



# Artificial intelligence for right whale photo identification: from data science competition to worldwide collaboration

Christin Khan<sup>1</sup> · Drew Blount<sup>2</sup> · Jason Parham<sup>2,4</sup> · Jason Holmberg<sup>2</sup> · Philip Hamilton<sup>3</sup> · Claire Charlton<sup>5</sup> · Fredrik Christiansen<sup>6,7,8</sup> · David Johnston<sup>9</sup> · Will Rayment<sup>9</sup> · Steve Dawson<sup>9</sup> · Els Vermeulen<sup>10</sup> · Victoria Rowntree<sup>11,12</sup> · Karina Groch<sup>13</sup> · J. Jacob Levenson<sup>14</sup> · Robert Bogucki<sup>15</sup>

Received: 7 December 2021 / Accepted: 6 April 2022

This is a U.S. government work and not under copyright protection in the U.S.; foreign copyright protection may apply 2022

## Abstract

Photo identification is an important tool in the conservation management of endangered species, and recent developments in artificial intelligence are revolutionizing existing workflows to identify individual animals. In 2015, the National Oceanic and Atmospheric Administration hosted a Kaggle data science competition to automate the identification of endangered North Atlantic right whales (*Eubalaena glacialis*). The winning algorithms developed by Deepsense.ai were able to identify individuals with 87% accuracy using a series of convolutional neural networks to identify the region of interest, create standardized photographs of uniform size and orientation, and then identify the correct individual. Since that time, we have brought in many more collaborators as we moved from prototype to production. Leveraging the existing infrastructure by Wild Me, the developers of Flukebook, we have created a web-based platform that allows biologists with no machine learning expertise to utilize semi-automated photo identification of right whales. New models were generated on an updated dataset using the winning Deepsense.ai algorithms. Given the morphological similarity between the North Atlantic right whale and closely related southern right whale (*Eubalaena australis*), we expanded the system to incorporate the largest long-term photo identification catalogs around the world including the United States, Canada, Australia, South Africa, Argentina, Brazil, and New Zealand. The system is now fully operational with multi-feature matching for both North Atlantic right whales and southern right whales from aerial photos of their heads (Deepsense), lateral photos of their heads (Pose Invariant Embeddings), flukes (CurvRank v2), and peduncle scarring (HotSpotter). We hope to encourage researchers to embrace both broad data collaborations and artificial intelligence to increase our understanding of wild populations and aid conservation efforts.

**Keywords** Automated matching · Automated pattern recognition · Computer vision · Individual identification · Machine learning · Right whale

## Introduction

This paper highlights the ongoing innovation in right whale (*Eubalaena sp.*) photo identification harnessing artificial intelligence (AI) in an attempt to increase efficiency in matching individual whales to further conservation efforts. We begin with an introduction to photo identification, and discuss previous work in this area as we continue to build on the success of the Deepsense algorithm that won the Kaggle competition (Bogucki et al. 2019). We then introduce the Flukebook platform which applies AI algorithms to the photo identification of whales and dolphins via a user-friendly website, and we go into detail about each of the right whale photo identification catalogs around the world that contributed to this project as we expanded from North

---

Handling editors: Leszek Karczmarski and Stephen C.Y. Chan.

---

This article is a contribution to the special issue on “Individual Identification and Photographic Techniques in Mammalian Ecological and Behavioural Research – Part 1: Methods and Concepts” — Editors: Leszek Karczmarski, Stephen C.Y. Chan, Daniel I. Rubenstein, Scott Y.S. Chui and Elissa Z. Cameron.

---

✉ Christin Khan  
christin.khan@noaa.gov

Extended author information available on the last page of the article

Atlantic right whales (*E. glacialis*) to include southern right whales (*E. australis*). Next we discuss detection, multi-feature matching, and algorithm evaluation and take a deeper dive into all of the different machine learning algorithms that we have implemented for both lateral (vessel, shore, cliffs) and aerial (aircraft, helicopter, uncrewed systems) viewpoints in this operational system for multi-feature matching of right whales. We conclude with a discussion of challenges faced, lessons learned, and achievements to date as we push the boundary forward on wildlife photo identification.

## Photo identification of right whales

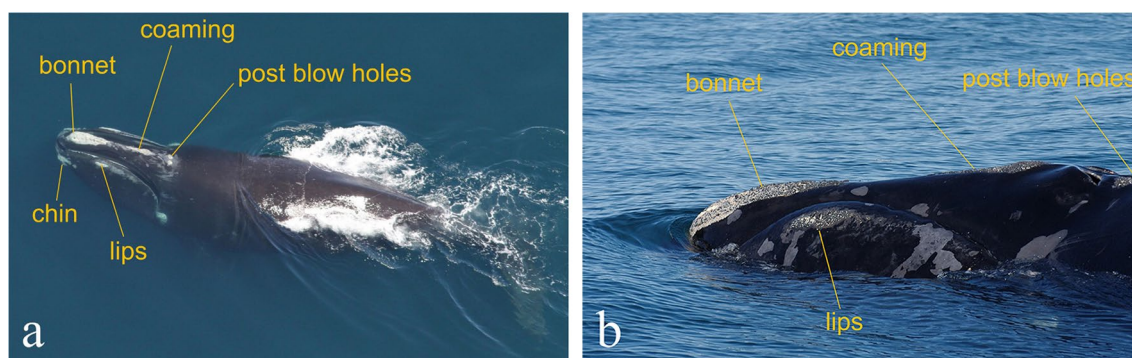
Although North Atlantic and southern right whales are geographically separated, they share many similar morphological and life-history traits (Fig. 1; Christiansen et al. 2020). Both species are well suited to photo identification because they can be individually identified by the callosity pattern on the top of their heads (Kraus et al. 1986; New England Aquarium 2019). Callosities are patches of rough skin colonized by cyamids (whale lice) that result in a distinctive white pattern against the otherwise black body (Payne 1976). The callosity pattern is unique for each animal and is used by researchers to identify each whale individually (primarily by the coaming, islands, lips, bonnet, chin, and mandibular islands) along with scars and other markings (Fig. 1; Appendix Figs. A1, A2).

Researchers take lateral photographs (from vessels, shore, cliffs) and aerial photographs (from aircraft, helicopters, uncrewed systems), and then match these images to existing photo identification catalogs. Individual photo identification is a crucial tool for estimating abundance and monitoring population trends over time (Calambokidis and Barlow 2006; Gormley et al. 2012; Pace et al. 2017; Koivu-niemi et al. 2019; Stamation et al. 2020; Crowe et al. 2021)

and makes nuanced understanding of individuals possible, including distribution, age at first calving, calving interval, adult and juvenile survival, individual health, entanglement rates, social relationships, site fidelity, and migratory patterns (see Karczmarski et al. 2022 for several relevant examples). The information gained from this research often has direct implications for the conservation and management of wild populations (e.g., in calculating Potential Biological Removal, drawing Seasonal and Dynamic Management areas, delineating Marine Protected Areas). Manually matching photographs to known individuals, however, can be very time-consuming and closely dependent on the skill of the researcher, the distinctiveness of the whale, the quality of the photographs, and the size of the dataset.

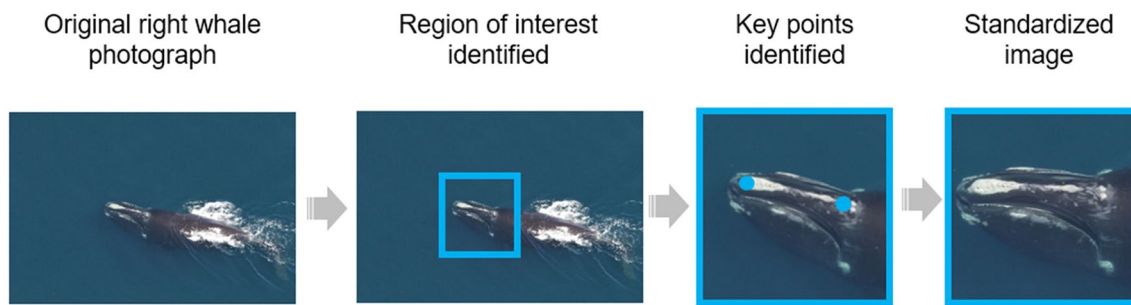
## Project background

Motivated by developments in the field of image recognition, the National Oceanic and Atmospheric Administration (NOAA) hosted a data science challenge on the crowdsourcing platform Kaggle in 2015 to automate the identification of endangered North Atlantic right whales. The winning solution by Deepsense.ai automatically identified individual whales with 87% accuracy using a series of convolutional neural networks to identify the region of interest on an image, rotate, crop, and create standardized photographs of uniform size and orientation and then identify the correct individual whale from these standardized photographs (Fig. 2). Recent advances in deep learning coupled with this semi-automated workflow have the potential to revolutionize traditional methods for the collection of data on populations with individually distinctive features (Bogucki et al. 2019). The Deepsense algorithm will be discussed in more detail in the “Machine learning” section of this paper.



