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

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.
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 (Loxodonta 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.
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
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, 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
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
**EAR Feature**

| Code | Position on the Ear |
|------|---------------------|
| 3    | Position 3          |
| 4    | Position 4          |
| 5    | Position 5          |
| 7    | Position 7          |
| 8    | Position 8          |
| 9    | Position 9          |
|     | Other               |
| 0    | Not present         |

**Position**: clock-face “3” to “9”

**(1)** = Position of the most prominent Tear
**(2)** = Position of the most prominent Hole
**(3)** = Position of the second most prominent Tear
**(4)** = Position of the second most prominent Hole

**RULES FOR CODING**

- Presence of two equally “obvious” features of the same type (tears or holes) → the deepest tear and biggest diameter hole will be coded for first.
- Presence of two equally “obvious” features of the same type and same size are present → the higher positioned feature will be coded for first.

**Code**

| Code | Definition |
|------|------------|
| B    | Elephant Bull |
| C    | Elephant Cow |

**TUSK Feature**

| Code | Definition |
|------|------------|
| 0    | Absence    |
| 1    | Presence   |

**EXTREME Feature**

A tear or a hole is classified as extreme when it extends 1/4 or more in length towards the inner ear and/or is 1/4 of the total ear margin width.

| Code | Definition |
|------|------------|
| 0    | Absence    |
| 1    | Presence   |

**AGE**

| Code | Birth Year Bracket |
|------|--------------------|
| 60   | 1900 - 1969        |
| 70   | 1970 – 1979        |
| 80   | 1980 – 1989        |
| 90   | 1990 – 1999        |
| 00   | 2000 – 2009        |
| 10   | 2010 – 2019        |

Refer to Appendix 2 for details

**SPECIAL Features**

| Code | Definition |
|------|------------|
| 0    | Absence    |
| 1    | Presence   |

**Note**: Broken tusks are not recorded as it can regrow or break again.

**Figure 5: The SEEK coding system**
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 our 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 annotators 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].
We expect additional sightings to improve accuracy, as it presents a pair of SEEK codes and 33 individuals with a triplet of codes. Optimizations that mimic a typical re-identification workflow. 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.

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 straightforward 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.

ACKNOWLEDGMENTS

We would like to thank the entire team at the Mara Elephant Project for their efforts in deploying this system. This work was supported, through funding, data storage, and computing resources, by Microsoft AI for Earth, the Caltech Resnick Sustainability Institute, and NSFGRFP Grant No. 1745301, the views are those of the authors and do not necessarily reflect the views of these organizations.

REFERENCES

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