ElephantBook: A Semi-Automated Human-in-the-Loop System for Elephant Re-Identification

Peter Kulits
kulits@caltech.edu
Caltech

Jake Wall
jake@maraelephantproject.org
Mara Elephant Project

Anka Bedetti
anka@elephantsalive.org
Elephants Alive

Michelle Henley
michelephant@savetheelephants.org
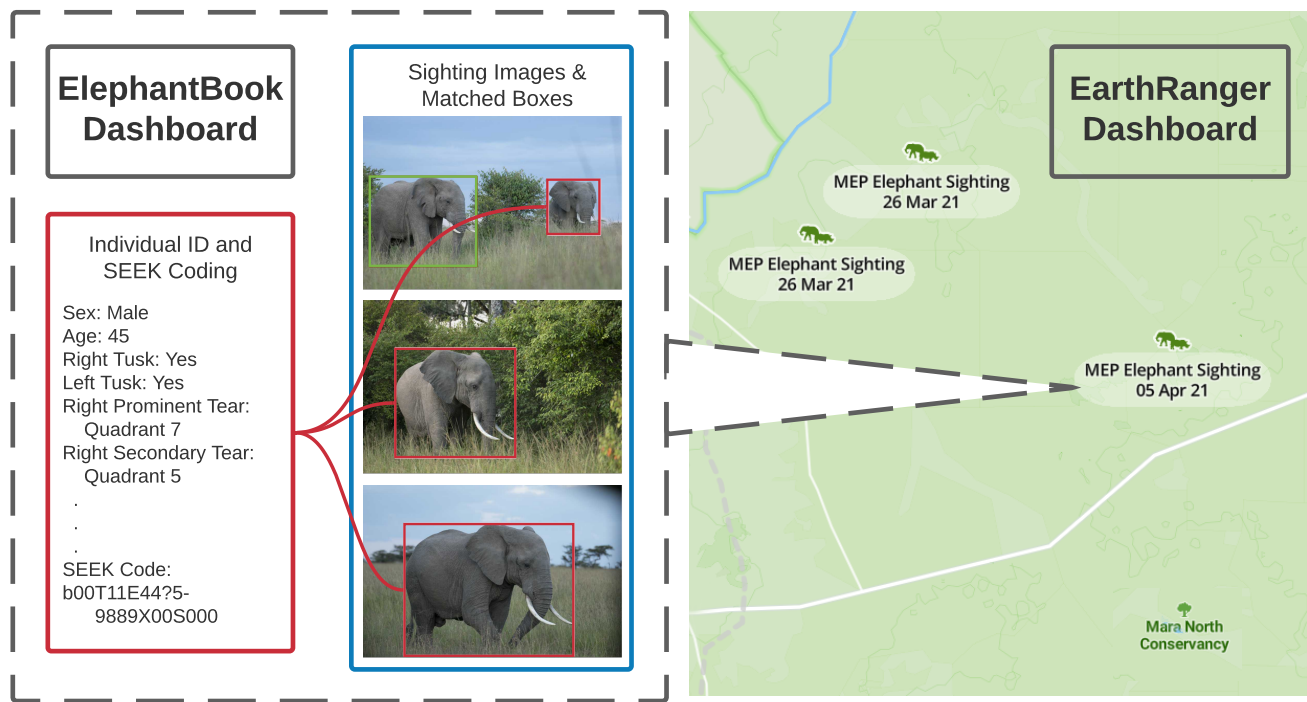Elephants Alive

Sara Beery
sbeery@caltech.edu
Caltech

Figure 1: ElephantBook: a system for human-in-the-loop elephant re-identification. Our system can be linked to the Earth-Ranger conservation land management platform [4], and it helps humans efficiently monitor elephant populations and locations from elephant sightings in the wild.

ABSTRACT

African elephants are vital to their ecosystems, but their populations are threatened by a rise in human-elephant conflict and poaching. Monitoring population dynamics is essential in conservation efforts; however, tracking elephants is a difficult task, usually relying on the invasive and sometimes dangerous placement of GPS collars. Although there have been many recent successes in the use of computer vision techniques for automated identification of other species, identification of elephants is extremely difficult and typically requires expertise as well as familiarity with elephants in the population. We have built and deployed a web-based platform and database for human-in-the-loop re-identification of elephants combining manual attribute labeling and state-of-the-art computer vision algorithms, known as ElephantBook. Our system is currently in use at the Mara Elephant Project, helping monitor the protected and at-risk population of elephants in the Greater Maasai Mara ecosystem. ElephantBook makes elephant re-identification usable by non-experts and scalable for use by multiple conservation NGOs.

CCS CONCEPTS

• **Computing methodologies** → Machine learning.

KEYWORDS

elephants, re-identification, conservation tech

1 INTRODUCTION

Reliable wildlife population monitoring is critical for effective conservation and species management. Accurate measurement of wildlife density and distribution across landscapes provides insight into trends and ecological processes such as population growth, fecundity, survival, mortality, and density-dependent regulation. A range of measurement techniques have been developed which include aerial surveys, camera trap networks, ground survey techniques, and individual-based re-identification (e.g., spatially explicit mark-recapture [56]). Individual-based recognition techniques can also be used in behavioral studies and human-wildlife conflict cases. The emergence of computational systems based on image algorithms has recently made traction enabling re-identification of certain species (e.g., whales, sharks, zebras, seals, lynx, and sea turtles) that present distinct morphology or patterns (e.g., contours, spots, or stripes) that facilitate visual separability among individuals [16]. However, many species are cryptic and difficult to observe, difficult even for experts to distinguish, or currently lack sufficient training data for application of computer-vision approaches.

Vital to their ecosystems, African elephants are especially important to monitor closely; they are considered ecosystem engineers who have the capacity to shape the environments in which they live, and their population density and distribution can impart multiple cascading effects on ecosystems, biodiversity, and tourism-based economies [26, 47, 53]. Both species of African elephants are threatened: the savanna elephant (*L. africana*) is endangered, and the forest elephant (*L. cyclotis*) was recently listed as critically endangered by the IUCN Red List [3]. Some populations have suffered as much as 62% population loss in recent years [43] with the ivory trade and associated poaching being the main drivers of their decline. Characterizing elephant population demographics across their range is therefore essential to conservation of the species.

Ecologists have recently attempted to create a general re-identification method that can be used by non-experts. The best known of these methods is System for Elephant Ear-pattern Knowledge (SEEK) coding, developed by Elephants Alive [8], which uses manual attribute labels such as sex and the presence/absence of tusks to improve the accuracy and efficiency of re-identification. The Mara Elephant Project, in collaboration with the California Institute of Technology and Elephants Alive, has developed a semi-automated ensemble visual-recognition system using photographs taken by rangers and research field teams along with manual SEEK attribute labeling. ElephantBook is a novel online software solution

with the goal of making elephant re-identification accessible by non-experts and scalable to multiple conservation NGOs.

2 BACKGROUND

2.1 The Greater Mara Ecosystem

The Greater Mara Ecosystem (GME) in Kenya is a critical ecosystem given its biodiversity, large wildlife populations, and rich cultural history. It forms the northern extent of the annual migration of 2.2 million wildebeest, zebra, and gazelle from the Serengeti, and it is the most-visited tourist destination in Kenya. The most recent census results estimate there are 2,493 elephants in the GME [61]. Elephants typically live in family units consisting of related females and their offspring. Adult male elephants roam alone or in bachelor herds after they've reached an age of sexual maturity. Despite its status as one of the most beautiful and important wildlife areas in the world, the GME faces significant conservation threats: 374 elephants have been illegally killed since 2012, and there has been a 60% increase in recorded incidents of human-elephant conflict since 2017 (Mara Elephant Project unpublished data). The expansion of agriculture, infrastructure, and human populations is infringing into current elephant ranging areas and severing movement corridors. 50% of elephant range now falls outside of protected areas [70].

