SC/68B/PH/06

Flukebook Continuing growth and technical advancement for cetacean photo identification and data archiving, including automated fin, fluke, and body matching

Drew Blount, Gianna Minton, Christin Khan, Jacob Levenson, Violaine Dulau, Shane Gero, Jason Parham, Jason Holmberg

INTERNATIONAL
WHALING COMMISSION

Flukebook – Continuing growth and technical advancement for cetacean photo identification and data archiving, including automated fin, fluke, and body matching

Drew Blount[1], Gianna Minton[2], Christin Khan[3], Jacob Levenson[4], Violaine Dulau[5], Shane Gero[6], Jason Parham[1], Jason Holmberg[1]

1. Wild Me
2. Megaptera Marine Conservation/Arabian Sea Whale Network
3. National Oceanic and Atmospheric Administration
4. Bureau of Ocean Energy Management, U.S. Department of the Interior
5. Globice-Reunion/ Indocet
6. Dominica Sperm Whale Project

Abstract:
Flukebook (flukebook.org) is a non-profit, open source cetacean data archiving and photo-identification tool developed under the larger Wildbook platform (wildbook.org) that uses computer vision and machine learning to facilitate automated identification of individual animals in the wild. In 2016, the IWC approved funding for the development of a regional data platform for the Arabian Sea Whale Network (ASWN) to be implemented in collaboration with Wild Me (wildme.org), the software and machine learning developers of Flukebook. This foundational collaboration expanded the capabilities of Flukebook and served as the springboard for subsequent years of growth in data and usage (e.g., by regional consortiums), as well as significant technical improvements in 2019-2020 in the application of computer vision and machine learning, specifically for North Atlantic and Southern right whales, humpback whales, sperm whales, and multiple species of dolphins. Ongoing improvements in our community support model and technical advances are bringing together industry, governmental, and NGO collaborators in a global-scale platform for cetacean research.

Background: Flukebook, the Wildbook for Whales and Dolphins

Wild Me (wildme.org) actively develops the Wildbook (wildbook.org) open source platform[1] to help scientists organize wildlife research, collect data from the public (e.g., photos and video), and integrate fully automated, multi-stage, and multi-modal machine learning (ML) to speed data curation. Through its web-based interface, Wildbook blends scientific collaboration and the growing "citizen scientist" movement, bringing the concepts of broad sector inclusion to wildlife conservation while retaining a focus on researchers and conservation authorities as the primary end-user for collected and reconciled data. The application of ML in Wildbook can help researchers determine population sizes faster (reducing human labor) and with greater specificity

[1] The latest Flukebook code is available from the Wildbook repository at: https://github.com/wildbookorg

(via increased data collected) and then subsequently adjust conservation action in shorter, iterative response cycles. Flukebook[2] (https://www.flukebook.org) is an instance of Wildbook tailored for regional- and global-scale research and collaboration on whales and dolphins. Additional background information about Wildbook and Flukebook can be found in past IWC publications (Blount et al. 2019; Blount et al. 2018).

Since 2015, Flukebook has grown and developed through two main drivers: 1) collaborations with research teams or projects that provide funding to develop new features or capability that help them address their own research, data management, or photo-ID matching priorities; and 2) internally driven improvements funded by grants, government contracts, and contributions managed centrally by the Flukebook team. Both of these sources contributed to the new developments discussed in this paper.

New Organizational Usage

The following organizations contributed data to Flukebook in 2019-2020:

- NOAA, USA
- The New England Aquarium, USA
- The Wild Dolphin Project, Bahamas
- Humpbacks and High Rises, Australia
- Channel Islands National Marine Sanctuary, USA
- Marine Conservation Research Inc, UK
- Norwegian Orca Survey, Norway
- Cedar Key Dolphin Project, USA
- CARI'MAM, Caribbean-wide marine mammal preservation network
- Cetacean Science Connections, Australia
- Outer Banks Center for Dolphin Research
- GLOBICE-Reunion, DOM France
- HDR, Inc., USA
- The Dominica Sperm Whale Project, Dominica

These collaborations have driven both the machine learning developments on Flukebook, like the NOAA-funded algorithms matching aerial photos of right whales, as well as the data management improvements benefitting the collaborative use of the platform, discussed in the next section.