**Fig. 1** Photographs of the North Atlantic right whale, *Eubalaena glacialis*, known as number ‘1706’ demonstrating some of the many features used in individual photo identification from **a** an aerial viewpoint from aircraft and **b** a lateral viewpoint from vessel. Aerial view-

point image collected under U.S. Marine Mammal Protection Act research permit number 775–1875 by NOAA/NEFSC/Christin Khan. Lateral viewpoint image collected by Lisa Conger/New England Aquarium



**Fig. 2** The winning solution in the NOAA “Right Whale Recognition” Kaggle Competition by Deepsense.ai automatically identified individual whales with 87% accuracy using a series of convolutional neural networks to identify the region of interest on an image, rotate, crop, and create standardized photographs of uniform size and orien-

tation and then identify the correct individual whale from these standardized photographs. Image collected under MMPA Research permit number 775–1875. Photo Credit: NOAA/NEFSC/Christin Khan. Figure adapted from Bogucki et al. (2019) with permission

## Flukebook

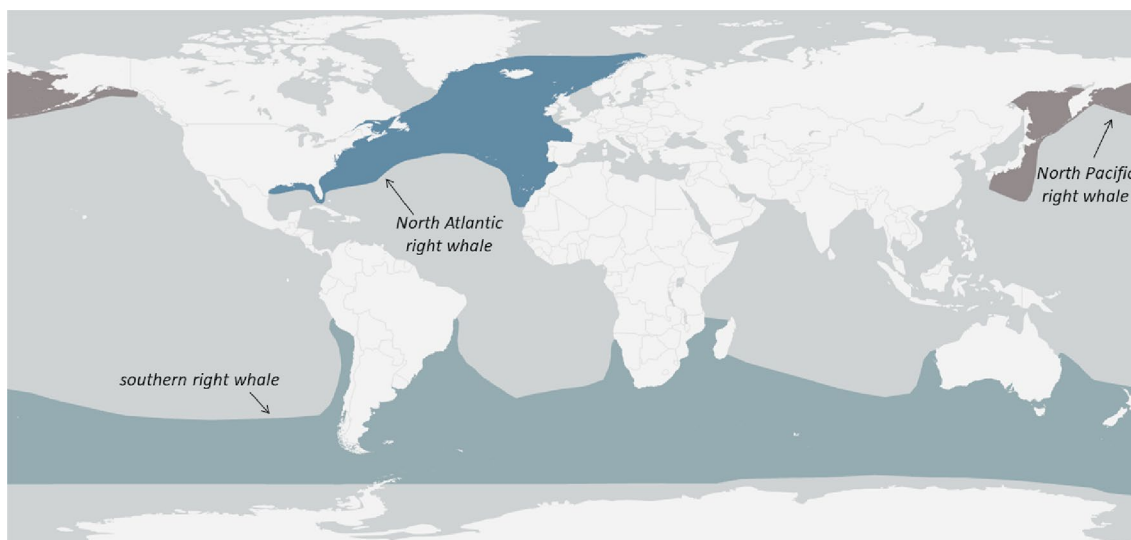
To make the AI algorithms for right whales broadly accessible to the research community, NOAA partnered with Wild Me, a 501(c) (3) non-profit organization. Their Wildbook project ([wildbook.org](http://wildbook.org)) is an AI-driven, open-source, and cloud-based software framework to support collaborative wildlife conservation. Wildbook uses computer vision to identify and detect animals in photos, creating data for individual and population analysis; this is integrated with genetic samples, habitat measurements, and geotags. Based on Wildbook, Flukebook (<https://www.flukebook.org>) is a flexible platform that integrates multiple computer vision algorithms for cetacean photo identification (Blount et al. 2022). Cetaceans often have large spatial distributions

spanning international boundaries, and therefore, cetacean conservation frequently requires collaboration between diverse stakeholders. Structured data sharing occurs on an opt-in, peer-approval basis, enabling inclusiveness across research teams and respect for data privacy and ownership.

## Right Whale catalogs

### Worldwide distribution of right whales

There are three species of right whales in the genus *Eubalaena*: the North Atlantic right whale, the North Pacific right whale (*E. japonica*), and the southern right whale (Fig. 3). Right whales received their name from the old whalers who considered them the “right” whale to hunt as they are



**Fig. 3** Historic worldwide distribution of right whales: the North Atlantic right whale (*Eubalaena glacialis*), the North Pacific right whale (*E. japonica*), and the southern right whale (*E. australis*). Map by Christin Khan, NOAA Fisheries using shapefiles from IUCN (2021)

generally slow and float when killed. Whaling pushed right whales to the brink of extinction and they received international protection under the International Convention for the Regulation of Whaling in 1935. Since this protection, the global population has increased to an estimated 14,000 southern right whales (IWC 2021) but less than 350 for the entire North Atlantic right whale species (Pace et al. 2017; NOAA Fisheries 2021). There are no reliable estimates for the North Pacific right whales but they likely number in the low 100s. While whaling is no longer a threat, right whales continue to face many anthropogenic threats including fishing gear entanglement, vessel strikes, climate change, habitat degradation, and noise pollution (Moore et al. 2004; Knowlton et al. 2012; Rolland et al. 2012; Meyer-Gutbrod et al. 2015; van der Hoop et al. 2017).

The three species of right whales are managed under various jurisdictions with regional photo identification catalogs housed by different organizations. Photographs ingested into Flukebook include the North Atlantic right whale photo identification records from the United States and Canada which are maintained in one comprehensive catalog curated by the New England Aquarium, and the southern right whale photo identification catalogs which are segregated by the location of their genetically distinct winter breeding grounds: Argentina/Brazil, South Africa, Australia, and New Zealand (Table 1). The North Pacific right whale is so rare and of such low abundance (Muto et al. 2021) that they have not yet been included in this effort to increase matching efficiency using AI. This effort to bring researchers from around the world together for collaboration and efficiency is tremendously valuable and a model for other conservation efforts. We go on to discuss each of these photo-identification catalogs in more detail in the following sections.

## Data ingestion

Ingesting this large volume of data from right whale catalogs worldwide nearly doubled the number of photos stored in “Flukebook”. Flukebook’s cloud-based architecture allowed the developers to scale up relatively easily by increasing the size of the server. Additionally, we significantly reduced data overhead by retraining on the entire 400,000-image North Atlantic right whale catalog but then only importing the first encounter and the most recent 10 encounters of each individual, so the most relevant photos for matching would still be available while conserving server resources.

One challenge in blending data from multiple catalogs was whether any of the datasets might have seen the same individual whale and given them different catalog names/numbers. While a major goal of this effort is to identify such cases, it is important to remove them from the training data to prevent giving an unreliable ground truth to the algorithm during training: including the same whale with multiple individual labels would confuse the system during training and result in a less reliable model. To address this, we ensured that only fully reconciled catalogs of distinct populations were included in training sets for the individual identifiers: Australasian Right Whale Photo-Identification Catalogue (ARWPIC) from Australasia, the University of Utah/Instituto de Conservacion de Ballenas Argentine dataset for South America, and the University of Pretoria dataset for South Africa. While there may still be a small number of these multi-label conflicts due to either missed matches in a given catalog or whales that travel between populations, this strategy minimized such cases without requiring full reconciliation between the catalogs which would be labor-intensive.

**Table 1** The number of individual whales and the number of images in each of the photo identification catalogs ingested into the Flukebook platform for machine learning

Species	Country	Institution	# Whales	# Images
North Atlantic Right Whale	United States and Canada	New England Aquarium	736	400,000
Southern Right Whale	Australia	Curtin University & Eubalaena Pty. Ltd (Head of the Bight)	1801	12,311
Southern Right Whale	Australia	Australian Antarctic Division (ARWPIC)	1932	6103
Southern Right Whale	Australia	Murdoch University/Aarhus University (Head of the Bight drone images)	578	8270
Southern Right Whale	South Africa	University of Pretoria	2083	8461
Southern Right Whale	Argentina	University of Utah, Instituto de Conservacion de Ballenas (Peninsula Valdez)	3350	8952
Southern Right Whale	Brazil	Instituto Australis	306	5473
Southern Right Whale	New Zealand	University of Otago	793	2913
<b>Total</b>			11,579	452,483

## North Atlantic right whales: United States and Canada

The North Atlantic Right Whale Catalog began as several separate catalogs housed by different investigators in the early 1980s. In 1986, the North Atlantic Right Whale Consortium was formed and a central catalog maintained at the New England Aquarium in Boston, MA was created. The consortium consisted of the New England Aquarium, University of Rhode Island, Center for Coastal Studies, Woods Hole Oceanographic Institution, and Marineland of Florida. The catalog houses all known photographed sightings of this species, which span from the Gulf of Mexico north along the eastern seaboard of North America, with additional sightings as far north and east as northern Norway and France. This range includes the calving grounds off the southeastern US as well as numerous feeding grounds to the north. The catalog includes opportunistic sightings dating back to 1935 and sightings from dedicated right whale surveys dating back to the late 1970s. Each year, multiple organizations perform aerial and vessel-based surveys throughout the species range and images and data from the subsequent 3000 to 5000 photographed sightings are submitted to the catalog.

As of July 2019, there were over 78,600 sightings of 742 different individuals—about 40% of those are known or suspected to be dead. The catalog originally consisted of slides and prints, but starting around 2003, the majority of the images submitted have been digital images. There are well over a million images of all types combined in the catalog; these include a mix of aerial and shipboard images (roughly 40/60) with some also taken from land and unmanned aerial vehicles (UAVs). A small subset of the prints and slides from the past have been digitized and the remainder are stored in filing cabinets. In 2005, the New England Aquarium developed the DIGITS software (Digital Image Gathering and Information Tracking System) to manage all the images and associated data. The images and data are all stored in a Microsoft SQL database and accessed and managed through the software. Each image is coded for body part and view direction, and each sighting is coded for identification information from that sighting (i.e., what the whale looks like) as well as for behaviors and associations. The images are used for many purposes in addition to individual identification including a detailed assessment of human-caused scarring, health assessments, and behavioral assessments. For this reason, many images from each sighting are kept—not just the best one or two, as is the case with some catalogs.

Manual matching is done using DIGITS. Sightings from a specific day, area, observer, or callosity pattern are shown on the left side of the screen. A detailed search screen allows the user to find sightings or whales that are coded in similar ways (whales are coded based on a compilation of

information from all their sightings whereas sightings are coded using information from that one sighting). Matching candidates are reviewed and logged as potential matches, definite matches, or definite not-matches. The user can opt to not see any candidate that a previous matcher has determined is not a match. Any match that is made is then confirmed by a second researcher using a confirmation console. That console will not let a researcher confirm their own work. Identifications are automatically entered in the database as soon as they are confirmed, and then that sighting becomes visible on the public website: <http://rwcatalog.neaq.org>.

The Flukebook interface was integrated into the public-facing website for the North Atlantic Right Whale Catalog (<http://rwcatalog.neaq.org/#/>) which contains the comprehensive list of known North Atlantic right whales along with associated metadata, such as their date of birth and death, sex, parentage, and sighting history. Photographers can choose to match the identity of the whale in their images by scrolling through all cataloged whales, searching by matching features, or using Flukebook's semi-automated matching system. When the Flukebook platform is used to identify the individual whale in the photograph, the top 12 most likely matches are displayed by default from the algorithm results, and the user can opt to view all results if desired. Next to the proposed individual matches is a link to each whale's

The screenshot shows a web interface for a whale catalog. At the top, it says 'Catalog #3560 - SNOW CONE'. Below this is a form with fields for 'Sex' (Female), 'Calving Female' (Yes), 'Mother' (1308), 'Death Year' (empty), 'Birth Year' (2005), 'Last Year Seen' (2020), and 'Father' (1155). Below the form is a section titled 'Whale Composite and Images'. It has dropdown menus for 'Body Parts' (All) and 'View Direction' (All), and a 'Filter' button. The main content area features a large composite drawing of a whale's head and flukes, labeled '2005 Calf Of 1308'. Below the drawing are several small thumbnail images of the whale in various poses and environments.