2.2 The Mara Elephant Project

The Mara Elephant Project (MEP), established in 2011, protects savanna elephants and works to conserve the greater Maasai Mara ecosystem (GME) in Kenya. MEP, in conjunction with the Kenyan Government, has deployed ranger teams to follow the locations of elephant groups fitted with real-time GPS tracking collars, which has led to the arrest of 373 poachers, the seizure of 1,676.5 kg of ivory, and the identification of core movement patterns of approximately 500 elephants [2]. MEP also frequently dispatches rangers to help mitigate conflicts involving "crop-raiding" elephants. Identifying which individual elephants are involved in crop-raiding is important because raiders are typically repeat offenders. Ongoing field monitoring, data analysis, and conservation efforts are needed to ensure the long-term survival of elephants and the overall GME.

2.3 Elephants Alive

Elephants Alive is a South Africa-based non-profit organization that operates across the Greater Limpopo Transfrontier Conservation Area and the southern part of Mozambique. Although officially registered with the Kruger National Park in 2003, Elephants Alive draws on data collected over a quarter of a century. Its work contributes to the long-term survival of African elephants through a greater understanding of the complex relationship between elephants and the ecosystems they occupy and by identifying science-based solutions that enable elephants and people to coexist.

2.4 EarthRanger

Vulcan's EarthRanger [4] is a real-time system for conservation-related data aggregation, storage, visualization, and dissemination [69]. It includes tracking data from wildlife, rangers, and vehicles, and it records "Events," which range from human-elephant conflict to poaching to illegal logging. Events are reported from the field
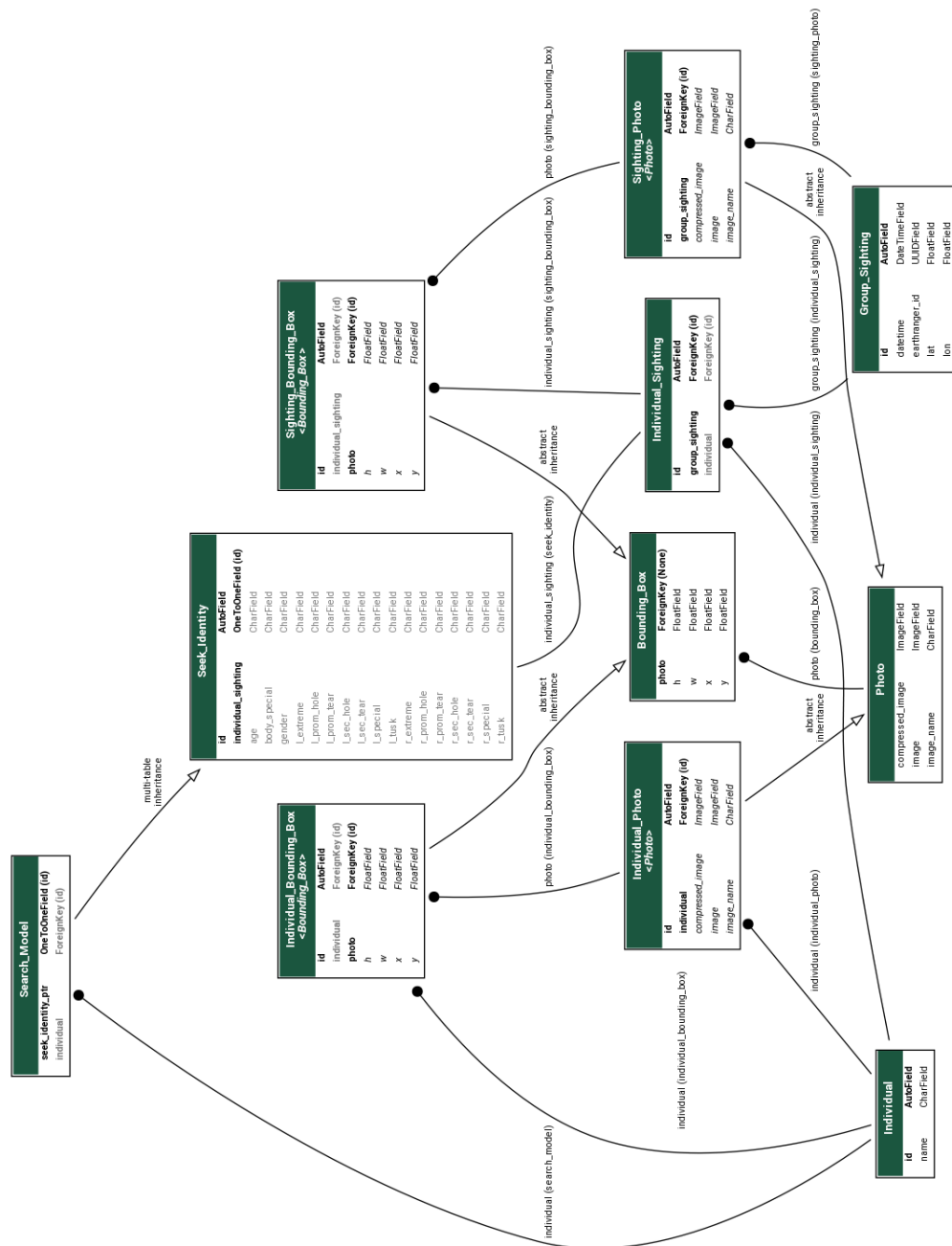
Figure 2: Database schema of major models in ElephantBook

using a mobile application called Cybertracker; these reports include the time, the location, and information specific to each event type. The Mara Elephant Project, along with many other NGOs in Sub-Saharan Africa, now uses EarthRanger daily to record elephant Group Sightings, including information about group size and composition. However, EarthRanger does not currently support or have any type of interface for individual-based elephant re-identification.

2.5 Human Expert Elephant Re-Identification

Elephant re-identification is a difficult task, and ecologists may spend thousands of hours over their careers cataloging and charting features that can be used to distinguish elephants. These approaches are often heavily subjective and based on the interpretation and skill of the observer, making the process difficult to replicate across multiple observers or elephant populations. Quantitative approaches
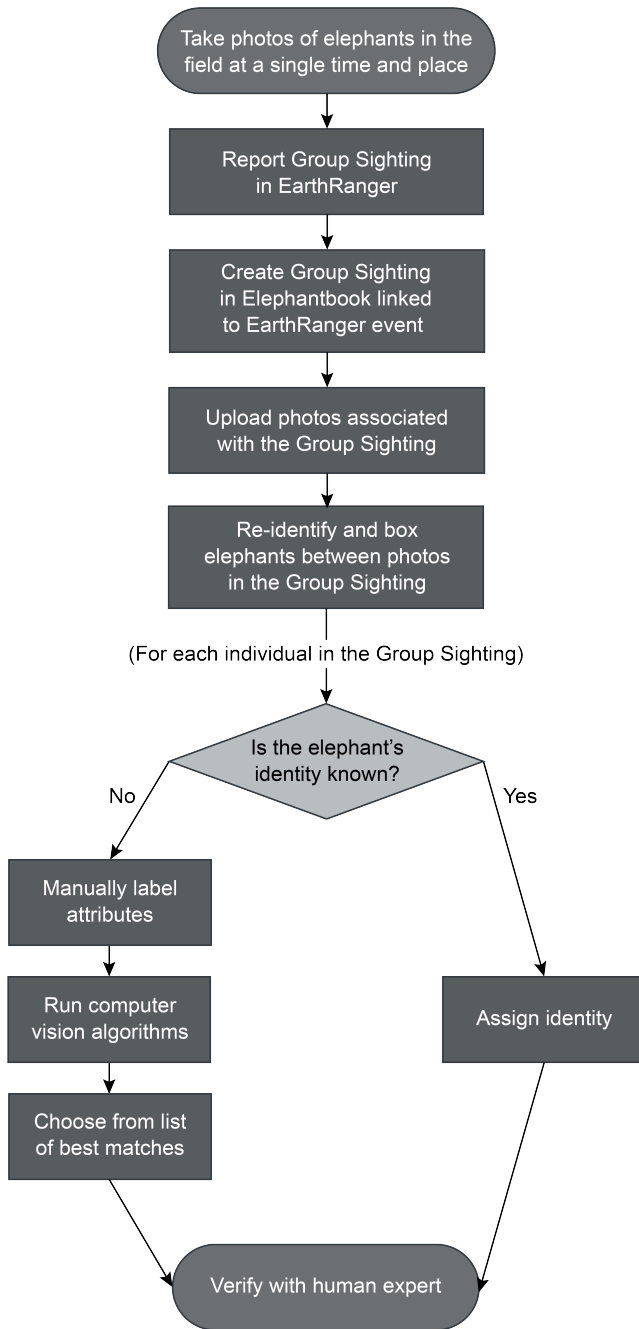
Take photos of elephants in the
field at a single time and place