Further Supporting Team and Organizational Data Archiving and Photo ID

A significant stage in the development of Flukebook was the IWC-funded collaboration with the Arabian Sea Whale Network (ASWN), which added a wide range of archiving and analysis features for cetacean sightings data both with and without photo identification. This resulted in a data architecture that allowed for the search and filtering of sighting records by species,

---

[2] Flukebook is managed by Wild Me (wildme.org), a U.S.-based 501(c)(3) non-profit organization. More information can be found at: https://www.wildme.org/contact/

behaviour category, group size and a range of other data fields. It also spurred the development of a ‘bulk upload’ template that allows users to format their own historical datasets for automatic import into Flukebook and a ‘metadata export’ that allows users to export their entire dataset into an Excel format that they can use to analyse and refine their own datasets after they have been uploaded to Flukebook. In combination, these tools allow researchers to link their existing data analysis workflows with the automatic matching, collaboration, and ecological tools available on Flukebook.

This work was followed by collaboration with the Indian Ocean Network for Cetacean Research (IndoCet), which has focused on developing and refining protocols and user interfaces for comparisons of photo-identification catalogues. Features developed under the Flukebook-IndoCet collaboration, which also involved representatives of the ASWN, include:

- organizationally-restricted data fields in the user interface (UI) that standardise data input and facilitate searching/filtering among organizations
- the ability to add and edit labelled keyworks on photoID images (e.g., for photo quality, distinctiveness, feature such as “dorsal fin”, “fluke”, etc.) and the ability to use this criteria to filter the database, which is critical for mark-recapture analysis
- robust support for multiple catalog names for a single whale, allowing easy cross-catalog comparison
- automatic generation of the “next name” for a given catalog when a new individual is ID’d
- the development of a merging interface that allows the user to reconcile data fields when it is discovered that two different individual records in a catalog are actually the same individual
- the ability to export the entire dataset
- tightened data security and protocols for granting access (edit or read only) to collaborators in pair-wise agreements to match catalogues

Technical Advancement: Plug-and-Play Machine Learning

2019 and 2020 have seen significant improvement in available ML techniques for cetaceans on Flukebook, with advancements in both the “detection” and “identification” phases of our automated pipeline, as shown in Figure 1.

