The AAAI 2017 Spring Symposium on
Artificial Intelligence for the Social Good
Technical Report SS-17-01

Animal Population Censusing at Scale with Citizen Science and Photographic Identification

Jason Parham, Jonathan Crall
Charles Stewart
Rensselaer Polytechnic Institute
Troy, New York 12180

Tanya Berger-Wolf
University of Illinois, Chicago
Chicago, Illinois 60607

Daniel Rubenstein
Princeton University
Princeton, New Jersey 08544

Abstract

Population censusing is critical to monitoring the health of an animal population. A census results in a population size estimate, which is a fundamental metric for deciding the demographic and conservation status of a species. Current methods for producing a population census are expensive, demanding, and may be invasive, leading to the use of overly-small sample sizes. In response, we propose to use volunteer citizen scientists to collect large numbers of photographs taken over large geographic areas, and to use computer vision algorithms to semi-automatically identify and count individual animals. Our data collection and processing are distributed, non-invasive, and require no specialized hardware and no scientific training. Our method also engages the community directly in conservation. We analyze the results of two population censusing events, the Great Zebra and Giraffe Count (2015) and the Great Grevy's Rally (2016), where combined we processed over 50,000 photographs taken with more than 200 different cameras and over 300 on-the-ground volunteers.

Introduction

Knowing the number of individual animals within a population (a population census) is one of the most important statistics for research and conservation management in wildlife biology. Moreover, a *current* population census is often needed repeatedly over time in order to understand changes in a population's size, demographics, and distribution. This enables assessments of the effects of ongoing conservation management strategies. Furthermore, the number of individuals in a population is seen as a fundamental basis for determining its conservation status. The IUCN Red List[1], which tracks the conservation status of species around the world, currently includes 83,000 species; of those a full 30% are considered threatened or worse. Therefore, it can be vital to perform a massive, species-wide effort to count every individual in a population. As it has recently been shown for the African Savannah elephant in mid-2016 (Chase et al. 2016), population censuses can be crucial in monitoring and protecting threatened species from extinction.

Unfortunately, producing a population census is difficult to do at scale and across large geographical areas us-

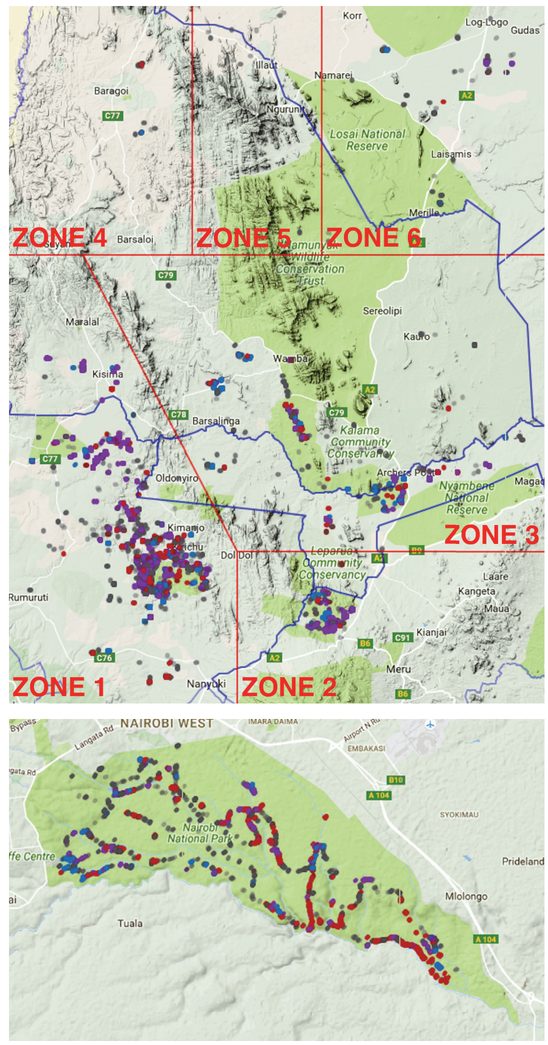

Figure 1: The locations of photographs taken during the GGR (top) and the GZGC (bottom). Colored dots indicate sightings during the two days of each census; red were from day 1 only, blue were day 2 only, purple were resightings, and gray were unused. Rendered with Google Maps. Best viewed in color.