**Fig. 4** The North Atlantic Right Whale Catalog page for the whale named *Snow Cone* (also known as #3560) including composite drawing and exemplar photos. Here we learn that she is a calving female born in 2005 to the mother #1308 and father #1155. There are also links to additional photos and sighting history (<http://rwcatalog.neaq.org/#/whales/3560>)

sighting page on the North Atlantic Right Whale Catalog website to compare photographs and confirm the ID (Fig. 4).

### Southern right whales: Australia

Aerial and lateral photo identification data are collected in Australia from cliff-based, vessel-based, and aerial platforms including fixed-wing aircraft and UAVs. The Australasian Right Whale Photo Identification Catalogue (ARWPIC) is managed by the Australian Antarctic Division and provides a national database for processing, curating, and storing southern right whale identification and sightings data. ARWPIC uses the BigFish Code Compare matching software and includes over 6,000 images of approximately 2000 individuals, curated between 1975 and 2018 (and 2011 for some datasets). Data are contributed by key researchers across Australia that manage regional research programs. Photo identification data are collected from across the southern right whale range of Australia from Exmouth, Western Australia in the west to Queensland in the east. However, data are primarily contributed from three long-term monitoring programs: (1) Annual aerial surveys of southern right whales in Australia were initiated in 1975 by John Bannister of the Western Australian Museum and are currently managed by the Western Australian Museum, Murdoch University, and the Australian Antarctic Division. Aerial surveys are completed annually between Cape Leeuwin Western Australia and Ceduna in South Australia and contribute aerial images from fixed-wing platforms (Smith et al. 2021); (2) Shore-based research at the major calving ground at the Head of the Great Australian Bight, South Australia has been undertaken annually since 1991 and managed by Eubalaena Pty. Ltd. and Curtin University and collects aerial and lateral images from cliff-top vantage points 33–53 m high (Charlton et al. 2019). Since 2016, UAV aerial images from the Head of Bight have been contributed by Murdoch University/Aarhus University (Christiansen et al. 2018). (3) The South Eastern Australia Southern Right Whale Photo Identification Catalogue (SEA SRW PIC) is managed by the Department of Environment, Land, Water, and Planning and includes aerial and lateral images from systematic aerial surveys and land-based opportunistic data collected from 1980–present in southeastern Australia including Victoria, Tasmania, New South Wales, and Queensland (Stamation et al. 2020; Watson et al. 2021).

### Southern right whales: Argentina

Annual aerial surveys of southern right whales off Peninsula Valdés, Argentina were initiated by Roger Payne in 1971 after he realized he could recognize individual whales from the pattern and shape of white markings (callosities) on their heads (Payne 1976), and have been conducted annually since

then by Ocean Alliance / Instituto de Conservación de Ballenas. Surveys are conducted using a high-wing light plane that flies along the perimeter of the Peninsula at the time of peak whale abundance and circles over every whale or group of whales encountered while their locations and presence of calves are recorded. A sequence of photographs is taken of each individual's callosity pattern and/or other distinctive markings on their bodies. Before 1990, the whale photographs were searched for manually in 3-ring notebooks of photographs sorted by types of callosity patterns according to the shape of the rear margin of the bonnet and the shape and placement of other callosities. As the catalog grew, the manual system became unwieldy, stimulating the development of a computerized catalog and photo-matching system (Hiby and Lovell 2001). The Hiby–Lovell system creates a 2-dimensional extract of a whale's 3-dimensional callosity pattern and searches the catalog of extracts for potential matches. The system then returns a file with photographs of potential matches ordered from best to worst match. During analyses, the first 70 whales are examined before concluding there is no match and it is a new whale. A unique feature of the Hiby–Lovell system is its ability to create distortions of the image to compensate for lateral views by projecting the extract as if on a balloon and then pulling and pushing to create different distorted perspectives of the pattern which are then searched for in the catalog. Boat-based photographs contributed to this study by whale watch boat photographers in Puerto Pirámides are also being identified using Australia's BigFish system. A total of 8952 images of 3350 individually identified whales taken from aircraft, cliffs and vessels in this catalog were ingested into Flukebook.

### Southern right whales: Brazil

Aerial surveys of right whales have been conducted off Brazil since 1987, irregularly at first (1987, 1988, 1992, 1993, and 1994) and then annually from 1997 to 2018 (except for 2014) during peak whale abundance in September with additional months in some years. Surveys are flown parallel to the shoreline along a 120 km stretch of the coast at 1000 feet altitude and 60 knots. Photographs are cataloged and analyzed using the Hiby–Lovell Right Whale Identification System (Hiby and Lovell 2001). A total of 5473 aerial images (aircraft) of 306 uniquely identified individuals in this catalog were ingested into Flukebook.

### Southern right whales: South Africa

Annual aerial surveys of southern right whales in South Africa were initiated in 1979 by Professor Peter Best, affiliated with the Sea Fisheries Research Institute of the South African government, and since 1985 affiliated with the Mammal Research Institute Whale Unit (MRIWU) of the

University of Pretoria, where it is still managed. These aerial surveys, and the resulting photo identification catalog comprising only aerial images, cover the area between Nature's Valley and Muizenberg, which is the main concentration area of the South African breeding ground. All photo identification was matched by manual comparison until 2005, when the Hiby–Lovell semi-automated computer-based image recognition system was adopted (Hiby and Lovell 2001). This system utilizes digitized extracts of the callosity patterns (automatically adjusted for tilt and inclination) to match individuals. Semi-automated comparisons of callosity patterns are rated for similarity using an index from 0–1, with the highest scoring candidate presented first. Manual comparison of photographs begins with the highest scored match and continues until a match is made or until the index has fallen to 0.50. The photograph is correctly identified within the first 3 candidates in over 90% of the cases. The dorsal pigmentation features are used to confirm a match, and can be particularly useful in linking photographs of adults to those taken when they were calves (even in the absence of a callosity pattern match). In its 41st year of continuous annual surveys, the South African southern right whale catalog includes 9492 images of 2321 uniquely identified individuals. A total of 8461 images of 2083 individually identified whales taken in this catalog were ingested into Flukebook.

### Southern right whales: New Zealand

There are two contemporary photo identification catalogs for southern right whales in New Zealand's subantarctic Auckland Islands; one for lateral images taken from boats, and one for aerial images taken from UAVs. All images have been taken during the austral winter in Port Ross, the only known consistent calving area for southern right whales in New Zealand waters (Rayment et al. 2012; Torres et al. 2017). The lateral catalog contains images gathered during annual expeditions to the Auckland Islands between 2006 and 2018 (although note a two-year hiatus from 2014–2015). Images have been contributed by teams from the University of Otago, University of Auckland, Department of Conservation, New England Aquarium, and Australian Antarctic Division. The aerial catalog contains images taken from UAVs during the expeditions from 2016 to 2018 (for methods see Dawson et al. 2017). The photo identification catalogs are stored as two separate custom Microsoft Access databases (BigFish; Pirzl et al. 2006). Comparison of images is facilitated by classification of each individual according to a suite of 17 distinguishing characteristics (e.g., nature of lip callosity, number of rostral islands). The database can then be queried each time a new image is compared to the existing photo-identification catalog. At the time of writing, there are 1064 individuals in the catalog of lateral images

(matching is yet to be completed for sub-adults and non-breeding adults from 2016 to 2018), and 639 individuals in the aerial catalog. Reconciliation between the lateral and aerial catalogs has yet to be attempted, so many individuals will feature in both. Currently ingested into Flukebook are all individuals from the lateral catalog, and all individuals encountered in years 2016–2017 from the aerial catalog.