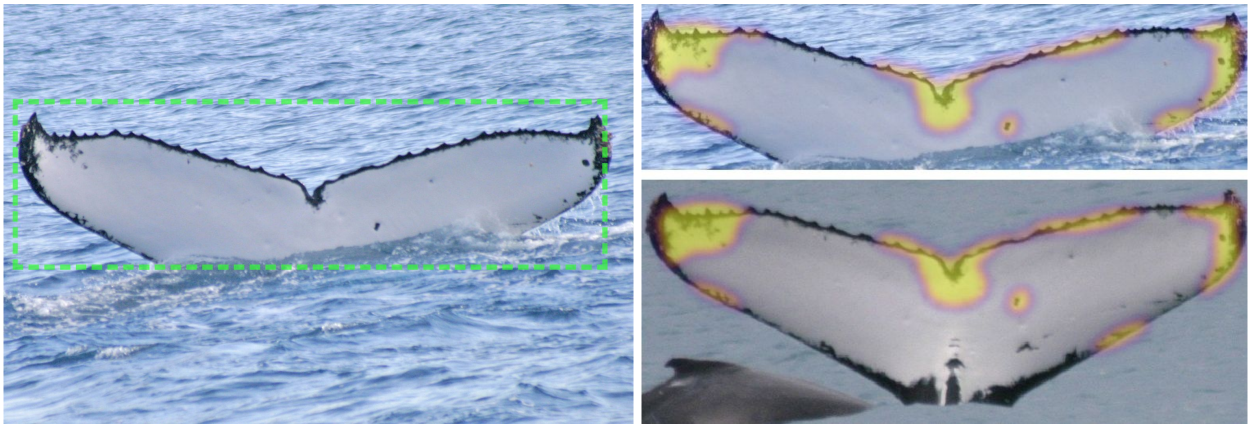

Figure 1. The Wildbook pipeline in Flukebook includes two primary phases: 1) "detection", which includes bounding box prediction around each animal and then species prediction, viewpoint labeling (e.g., left, right, etc.), and background subtraction (i.e. removing non-animal pixels) for each bounding box followed by 2) "identification" using one or more algorithms or machine learning models to identify each individual represented based on visual feature similarity.

Flukebook currently provides automated computer vision for multiple species of cetaceans, as presented in Table 1.

Species	**Feature Matched**	**Computer Vision Techniques**
Megaptera novaeangliae	Fluke	• Kaggle7 • HotSpotter • CurvRank • OC/DTW
Physeter macrocephalus	Fluke	• CurvRank • OC/DTW
Tursiops truncatus	Dorsal fin	• CurvRank • finFindR
Tursiops aduncus	Dorsal fin	• CurvRank • finFindR
Delphinus delphis	Dorsal fin	• CurvRank • finFindR
Eubalaena glacialis	Callosity pattern (aerial)	• Deepsense.ai
Eubalaena australis	Callosity pattern (aerial)	• Deepsense.ai • HotSpotter
Stenella frontalis	Flank patterning (adults)	• HotSpotter

Table 1. Summary of Species Supported by Computer Vision Matching in Flukebook

In 2019, the Wild Me machine learning team integrated the Kaggle7+, finFindR and Deepsense.ai matching algorithms to Flukebook. Combined with a dedicated machine learning staff, Flukebook's modular, Python-based computer vision architecture (Parham et al. 2018) allows for rapid integration of new algorithms. For example, integration of a Kaggle competition algorithm[3] for humpback whales was accomplished in only two weeks, including modifications and improvements from real-world application and training data provided by Cascadia Research Collective. Table 2 summarizes the basic approach and type of all of the matching approaches now employed in Flukebook. Figure 2 shows their orchestration as of May 2020.

Matching Technology	**Type**	**Approach**	**Advant.**	**Disadv.**
HotSpotter[4]	image similarity metric	SIFT-based comparison of areas of significant visual texture	no retraining needed for new IDs; broadly reusable across species; easy to choose match-against set	lower power matching (70% top-1) than deep learning classifiers
CurvRank[5,6]	image similarity metric	Edge extraction and comparison with learned areas of the fin/fluke edge weighted	no retraining needed for new IDs; good comparative matching across high quality photos (74% top-1 flukes; 95% top-1 fins); easy to choose match-against set	sensitive to extracted edge clarity
OC/DTW[7]	image similarity metric	Edge extraction and comparison using a modified form of Dynamic Time Warping	no retraining needed for new IDs; reusable across species with flukes; excellent for new study sites; easy to choose match-against set	sensitive to extracted edge clarity; lower power matching (70-80% top-1) than deep learning classifiers
NEW: finFindR[8,6]	image similarity metric	Edge extraction and comparison	no retraining needed for new IDs; good comparative matching	sensitive to extracted edge clarity

[3] https://www.kaggle.com/c/humpback-whale-identification
[4] Crall et al. 2013
[5] Weideman et al. 2017
[6] Moore et al. 2019
[7] Jablons 2016
[8] Thompson 2019

		using a Triplet Loss Network	across high quality photos (96% top-1); easy to choose match-against set	
NEW: Deepsense.ai Right Whale Matcher[9]	static classifier	Convolutional Neural Network (CNN)-based feature extraction and classification	89% top-1 accuracy; fast execution. First successful algorithm for right whales.	needs to be retrained to add new individuals (multi-day process); accuracy is dependent on large volume of training images per individual
NEW: Kaggle7+[10]	static classifier	Ensembled CNNs	93% top-1 accuracy in practice; 96% top-1 in competition; fast execution	Needs to be retrained to add new individuals (multi-day process); humpback whales only

Table 2. Summary of Computer Vision Techniques for Matching Individual Cetaceans in Flukebook.

