatly from using this framework as a guide. The research also recognises the need of using appropriate terminology for citizen science and giving citizen scientists adequate recognition and tries to exemplify this. Additionally, the research effectively combines viewpoints from several academic fields, such as marine biology, citizen science, and artificial intelligence.

However, the study of the potential combination of citizen science and AI in whale shark research and conservation is not without limitations. The main drawback of this study is its reliance on a small number of data sources, namely the WildMe database and Google Scholar. Due to time and resource restrictions, this method may have led to data coverage and representation limitations, with potentially relevant sources like grey literature or specialised databases not being fully reviewed. Furthermore, different sources of relevant literature are available in different ways; some required special methods to obtain them (i.e. RUG licence privileges). Due to the confusing way in which its algorithm works and the errors in its logical operators, using Google Scholar also causes doubt in the data collection. There is no clear description on what basis Google Scholar presents its results and in what order. Interestingly, the literature extracted from the WildMe archive and Google Scholar do not coincide. This further exemplifies the lack of completeness of the Google Scholar algorithm. Additionally, even though the logical operators AND and OR were used, the majority of the literature did not mention both search terms connected by the operator. Furthermore, it was difficult to accurately identify and categorise material because of the ambiguity surrounding the terms "citizen science" and "AI". Lastly, as mentioned before, literature fails to recognise citizen scientists' contributions. Therefore, limited research on citizen science and its added value to the research and conservation of whale sharks may have constrained the representation in this research.

In terms of the found literature of this research, the lack of studies that thoroughly discuss both citizen science and AI in the context of whale sharks is a severe constraint. Although there are broad publications on the application of both citizen science and AI, except for the included literature, they were applicable to marine habitats, let alone particularly whale sharks. On a lot of land-based monitoring programmes, projects can be put in place that do not require humans at all. In whale shark research, humans will always need to collect data in the form of physical photography. At least, until advanced technology emerges which allows for observation without human intervention. Furthermore, the geographic emphasis of the study on whale sharks may restrict the applicability of the findings to other species or areas. Due to their distinctive traits and behaviours, whale sharks may require special considerations that other marine species may not require. Another notable limitation is the lack of studies addressing ethical issues specifically associated with combining citizen science and AI in whale shark research and conservation. Considering ethical issues in citizen science and AI is crucial when applying them to whale shark research and conservation to maintain the integrity of scientific research, guarantee ethical and sustainable practices, and safeguard the welfare of both whale sharks and the associated communities.

Research Gaps

The research can identify several gaps that exist in the field of whale shark research and conservation. Evidently, there is a research gap concerning the integration of citizen science and AI methodologies to maximise their respective advantages and get over drawbacks. There is a lack of knowledge on how to standardise methodologies and ensure quality control, which is essential in addressing the concerns about data consistency and dependability. Additionally, there is room for improvement in the contextual insights offered by AI-based computer vision algorithms, which may fall short in giving thorough contextual information yet excel in producing correct data. Another research gap includes the lack of knowledge in how to align the necessary resources for both citizen science and AI techniques is another area of study that needs improvement because their present demands for resources differ. The evaluation of the long-term effects of citizen science activities and their contributions to whale shark study and conservation can also be identified as a research gap. Researchers may advance the field and improve the success of whale shark conservation initiatives by filling in these research gaps.

Additionally, what is not yet clear are the ethical considerations of implementing citizen science and AI, both separately and combined. One important factor is giving citizen scientists the credit and respect they deserve for their work. There is no description of how citizen scientists and their generated output are accredited which can lead to discontinued engagement and motivation. This also comes into play when engaging with local communities. The exclusion or exploitation of citizen scientists and local communities can exacerbate existing power disparities and injustices and harm their long-term viability. When taking into account the possible dangers and effects on the wellbeing of whale sharks and their ecosystems, ethics also come into play. In order to ensure that data collecting techniques and activities do not hurt or disturb the animals' normal behaviour, citizen science initiatives should put the welfare and protection of the species first. Additionally, concerns about data confidentiality and privacy have not been addressed in the literature. To avoid potential injury or exploitation, it is crucial to preserve sensitive information and maintain privacy, such as the locations of whale shark aggregations or specific persons.

It is crucial for AI to be transparent about the constraints and unpredictabilities that come with analysis based on the technology, as well as to give detailed explanations of how AI is used in decision-making. The ethical usage of AI in connection to privacy and data protection, as well as the possible social and economic effects on nearby people that may be impacted by its use in whale shark study and conservation, are additional issues. To ensure the ethical and sustainable implementation of this technology, which is insufficiently represented in the literature, it is crucial to strike a balance between its advantages and these ethical issues.