[1]redlist.org

ing traditional, manual methods. One of the most popular and prevalent techniques for producing a population size estimate is mark-recapture (Robson and Regier 1964; Pradel 1996) via a population count. However, performing a mark-recapture study can be prohibitively demanding when the number of individuals in a population grows too large (Seber 2002), the population moves across large distances, or the species is difficult to capture due to evasiveness or habitat inaccessibility. More importantly, however, a population *count* is not as robust as a population *census*; a count tracks sightings whereas a census tracks individuals. A census is stronger because it can still produce a population size estimate implicitly but also unlocks more powerful ecological metrics that can track the long-term trends of individuals. In recent years, technology has been used to help improve censusing efforts towards more accurate population size estimates (Chase et al. 2016; Forrester et al. 2014; Simpson, Page, and De Roure 2014; Swanson et al. 2015) and scale up[2]. However, these types of population counts are still typically custom, one-off efforts, with no uniform collection protocols or data analysis, and do not attempt to accurately track *individuals* within a population across time.

To address the problems with collecting data and producing a scalable population census, we propose:

- using citizen scientists (Irwin 1995; Cohn 2008) to rapidly collect a large number of photographs over a short time period (e.g. two days) and over an area that covers the expected population, and
- using computer vision algorithms to process these photographs semi-automatically to identify all seen animals.

We show that this proposed process can be leveraged at scale and across large geographical areas by analyzing the results of two completed censuses. The first census is the Great Zebra and Giraffe Count (GZGC) held March 1-2, 2015 at the Nairobi National Park in Nairobi, Kenya to estimate the resident populations of Masai giraffes (*Giraffa camelopardalis tippelskirchi*) and plains zebras (*Equus quagga*). The second is the Great Grevy's Rally (GGR) held January 30-31, 2016 in a region of central and northern Kenya covering the known migratory range of the endangered Grevy's zebra (*Equus grevyi*). See Figure 1 for a map of the collected photographs during these rallies.

While our method relies heavily on collecting a large number of photographs, it is designed to be simple enough for the average person to help collect them. Any volunteers typically must only be familiar with a digital camera and be able to follow a small set of collection guidelines. The ease of collecting photographs means a large population can be sampled simultaneously over a large geographical area. Furthermore, our method requires no special equipment other than a standard, mid-range camera and some form of ground transportation. By not requiring specialized hardware (i.e. camera traps, radio collars, drones), transportation (i.e. planes), or special education (i.e. ecology researchers, veterinarians), our method allows for the community to engage in conservation.

[2] penguinwatch.org, mturk.com

Methods

Our censusing rallies are structured around the traditional protocols of mark-recapture and use citizen scientists to collect a large volume of data. These photographs are processed by computer vision algorithms that detect animals of desired species and determine which photographs show the same animal. Results are then reviewed by trained individuals, resulting in the semi-automatic generation of the data needed for a population size estimate. Expert ecologists add final meta-data about individuals, such as age and sex, to generate demographics for a population.

Transforming Mark-Recapture into Sight-Resight

Mark-recapture is a standard way of estimating the size of an animal population (Chapman and Chapman 1975; Pradel 1996; Robson and Regier 1964). Typically, a portion of the population is captured at one point in time and the individuals are marked *as a group*. Later, another portion of the population is captured and the number of previously marked individuals is counted and recorded. Since the number of marked individuals in the second sample should be proportional to the number of marked individuals in the entire population (assuming consistent sampling processes and controlled biases), the size of the entire population can be estimated.

The population size estimate is calculated by dividing the total number of marked individuals during the first capture by the proportion of marked individuals counted in the second. The formula for the simple Lincoln-Peterson estimator (Pacala and Roughgarden 1985) is:

$$N_{\text{est}} = \frac{K * n}{k}$$

where N_{est} is the population size estimate, n is the number of individuals captured and marked during the first capture, K is the number of individuals captured during the second capture, and k is the number of *recaptured* individuals that were marked from the first capture.