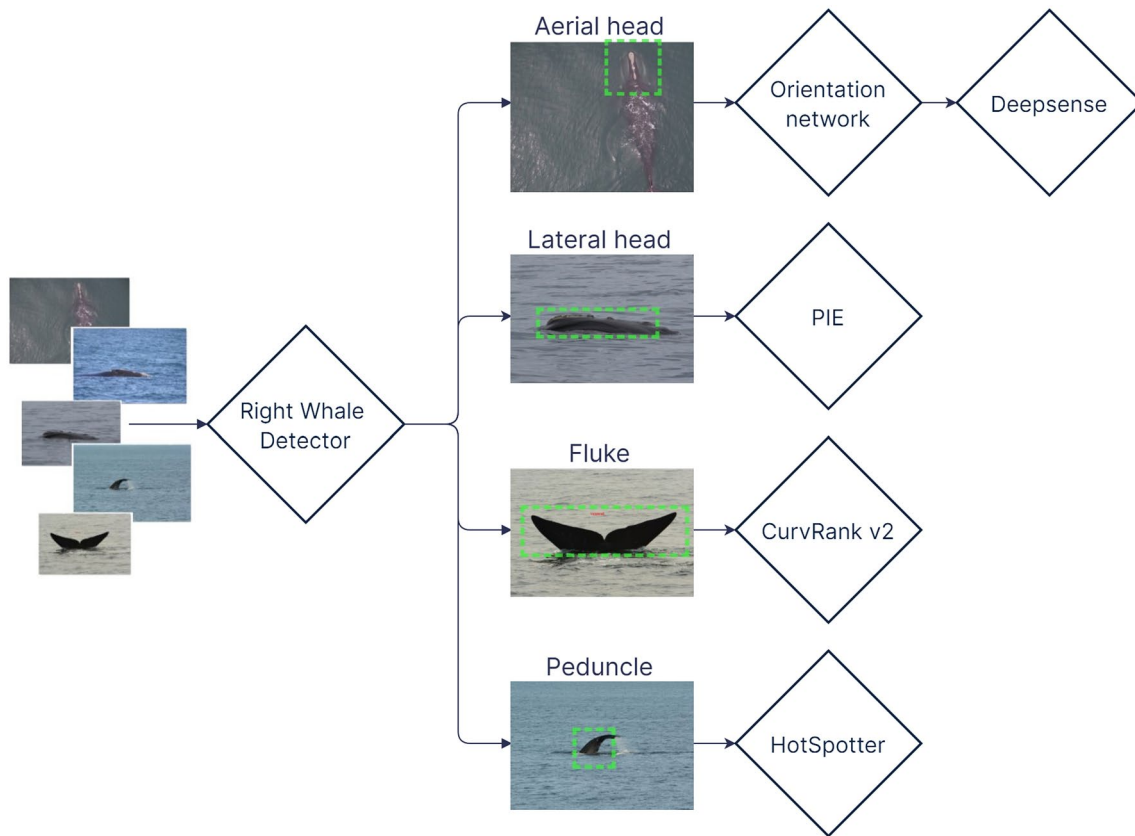
## Machine learning

### Detection

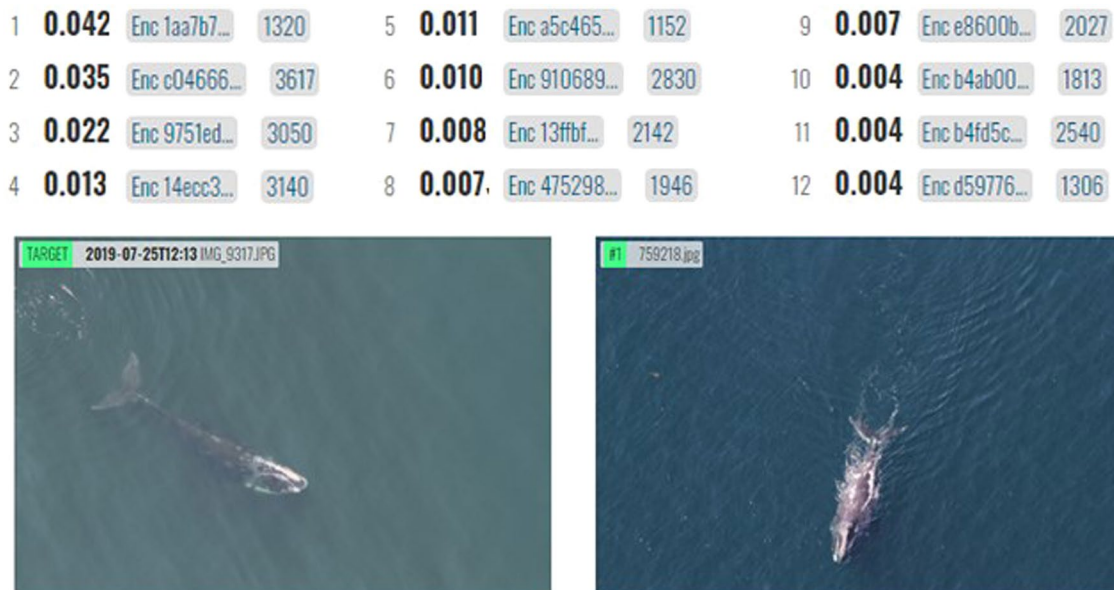
The first step in the Flukebook image processing pipeline is *detection* (architecture described in detail in Parham et al. 2018), which is the automated drawing of bounding boxes (rectangles overlaid on an image) around features of interest and then labeling those features (right whale head, body, etc.). The method here uses deep learning, a popular tool for animal detection (Yousif et al. 2017; Verma and Gupta 2018; Schneider et al. 2018). The detection models are implemented as a You Only Look Once (YOLO) V2 detector (Redmon et al. 2017) and are trained on NVIDIA Graphics Processing units (GPUs) over approximately 20 h. Each detector produces a large, fixed-length collection of bounding boxes with species labels, and then non-maximum suppression (NMS) is used to duplicate similar detected regions. The detector can be trained to detect any axis-aligned box (i.e., parallel to the image boundaries), representing an entire animal or a feature of interest. The correct NMS configuration and score threshold are selected with held-out (20%) validation images from the training set. After the boxes are created, a second convolutional neural network predicts aerial viewpoints (heads) and lateral viewpoints (lateral heads, full bodies, flukes, peduncles).

### Multi-feature matching

This AI-powered photo identification system was expanded to allow identification via multiple features (head, peduncle, fluke) and viewpoints (overhead from an airplane or lateral from a vessel) in 2021. Matches can be restricted to a dataset or geographic location if desired to speed up processing time. The platform can now support this multi-feature, multi-algorithm matching with a new image intake platform (Fig. 5) that assigns annotations (viewpoint, body part) and passes them to one or more appropriate algorithms. Model results are presented independently for each pipeline (aerial head, lateral head, fluke, peduncle), and for each image the user is presented with the most likely candidate matches based on the appropriate algorithm for that viewpoint and body part (Fig. 6). This multi-modal, multi-feature, and multi-species machine learning architecture



**Fig. 5** Pipeline for the annotation and identification of multi-feature matching of right whales in Flukebook (<https://www.flukebook.org/>) with the algorithm of choice for each body part shown in the node on the far right



**Fig. 6** Example output from running one of the algorithms in Flukebook, in this case, the Deepsense algorithm that matches based on the head from an aerial viewpoint. By default, the top 12 candidate matches are shown with match scores in bold on the left, followed by

two gray boxes containing the encounter IDs and the 4-digit whale identification numbers. The image on the left side is the one submitted to the system for matching and the image on the right side is another image of the candidate whale being suggested as a match



**Table 2** Model performance for the different matching algorithms applied to North Atlantic right whales (results not yet available for southern right whales)

	Model accuracy (%)		Data used to evaluate model		
	Top 1	Top 5	# Query images	# Database images	# Individual whales
Head: Deepsense (aerial)	89	98	6347	6347	668
Head: PIE (lateral)	55	81	531	3508	64
Peduncle: HotSpotter	17	18	1334	1334	339

The ‘top 1’ accuracy is the number of times the model suggests the correct individual when only allowed to output one suggestion. The ‘top 5’ accuracy is the number of times that the correct individual can be found in the top five suggested matches returned. The Deepsense accuracy here reflects the retraining by the Flukebook team which resulted in slightly higher accuracy than the original Kaggle competition result of 87% top 1. “Query images” are the number of queries computed to produce accuracy results, “database images” are the number of images being matched against in those queries. In cases where the two numbers are the same, every image in the evaluation set was queried against all other images. In all cases, data were taken from held-out test sets, i.e., images that were not included in training

(“Wildbook Image Analysis; WBIA”) has subsequently enabled new advancements in photo identification for other species, including simultaneous killer whale (*Orcinus orca*) saddle patch and dorsal matching from a single photo using the Pose Invariant Embeddings (PIE) and CurvRank v2 algorithms.

## Evaluation

When training machine learning models and then evaluating their accuracy, we followed the standard practice of dividing the data into a training set of 80% of the data size and a test set with the remaining 20%. This allowed us to evaluate algorithm performance using data the models had not used during training. Top 1 and top 5 accuracy are commonly used criteria for evaluating individual identification algorithms which return a list of candidate matches. Top 1 and top 5 accuracy are the number of times that the model outputs the correct identification when allowed to output only one and only five possible whale identifications, respectively. Table 2 presents the top 1 and top 5 accuracy of the algorithms applied to North Atlantic right whale photo identification. Note that for the trained algorithms (Deepsense and PIE), these figures were computed on held-out test datasets which were not included in training data.

## Deepsense: aerial photos of the head

The Deepsense algorithm that won the Kaggle data science competition consists of several convolutional neural networks and is both fast and accurate. Neural networks are an established family of machine learning models able to learn and perform tasks by adjusting a network of artificial neurons that respond to an input layer, hidden layer(s), and then output to the next layer. Convolutional neural networks (CNN) are a variant of neural networks that have been successfully applied to computer vision problems, such as

image classification (He et al. 2016; Ma et al. 2018; Tan et al. 2019), image segmentation (Long et al. 2015; Ronneberger et al. 2015; Girshick et al. 2018), object detection (Ren et al. 2015; Redmon et al. 2016; Hu et al. 2017), microscopy (Xing et al. 2017), and for photo identification of humans (Taigman et al. 2014; Zhang et al. 2017; Phillips et al. 2018) and other animals (Crall et al. 2013; Bogucki et al. 2019; Moskvayak et al. 2019; Weideman et al. 2020; Cheeseman et al. 2022; Clapham et al. 2022; de Silva et al. 2022; Blount et al. 2022). As the number of layers in the neural networks increased, the term “deep learning” was coined (Dechter 1986; Aizenberg et al 2000).

In the Deepsense algorithm, multiple CNNs are used to produce a standardized image of each whale by identifying the region of interest from key points on the head, cropping the image, and aligning the original photograph to those keypoints, and then finally identifying the individual whale from this cropped and rotated photograph. Identifying the region of interest was an important first step because the whale occupies only a tiny fraction of the image. Three different CNNs make up the Deepsense model—the first CNN identifies the region of interest, the 2nd CNN identifies key points on the whale, and the 3rd identifies the correct individual. The CNN used to identify the region of interest used 5 convolutional layers interspersed with 5 pooling layers, followed by a fully connected layer. The CNN used to identify the key points on the whale (the tip of the bonnet and just below the blowholes on the head) used 9 convolutional layers, most of them followed by pooling layers, and a fully connected layer. Once the key-point positions were identified, the image was scaled, rotated, and cropped so that the whale's head occupied a predefined position. CNNs do not have scale and rotation invariance, so rotating and cropping the images into a standardized format improved performance much in the same way that human passport photos are easier to identify than random snapshots with more variable conditions. Finally, the most complex CNN was used to perform

actual whale identification, which had 11 convolutional layers, 6 pooling layers and a fully connected layer. More details on the architecture of the Deepsense algorithm can be found in Bogucki et al. (2016, 2019). This model was able to match the photograph to the correct individual right whale in 87.44% of the 2493 withheld images during the Kaggle data science competition. Challenges identified included the difficulty of discriminating such similar images. Accuracy could likely be improved by excluding challenging images (sun glare, poor angle, partially visible whale) however that would undermine the utility of the system for the end user.