Conclusion

This research has explored the current application of citizen science and AI on whale shark research and conservation. Through the lens of 9 scientific publications relevant to whale sharks, the research has identified advantages, disadvantages and implications of utilising citizen science and AI. Summarising, both citizen science and AI have positive implications and benefits for studying and protecting whale sharks when comparing the conservation implications. AI and citizen science complement one another and may be used together to improve whale shark research and conservation. While AI may enhance data processing, population monitoring, and decision-making through effective identification and analysis, citizen science can offer vital ecological insights, local expertise, and large-scale data gathering. This final section will conclude the research by answering the research question and recommending future research possibilities.

To answer the research question: "How can the combination of citizen science and artificial intelligence help the conservation efforts of whale sharks?" the following conclusion can be drawn from the discussion: AI and citizen science can improve whale shark conservation efforts together. Both strategies help to acquire more data, which improves our knowledge of the population dynamics, migratory patterns, and habitat use of whale sharks. While AI helps with data processing, validation, and standardisation, citizen science offers a more extensive range of data and community interaction. They can advance conservation planning, decision-making, and research for the benefit of the whale shark.

Future Research Recommendations

The research can recommend certain areas of future research from the limitations and research gaps. In general, future research should include a broader scope to further assess the implications of citizen science and AI. While the discussion highlights the complementary nature of citizen science and AI, further research is needed to explore how these approaches can be effectively integrated. This includes investigating how citizen science data can be utilised to enhance AI algorithms and how AI can provide guidance and feedback to citizen scientists for improved data collection and analysis. Another important aspect is the research on ethical considerations. Establishing channels for proper attribution, such as authorship or recognition in scientific papers, is crucial to recognising the significant contribution that citizen scientists make to producing data and ideas. Moreover, citizen scientists should be taught and instructed on ethical practices to guarantee appropriate encounters with whale sharks, and protocols and procedures should be devised to reduce any potentially harmful effects. With these recommendations, future research could benefit whale shark research and conservation even further than has been achieved in this research.

Bibliography

Adler, F. R., Green, A. M., & Şekercioğlu, Ç. H. (2020). Citizen science in ecology: a place for humans in nature. Annals of the New York Academy of Sciences, 1469(1), 52–64. https://doi.org/10.1111/nyas.14340

Araujo, G., Agustines, A., Tracey, B. H., Snow, S. K., Labaja, J., & Ponzo, A. (2019). Photo-ID and telemetry highlight a global whale shark hotspot in Palawan, Philippines. *Scientific Reports*, *9*(1). https://doi.org/10.1038/s41598-019-53718-w

Araujo, G., Ismail, A. G., McCann, C., McCann, D. R., Legaspi, C., Snow, S. K., Labaja, J., Manjaji-Matsumoto, B. M., & Ponzo, A. (2020). Getting the most out of citizen science for endangered species such as whale shark. *Journal of Fish Biology*, *96*(4), 864–867. https://doi.org/10.1111/jfb.14254

Araujo, G., Kwong, K. Z., Jones, I., Holmberg, J., Pierce, S., & Manjaji-Matsumoto, B. M. (2022). Citizen science as a key tool in whale shark conservation. Aquatic Conservation-Marine and Freshwater Ecosystems, 32(6), 1099–1100. https://doi.org/10.1002/aqc.3806

Araujo, G., Snow, S. K., So, C. L., Labaja, J., Murray, R., Colucci, A., & Ponzo, A. (2017). Population structure, residency patterns and movements of whale sharks in Southern Leyte, Philippines: results from dedicated photo-ID and citizen science. *Aquatic Conservation-marine and Freshwater Ecosystems*, *27*(1), 237–252. https://doi.org/10.1002/aqc.2636

Arzoumanian, Z., Holmberg, J., & Norman, B. (2005). An astronomical pattern-matching algorithm for computer-aided identification of whale sharks Rhincodon Typus. Journal of Applied Ecology, 42(6), 999–1011. https://doi.org/10.1111/j.1365-2664.2005.01117.x

Bonney, R., Cooper, C. B., Dickinson, J. L., Kelling, S., Phillips, T. B., Rosenberg, K. V., & Shirk, J. L. (2009). Citizen Science: A Developing Tool for Expanding Science Knowledge and Scientific Literacy. *BioScience*, *59*(11), 977–984. https://doi.org/10.1525/bio.2009.59.11.9