Report Group Sighting
in EarthRanger

Create Group Sighting
in Elephantbook linked
to EarthRanger event

Upload photos associated
with the Group Sighting

Re-identify and box
elephants between photos
in the Group Sighting

(For each individual in the Group Sighting)

Is the elephant's
identity known?

No   Yes

Manually label
attributes

Run computer
vision algorithms

Assign identity

Choose from list
of best matches

Verify with human expert

Figure 3: Workflow of elephant re-identification

are needed to reliably re-identify elephants without dependence on the one or two experts typically available within an organization. One of the most successful existing methods of differentiating between elephants relies on comparison of the elephants' ears including notches, tears, holes, and other identifiable patterns. Several organizations (e.g., Save the Elephants, Elephants Alive, Elephant

Voices, Amboseli Elephant Trust) use this expert-based approach for elephant identification.

Elephants Alive developed SEEK [8], which involves a comprehensive identification dataset comprised of photos, drawings, and codes of elephant ear patterns that were collected over 25 years (since 1996). The identification system has been refined over time to exclude observer bias and accelerate the photographic identification process. We believe SEEK is the least subjective or expert-reliant elephant re-identification system in-use by any organization to date.

2.6 Automated Animal Re-Identification

The most commonly studied re-identification problems in computer vision focus on humans, with popular benchmarks and vast literature for human facial re-identification [38, 73, 74]. There have been many recent successes in computer vision for automated species identification, in both camera trap data [9–14, 48, 57, 63] and human-captured community science data [21, 22, 42, 65–67]. Automated re-identification of *individual animals* using computer vision is an increasingly popular topic, with publications and workshops on the subject at major computer vision conferences [1]. There are several excellent reviews of computer vision for animal re-ID [54, 59, 68].

One of the main, and significant, differences between animal re-identification and other fine-grained categorization tasks is that populations are not fixed, making re-ID an open-set categorization problem [75]. You must be able to recognize if and when an individual does not already exist in your database. The set of individuals might also be quite large: even for the relatively small global population of Grevy's zebra, your full set of identities would be 8,000 individuals [49, 50].

The earliest proposed semi-automated re-identification systems go back as far as 1990, with works on whale re-identification based on human-annotated attribute similarity [46]. The next big breakthroughs in the field relied on traditional feature-engineered computer vision techniques for pattern matching (including SIFT-based feature matching) [7, 16, 18, 29, 30, 34, 41, 45, 64] and numerical representations of unique contours [6, 31, 32]. Animal re-identification, like most of computer vision, has seen significant advances with the onset of deep learning, including several neural-network-based approaches [17, 19, 20, 25, 28, 35–37, 60]. The field has recently explored metric-learning-based methods [24, 58, 75], inspired by the success of these methods for human re-identification [38, 73, 74]. Metric-learning methods are also more robust to open-set categorization, as they are similarity-based and require only a single example of an individual with which to compare, as opposed to the tens or hundreds of examples needed by data-hungry CNNs. Another common tactic for handling the open-set and data-scarce nature of re-identification is hybridizing deep networks for notable part localization with previous pattern or contour feature-based matching methods which do not require large amounts of training data per individual [62, 72].

2.7 Automated Elephant Re-Identification

In 2010, Dabarera and Rodrigo proposed an image-based algorithm to identify individual elephants based on full-frontal facial images [23]. Korschens et al. proposed a matching algorithm based on
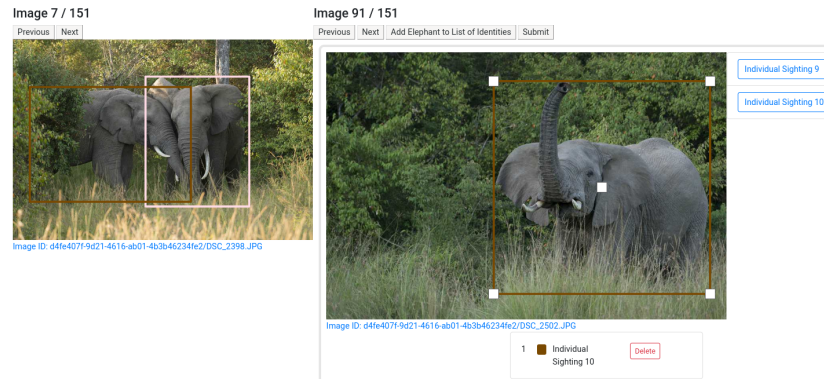
Figure 4: Bounding-box annotation interface

human-labeled whole-head annotations, including the elephant's ears and tusks, where present [35]. In an extension, they released a dataset, ELPephants, and demonstrated good results on a closed set of individuals with localized feature extraction using deep nets and SVM-based feature discrimination [36]. Recent methods of robustly differentiating between elephant images in an open-set population rely on finding and matching the contours of the ear (Figure 9), similar to many human-expert re-identification methods like SEEK. Multi-curve matching algorithms based on human-annotated contours of elephant ears were proposed by Ardovini et al. [6] and Weideman et al. [71]. Weideman's CurvRank algorithm was originally designed for re-identification of whale flukes and dorsal fins. Recently, Weideman et al. [72] proposed an extension of CurvRank that is capable of automatically extracting matchable contours from images, and report strong results matching contours of elephant ears.

3 ELEPHANTBOOK

Our solution, which we call ElephantBook, by default integrates with EarthRanger through its REST API to consume Group Sightings recorded by field teams. ElephantBook can also be reconfigured for use without EarthRanger if needed. It is web-based and built primarily with the Django Python package [5]. This configuration allows our system to be both lightweight and easily reconfigurable.

3.1 Human-in-the-Loop Re-Identification Pipeline

3.1.0 Data Collection in the Field. Rangers at the Mara Elephant Project routinely survey the Maasai Mara in search of elephants. Rangers record the time and location of every elephant sighting and submit it to EarthRanger. When possible, rangers photograph each elephant from multiple angles. If no photographs are taken, the event is resolved in EarthRanger, and no re-identification occurs.

3.1.1 Adding a Group Sighting to ElephantBook via EarthRanger. ElephantBook pulls a list of active elephant sighting events from EarthRanger. Users select the appropriate EarthRanger event and create a corresponding ElephantBook "Group Sighting." A Group

Sighting is one or more elephants spotted at the same time and place. This step is usually performed after returning from the field.

3.1.2 Uploading Photos. All photos taken at the same time and place of the Group Sighting are uploaded to ElephantBook. However, only photos labeled with boxes (in the next step) are used for re-identification.

3.1.3 Boxing Elephants. Once all photos are uploaded, elephants in each photo are boxed with an image annotation tool. While the human annotator likely will not know the name of each individual elephant in the Group Sighting photos, the annotator should be able to differentiate between elephants and identify the same elephant across multiple images. If it is impossible to tell elephants apart in a single instance, matching over a period of months is unlikely. Boxes are marked with numbers unique to each specific elephant within the Group Sighting. This identification marking reduces the number of matches we must make from the sum of the number of elephants in all photos to the number of actual elephants sighted. See Section 4 for more details.