[9] Bogucki et al. 2018.
[10] Mishkin et al. 2019

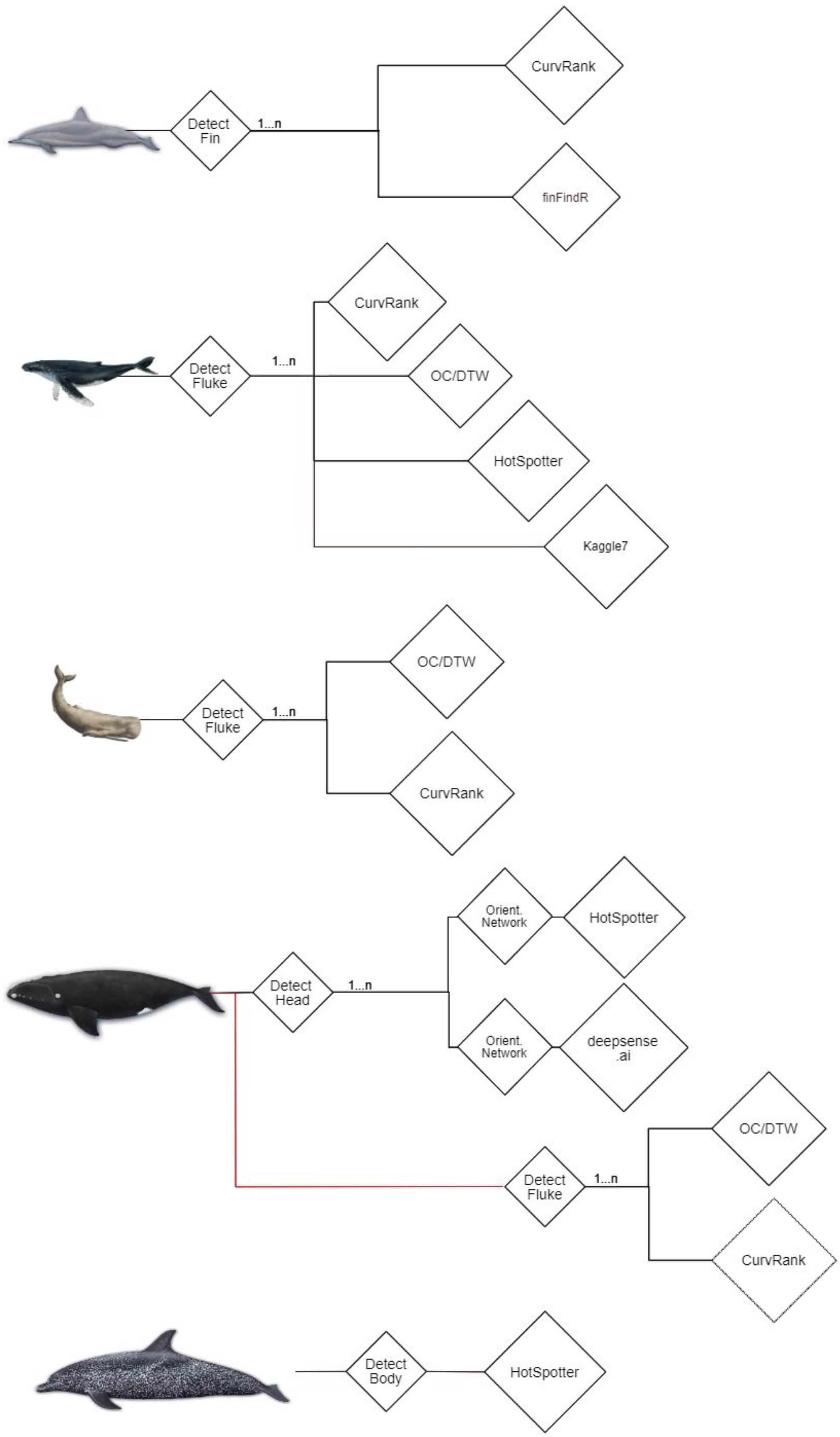

Figure 2. The Flukebook machine learning pipeline for identification of individuals from multiple species with multiple, pluggable matchers and configurations.