Chandler, M. A., See, L., Copas, K., Bonde, A. M., López, B. C., Danielsen, F., Legind, J. K., Masinde, S., Miller-Rushing, A. J., Newman, G., Rosemartin, A., & Turak, E. (2017). Contribution of citizen science towards international biodiversity monitoring. Biological Conservation, 213, 280–294. https://doi.org/10.1016/j.biocon.2016.09.004

Cooper, C. B., Shirk, J. L., & Zuckerberg, B. (2014). The Invisible Prevalence of Citizen Science in Global Research: Migratory Birds and Climate Change. PLOS ONE, 9(9), e106508. https://doi.org/10.1371/journal.pone.0106508

Dove, A. D. M., & Pierce, S. J. (2021). Whale Sharks: Biology, Ecology, and Conservation. CRC Press.

Hoffmayer, E. R., McKinney, J., Franks, J. S., Hendon, J. M., Driggers, W. B., Falterman, B. J., Galuardi, B., & Byrne, M. (2021). Seasonal Occurrence, Horizontal Movements, and Habitat Use Patterns of Whale Sharks (Rhincodon typus) in the Gulf of Mexico. *Frontiers in Marine Science*, *7*. https://doi.org/10.3389/fmars.2020.598515

Holmberg, J., Norman, B. M., & Arzoumanian, Z. (2008). Robust, comparable population metrics through collaborative photomonitoring of whale sharks Rhincodon typus. *Ecological Applications*, *18*(1), 222–233. https://doi.org/10.1890/07-0315.1

Magson, K., Monacella, E., Scott, C. M., Buffat, N., Arunrugstichai, S., Chuangcharoendee, M., Pierce, S., Holmberg, J., & Araujo, G. (2022). Citizen science reveals the population structure and seasonal presence of whale sharks in the Gulf of Thailand. *Journal of Fish Biology*, *101*(3), 540–549. https://doi.org/10.1111/jfb.15121

Meho, L. I., & Yang, K. (2007). Impact of data sources on citation counts and rankings of LIS faculty: Web of science versus scopus and google scholar. Journal of the Association for Information Science and Technology, 58(13), 2105–2125. https://doi.org/10.1002/asi.20677

Norman, B. M., Holmberg, J., Arzoumanian, Z., Reynolds, S. D., Wilson, R. P., Rob, D., Pierce, S., Gleiss, A. C., De La Parra, R. B., Galvan, B., Ramírez-Macías, D., Robinson, D., Fox, S., Graham, R. L., Rowat, D., Potenski, M., Levine, M., McKinney, J., Hoffmayer, E. R., . . . Morgan, D. J. (2017). Undersea Constellations: The Global Biology of an Endangered Marine Megavertebrate Further Informed through Citizen Science. *BioScience*, *67*(12), 1029–1043. https://doi.org/10.1093/biosci/bix127

Norman, B. M., Holmberg, J., Arzoumanian, Z., Reynolds, S. D., Wilson, R. P., Gleiss, A. C., Rob, D., Pierce, S., De La Parra, R. B., Galvan, B., Ramírez-Macías, D., Robinson, D., Fox, S., Graham, R. L., Rowat, D., Potenski, M., Levine, M., McKinney, J., Hoffmayer, E. R., . . . Morgan, D. J. (2016). Understanding constellations: The use of ‘citizen science’ to elucidate the global biology of a threatened marine mega-vertebrate. Hamad Bin Khalifa University Press (HBKU Press). https://doi.org/10.5339/qproc.2016.iwsc4.38

Pierce, S.J., Grace, M. & Araujo, G. (2021). Rhincodon typus (Green Status assessment). The IUCN Red List of Threatened Species 2021, e. T19488A1948820213.