Applying the Lincoln-Peterson estimator requires that several assumptions be met. Chiefly, no births, deaths, immigrations or emigrations should take place and the sightability of individuals must be equal between sightings. Sampling on consecutive days reduces the likelihood of violating the first two assumptions for most large mammal species. Furthermore, by assigning multiple teams of volunteers to traverse the survey area, the number of overall sightings can be increased. More sightings on the first day means better population coverage and more resightings on the second day gives a better population size estimate. By intensively sampling a survey area (that may haphazardly overlap), the confidence for equal sightability is high and identical for any given individual in the population. Therefore, all of the principle assumptions for the Lincoln-Peterson estimator can be satisfied. Finally, the coordination of volunteers for a two-day collection can be structured into a "rally" that focuses specifically on upholding these sampling assumptions. The number of cars, volunteers, and the number of photographs taken for both rallies can be seen in Table 1. Importantly, since the volunteers taking photographs are *mobile* they are able to go

	Cars	Cameras	Photographs
GZGC	27	55	9,406
GGR	121	162	40,810

Table 1: The number of cars, participating cameras (citizen scientists), and photographs collected between the GZGC and the GGR. The GGR had over 3-times as many citizen scientists who contributed 4-times the number of photographs for processing.

	Annots.	Individuals	Estimate
GZGC Masai	466	103	119 ± 4
GZGC Plains	4,545	1,258	$2{,}307 \pm 366$
GGR Grevy's	16,866	1,942	$2{,}250 \pm 93$

Table 2: The number of annotations, matched individuals, and the final mark-recapture population size estimates for the three species. The Lincoln-Peterson estimate has a 95% confidence range.

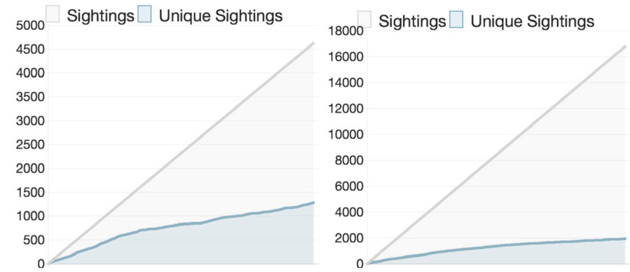

Figure 2: The convergence of the identification algorithm during the GZGC (left) and the GGR (right). The x-axis shows all collected photographs in chronological order and the y-axis shows the rate of new sightings. As photos are processed over time, the rate of new sightings decreases. The smaller slope of the GGR indicates that the rate of resightings for the GGR was higher than the GZGC.

where the animals are, in contrast to static camera traps or fixed-route surveys.

For distinctively marked species (e.g. zebras, giraffes, leopards) a high-quality photograph can serve as a non-invasive way of "capturing" and cataloging the natural markings of the animal. In this way the mark-recapture technique is transformed into a minimally-disturbing sight-resight approach (Bolger et al. 2012; Hiby et al. 2013). A sight-resight study can be used to estimate a population's size, but it provides photographic-based evidence for the seen individuals in a population. This evidence allows more detailed analysis of the population, access to more insightful metrics (e.g. individual life-expectancy), and allows for performing recounts. Furthermore, our method is not crippled by duplicate sightings; rather it depends crucially on them. In contrast to population size estimates that rely merely on sighting counts taken in counting blocks, we embrace resightings as they do not cause double-counting.

By giving the collected photographs to a computer vision pipeline, a semi-automated and more sophisticated census can be made. The speed of processing large qualities of photographs allows for a more thorough analysis of the age-structure of a population, the distribution of males and females, and the movements of individuals and groups of animals, etc. By tracking individuals, related to (Jolly 1965; Seber 1965), our method is able to make more confident claims about statistics for the population. The more individuals that are sighted *and* resighted, the more robust the estimate and ecological analyses will be.

Citizen Scientists and Data Collection Biases

The photographers for the GZGC were recruited both from civic groups and by asking for volunteers at the entrance gate in Nairobi National Park on the two days of the rally. Photographers for the GGR were partially from civic and conservation groups as well as field scouts, technicians, and scientists. All volunteers were briefly trained in a collection protocol and tasked to take pictures of animals within specific, designated regions. These regions helped to enforce better coverage and prevent a particular area from becoming uselessly over-sampled.