General Flukebook Upgrades

User-selectable Multi-site Matching Support

In 2019, Flukebook added user-selectable multi-site matching, allowing users to select the subset of data they want to inspect for matches using one or more computer vision algorithms. Options to match at the Consortium/project level will be created in 2020.

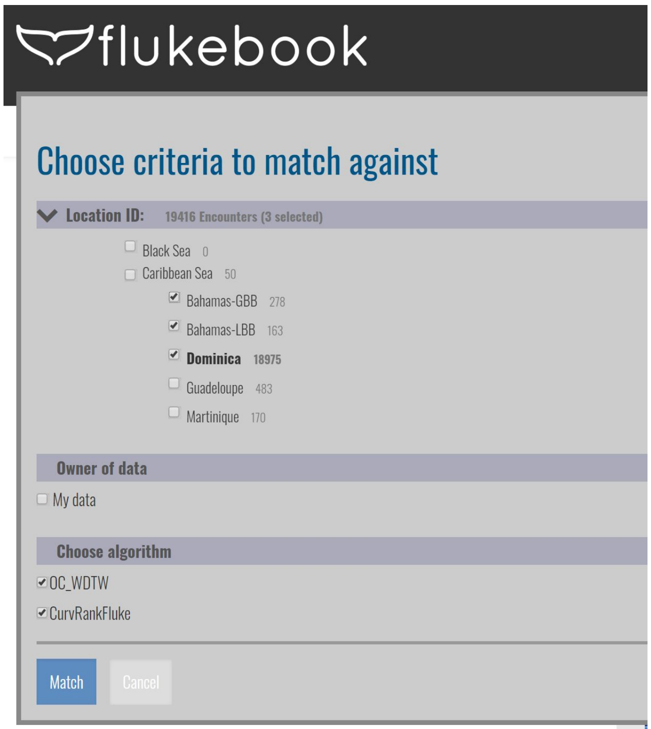

Figure 3. When looking for matching cetacean IDs, Flukebook users can subselect regional catalogs to match against.

Genus-level Matching Support

In March 2020, Flukebook expanded its multi-species pipeline, which defaults to comparing IDs only at the species level, to also support matching at the genus level to address situations in which exact species may not be known at the time of observation, such as photographing dolphin fins in Port Phillip Bay, Australia in which *Tursiops aduncus* and *Tursiops truncatus* are both present. Thus Flukebook now supports setting species to "sp." (e.g., "*Tursiops sp.*") to expand the customizable match-against set to include candidates from a common genus.

Improving Social Visualization

For social cetacean species, Flukebook now displays new visualizations of social structure, including co-occurrence, community membership, and direct relationships (e.g., "mother-calf"), as shown in Figure 4. This is a new area of data visualization on the profile page of identified individuals, and work to refine this will continue throughout 2020.