Pierce, S. J. & Norman, B. (2016). Rhincodon typus. *The IUCN Red List of Threatened Species 2016*: e.T19488A2365291. https://dx.doi.org/10.2305/IUCN.UK.2016-1.RLTS.T19488A2365291

Theobald, E. J., Ettinger, A. K., Burgess, H. K., DeBey, L. B., Schmidt, N., Froehlich, H. E., Wagner, C. E. M., HilleRisLambers, J., Tewksbury, J. J., Harsch, M. A., & Parrish, J. K. (2015). Global change and local solutions: Tapping the unrealized potential of citizen science for biodiversity research. Biological Conservation, 181, 236–244. https://doi.org/10.1016/j.biocon.2014.10.021

Appendix I: Data Extraction Format

	Authors	Year	Journal	Open source	Context	Participants	Extent of citizen involvement	Extent of AI/machine learning involvement	Scale and location	Primary aim	Summary of results
1	Araujo et al.	2019	Scientific Reports	Open	The study investigates a global whale shark hotspot in Palawan, Philippines.	Researchers, citizen scientists	Citizen scientists contributed significantly (36%) to photo-ID data collection.	None	Honda Bay, Palawan, Philippines	To understand the movements and international migration of whale sharks in A sia.	Honda Bay is a global hotspot for whale sharks, with seasonal visits for feeding. Social media platforms were used to generate data, and citizen science programs were encouraged for monitoring efforts. The study confirmed the importance of Honda Bay as a critical habitat and identified the strait between Tawi-tawi (PH) and Lahad Datu (MY) as an important migratory corridor. International cooperation is needed to manage the endangered species.
2	Araujo et al.	2020	Journal of Fish Biology	Closed (RUG)	Conservation of endangered species (specifically whale sharks) through citizen science	Researchers, citizen scientists	High involvement in data collection, photo identification, and reporting	None	Global scale; Locations not specified	Opinion piece, inform, provide empirical evidence	Development of participatory science programmes involving citizen scientists following standardised protocols is encouraged. The use of the free global database Wildbook for Whale Sharks is an advantageous way of centralising data that facilitates citizen scientists' engagement.
3	Araujo et al.	2022	Aquatic Conservation-marine	Closed (RUG)	The study focuses on using citizen	Citizen scientists	Citizen science is actively involved in data collection, monitoring, and	None	Not specified	Enhancing whale shark conservation	The study demonstrates the valuable role of citizen science in whale shark conservation. It highlights how citizen scientists actively contribute to data

			and Freshwater Ecosystems		science as a key tool in whale shark conservation efforts. It highlights the importance of involving the public in data collection, monitoring, and conservation initiatives related to whale sharks.		conservation efforts related to whale sharks. Participants contribute sightings, photographs, and other relevant data to enhance scientific understanding and inform conservation strategies.				collection, monitoring, and conservation efforts. Their involvement enhances scientific knowledge and informs conservation strategies.
4	Araujo et al.	2017	Aquatic Conservation-marine and Freshwater Ecosystems	Closed (RUG)	The study is conducted in Southern Leyte, Philippines, focusing on the population structure, residency patterns,	Researchers, citizen scientists	Citizen science plays a significant role in the study. The researchers leverage the participation of citizen scientists to collect photo-ID data of whale sharks. The citizen scientists contribute by	It seems that the photo-ID analysis in the study was conducted manually by the researchers rather than relying on automated algorithms or software.	The study is conducted in Southern Leyte, Philippines, focusing on the local population of whale	The primary aim of the study is to investigate the population structure, residency patterns, and movements of whale sharks in Southern Leyte, Philippines, using dedicated	The study provides insights into the population structure and movement patterns of whale sharks in Southern Leyte, Philippines. The dedicated photo-ID efforts and citizen science participation contributed to the understanding of the residency patterns and movements of individual whale sharks. The study emphasises the importance of citizen science in expanding data collection efforts and fostering public engagement in

					and movements of whale sharks.		capturing photographs of whale sharks encountered during their diving and snorkelling activities.		sharks.	photo-ID efforts and citizen science participation.	conservation initiatives. The results highlight the value of photo-ID as a non-invasive method for studying whale sharks and provide information for their conservation in the study area.
5	Arzoumanian et al.	2005	Journal of Applied ecology	Open	The study is conducted to develop an astronomical pattern-matching algorithm for the computer-aided identification of whale sharks based on their unique spot patterns.	Researchers, citizen scientists	The study mentions the involvement of citizen scientists in the data collection process. The authors state that photographs of whale sharks were obtained from dedicated citizen science programs. The exact extent and details of the citizen science involvement in data collection are not provided in the available information.	The study focuses on the development of an algorithm using pattern-matching techniques, which can be considered a form of machine learning/AI involvement. The algorithm helps in the automated identification of whale sharks based on their spot patterns.	Not specified	The primary aim of the study is to develop and test an algorithm that can assist in the automated identification of whale sharks using their spot patterns.	The study successfully developed an astronomical pattern-matching algorithm that aids in the computer-aided identification of individual whale sharks based on their spot patterns. The involvement of citizen scientists in data collection highlights the collaborative effort between researchers and the public in gathering photographs of whale sharks. The algorithm shows promising results and can potentially contribute to the field of whale shark research and conservation efforts. However, specific details of the results are not mentioned in the provided information.
6	Hoffmayer et al.	2021	Frontiers in Marine	Open	The study aimed to investigate	Researchers, boat captains,	The study used a combination of data sources,	None	Gulf of Mexico	To determine the spatiotemporal patterns of whale	The study found evidence of seasonal patterns in the occurrence and movements of whale sharks in the Gulf of