All photographers for the GZGC were requested to take pictures of the left sides of plains zebras and Masai giraffes, while photographers for the GGR were requested to take pictures of the right sides of Grevy's zebras. Having a consistent viewpoint (left or right) allows for effective sight-resight and minimizes bias; the distinguishing visual markings for the three species of focus are not left-right symmetrical and the animal's appearance differs (sometimes significantly) from side to side.

Photographers were shown examples of good/poor quality photographs emphasizing (a) the side of the animal, (b) getting a large enough and clear view, and (c) seeing the animal in relative isolation from other animals. GGR photographers were requested to take about three pictures of the right side of each Grevy's they saw. In both the GZGC and GGR photographers were invited to take other pictures once they had properly photographed each animal encountered, causing miscellaneous photographs to be collected.

Like all data, photographic samples of animal ecology are biased. To administer a correct population census, we must take these biases into account explicitly as different sources of photographs inherently come with different forms of bias. For example, stationary camera traps, cameras mounted on moving vehicles, and drones are each biased by their location, by the presence of animals at that location, by photographic quality, and by the camera settings (such as sensitivity of the motion sensor) (Hodgson, Kelly, and Peel 2013; Hombal, Sanderson, and Blidberg 2010; Foster and Harmsen 2012; Rowcliffe et al. 2013). These factors result in biased samples of species and spatial distributions, which recent studies are trying to overcome (Ancrenaz et al. 2012; Maputla, Chimimba, and Ferreira 2013; Xue et al. 2016).

Any human observer, including scientists and trained field assistants, is affected by observer bias (Marsh and Hanlon 2004; 2007). Specifically, the harsh constraint of being at a

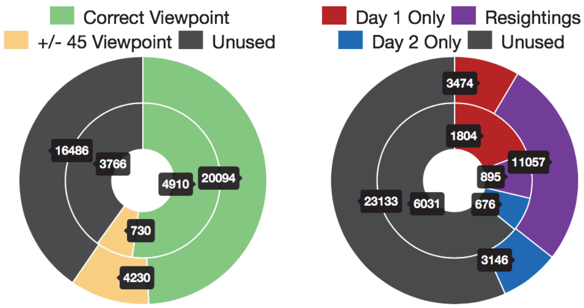

Figure 3: The breakdown (left) of photographs that adhered to the collection protocol for the GZGC (inner-ring) and the GGR (outer-ring). The number of photographs that adhered exactly to the viewpoint collection protocol was around 50% (green) for both the GGR and the GZGC. The breakdown (right) of which photographs had sightings on day 1 only, day 2 only, and resightings for the GZGC (inner-ring) and GGR (outer-ring); the sightings data and its colors are meant to mirror that of Figure 1. Note that any photographs with no sightings are grouped with unused.

single given location at a given time makes sampling arbitrary. Citizen scientists, as the foundation of the data collection, have additional variances in a wide range of training, expertise, goal alignment, sex, age, etc. (Dickinson, Zuckerberg, and Bonter 2010). Nonetheless, recent ecological studies are starting to successfully take advantage of this source of data, explicitly testing and correcting for bias (van Strien, van Swaay, and Termaat 2013); recent computational approaches address the question of if and how data from citizen scientists can be considered valid (Wiggins et al. 2011), which can be leveraged with new studies in protocol design and validation. Moreover, combining these differently biased sources of data mutually constrains these biases and allows much more accurate statistical estimates than any one source of data would individually allow (Bird et al. 2014). In this study, we explicitly test for biases that may affect the end results, detailed in the Results section.

IBEIS Computer Vision Pipeline

Our computer vision pipeline includes two components: detection and identification. Detection locates animals in photographs, determines their species, and draws a bounding box around each animal to produce an *annotation*. Importantly, detection also includes labeling the viewpoint on the animal relative to the camera and determining the photographic quality. An annotation is "low quality" if it is too small, blurry, or poorly illuminated, or if the animal is occluded by vegetation or other animals – anything that makes the animal hard to identify. The viewpoint is the "side" of the animal photographed; our implementation includes eight possible labels for viewpoint: left, back-left, back, back-right, right, front-right, front, and front-left. The number of annotations collected for each species can be seen in Table 2.