The software and machine learning engineers at Wild Me retrained the Deepsense algorithm on an updated dataset from the North Atlantic Right Whale Consortium in collaboration with the original algorithm development team at Deepsense who provided the source code and documentation. Expanding the pool of training data allowed for new individuals to be added to the system as well as to create a more robust and generalizable model with a wider variety of images (Appendix Fig. A3). Adding the Deepsense matcher to Flukebook involved wrapping it into Flukebook's backend computer vision system, which like the Deepsense algorithm itself is written in the Python programming language; this made both retraining the algorithm and its integration into Flukebook relatively straightforward. Further, the Deepsense result was *containerized*, which means the algorithm was provided not just as source code but as a stand-alone program bundled with its various software requirements that can be easily integrated into other servers. The integration into Flukebook then consisted of launching and maintaining the Deepsense container from the Flukebook backend, then sending and receiving matching jobs between the two. Once fully integrated, using the Deepsense matcher is the same as using any other algorithm on Flukebook; users can send match queries via an overlay menu on their images. The Deepsense algorithm went live on Flukebook in 2019 for semi-automated matching of aerial photographs of right whale head callosity patterns for North Atlantic right whales.

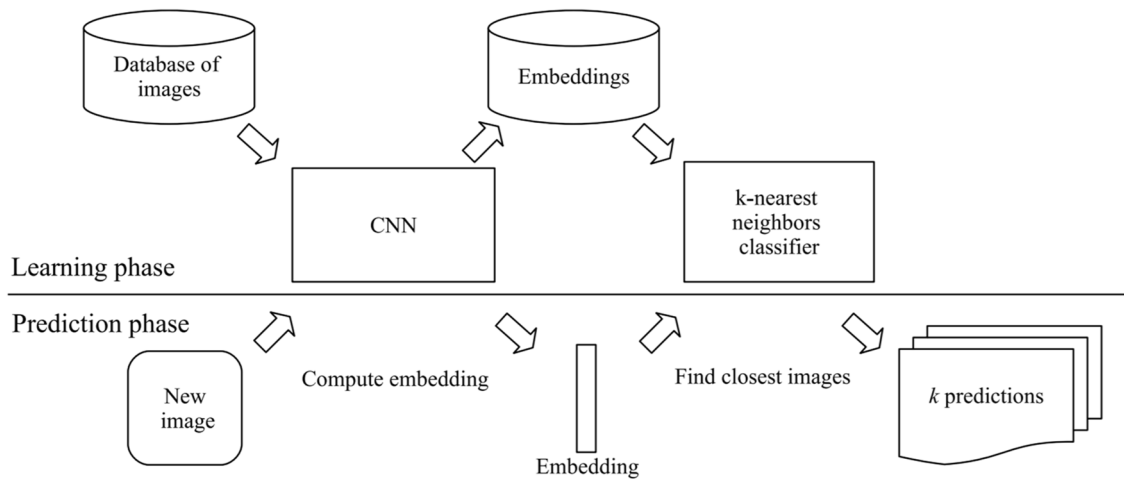
The AI infrastructure built for North Atlantic right whales was expanded to include southern right whales in 2020 (Appendix Fig. A4, Table 1). After filtering for known individuals with at least 2 sightings, the total dataset of 10,451 photographs was used to train a southern right whale model of the Deepsense algorithm. The average resight rate of southern right whales was only 4 sightings per individual (compared to 88 in the North Atlantic), which resulted in significantly less accurate models. Further funding and research are required to make the algorithm more generalizable so that the southern right whale model can more closely approach the North Atlantic model in accuracy. Generalization refers to a model's ability to draw from the data

used to create the model and apply that to new, previously unseen data.

The Deepsense algorithm is a *classifier*, meaning that it processes each input query (photo of a right whale) into one of its predefined classes (individual whales). Classifiers aim to achieve high accuracy and support a very large number of classes, however, they are architecturally limited to assigning each input to only those classes on which they were trained. This means that the Deepsense algorithm can identify only the individuals that were present in its training set. Newborn individuals or those that have been seen for the first time cannot be matched without retraining the algorithm, which is a computationally expensive process and can take up to a week of server time. Since the Deepsense algorithm is optimized for right whales (*Eubalaena* sp.), it cannot be cross-applied to other species without a new training procedure and tweaking of the architecture. These challenges are discussed further in the 'Lessons Learned' section where we recommend 'Image Similarity Algorithms' for future wildlife photo identification models.

### PIE: lateral photos of the head

The Pose Invariant Embeddings (PIE) algorithm was developed by Olga Moskvayak at Queensland University of Technology originally to identify manta rays (previously genus *Manta* but now included in genus *Mobula*) based on the pattern on their ventral surface as well as humpback whale flukes (*Megaptera novaeangliae*) (Moskvayak et al. 2019; Fig. 7). Unlike Deepsense, the PIE neural network is not trained to classify images into bins of individuals. Instead, its deep neural network is trained to extract embeddings (abstract, numerical representations of annotations from images). Given an annotation, the PIE algorithm returns a list of 256 numbers between 0 and 1. PIE is trained with the simple concept that images of the same individual should produce similar embeddings, and images of different individuals should produce different embeddings; the distance between any two images in embedding space corresponds to the similarity between those two images for individual identification (this is an older concept of similarity in machine learning, see e.g. Bromley et al. 1993). The photos of a single individual then represent a cluster in the space of all embeddings, and individual identification is performed by computing the embedding for a new query image and finding which cluster is nearest that embedding using the k-nearest neighbor algorithm. An advantage of this approach is that PIE can gracefully add new individuals to its catalog without being retrained: it learns the general task of mapping images into embeddings that represent individuals, rather than the specific task of sorting images into a fixed number of IDs. PIE strikes a balance between a flexible general-purpose identifier and one that can be trained and refined on



**Fig. 7** Pose Invariant Embeddings (PIE) algorithm workflow for matching lateral photographs of right whale heads taken from a vessel on the Flukebook platform. From Olga Moskvayak's open-source GitHub repository: <https://github.com/olgamoskvayak/reid-manta>

a given problem. Moskvayak's publicly available PIE algorithm (<https://github.com/olgamoskvayak/reid-manta>) was trained and tuned on lateral photos of right whale heads by D. Blount at Wild Me, resulting in a model with a top 1 accuracy of 55% and top 5 accuracy of 81% on a held-out test dataset. Like the training data, these test data included only individuals with multiple photos so that a correct match was possible. Top 1 accuracy of 55% means that in 55 out of 100 attempts to match a photograph, the model returns the correct individual identification when given only one output or, in other words, the correct whale is listed as the first choice. Top 5 accuracy of 81% means that in 81 out of 100 attempts to match a photograph, the model returns the correct individual identification as one of the whales listed in the first five results. This accuracy level reflects both the challenging nature of these photographs and that this is the first-ever automatic system for matching boat-based photos of right whales. Application of PIE to North Atlantic right whales has also subsequently enabled new advancements in gray whale (*Eschrichtius robustus*) and killer whale lateral matching, as well as sperm whale (*Physeter macrocephalus*) fluke matching, in [Flukebook.org](https://www.flukebook.org).

Recently, the PIE neural network algorithm has been applied to lateral photos of southern right whales in New Zealand. The user experience was satisfying, with the process of bulk uploading and matching relatively straightforward. Even with large numbers of images per bulk upload, the processing time from upload to match results was rapid (e.g., bulk upload of 200+ images was complete within hours). Though not as accurate as initial applications to North Atlantic right whales, the matching algorithms reduced the time spent on matching by 25–33% of that required by manual matching. When potential matches were

expanded beyond top 5, overall top 24 accuracy was 56%, top 12 accuracy was 44%, top 10 accuracy was 41%, top 5 accuracy was 31% and top 1 accuracy was 16%. The lower accuracy observed for southern right whales is likely due to the vastly larger number of individuals compared to North Atlantic right whales, with a highly skewed dataset toward fewer photos per individual that could be matched against (e.g., 51% had only one photo). With continued application and more balanced machine learning training, accuracy should increase for this and other less-studied populations.

### CurvRank v2: fluke photos

For fluke matching, the Wild Me team implemented the machine learning-based CurvRank v2 algorithm, which has been recently advanced (Charles V. Stewart pers. comm; Alex Mankowski pers. comm.) at Rensselaer Polytechnic Institute beyond the originally published implementation for humpback whale flukes (Weideman et al. 2020). CurvRank v2 uses AI to learn how to extract the trailing edge of a fluke and then searches for matches to other edges by measuring the differences between them, learning to focus on sections that contain more individually identifiable information. The CurvRank v2 model is a multi-species model that was trained on a wide variety of whale flukes, including humpback and sperm whales in addition to right whales. The CurvRank algorithm is now available for fluke photographs of both North Atlantic and southern right whales on the Flukebook platform. In the original CurvRank v2 paper (Weideman et al. 2020), the algorithm was shown to have a top 1 accuracy of 85% and a top 5 accuracy of 89% on humpback whale flukes.

## HotSpotter: peduncle scar photos

While matching success continues to be strongest when using the identifying features of the head (DeepSense and PIE) or fluke (CurvRank), we have also pushed forward on the development of AI matching of the peduncle with the HotSpotter algorithm for those instances where that is the only photograph available. Hotspotter is a SIFT-based computer vision algorithm initially applied to Grevy's zebras (*Equus grevyi*), plains zebras (*E. quagga*), giraffes (*Giraffa sp.*), leopards (*Panthera pardus*), and lionfish (*Pterois sp.*) (Crall et al. 2013). HotSpotter analyzes the image to be matched for distinct patterns or "hotspots", and then compares that image against all the other images of that species and viewpoint in the Flukebook database (including recent uploads), and produces a ranked list of potential matches based on "hotspot" similarity. One advantage of this approach is that new individuals can be identified without the need for network retraining. HotSpotter was originally applied to the identification of right whale heads from aerial photographs but given the superior performance of the DeepSense algorithm for this task, was turned off to maximize server time efficiency. However, the HotSpotter algorithm has been effective in opportunistically matching peduncle scar patterns of right whales and is available for this application as needed (Fig. 8). HotSpotter has a top 1 accuracy of 17% and top 5 accuracy of 18% for photographs of the peduncle.

## Challenges

### Training data

Gathering sufficient volumes of training data for machine learning can be particularly challenging in collaborative studies, wherein different extant research protocols and project histories have resulted in different data collection standards. Differences as little as cropped right whale heads

facing upward in an image versus cropped heads facing downward can severely limit machine learning's ability to find animals in images and then individually identify them. The Wild Me team retrained its machine learning detector at least four times (to date) as data from heterogeneous North Atlantic right whale and southern right whale catalogs were added, ultimately resulting in a detector that can now handle many more aspects and angles in real-world data collection. Additional data labeling is still needed to better distinguish the dorsal and ventral surfaces in full-body aerial images.