3.1.4 Human Attribute Labeling. An "Individual Sighting" is created for each elephant identified in the previous step. An Individual Sighting is an elephant encounter at a single time and place, and it is always connected to a parent Group Sighting. Manual attribute-labeling is performed for each Individual Sighting. We use the recently-developed SEEK coding system [8]. See Section 5 for more details.

3.1.5 Computer Vision. Confidence-producing computer vision matching algorithms are run to identify potential matches. See Section 6 for more details.

3.1.6 Matching. Manual attributes are combined with the output from the computer vision algorithms to provide a list of the most likely previously identified elephants. See Section 7 for more details.

4 BOUNDING-BOX ANNOTATION

We customized an open-source online bounding-box annotation tool from the Visipedia project [15]. Because annotators need to match individual elephants across photos taken at a sighting, a
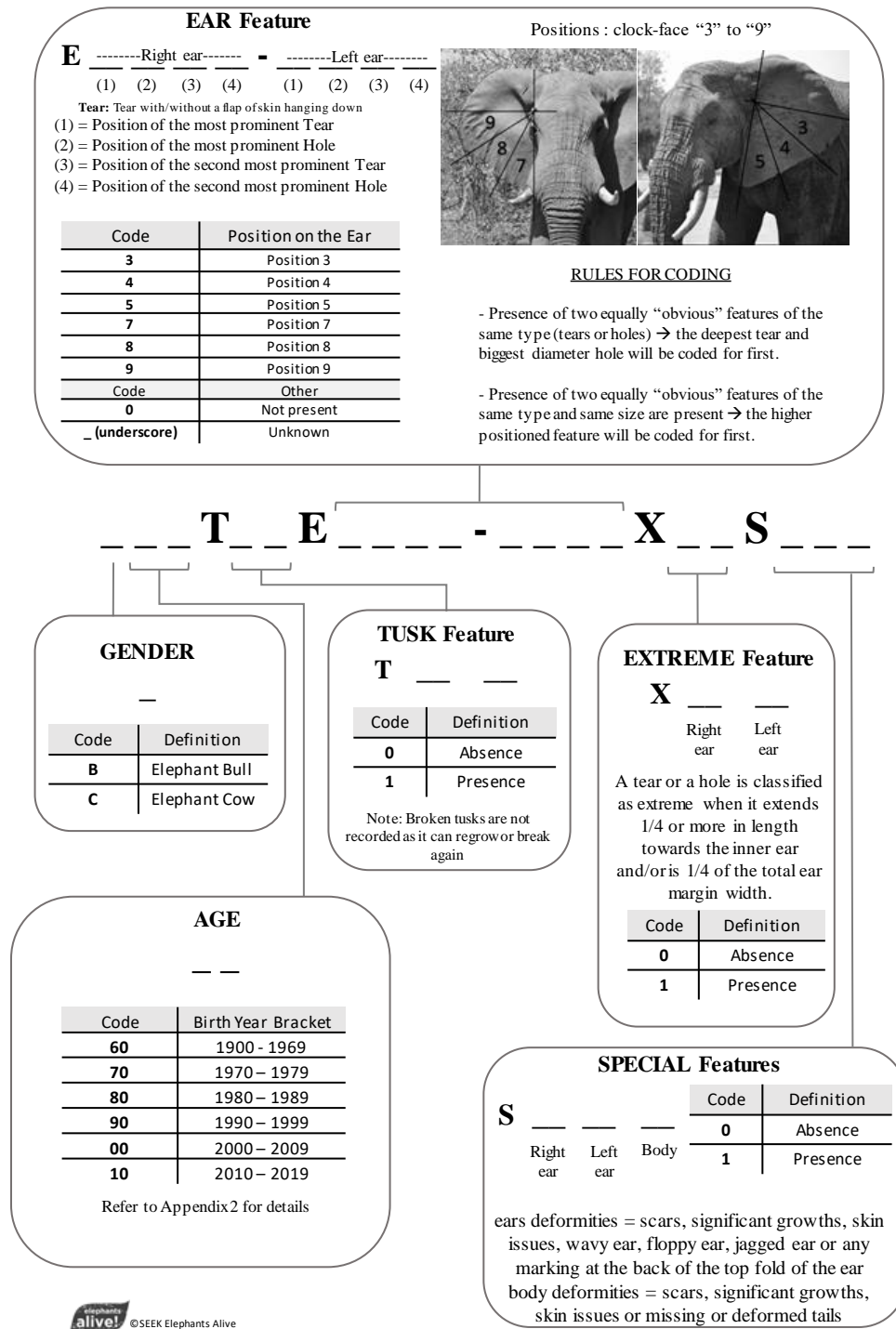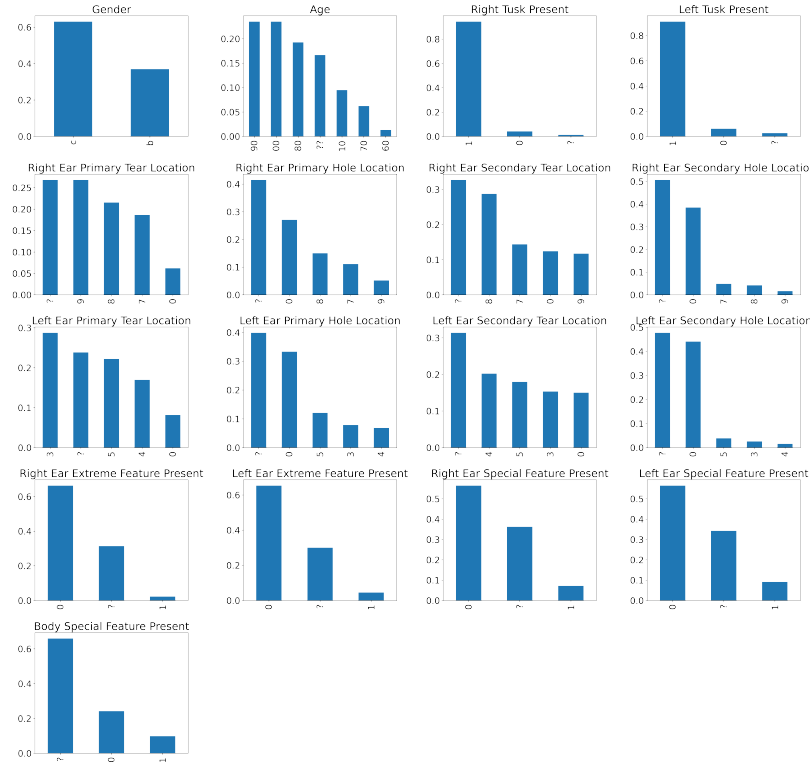
Figure 5: The SEEK coding system

Figure 6: Frequency of SEEK attributes

second pane was added to allow annotators to compare multiple photos at once.

5 SEEK

In SEEK, each elephant is assigned a unique descriptive code which is used to narrow the set of potential matches that must be considered by human experts. The code begins with the elephant's sex and age, followed by the presence or absence of tusks (Figure 5). The code further defines the type and position of prominent and secondary tears and holes found on the right and left ears. Finally, it notes the presence of any extreme features on the ears and body, such as a missing tail.

6 COMPUTER VISION

6.1 Elephant Ear Localization

To allow CurvRank to focus on the ear, the localization of which is key to extracting accurate ear contours, we trained a simple elephant-ear detector.

The ELPephants dataset from the Elephant Listening Project [36] – consisting of images of African forest elephants visiting the Dzanga bai clearing in the Dzanga-Ndoki National Park in the Central African Republic – was used for training and validation. After removing duplicates, the dataset consisted of 1935 images of 276 unique individuals. The dataset is provided with the identities of the elephants, but each image was manually annotated for bounding-boxes of left and right ears. Only ears where the contours

were fully visible were annotated. Annotations were made on 910 left ears and 1,045 right ears (Figure 8). Two-hundred randomly sampled images were reserved for object detection validation.