Our detection pipeline is a cascade of deep convolutional neural networks (DCNNs) that applies a fully-connected classifier on extracted convolutional feature vectors. Three separate networks produce: (1) whole-scene classifications looking for specific species of animals in the photograph, (2) object annotation bounding box localizations, and (3) the viewpoint, quality, and *final* species classifications for the candidate bounding boxes proposed by network 2. Networks 1 and 3 are custom networks based on a structure similar to OverFeat (Sermanet et al. 2013) whereas network 2 uses the structure of the YOLO network by (Redmon et al. 2015). Importantly, the species classifications provided by network 2 are replaced by the species classification from network 3, which results in an increase in performance for our species of interest, as shown in (Parham and Stewart 2016) .

The three networks are trained separately with different data. Training data for the detection pipeline is collected and verified using a web-based interface through which reviewers can draw bounding boxes, label species, and determine quality and viewpoint. A similar interface is used to verify and modify the results when processing the photographs from a census' data collection. See (Parham 2015) for implementation details.

The second major computer vision step is identification, which assigns a name label to each annotation or, viewed conversely, forms clusters of annotations that were taken of the same individual. To do this, SIFT descriptors (Lowe 2004) are first extracted at keypoint locations (Perdoch, Chum, and Matas 2009) from each annotation. Descriptors are gathered into an approximate nearest-neighbor (ANN) search data structure (Muja and Lowe 2009). Each annotation is then, in turn, treated as a query annotation against this ANN index. For each descriptor from the query, the closest matching descriptors are found. Matches in the sparser regions of descriptor space (i.e. those that are most distinctive) are assigned higher scores using a "Local Naive Bayes Nearest Neighbor" method (McCann and Lowe 2012). The scores from the query that match the same individual are accumulated to produce a single score for each animal. The animals in the database are then ranked by their accumulated scores. A post-processing step spatially verifies the descriptor matches and then re-scores and re-ranks the database individuals (Philbin et al. 2007). These ranked lists are merged across all query annotations and ordered by scores.

At this point, the potentially-matching pairs of annotations are shown in rank order to human reviewers who make the final decisions about pairs of annotations that are indeed of the same animal. After processing a number of pairs (typically around 10-15%) the matching process is repeated (with extensive caching to prevent recomputing unchanged data) using the decisions made by reviewers to avoid unnecessary scoring competition between annotations from the same animal. Looking at Figure 2, as matching is performed, the rate of finding new animals slows. The ranking, display, and human-decision making processes are repeated, with previously-made decisions hidden from the reviewer. This overall process is repeated until no new matches need to be reviewed. Final consistency checks are applied, again using the basic matching algorithm, to find and "split" clusters of annotations falsely marked as all the same animal.

The semi-automatic nature of these algorithms comes from

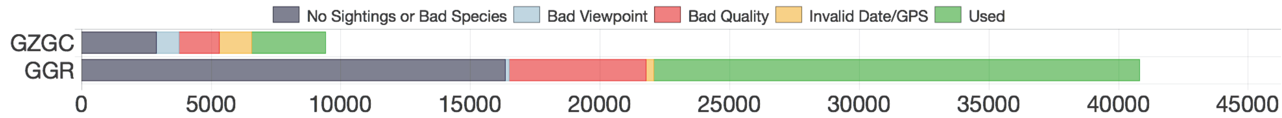

Figure 4: The breakdown of collected photographs and how they were used for the two censuses. A large number (gray) were filtered out simply because they had no sightings or captured miscellaneous species. We further filtered the photographs taken of undesired viewpoints and had poor quality. Lastly, we filtered photographs that were not taken during the two days of each rally (some volunteers brought their own cameras with non-empty personal memory cards) or had corrupt/invalid GPS.

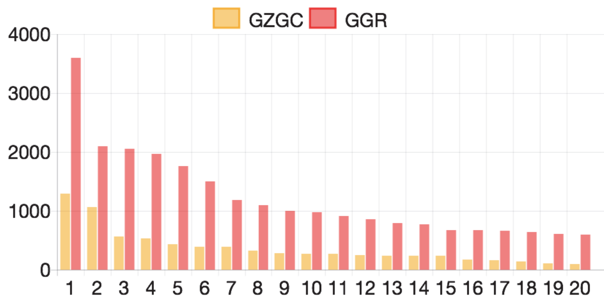

Figure 5: The number of photographs taken by the top 20 cars during the GZGC and the GGR

the fact that some detection decisions and all positive identification results are reviewed before they are accepted into the analysis. This is discussed in more depth at the end of the results section.