Data balance in training machine learning-based ID algorithms is also a challenge. Wild Me's cross-application of PIE, which was trained entirely on a very large North Atlantic right whale catalog, to smaller southern right whale data sets resulted in decreased matching power of PIE for southern right whale versus North Atlantic right whale. This suggests exploration of the hypothesis that either southern right whale-specific machine learning training or balanced (in terms of data size and photos per individual), combined North Atlantic right whale and southern right whale datasets in machine learning training could yield better performing models for both species.

### Difficult matches

Some photographic data remain too challenging for semi-automated photo identification of right whales at this time, such as dead or decomposed animals, those showing only a portion of the body while engaged in a surface active group, or newborns with features that have not yet stabilized (Hamilton et al. 2022).

### Slow user adoption

AI is often portrayed in popular culture as a seamless and instantaneous user experience producing seemingly magical insights at the click of a button—Facebook facial recognition, Zillow home values, Amazon product recommendations, Google Maps driving directions, Gmail smart replies.



**Fig. 8** Example of a correct peduncle insertion scar match using the HotSpotter algorithm in the Flukebook platform (<https://www.flukebook.org/>). Photos of catalog #2223, Calvin, by the New England Aquarium

The use of machine learning for wildlife conservation has not yet matured to this level of ease and still requires user engagement and data stewardship in uploading the photographs into the system, associating relevant metadata, and synthesizing results. The speed of matching results can vary depending on the load on the servers and can range from nearly instantaneous to slow depending on the number of jobs in the queue at any given time. A particular challenge to broad adoption is in migrating from an existing user workflow, which often has complicated layers of dependencies for formatting data to accommodate existing queries, reporting, and data submissions. Further, it has yet to be seen how best to integrate the data in existing, well-established catalogs within the Flukebook platform. Change can be hard, and the barrier to entry remains an obstacle in transitioning to new systems. Finding ways to bridge the gap between the legacy system and the new methodology can be time-consuming, especially for biologists already working under high workloads with low staff. Having engaged early adopters with the persistence and time to manage these transitions are essential for the success of the project.

## Lessons learned

### Collaboration is key

For researchers considering the adoption of machine learning for photo identification, we strongly encourage close collaboration between conservation biologists and data scientists with expertise in machine learning and automated workflows. These conversations create both an opportunity for innovative approaches for the application of AI to conservation biology as well as some context for the limitations on what problems AI can solve. Similarly, collaboration among biologists can yield larger and more diverse data sets that maximize the impact of machine learning models in application to real-world photographs and observational conditions. Persistence, communication, and strong project management skills are essential for success. This project was able to leverage an enormous amount of data collected over five decades from researchers spanning the global distribution of right whales in both hemispheres into one training dataset for machine learning development.

### Algorithm recommendations

The significant overhead and drawbacks of retraining classifiers like Deepsense (Bogucki et al. 2019) lead us to strongly encourage future data scientists to develop algorithms that do not require retraining to match new individuals. Matching

algorithms based on image similarity scores, such as HotSpotter (Crall et al. 2013) and Pose Invariant Embeddings (Moskvyak 2019) achieve this goal by generating a similarity score between any two images, and then matching queries by calculating the similarity of a new image to each image in the existing catalog (or a subset) and then aggregating those scores against all the individual labels in the image catalog to find the most likely match. These image similarity algorithms naturally accommodate the addition of new photographs and even new individuals added to the catalog and therefore do not require retraining to remain useful for years at a time. We believe flexible algorithms such as these are the future of re-identification for conservation. Future research goals include training a PIE model to identify right whale aerial head photos. We anticipate competitive accuracy to Deepsense, with the benefit of applying to all populations and individuals that might be studied without successive retraining.

## Achievements

The use of AI for right whale photo identification has moved from prototype to production, building on the initial efforts of NOAA in 2015 to host a Kaggle data science competition and the winning Deepsense algorithm. Flukebook successfully implemented new machine learning algorithms and multi-feature matching, incorporating them into their existing operational platform for both the North Atlantic and southern right whales. Multi-feature matching allows right whales to be matched by aerial photos of their heads (Deepsense), lateral photos of their heads (Pose Invariant Embeddings), flukes (new CurvRank v2), and peduncle scarring (HotSpotter). Photo identification matching of individual wild animals from multiple angles and using various body parts is a significant advancement in the use of AI for wildlife research and conservation.

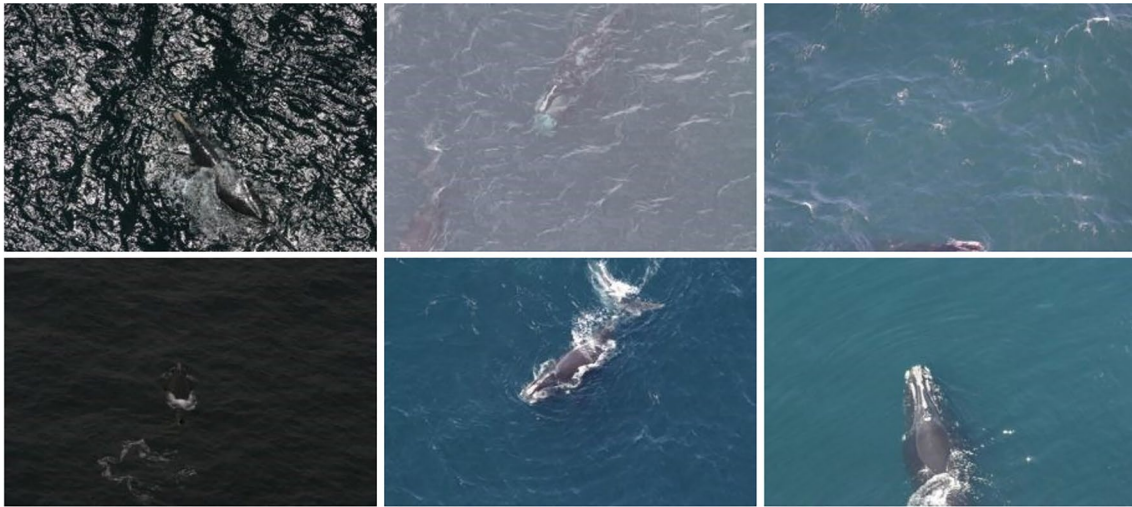
While human verification of the match remains the final step in the process, these semi-automated techniques hold tremendous potential to speed up the process of identifying individual whales, particularly for larger populations as in the southern right whale. Early work applying these techniques to research in New Zealand has shown significant time savings. This is particularly important as the volume and size of images submitted to photo identification catalogs have been expanding exponentially with contributions coming from vessels, aircraft, and unmanned systems. Managing this tsunami of images and curating the metadata to allow for detailed image analysis still requires manual inspection and data entry and thus remains a tremendous undertaking, particularly in the North Atlantic where images are used for

**Fig. A1** Photograph of a North Atlantic right whale, *Eubalaena glacialis*, known as number '2479' with a continuous callosity pattern subsurface feeding as seen from an aerial survey. Image collected under U.S. Marine Mammal Protection Act research permit number 17355. Photograph by National Oceanic and Atmospheric Administration/Northeast Fisheries Science Center/Christin Khan



**Fig. A2** Photograph of a mother calf pair of North Atlantic right whales, *Eubalaena glacialis*, as seen from an aerial survey and demonstrating a more broken callosity pattern. The mother is known as Naevus or number '2040' and her calf born in 2014 is known as Avalanche or number '4440'. Image collected under U.S. Marine Mammal Protection Act research permit number 17355. Photograph by National Oceanic and Atmospheric Administration/Northeast Fisheries Science Center/Christin Khan





**Fig. A3** Example of the diversity of images that improve model performance including factors, such as the orientation of the whale as well as environmental conditions, such as sun glare and wind. Figure from Bogucki et al. 2019 with permission



**Fig. A4** An image of a southern right whale, *Eubalaena australis*, which shares morphological similarities to the North Atlantic right whale. Photo by Claire Charlton

health assessments, scarring analysis to better understand entanglement rates, and to match poor quality photographs to identify dead or injured whales none of which is automated. High-quality photo identifications are crucial to the management of this species, and many others, and to better understand the human impacts that threaten their survival.

## Appendix

See Figs. A1, A2, A3 and A4

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s42991-022-00253-3>.

**Acknowledgements** We thank Dr. Christopher Clark for pushing this concept into reality by singing the praises of data science competitions, and to Sean Hardison at NOAA Fisheries for providing the push to “go big” with this project and make it operational for the Southern Hemisphere as well. Simon Childerhouse and Scott Baker helpfully provided access to the Auckland Islands, New Zealand, photo identification datasets for years 2006-2009. We would also like to thank the three anonymous reviewers for their detailed and insightful comments that greatly improved the quality of this manuscript. We are deeply grateful to the data contributors to the ARWPIC catalog: John Bannister, Western Australian Museum; Mike Double, Australian Antarctic Division; Mandy Watson, Department of Environment, Land, Water & Planning, Victoria; Stephen Burnell, Eubalaena Pty. Ltd.; Rachael Alderman, Department of Primary Industries, Parks, Water and Environment Tasmania; Kris Carlyon, Department of Primary Industries, Parks, Water and Environment Tasmania. We thank the editors, L. Karczmarski and S.C.Y. Chan for their meticulous handling of this manuscript.

**Author Contributions** CK initiated the vision to use AI for right whale photo identification. RB led the Deepsense team in the Kaggle data science competition which created the first right whale algorithm. CK, JH, and JLL led the vision to transition into an operational system. Machine learning and ingestion into the Flukebook platform was performed by DB, JP, and JH. Right whale photo-identification catalog curation, contribution, and writing was done by PH, CC, FC, DJ, WR, SD, EV, VR, and KG. All authors commented on and approved the final manuscript.

**Funding** We received funding support from NOAA Fisheries High Performance Computing IT Incubator (contract 1333MF18P-NFFM0103), NOAA Fisheries NERAP/EBFM (contract 1333MF19P-NFFM0139) and the Bureau of Ocean Energy Management (contract M17AC00013). Each of the photo-identification catalogs were supported by multiple funding sources over the many decades of their curation.