We trained a Faster R-CNN object detection model [55] with a ResNet-50 backbone [27] and added Feature Pyramid Networks (FPN) [39] in Pytorch [51]. Beginning with a model checkpoint pretrained on the Microsoft COCO dataset [40], we trained our model on 1,735 images to detect and categorize left and right ears. Our detector achieves a Mean Average Precision (mAP) [40] of 95% on our held-out test dataset of 200 randomly-selected images.

6.2 Matching Ear Contours with CurvRank

After extracting ear images from out ear detection model, we use CurvRank [72] to filter possible matches.

CurvRank was initially developed to recognize individual cetaceans based on contours of flukes and dorsal fins [71]. As elephant ears are also a strong re-identifiable feature and are delineated by a contoured edge, applying CurvRank to elephant re-identification was an intuitive next step, and the authors determined the transferability of the matching algorithm from cetaceans to elephants by analyzing results on hand-drawn contours.

Recently the CurvRank authors proposed a deep-learning based algorithm to automatically extract the contours used as input to their matching algorithm [72]. They evaluated this automated approach on humpback whales and African savanna elephants, with impressive results. Their method relies on two fully convolutional
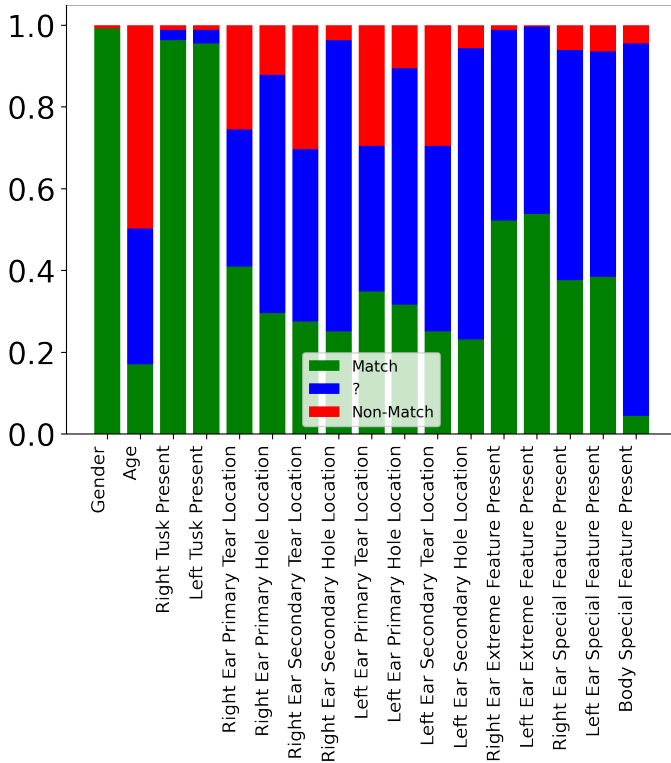
Figure 7: Agreement of an notators by attribute for pairs of SEEK codes on the same individual

neural networks for curve extraction, one coarse-grained and one fine-grained (CG-FCNN and FG-FCNN). Annotators initially traced the identifying contour in cropped images with a broad line, using a single brushstroke, to produce coarse-grained training data for the CG-FCNN. In the second step, the FG-FCNN is trained to predict for each pixel in the (ear or fluke) image the probability that it would be covered by the coarse brush stroke, producing a probability image at the same resolution as the initial image. By using the coarse, easily extracted training data to train the FG-FCNN, tedious manual effort is avoided. These pixel-level probability maps guide the third step: a shortest path contour extraction algorithm.

Once the contour is extracted, it is represented as an ordered sequence of (x, y) coordinate pairs. Then CurvRank builds an integral curvature by sliding multiple disks of increasing radius along the contour [72]. For each scale, every point is represented as the ratio of the areas of the disk for that scale on either side of the contour [72]. Feature keypoints are defined at local extrema of the integral curvature representation [33, 72], and feature descriptors are extracted from the regions between all pairs of keypoints. This set of feature descriptors forms a densely sampled, overlapping representation of the entire individual contour across multiple scales. Match similarity is determined and possible matching individuals ranked via the local naive Bayes nearest neighbors (LNBNN) algorithm [44].

The method reported a top-1 matching accuracy of 84% for high-quality, high-resolution images of elephant ears on the closed set of 132 individuals on which the model is also trained [72]. The authors remark that elephant ear recognition is more difficult than whale fluke identification, due to challenges of contour extraction against more-highly textured image backgrounds and because the identifying information is more localized and subtle.
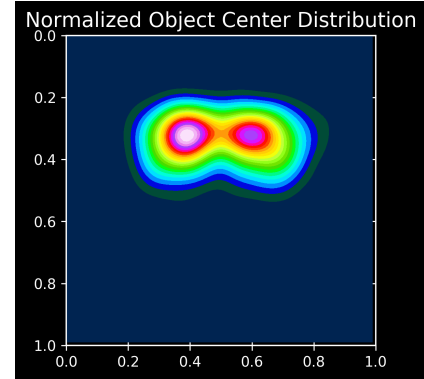

Figure 8: We visualize the center of the ground truth annotated boxes across our training set, and see that there is a strong bias in the imagery towards ears being in the upper center of the image, with modes slightly to the right and left

7 MATCHING

To allow rangers to efficiently identify an individual from the large set of previously encountered elephants, rangers are presented with a ranked list of possible matches to visually examine. These matches are computed with a score function that is a linear combination of manual attribute differences and computer vision matching confidence. Each SEEK attribute of the new Individual Sighting is compared to the SEEK code of all known Individuals. For each attribute in the codes, the distance is zero if the attributes match, one if they differ, and 0.6 if either of them contain a wildcard character. Additionally, the weight of the age component of the distance is set to 0.4 because of the known difficulty in accurately aging elephants (Figure 7). The mean of these differences is taken. The weighting parameters were learned separately on a training set of codes to optimize matching accuracy.

CurvRank produces an unbounded matching score between the new Individual Sighting and all Individuals. A greater score indicates greater likelihood of a true match. CurvRank scores are subtracted from the SEEK score and multiplied by 0.1. The constant 0.1 was learned in a training set of SEEK codes and CurvRank contours.

7.1 Evaluating Matching Accuracy

To evaluate the robustness of SEEK, CurvRank, and our proposed combination of the two, we trained a non-expert team of seven college undergraduates to perform SEEK labeling, and we collected labels from two to three students annotator for a set of Individual Sightings from the Elephant Voices collected by Joyce Poole [52].

(a) Successful curve extraction

(b) Failed curve extraction

Figure 9: Here we show successful and failed CurvRank examples. CurvRank is highly successful in high-quality, high-resultion imagery, but performance drops off in lower-resolution or blurry data as the edge of the ear is harder to distinguish.

In total, we have three annotations for 75 Individual Sightings and two annotations for 26 Individual Sightings.

We held out individuals that had at least two SEEK code annotations and two right-ear CurvRank contours. There are 45 individuals with a pair of SEEK codes and 33 individuals with a triplet of codes. Comparisons of top-k matching performance for SEEK, CurvRank, and our ensembled approach can be seen in Figure 10. We observe that matching performed with SEEK codes generally outperforms that of CurvRank alone, but that a combined approach is able to leverage the best of both, leading to more accurate matching. Using our combined system, with only two previous sightings of an individual in our database, we are able to match to the correct individual within the top 15 for 92.9% of sightings, and within the top 5 for 66.7%, helping rangers reduce the time needed to find the correct matched individual in the database. As Mara Elephant Project continues to collect and label Individual Sightings we will continue to analyze and hopefully improve matching performance. We expect additional sightings to improve accuracy, as it presents more potential sightings per individual to match with correctly. However, as we collect additional sightings we will also be increasing the number of individuals in the database, making the matching task more nuanced and potentially more challenging.