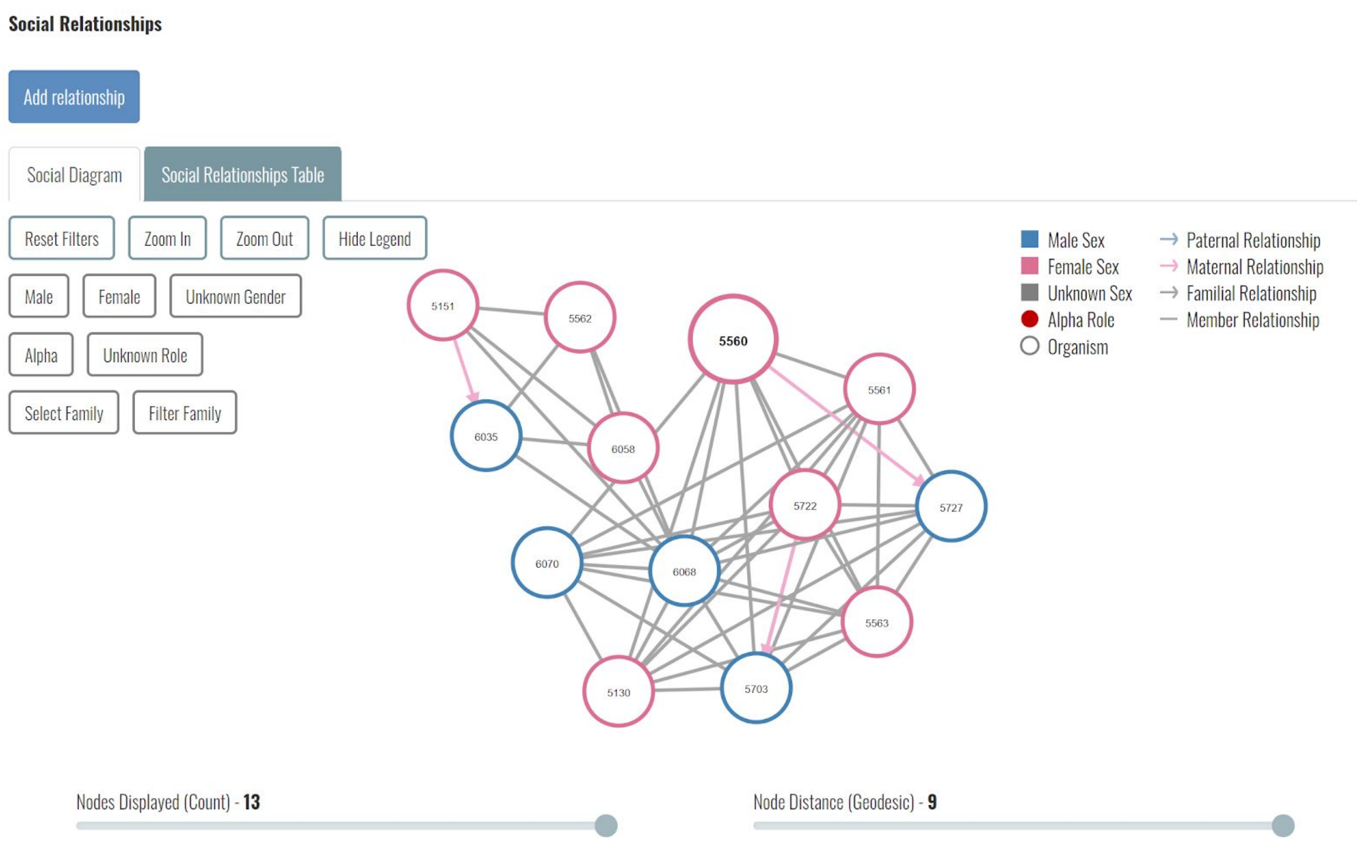

Figure 4. Social relationships within Units F and U visualized for "Pinchy", #5560, a sperm whale in the Eastern Caribbean. Data courtesy of The Dominica Sperm Whale Project.

Improving Broad Community Support

To better support its growing Wildbook community, especially Flukebook users, Wild Me recently added a professionally staffed, community support site at https://community.wildbook.org and has started publishing monthly release notes[11] to communicate the monthly advancements going into the platform. Bug reports and enhancement requests are tracked through a JIRA system, prioritized in weekly meetings, and assigned to Wild Me team members for resolution. Our entire support team can be reached directly at support@wildme.org.