## Declarations

**Conflict of interest** The authors declare no conflicting or competing interests.

**Research permits and ethics approval** United States data were collected under Marine Mammal Protection Act research permits. Canadian data were collected under SARA permits. New Zealand data were collected under the New Zealand Department of Conservation Marine Mammal Research permit and University of Auckland Animal Ethics Committee approved protocol AEC/02/2005/R334 to C.S. Baker, and the New Zealand Department of Conservation Marine Mammal Research permit 50094-MAR to the University of Otago.

**Code** DeepSense algorithm GitLab repo: <https://gitlab.com/deepsense.ai/whales/>. Wildbook Image Analysis: <https://github.com/WildMeOrg/wildbook-ia>. Wildbook PIE algorithm Github repo: <https://github.com/WildbookOrg/wbia-plugin-pie>. Wildbook CurvRank v2 Github repo: <https://github.com/WildMeOrg/wbia-tpl-curvrnk-v2>.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Aizenberg I, Aizenberg NN, Vandewalle JPL (2000) Multi-valued and universal binary neurons: theory, learning and applications. Springer, Boston. <https://doi.org/10.1007/978-1-4757-3115-6>
- Blount D, Gero S, Van Oast J, Parham J, Kingen C, Scheiner B, Stere T, Fisher M, Minton G, Khan C, Dulau V, Thompson J, Moskyvak O, Berger-Wolf T, Stewart CV, Holmberg J, Levenson JJ (2022) Flukebook: an open-source AI platform for cetacean photo identification. *Mamm Biol* (special Issue). <https://doi.org/10.1007/s42991-021-00221-3>
- Bogucki R, Cygan M, Khan CB, Klimek M, Milczek JK, Mucha M (2019) Applying deep learning to right whale photo identification. *Conserv Biol* 33:676–684. <https://doi.org/10.1111/cobi.13226>
- Bogucki R, Cygan M, Klimek M, Milczek JK, Mucha M (2016) Which whale is it, anyway? Face recognition for right whales using deep learning. <https://deepsense.ai/deep-learning-right-whale-recognition-kaggle/>. Accessed 28 Jan 2022
- Bromley J, Bentz J, Bottou L, Guyon I, Lecun Y, Moore C, Sackinger E, Shah R (1993) Signature verification using a “siamese” time delay neural network. *Intern J Pattern Recognit Artif Intell* 7:669–688. <https://doi.org/10.1142/S0218001493000339>
- Calambokidis J, Barlow J (2006) Abundance of blue and humpback whales in the eastern North Pacific estimated by capture-recapture and line-transect methods. *Mar Mamm Sci* 20(1):63–85. <https://doi.org/10.1111/j.1748-7692.2004.tb01141.x>
- Charlton C, Ward R, McCauley RD, Brownell RL, Kent CS, Burnell S (2019) Southern right whale (*Eubalaena australis*), seasonal abundance and distribution at Head of Bight, South Australia. *Aquat Conserv* 29:576–588. <https://doi.org/10.1002/aqc.3032>
- Cheeseman T, Southerland K, Park J, Olio M, Flynn K, Calambokidis J, Jones L, Garrigue C, Jordán AF, Howard A, Reade W, Neilson J, Gabriele C, Clapham P (2022) Advanced image recognition: a fully automated, high-accuracy photo-identification matching system for humpback whales. *Mamm Biol*. <https://doi.org/10.1007/s42991-021-00180-9>
- Christiansen F, Vivier F, Charlton C, Ward R, Amerson A, Burnell S, Bejder L (2018) Maternal body size and condition determine calf growth rates in southern right whales. *Mar Ecol Prog Ser* 592:267–282. <https://doi.org/10.3354/meps12522>
- Christiansen F, Dawson SM, Durban JW, Fearnbach H, Miller CA, Bejder L, Uhart M, Sironi M, Corkeron P, Rayment W, Leunissen E, Haria E, Ward R, Warick HA, Kerr I, Lynn MS, Pettis HM, Moore MJ (2020) Population comparison of right whale body condition reveals poor state of the North Atlantic right whale. *Mar Ecol Prog Ser* 640:1–16. <https://doi.org/10.3354/meps13299>
- Clapham M, Miller E, Nguyen M, Van Horn RC (2022) Multispecies facial detection for individual identification of wildlife: a case study across ursids. *Mamm Biol*. <https://doi.org/10.1007/s42991-021-00168-5>
- Crall JP, Stewart CV, Berger-Wolf TY, Rubenstein DI, Sundaresan SR (2013) HotSpotter — Patterned species instance recognition. In: *IEEE Workshop on Applications of Computer Vision (WACV)* 230–237. <https://doi.org/10.1109/WACV.2013.6475023>
- Crowe LM, Brown MW, Corkeron PJ, Hamilton PK, Ramp C, Ratelle S, Vanderlaan ASM, Cole TVN (2021) In plane sight: a mark-recapture analysis of North Atlantic right whales in the Gulf of St. Lawrence Endang Species Res 46:227–251. <https://doi.org/10.3354/esr01156>
- Dawson SM, Bowman MH, Leunissen E, Sirguyev P (2017) Inexpensive aerial photogrammetry for studies of whales and large marine animals. *Front Mar Sci* 4:366. <https://doi.org/10.3389/fmars.2017.00366>
- de Silva EMK, Kumarasinghe P, Indrajith KKDAK, Pushpakumara TV, Vimukthi RDY, de Zoysa K, Gunawardana K, de Silva S (2022) Feasibility of using convolutional neural networks for individual-identification of wild Asian elephants. *Mamm Biol*. <https://doi.org/10.1007/s42991-021-00206-2>
- Dechter R (1986) Learning while searching in constraint-satisfaction-problems. In: *AAAI'86: Proceedings of the Fifth AAAI National Conference on Artificial Intelligence*, 178–183. AAAI Press, Philadelphia, PA. <https://dl.acm.org/doi/https://doi.org/10.5555/2887770.2887799>
- Girshick R, Radosavovic I, Gkioxari G, Dollar P, He K (2018) Detectron. <https://github.com/facebookresearch/detectron>. Accessed 28 Jan 2022
- Gormley AM, Slooten E, Dawson S, Barker RJ, Rayment W, du Fresne S, Bräger S (2012) First evidence that marine protected areas can work for marine mammals. *J Appl Ecol* 49:474–480. <https://doi.org/10.1111/j.1365-2664.2012.02121.x>
- Hamilton PK, Frasier BA, Conger LA, George RC, Jackson KA, Frasier TR (2022) Genetic identifications challenge our assumptions of physical development and mother–calf associations and separation times: a case study of the North Atlantic right whale (*Eubalaena glacialis*). *Mamm Biol*. <https://doi.org/10.1007/s42991-021-00177-4>
- He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. *The IEEE conference on computer vision and pattern recognition*. *CVPR* 2016:770–778. <https://doi.org/10.1109/CVPR.2016.90>
- Hiby L, Lovell P (2001) A note on an automated system for matching the callosity patterns on aerial photographs of southern right whales. *J Cetacean Res Manag* 2:291–295. <https://doi.org/10.47536/jcrm.vi.278>
- Hu H, Lan S, Jiang Y, Cao Z, Sha F (2017) Fastmask: Segment multi-scale object candidates in one shot. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* 991–999. <https://arxiv.org/abs/1612.08843>
- IUCN 2021. The IUCN Red List of Threatened Species. Version 2021–2. <https://www.iucnredlist.org>. Accessed 28 Jan 2022