8 MARA ELEPHANT PROJECT INITIAL DEPLOYMENT

The Mara Elephant Project began using ElephantBook in January 2021 after a six-month prototyping period and so far has logged 140 Group Sightings and 251 Individual Sightings and has ingested and boxed 10,462 images of elephants. Beginning in March 2021, the organization has hired and trained a full-time team of four research assistants for collecting elephant sightings in the field, processing photos, and developing SEEK codes for individual elephants. Initial training on both field methodology for cataloging elephant Group Sightings and in the use of ElephantBook and SEEK labeling took one week. MEP's goal is to characterize and document the majority of the Mara's 2500 individuals. Extension of the ElephantBook system with partner organizations in Tanzania would further enable documenting the greater, connected elephant population stretching south into the Serengeti and consisting of >7000 individuals.

Initial experience using ElephantBook is that it is an intuitive system that mimics a typical re-identification workflow. Optimizations

for low-bandwidth connections, such as compression of photos before viewing them, but also keeping original full-resolution versions available for detailed scrutiny by a SEEK coder, have greatly improved the user experience. Boxing individuals has been relatively straight-forward even for novice users. Accurately labeling SEEK codes is perhaps the most challenging component of the ElephantBook system, particularly the correct estimation of age category which requires considerable expertise and, to a lesser degree, the determination of sex.

9 CONCLUSION AND FUTURE WORK

We have built a robust semi-automated system for human-in-the-loop elephant re-identification, and we have deployed our system on the ground in the Greater Mara Ecosystem. This system allows the Mara Elephant Project to track a much larger population of elephants over time, as they will no longer need to collar an elephant to track its movements. The system is a needed tool to assist in their vital elephant conservation efforts. As we move forward, we will expand ElephantBook to additional parks, including the Grumeti Game Reserve in Serengeti National Park in Tanzania and Greater Limpopo Transfrontier Conservation Area in South Africa and Southern Mozambique.

In the coming months, we will continue to collect new elephant sightings and refine our matching system to further reduce the human effort needed for re-identification. We plan to investigate automating SEEK coding and integrating additional computer vision methodology into our system, building learned representations of individual elephants beyond their ear contours. The data collected will also allow us to further analyze how these elephant features change over time and allow us to conduct deeper analysis of our current system on an expanding set of known elephants.

REFERENCES
[1] [n.d.].   Deep Learning Methods and Applications for Animal Re-Identification Workshop at WACV 2020.   https://sites.google.com/corp/

(a) One code per individual
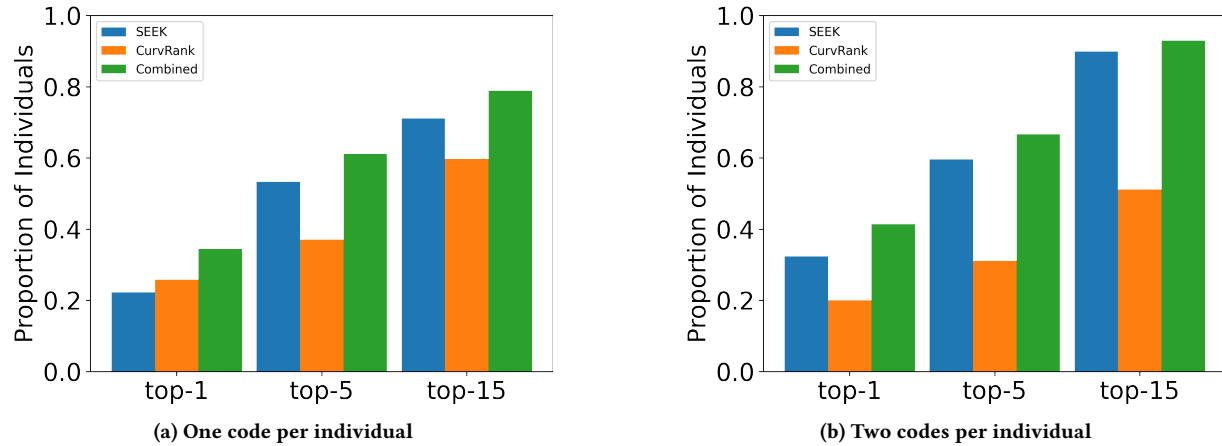
(b) Two codes per individual

Figure 10: Comparing matching accuracy for our SEEK-based matching algorithm, CurvRank, and our hybrid SEEK-CurvRank aggregated approach. We see that SEEK and CurvRank are complementary, with the combined approach outperforming either method on its own for tests with both one and two database example for each individual.

view/wacv2020animalreid/.

[2] [n.d.]. Increased Security. https://maraelephantproject.org/results/increased-security/

[3] [n.d.]. The IUCN Red List of Threatened Species. https://www.iucnredlist.org/en

[4] [n.d.]. Vulcan EarthRanger: A Domain Awareness System. https://earthranger.com/

[5] [n.d.]. The Web framework for perfectionists with deadlines | Django. https://www.djangoproject.com/

[6] A. Ardovini, L. Cinque, and E. Sangineto. 2008. Identifying elephant photos by multi-curve matching. *Pattern Recognition* 41 (2008), 1867–1877. https://doi.org/10.1016/j.patcog.2007.11.010

[7] Z Arzoumanian, Jason Holmberg, and Brad Norman. 2005. An astronomical pattern-matching algorithm for computer-aided identification of whale sharks Rhincodon typus. *Journal of Applied Ecology* 42, 6 (2005), 999–1011.

[8] Anka Bedetti, Cathy Greyling, Barry Paul, Jennifer Blondeau, Amy Clark, Hannah Malin, Jackie Horne, Ronny Makukule, Jessica Wilmot, Tammy Eggeling, et al. 2020. System for elephant ear-pattern knowledge (SEEK) to identify individual African elephants. *Pachyderm* 61 (2020), 63–77.

[9] Sara Beery, Elijah Cole, and Arvi Gjoka. 2020. The iWildCam 2020 Competition Dataset. *arXiv preprint arXiv:2004.10340* (2020).

[10] Sara Beery, Yang Liu, Dan Morris, Jim Piavis, Ashish Kapoor, Neel Joshi, Markus Meister, and Pietro Perona. 2020. Synthetic examples improve generalization for rare classes. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*. 863–873.

[11] Sara Beery, Dan Morris, and Pietro Perona. 2019. The iWildCam 2019 Challenge Dataset. *ArXiv* abs/1907.07617 (2019).

[12] Sara Beery, Dan Morris, and Siyu Yang. 2019. Efficient Pipeline for Camera Trap Image Review. *arXiv preprint arXiv:1907.06772* (2019).

[13] Sara Beery, Grant van Horn, Oisin MacAodha, and Pietro Perona. 2019. The iWildCam 2018 Challenge Dataset. *arXiv preprint arXiv:1904.05986* (2019).

[14] Sara Beery, Guanhang Wu, Vivek Rathod, Ronny Votel, and Jonathan Huang. 2020. Context r-cnn: Long term temporal context for per-camera object detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 13075–13085.

[15] Serge Belongie and Pietro Perona. 2016. Visipedia circa 2015. *Pattern Recognition Letters* 72 (2016), 15–24. https://doi.org/10.1016/j.patrec.2015.11.023

[16] Tanya Y. Berger-Wolf, Daniel I. Rubenstein, Charles V. Stewart, Jason A. Holmberg, Jason Parham, Sreejith Menon, Jonathan Crall, Jon Van Oast, Emre Kiciman, and Lucas Joppa. 2017. Wildbook: Crowdsourcing, computer vision, and data science for conservation. *arXiv:1710.08880 [cs]* (2017). http://arxiv.org/abs/1710.08880 arXiv: 1710.08880.

[17] Clemens-Alexander Brust, Tilo Burghardt, Milou Groenenberg, Christoph Kading, Hjalmar S Kuhl, Marie L Manguette, and Joachim Denzler. 2017. Towards automated visual monitoring of individual gorillas in the wild. In *Proceedings of the IEEE International Conference on Computer Vision Workshops*. 2820–2830.