			Science		the seasonal occurrence, horizontal movements, and habitat use patterns of whale sharks (Rhincodon typus) in the Gulf of Mexico.	and local fishermen (= citizen scientists)	including opportunistic sightings reported by boat captains and local fishermen and photo-identification data contributed by citizen scientists through the Wildbook for Whale Sharks platform.			shark occurrence and movements, as well as the environmental factors influencing their distribution.	Mexico, with higher abundance during the summer months. Citizen science involvement through photo-identification data provided valuable information on individual whale shark movements and helped expand the geographic coverage of the study. The study highlights the importance of collaboration between researchers, boat captains, local fishermen, and citizen science initiatives for the conservation and management of this endangered species.
7	Holmberg et al.	2008	Ecological Applications	Closed (Sci-Hub)	The study focuses on collaborative photo monitoring of whale sharks for robust and comparable population metrics.	Researchers	None	AI photographic identification	Not specified	Evaluate accuracy of AI photographic identification and mark-recapture model	Using computer-assisted photo-identification and opportunistic data collected from ecotourism is an effective approach to analyse and model mark-recapture data to ensure quality and volume of data required for population analysis.
8	Magson et al.	2022	Journal of Fish Biology	Closed (RUG)	This study explores the occurrence and seasonality	Divers, other marine resource users (= citizen	Citizen science is extensively involved in this study as a primary approach for collecting data on	None	The scale of this study was focused on the waters	The aim of this study is to investigate the occurrence, seasonality, and population	The study identified a total of 178 individual whale sharks from 249 sightings in the waters around Koh Tao, Thailand, and neighbouring islands between 2004 and 2019. The majority of these sightings (84%) occurred during the

					of whale sharks in the waters around Koh Tao, Thailand and neighbouring islands in the Gulf of Thailand using a citizen science-based approach.	scientists)	whale shark sightings. Divers, other marine resource users, and members of the general public contributed sighting reports, which were facilitated by social media and direct marketing. These citizen science contributions played a crucial role in providing information on the occurrence, seasonality, and population dynamics of whale sharks.		around Koh Tao, Thailand, and neighbouring islands in the Gulf of Thailand.	characteristics of whale sharks, using a citizen science-based approach and anecdotal reports to supplement the limited quantitative data available.	period of 2015-2019, likely due to increased reporting facilitated by social media and direct marketing. Size estimates were available for 102 sightings, ranging from 2 to 6 metres with an average of 3.7 metres. Sex information was reported for 27% of sightings, revealing a female-to-male ratio of 2:1. Modified maximum likelihood methods indicated that whale sharks are transient to Koh Tao and the surrounding areas, with their sightings following the regional monsoon cycle. Additionally, one international resighting was recorded from Malaysian waters, approximately 700 km away. The study highlights the value of citizen science participation in data-poor regions like the Gulf of Thailand, despite limitations in the reliability of size and sex estimation.
9	Norman et al.	2017	BioScience	Open	The study focuses on understanding the global biology of an endangered marine	Researchers, citizen scientists	Citizen science involvement is significant, with participants contributing photographic data and sightings to enhance the understanding of	None	The study has a global scale, covering multiple locations where whale	The primary aim is to utilise citizen science and collaborative efforts to gather data and enhance knowledge of whale shark biology,	The results highlight the value of citizen science in expanding the global dataset on whale sharks, providing insights into their movements, habitat preferences, and population structure, which can inform conservation strategies and management efforts.

					mega vertebrate.		whale shark biology and conservation. The study highlights whale sharks through the use of citizen science initiatives, particularly the 'Ecocean Whale Shark Photo-Identification Library' and the 'Whale Shark Watch' program. These initiatives involve the collection of photographic data and sightings contributed by citizen scientists and researchers.		sharks occur, including Australia, the Maldives, Mozambique, the Philippines, and the United States.	population dynamics, migration patterns, and habitat use for conservation purposes.	