- IWC (International Whaling Commission) (2021) Population status. <https://iwc.int/about-whales/status>. Accessed 26 Apr 2022
- Karczmariski L, Chan SCY, Chui SYS, Cameron EZ (2022) Individual identification and photographic techniques in mammalian ecological and behavioural research – Part 2: Field studies and applications. *Mamm Biol (Special Issue)* 102(4). <https://link.springer.com/journal/42991/volumes-and-issues/102-4>
- Knowlton AR, Hamilton PK, Marx MK, Pettis HM, Kraus SD (2012) Monitoring North Atlantic right whale *Eubalaena glacialis* entanglement rates: a 30 yr retrospective. *Mar Ecol Prog Ser* 466:293–302. <https://doi.org/10.3354/meps09923>
- Koivuniemi M, Kurkilahti M, Niemi M, Auttila M, Kunasranta M (2019) A mark–recapture approach for estimating population size of the endangered ringed seal (*Phoca hispida saimensis*). *PLoS ONE* 14(3):e0214269. <https://doi.org/10.1371/journal.pone.0214269>
- Kraus SD, Moore KE, Price CA, Crone MJ, Watkins WA, Winn HE, Prescott JH (1986) The use of photographs to identify individual North Atlantic right whales (*Eubalaena glacialis*). *Rep Int Whal Comm* 10:145–151
- Long J, Shelhamer E, Darrell T (2015) Fully convolutional networks for semantic segmentation. In: *The IEEE Conference on Computer Vision and Pattern Recognition*, Boston, Massachusetts, pp 3431–3440. <https://doi.org/10.1109/CVPR.2015.7298965>
- Ma N, Zhang X, Zheng HT, Sun J (2018) Shufflenet v2: Practical guidelines for efficient CNN architecture design. In: *Proceedings of the European conference on computer vision (ECCV)* 116–131. <https://arxiv.org/abs/1807.11164v1>
- Meyer-Gutbrod EL, Greene CH, Sullivan PJ, Pershing AJ (2015) Climate-associated changes in prey availability drive reproductive dynamics of the North Atlantic right whale population. *Mar Ecol Prog Ser* 535:243–258. <https://doi.org/10.3354/meps11372>
- Moore MJ, Knowlton AR, Kraus SD, McLellan WA, Bonde RK (2004) Morphometry, gross morphology and available histopathology in North Atlantic right whale (*Eubalaena glacialis*) mortalities (1970–2002). *J Cetacean Res Manag* 6:199–214
- Moskvyak O, Maire F, Armstrong AO, Dayoub F, Baktashmotlagh M (2019) Robust re-identification of manta rays from natural markings by learning pose invariant embeddings. [arXiv:1902.10847](https://arxiv.org/abs/1902.10847)
- Muto MM, Helker VT, Delean BJ, Young NC, Freed JC, Angliss RP, Friday NA, Boveng PL, Breiwick JM, Brost BM, Cameron MF, Clapham PJ, Crance JL, Dahle SP, Dahlheim ME, Fadely BS, Ferguson MC, Fritz LW, Goetz KT, Hobbs RC, Ivashchenko YV, Kennedy AS, London JM, Mizroch SA, Ream RR, Richmond EL, Shelden KEW, Sweeney KL, Towell RG, Wade PR, Waite JM, Zerbini AN (2021) Alaska marine mammal stock assessments, 2020. U.S. Dep. Commer., NOAA Tech. Memo. NMFS-AFSC-421, 398 p. <https://repository.library.noaa.gov/>
- New England Aquarium (2019) North Atlantic right whale catalog. <http://rwcatalog.neaq.org/#/>. Accessed 28 Jan 2022
- NOAA Fisheries (2021) North Atlantic right whale. <https://www.fisheries.noaa.gov/species/north-atlantic-right-whale>. Accessed 29 Nov 2021
- Pace RM, Corkeron PJ, Kraus SD (2017) State–space mark–recapture estimates reveal a recent decline in abundance of North Atlantic right whales. *Ecol Evol* 7:8730–8741. <https://doi.org/10.1002/ece3.3406>
- Parham J, Stewart C, Crall J, Rubenstein D, Holmberg J, Berger-Wolf T (2018) An animal detection pipeline for identification. In: *2018 IEEE Winter Conference on Applications of Computer Vision (WACV)* (pp. 1075–1083). IEEE. <https://doi.org/10.1109/WACV.2018.00123>
- Payne R (1976) At home with right whales. *Natl Geogr Mag* 149:322–339
- Phillips PJ, Yates AN, Hu Y, Hahn CA, Noyes E, Jackson K, Cavazos JG, Jeckeln G, Ranjan R, Sankaranarayanan S, Chen JC (2018) Face recognition accuracy of forensic examiners, super-recognizers, and face recognition algorithms. *Proc Natl Acad Sci* 115(24):6171–6176. <https://doi.org/10.1073/pnas.1721355115>
- Pirzl R, Murdoch G, Lawton K (2006) BigFish: Computer assisted matching software and data management system for photo-identification. Skadia Pty Ltd, Horsham, Australia
- Rayment W, Davidson A, Dawson S, Slooten E, Webster T (2012) Distribution of southern right whales on the Auckland Islands calving grounds. *N Z J Mar Freshw Res* 46:431–436. <https://doi.org/10.1080/00288330.2012.697072>
- Redmon J, Farhadi A (2017) YOLO9000: better, faster, stronger. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, 7263–7271. <https://doi.org/10.1109/CVPR.2017.690>
- Redmon J, Divvala S, Girshick R, Farhadi A (2016) You only look once: unified, real-time object detection. In: *The IEEE Conference on Computer Vision and Pattern Recognition* 779–788. <https://arxiv.org/abs/1506.02640v5>
- Ren S, He K, Girshick R, Jian S (2015) Faster R-CNN: Towards real-time object detection with region proposal networks. *Adv Neural Inf Process Syst* 91–99. <https://arxiv.org/abs/1506.01497>
- Rolland RM, Parks SE, Hunt KE, Castellote M, Corkeron PJ, Nowacek DP, Wasser SK, Kraus SD (2012) Evidence that ship noise increases stress in right whales. *Proc R Soc B* 279:2363–2368. <https://doi.org/10.1098/rspb.2011.2429>
- Ronneberger O, Fischer P, Brox T (2015) U-net: Convolutional networks for biomedical image segmentation. In: *International Conference on Medical image computing and computer-assisted intervention*, pp 234–241 Springer, Cham. <https://arxiv.org/abs/1505.04597>
- Schneider S, Taylor GW, Kremer S (2018) Deep learning object detection methods for ecological camera trap data. In: *15th Conf Comp and Robot Vis*, 321–328. <https://doi.org/10.1109/CRV.2018.00052>
- Smith JN, Kelly N, Double M, Bannister JL (2021) Population trend in right whales off southern Australia 1993–2020. In: *Report presented to the 68C IWC scientific committee (Southern Hemisphere Subcommittee) SC\_68C\_SH\_18*. <https://iwc.int/documents>
- Stamation K, Watson M, Moloney P, Charlton C, Bannister J (2020) Population estimate and rate of increase of southern right whales *Eubalaena australis* in southeastern Australia. *Endang Species Res* 41:373–383. <https://doi.org/10.3354/esr01031>
- Taigman Y, Yang M, Ranzato MA, Wolf L (2014) DeepFace: closing the gap to human-level performance in face verification. In: *The IEEE Conference on Computer Vision and Pattern Recognition*, 1701–1708. <https://doi.org/10.1109/CVPR.2014.220>
- Tan M, Chen B, Pang R, Vasudevan V, Sandler M, Howard A, Le QV (2019) Mnasnet: Platform-aware neural architecture search for mobile. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp 2820–2828. [https://openaccess.thecvf.com/content\\_CVPR\\_2019/papers/Tan\\_MnasNet\\_Platform-Aware\\_Neural\\_Architecture\\_Search\\_for\\_Mobile\\_CVPR\\_2019\\_paper.pdf](https://openaccess.thecvf.com/content_CVPR_2019/papers/Tan_MnasNet_Platform-Aware_Neural_Architecture_Search_for_Mobile_CVPR_2019_paper.pdf)
- Torres L, Rayment W, Olavarria C, Thompson D, Graham B, Baker C, Patenaude N, Bury S, Boren L, Parker G, Carroll E (2017) Demography and ecology of southern right whales wintering at sub-Antarctic Campbell Island, New Zealand. *Polar Biol* 40:95–106. <https://doi.org/10.1007/s00300-016-1926-x>
- van der Hoop J, Corkeron P, Moore M (2017) Entanglement is a costly life-history stage in large whales. *Ecol Evol* 7:92–106. <https://doi.org/10.1002/ece3.2615>
- Verma GK, Gupta P (2018) Wild animal detection using deep convolutional neural network. In: *Proc 2nd Intl Conf on Comp Vis and Image Processing*, 327–338. [https://doi.org/10.1007/978-981-10-7898-9\\_27](https://doi.org/10.1007/978-981-10-7898-9_27)
- Watson M, Stamation K, Charlton C, Bannister J (2021) Calving rates, long-range movements and site fidelity of southern right whales

- (*Eubalaena australis*) in south-eastern Australia. *J Cetacean Res Manag* 22(1):17–28. <https://doi.org/10.47536/jcrm.v22i1.210>
- Weideman HJ, Stewart CV, Parham JR, Holmberg J, Flynn K, Calambokidis J, Paul DB, Bedetti A, Henley M, Pope FG, Lepirei J (2020) Extracting identifying contours for African elephants and humpback whales using a learned appearance model. In: Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), 2020, pp. 1276–1285. [https://openaccess.thecvf.com/content\\_WACV\\_2020/papers/Weideman\\_Extracting\\_identifying\\_contours\\_for\\_African\\_elephants\\_and\\_humpback\\_whales\\_using\\_WACV\\_2020\\_paper.pdf](https://openaccess.thecvf.com/content_WACV_2020/papers/Weideman_Extracting_identifying_contours_for_African_elephants_and_humpback_whales_using_WACV_2020_paper.pdf)
- Xing F, Xie Y, Su H, Liu F, Yang L (2017) Deep learning in microscopy image analysis: a survey. *IEEE Trans Neural Netw Learn Syst* 99:1–19. <https://doi.org/10.1109/TNNLS.2017.2766168>
- Yousif H, Yuan J, Kays R, He Z (2017) Fast human-animal detection from highly cluttered camera-trap images using joint background modeling and deep learning classification. *IEEE Int Symp Circuits Syst.* <https://doi.org/10.1109/ISCAS.2017.8050762>
- Zhang X, Luo H, Fan X, Xiang W, Sun Y, Xiao Q, Jiang W, Zhang C, Sun J (2017) AlignedReID: Surpassing human-level performance in person re-identification. [arXiv:1711.08184](https://arxiv.org/abs/1711.08184).

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Authors and Affiliations

Christin Khan<sup>1</sup>  · Drew Blount<sup>2</sup>  · Jason Parham<sup>2,4</sup>  · Jason Holmberg<sup>2</sup>  · Philip Hamilton<sup>3</sup>  · Claire Charlton<sup>5</sup>  · Fredrik Christiansen<sup>6,7,8</sup>  · David Johnston<sup>9</sup>  · Will Rayment<sup>9</sup>  · Steve Dawson<sup>9</sup>  · Els Vermeulen<sup>10</sup>  · Victoria Rowntree<sup>11,12</sup>  · Karina Groch<sup>13</sup>  · J. Jacob Levenson<sup>14</sup>  · Robert Bogucki<sup>15</sup>

<sup>1</sup> NOAA Northeast Fisheries Science Center, Woods Hole, MA, USA

<sup>2</sup> Wild Me, Portland, OR, USA

<sup>3</sup> Anderson Cabot Center for Ocean Life at the New England Aquarium, Boston, MA, USA

<sup>4</sup> Rensselaer Polytechnic Institute, Troy, NY, USA

<sup>5</sup> Centre for Marine Science and Technology, Curtin University, Bentley, WA, Australia

<sup>6</sup> Aarhus Institute of Advanced Studies, Høegh-Guldbergs Gade 6B, 8000 Aarhus C, Denmark

<sup>7</sup> Zoophysiology, Department of Biology, Aarhus University, C.F. Møllers Allé 3, 8000 Aarhus C, Denmark

<sup>8</sup> Centre for Sustainable Aquatic Ecosystems, Harry Butler Institute, Murdoch University, Murdoch, Western Australia 6150, Australia

<sup>9</sup> Marine Science Department, Te Tari Putaiao Taimoana, University of Otago, Otago, New Zealand

<sup>10</sup> Mammal Research Institute Whale Unit, University of Pretoria, Hatfield, South Africa

<sup>11</sup> School of Biological Sciences, Department of Biology, University of Utah, Salt Lake City, Utah, USA

<sup>12</sup> Instituto de Conservacion de Ballenas, Capital Federal, Buenos Aires, Argentina

<sup>13</sup> Southern Right Whale Project / Instituto Australis, Imbituba, Santa Catarina, Brazil

<sup>14</sup> U.S. Department of Interior, Bureau of Ocean Energy Management, Sterling, VA, USA

<sup>15</sup> deepsense.ai, Al. Jerozolimskie 162A, 02-342 Warszawa, Poland