[18] T Burghardt, PJ Barham, N Campbell, IC Cuthill, RB Sherley, and TM Leshoro. 2007. A fully automated computer vision system for the biometric identification of african penguins (Spheniscus demersus) on robben island. In *6th International Penguin Conference (IPC07)*. 19.

[19] Steven JB Carter, Ian P Bell, Jessica J Miller, and Peter P Gash. 2014. Automated marine turtle photograph identification using artificial neural networks, with application to green turtles. *Journal of experimental marine biology and ecology* 452 (2014), 105–110.

[20] Gullal Singh Cheema and Saket Anand. 2017. Automatic Detection and Recognition of Individuals in Patterned Species. *Lecture Notes in Computer Science* (2017), 27–38. https://doi.org/10.1007/978-3-319-71273-4_3

[21] Yin Cui, Menglin Jia, Tsung-Yi Lin, Yang Song, and Serge Belongie. 2019. Class-balanced loss based on effective number of samples. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 9268–9277.

[22] Yin Cui, Yang Song, Chen Sun, Andrew Howard, and Serge Belongie. 2018. Large scale fine-grained categorization and domain-specific transfer learning. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 4109–4118.

[23] Ranga Dabarera and Ranga Rodrigo. 2010. Vision Based Elephant Recognition for Management and Conservation. *Fifth International Conference on Information and Automation for Sustainability* (2010), 163–166. https://doi.org/10.1109/ICIAFS.2010.5715653

[24] Debayan Deb, Susan Wiper, Sixue Gong, Yichun Shi, Cori Tymoszek, Alison Fletcher, and Anil K Jain. 2018. Face recognition: Primates in the wild. In *2018 IEEE 9th International Conference on Biometrics Theory, Applications and Systems (BTAS)*. IEEE, 1–10.

[25] Alexander Freytag, Erik Rodner, Marcel Simon, Alexander Loos, Hjalmar S Kühl, and Joachim Denzler. 2016. Chimpanzee faces in the wild: Log-euclidean CNNs for predicting identities and attributes of primates. In *German Conference on Pattern Recognition*. Springer, 51–63.

[26] Gary Haynes. 2012. Elephants (and extinct relatives) as earth-movers and ecosystem engineers. *Geomorphology* 157 (2012), 99–107.

[27] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2015. Deep Residual Learning for Image Recognition. arXiv:1512.03385 [cs.CV]

[28] Qi He, Qijun Zhao, Ning Liu, Peng Chen, Zhihe Zhang, and Rong Hou. 2019. Distinguishing individual red pandas from their faces. In *Chinese Conference on Pattern Recognition and Computer Vision (PRCV)*. Springer, 714–724.

[29] Lex Hiby and Phil Lovell. 1990. Computer aided matching of natural markings: a prototype system for grey seals. *Report of the International Whaling Commission* 12 (1990), 57–61.

[30] Lex Hiby, Phil Lovell, Narendra Patil, N Samba Kumar, Arjun M Gopalaswamy, and K Ullas Karanth. 2009. A tiger cannot change its stripes: using a three-dimensional model to match images of living tigers and tiger skins. *Biology letters* 5, 3 (2009), 383–386.

[31] GR Hillman, B Wursig, GA Gailey, N Kehtarnavaz, A Drobyshevsky, BN Araabi, HD Tagare, and DW Weller. 2003. Computer-assisted photo-identification of individual marine vertebrates: a multi-species system. *Aquatic Mammals* 29, 1 (2003), 117–123.

[32] Ruben Huele and Helias Udo de Haes. 1998. Identification of individual sperm whales by wavelet transform of the trailing edge of the flukes. *Marine Mammal Science* 14, 1 (1998), 143–145.

[33] Benjamin Hughes and Tilo Burghardt. 2017. Automated visual fin identification of individual great white sharks. *International Journal of Computer Vision* 122, 3

(2017), 542–557.

[34] Marcella J Kelly. 2001. Computer-aided photograph matching in studies using individual identification: an example from Serengeti cheetahs. *Journal of Mammalogy* 82, 2 (2001), 440–449.

[35] Matthias Körschens, Björn Barz, and Joachim Denzler. 2018. Towards Automatic Identification of Elephants in the Wild. *arXiv:1812.04418 [cs]* (2018). http://arxiv.org/abs/1812.04418 arXiv: 1812.04418.

[36] Matthias Korschens and Joachim Denzler. 2019. Elpephants: A fine-grained dataset for elephant re-identification. In *Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops*. 0–0.

[37] Zhangyong Li, Chao Ge, Siwan Shen, and Xinwei Li. 2018. Cow individual identification based on convolutional neural network. In *Proceedings of the 2018 International Conference on Algorithms, Computing and Artificial Intelligence*. 1–5.

[38] Shengcai Liao, Yang Hu, Xiangyu Zhu, and Stan Z Li. 2015. Person re-identification by local maximal occurrence representation and metric learning. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2197–2206.

[39] Tsung-Yi Lin, Piotr Dollár, Ross Girshick, Kaiming He, Bharath Hariharan, and Serge Belongie. 2017. Feature pyramid networks for object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2117–2125.

[40] Tsung-Yi Lin, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C Lawrence Zitnick. 2014. Microsoft coco: Common objects in context. In *European conference on computer vision*. Springer, 740–755.

[41] Alexander Loos and Andreas Ernst. 2013. An automated chimpanzee identification system using face detection and recognition. *EURASIP Journal on Image and Video Processing* 2013, 1 (2013), 1–17.

[42] Oisin Mac Aodha, Elijah Cole, and Pietro Perona. 2019. Presence-only geographical priors for fine-grained image classification. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 9596–9606.

[43] Fiona Maisels, Samantha Strindberg, Stephen Blake, George Wittemyer, John Hart, Elizabeth A. Williamson, Rostand Aba'a, Gaspard Abitsi, Ruffin D. Ambahe, Fid??l Amsini, Parfait C. Bakabana, Thurston Cleveland Hicks, Rosine E. Bayogo, Martha Bechem, Rene L. Beyers, Anicet N. Bezangoye, Patrick Boundja, Nicolas Bout, Marc Ella Akou, Lambert Bene Bene, Bernard Fosso, Elizabeth Greengrass, Falk Grossmann, Clement Ikamba-Nkulu, Omari Ilambu, Bila Isia Inogwabini, Fortune Iyenguet, Franck Kiminou, Max Kokangoye, Deo Kujirakwinja, Stephanie Latour, Innocent Liengola, Quevain Mackaya, Jacob Madidi, Bola Madzoke, Calixte Makoumbou, Guy Aim?? Malanda, Richard Malonga, Olivier Mbani, Valentin A. Mbendzo, Edgar Ambassa, Albert Ekinde, Yves Mihindou, Bethan J. Morgan, Prosper Motsaba, Gabin Moukala, Anselme Mounguengui, Brice S. Mowawa, Christian Ndzai, Stuart Nixon, Pele Nkumu, Fabian Nzolani, Lilian Pintea, Andrew Plumptre, Hugo Rainey, Bruno Bokoto de Semboli, Adeline Serckx, Emma Stokes, Andrea Turkalo, Hilde Vanleeuwe, Ashley Vosper, and Ymke Warren. 2013. Devastating Decline of Forest Elephants in Central Africa. *PLoS ONE* 8 (2013). https://doi.org/10.1371/journal.pone.0059469

[44] Sancho McCann and David G Lowe. 2012. Local naive bayes nearest neighbor for image classification. In *2012 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 3650–3656.

[45] Vincent Miele, Gaspard Dussert, Bruno Spataro, Simon Chamaillé-Jammes, Dominique Allainé, and Christophe Bonenfant. 2020. Revisiting giraffe photo-identification using deep learning and network analysis. *bioRxiv* (2020).