Discussion

Flukebook provides a dedicated platform for collaborative photo-identification of cetaceans with multiple applications of proven and emerging machine learning techniques to generate significant time and cost savings in data curation and reconciliation across catalogs. Investment by the governments of France, the European Union, and the United States, as well as NGOs and industry partners, has advanced its technical capabilities and furthered adoption in the research community in 2019-2020. As an open source platform, each additional collaboration contributes to its further development and improvement. Investments made by one research project or consortium benefit

[11] April 2020 Wildbook release notes: https://community.wildbook.org/t/wildbook-release-notes-april-2020/68

the wider cetacean research community, as the newly developed algorithms, data archiving and analytical functions become available to all users. Importantly, our shared experience as an online community studying a total of 38000 cetaceans across 200000 sightings and 1.2 million photos is significant, and our application of multiple techniques of computer vision in a comparable, common architecture and workflow has generated unique experience and new insights for advancement.

Recommendation: Prioritizing Research on Image Similarity-based Matching Algorithms

We include this brief discussion on two types of matching algorithms to make an important distinction that should be recognized in the community, between 1) image classifiers and 2) image similarity-based matching algorithms.

Competitions in machine learning can serve a valuable role in the photo ID ecosystem by 1) creating purpose-built models where more general techniques may not yet exist and 2) by raising the bar for accuracy in a problem space (e.g., humpback fluke matching). However, due to the winning criteria that competitors work to optimize in these competitions, they often result in static, inflexible systems that have significant shortcomings in deployment. Future competitions should be carefully constructed to encourage competitors to make matching systems based on image similarity metrics rather than static classifiers.

In Kaggle-style competitions, competitors train their machine learning models on a training set of labeled images, and after training they are scored on their accuracy labeling a distinct validation set of images. In most competitions we have seen, the validation set includes only those individuals who were present in the training set. In other words, there is a static set of matchable individuals for the purposes of the competition. In this framing, competitors usually make a machine learning *classifier*: a system that takes an input (in this case an image) and categorizes it as one of a defined set of finite classes. Classifiers are quick and powerful and are used for example by the U.S. Post Office to automatically read the digits in handwritten postal codes on parcels. The relevant architectural constraint in a classifier model is that the number of classes is fixed in the neural networks trained for that task. So for matching purposes, a classifier can only recognize the fixed set of whales that were in its training data.

In contrast, many of the algorithms used in deployed automatic photo ID systems, such as Hotspotter and finFindR, are not fixed classifiers at all but rather image similarity-based matching systems. These algorithms use a similarity metric which is constructed such that highly-similar images depict the same individual. In mathematical terms, images are embedded into a feature space where proximity corresponds to similarity, and individual animals are represented by clusters of similar images. Image-similarity-based machine learning algorithms can achieve comparable accuracy to classifiers (Weideman et al. 2017; Thompson et al. 2019; Moore et al. 2019) yet match against an ever-growing reference catalog, or even an entirely new, never-before-seen catalog, without needing to be retrained.

Adding high ranking Kaggle classifiers (Bogucki et al. 2018; Mishkin et al. 2019) for humpback whales and North Atlantic right whales to Flukebook yielded a general insight: the classifiers' fixed matching set is a significant shortcoming for a deployed matching tool that can only be overcome with frequent and computationally expensive retraining, a process that takes from several hours to several days to complete. There are two common scenarios where a fixed classifier becomes invalid and needs retraining. In both cases, the problem arises when the size of the reference catalog changes: first, when a new individual is photographed for the first time; this individual is not matchable until the system can been re-trained. Second, the perennial catalog curation problems of “merges” and “splits”---when two individuals in the catalog are found to actually be the same real-world individual, or the inverse, when one individual record is found to actually depict two distinct animals.

Considering these scenarios where the classifier becomes invalid, one can see that it is almost impossible to use a static classifier to compare two separate catalogs to each other, for example reconciling SPLASH with a single researcher’s catalog, without retraining the system from scratch every time a match is made: we cannot simply combine the training data for both catalogs without first finding all of the merges between them, lest we introduce errors to the training set. Yet all of these cases are handled gracefully by image similarity algorithms: as long as new photos show the same types of distinctive features that were in the original training set, the similarity metric and thus matching algorithm remain valid.