[46] Sally A Mizroch, Judith A Beard, and Macgill Lynde. 1990. Computer assisted photo-identification of humpback whales. *Report of the International Whaling Commission* 12 (1990), 63–70.

[47] Robin Naidoo, Brendan Fisher, Andrea Manica, and Andrew Balmford. 2016. Estimating economic losses to tourism in Africa from the illegal killing of elephants. *Nature communications* 7, 1 (2016), 1–9.

[48] Mohammad Sadegh Norouzzadeh, Dan Morris, Sara Beery, Neel Joshi, Nebojsa Jojic, and Jeff Clune. 2019. A deep active learning system for species identification and counting in camera trap images. *arXiv preprint arXiv:1910.09716* (2019).

[49] Joy Omulupi Ontita, Jennifer Weston10, George Anyona11, Geoffrey Chege12, David Kimiti12, Kaia Tombak, Andrew Gersick, and Nancy Rubenstein. [n.d.]. The State of Kenya's Grevy's Zebras and Reticulated Giraffes: Results of the Great Grevy's Rally 2018. ([n. d.]).

[50] Jason Parham, Jonathan Crall, Charles Stewart, Tanya Berger-Wolf, and Daniel I Rubenstein. 2017. Animal population censusing at scale with citizen science and photographic identification. In *AAAI Spring Symposium-Technical Report*.

[51] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, et al. 2019. Pytorch: An imperative style, high-performance deep learning library. *arXiv preprint arXiv:1912.01703* (2019).

[52] Joyce Poole. [n.d.]. How to identify African elephants. https://elephantvoices.org/multimedia-resources/how-to-identify-african-elephants.html

[53] Robert M Pringle. 2008. Elephants as agents of habitat creation for small vertebrates at the patch scale. *Ecology* 89, 1 (2008), 26–33.

[54] Prashanth C Ravoor and TSB Sudarshan. 2020. Deep Learning Methods for Multi-Species Animal Re-identification and Tracking–A Survey. *Computer Science Review* 38 (2020), 100289.

[55] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. 2015. Faster r-cnn: Towards real-time object detection with region proposal networks. *arXiv preprint arXiv:1506.01497* (2015).

[56] J. Andrew Royle, Richard B. Chandler, Rahel Sollmann, and Beth Gardner. 2013. *Spatial Capture-Recapture*. Academic Press.

[57] Stefan Schneider, Graham W Taylor, and Stefan Kremer. 2018. Deep learning object detection methods for ecological camera trap data. In *2018 15th Conference on Computer and Robot Vision (CRV)*. IEEE, 321–328.

[58] Stefan Schneider, Graham W Taylor, and Stefan C Kremer. 2020. Similarity learning networks for animal individual re-identification-beyond the capabilities of a human observer. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision Workshops*. 44–52.

[59] Stefan Schneider, Graham W Taylor, Stefan Linquist, and Stefan C Kremer. 2019. Past, present and future approaches using computer vision for animal re-identification from camera trap data. *Methods in Ecology and Evolution* 10, 4 (2019), 461–470.

[60] Daniel Schofield, Arsha Nagrani, Andrew Zisserman, Misato Hayashi, Tetsuro Matsuzawa, Dora Biro, and Susana Carvalho. 2019. Chimpanzee face recognition from videos in the wild using deep learning. *Science advances* (2019).

[61] Kenya WildlIfe Service. [n.d.]. AERIAL TOTAL COUNT OF ELEPHANTS, BUFFALOES AND GIRAFFES IN THE MASAI MARA ECOSYSTEM (MAY 2017) - Kenya Wildlife Service. http://www.kws.go.ke/content/aerial-total-count-elephants-buffaloes-and-giraffes-masai-mara-ecosytem-may-2017-0

[62] Ankita Shukla, Gullal Sigh Cheema, Pei Gao, Suguru Onda, Divyam Anshumaan, Saket Anand, Ryan Farrell, et al. 2019. A hybrid approach to tiger re-identification. In *Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops*. 0–0.

[63] Michael A Tabak, Mohammad S Norouzzadeh, David W Wolfson, Erica J Newton, Raoul K Boughton, Jacob S Ivan, Eric A Odell, Eric S Newkirk, Reesa Y Conrey, Jennifer L Stenglein, et al. 2020. Improving the accessibility and transferability of machine learning algorithms for identification of animals in camera trap images: MLWIC2. *bioRxiv* (2020).

[64] Christopher Town, Andrea Marshall, and Nutthaporn Sethasathien. 2013. Manta Matcher: automated photographic identification of manta rays using keypoint features. *Ecology and evolution* 3, 7 (2013), 1902–1914.

[65] Grant Van Horn, Steve Branson, Ryan Farrell, Scott Haber, Jessie Barry, Panos Ipeirotis, Pietro Perona, and Serge Belongie. 2015. Building a bird recognition app and large scale dataset with citizen scientists: The fine print in fine-grained dataset collection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 595–604.

[66] Grant Van Horn, Elijah Cole, Sara Beery, Kimberly Wilber, Serge Belongie, and Oisin Mac Aodha. 2021. Benchmarking Representation Learning for Natural World Image Collections. *arXiv preprint arXiv:2103.16483* (2021).

[67] Grant Van Horn, Oisin Mac Aodha, Yang Song, Yin Cui, Chen Sun, Alex Shepard, Hartwig Adam, Pietro Perona, and Serge Belongie. 2018. The inaturalist species classification and detection dataset. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 8769–8778.

[68] Maxime Vidal, Nathan Wolf, Beth Rosenberg, Bradley P Harris, and Alexander Mathis. 2021. Perspectives on individual animal identification from biology and computer vision. *arXiv preprint arXiv:2103.00560* (2021).

[69] Jake Wall, George Wittemyer, Brian Klinkenberg, and Iain Douglas-Hamilton. 2014. Novel Opportunities for Wildlife Conservation and Research with Real-Time Monitoring. *Ecological Applications* 24, 4 (2014), 593–601.

[70] Jake Wall, George Wittemyer, Brian Klinkenberg, Valerie LeMay, Stephen Blake, Samantha Strindberg, Michelle Henley, Fritz Vollrath, Fiona Maisels, Jelle Ferwerda, and Iain Douglas-Hamilton. 2021. Human footprint and protected areas shape elephant range across Africa. *Current Biology* (April 2021). https://doi.org/10.1016/j.cub.2021.03.042

[71] Hendrik J. Weideman, Zachary M. Jablons, Jason Holmberg, Kiirsten Flynn, John Calambokidis, Reny B. Tyson, Jason B. Allen, Randall S. Wells, Krista Hupman, Kim Urian, and Charles V. Stewart. 2017. Integral Curvature Representation and Matching Algorithms for Identification of Dolphins and Whales. *arXiv:1708.07785 [cs]* (2017). http://arxiv.org/abs/1708.07785 arXiv: 1708.07785.

[72] Hendrik J. Weideman, Charles V. Stewart, Jason R. Parham, Jason Holmberg, Kiirsten Flynn, John Calambokidis, D. Barry Paul, Anka Bedetti, Michelle Henley, Jerenimo Lepirei, and Frank G. Pope. 2020. Extracting identifying contours for African elephants and humpback whales using a learned appearance model. *The IEEE Winter Conference on Applications of Computer Vision (WACV)* (2020), 1276–1285.

[73] Fei Xiong, Mengran Gou, Octavia Camps, and Mario Sznaier. 2014. Person re-identification using kernel-based metric learning methods. In *European conference on computer vision*. Springer, 1–16.

[74] Dong Yi, Zhen Lei, Shengcai Liao, and Stan Z Li. 2014. Deep metric learning for person re-identification. In *2014 22nd International Conference on Pattern Recognition*. IEEE, 34–39.

[75] Zhi-Hua Zhou. 2018. A brief introduction to weakly supervised learning. *National science review* 5, 1 (2018), 44–53.