Additionally, classifier-based models operate as black box systems that predict IDs without inspectability, whereas image-similarity-based systems often extract features that make sense to humans and can be used to illustrate the “why” of proposed matches (Figure 1 shows an example of corresponding areas of feature similarity between two humpback fluke photos). Thus, feature-based matching can provide reviewing researchers with important “explainability” for interpreting results, a subject of growing importance in ML research overall (Roscher et al. 2020), especially for reducing bias, understanding errors, and extracting novel scientific insights.

Advancing Onward Together

The Flukebook community welcomes new collaborations that will enable the addition of new species and functional dimensions to the Flukebook Platform, and we hope that existing functions can continue to be refined and applied to additional cetacean research and conservation efforts around the globe. Further developments for broad, organization-level support within the Flukebook platform are underway, and we are continuing to refine, cross-apply, and advance new machine learning techniques in support of cetacean research.

Acknowledgements
In 2019-2020, Wild Me advanced Flukebook with financial or in-kind support from:

- The Gordon and Betty Moore Foundation
- Microsoft AI for Earth Program
- U.S. Department of the Interior, Bureau of Ocean Energy Management (BOEM)
- U.S. National Oceanic and Atmospheric Administration (NOAA)
- CARI’MAM
- The European Union and the Conseil Regional Reunion (via Globice and IndoCet)

References

Blount D, Holmberg J, and Minton G. (2019) Flukebook – Recent advances for cetacean photo identification and data archiving including automated fluke matching. International Whaling

Commission report SC/68A/SH/07.

Blount D, Holmberg J, and Minton G. (2018) Flukebook – A tool for cetacean photo identification, data archiving and automated fluke matching. International Whaling Commission report SC/67B/PH/03.

Bogucki, R. , Cygan, M. , Khan, C. B., Klimek, M. , Milczek, J. K. and Mucha, M. (2018), Applying deep learning to right whale photo identification. Conservation Biology. doi:10.1111/cobi.13226

Crall JP, Stewart CV, Berger-Wolf T, Rubenstein DI and Sundaresan SR, "HotSpotter — Patterned species instance recognition," 2013 IEEE Workshop on Applications of Computer Vision (WACV 2013), pp. 230-237.

Jablons, Z. Identifying humpback whale flukes by sequence matching of trailing edge curvature. Master's thesis, Rensselaer Polytechnic Institute, 2016.

Mishkin D, Mishchuk A, Krashenyi I. (2019) Kaggle: Humpback Whale Identification Challenge. Solution: https://github.com/ducha-aiki/whale-identification-2018
Modified for Flukebook by Parham J: https://github.com/WildbookOrg/ibeis-kaggle7-module

Moore RT, Allen J, Urian K (2019) "Advances in automated dorsal fin matching" workshop in the World Marine Mammal Conference 2019 (WMMC'19), Barcelona, Spain, Dec. 2019.

Parham, Jason & Stewart, Charles & Crall, J.P. & Rubenstein, Daniel & Holmberg, Jason & Berger-Wolf, Tanya. (2018). An Animal Detection Pipeline for Identification. 1075-1083. 10.1109/WACV.2018.00123.

Roscher R, Bohn B, Duarte and J. Garcke (2020) "Explainable Machine Learning for Scientific Insights and Discoveries," in IEEE Access, vol. 8, pp. 42200-42216.

Thompson, J. (2019) finFindR: https://github.com/haimeh/finFindR. Downloaded April 4, 2020.

Weideman, H.J, Z.M. Jablons, J. Holmberg, K. Flynn, J. Calambokidis, R. B. Tyson, J.B. Allen, R.S. Wells, K. Hupman, K. Urian, C.V. Stewart. 2017. Integral Curvature Representation and Matching Algorithms for Identification of Dolphins and Whales. doi: 10.1002/ar.2365