Results

The IDs of the animals in the photographs – the final output of the computer vision pipeline – are combined with their date/time-stamps to determine when and where an animal was seen. Knowing the number of sightings on day 1 only, day 2 only, and resightings between both days allows a Peterson-Lincoln estimate to be calculated as in a traditional mark-recapture study (Table 2). We can use embedded GPS meta-data, along with knowledge of which cameras and cars photographs were taken from, to analyze the spatial and temporal distributions of the data and the distributions by car and photographer.[3]

Sampling with Citizen Scientists

First, we analyze how well the citizen scientists followed the data collection protocols. As discussed earlier, citizen scientists were instructed to first take photographs from specific viewpoints on the animals – left side during the GZGC and right sides for Grevy's zebras (GGR) – and then take additional photographs if they desired. Hence, the distribution of viewpoints is a strong indicator of the adherence to the protocol. Figure 3 (left) shows that for both the GZGC and the GGR around 50% of the photographs had an annotation from the desired viewpoint (green). Furthermore, when the photographs of neighboring viewpoints (yellow) are taken into account, the percentage grows to 60%. A side note: the computer vision algorithms for identification can allow for up to a 30-degree deviance in perspective (Mikolajczyk et al. 2005), which means that these slightly-off viewpoints can still yield usable information in the photographic census.

The argument of good adherence is reinforced by the bar graph in Figure 4, which shows how the photographs were used during the analysis. The largest percentage of photographs filtered out did not include animals of the desired species. The next highest percentage was from poor photograph quality. Even so, the number of photographs used is still around 50% for the GGR. One can consider our data collection process to be equivalent to high throughput data in noise processing, where our signal-to-noise ratio is clearly above the generally-accepted threshold of 30%. Encouragingly, the number of sightings and resightings also exceeds this threshold, seen in Figure 3 (right).

Figure 5 plots the distribution of photographs per camera. There are several possible reasons for the observed drop-off. In the GZGC, since some photographers were professional ecologists and conservationists, while others were volunteers recruited on site, we expected a significant difference in the commitment to take large numbers of photographs. For GGR, where volunteers were recruited in advance, the expertise was more uniform, but each car had an assigned region, and the regions differed significantly in the expected density of animals.

Despite this skewed distribution we still had strong area coverage, as the maps in Figure 1 show. Note that the maps shown in the figures are at vastly different scales and the coverage plots in the GZGC essentially show the roads through the Nairobi National Park. We split the park into 5 zones to help enforce coverage, which was very good in most cases. For the GGR, the 25,000 km^2 survey area was broken into 45 counting blocks with heavy variation in the animal density due to the presence of human settlements and the availability of habitat and resources to sustain Grevy's zebras. The spatial distributions of resightings are fairly uniform for both rallies and it indicates that the respective counting block partitioning schemes accomplished their intended goals.

Photographic Identification (Sight-Resight)

Next, we examine the reliability of the sight-resight population size estimate. Figure 2 plots the number of new animals identified vs. the number of processed photographs, ordered

[3]Portions of the results in this section were previously reported in two technical reports: (Rubenstein et al. 2015) for the GZGC and (Berger-Wolf et al. 2016) for the GGR.

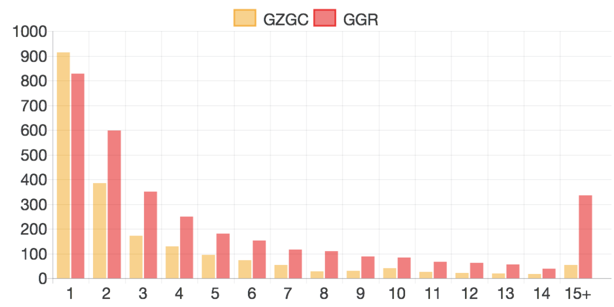

Figure 6: Histogram of the number of photos per animal for the GZC and the GGR. The total number of photos from the GGR is much higher than the GZGC and the number of 15+ photos much more saturated, indicating better coverage and that the number of resights should be much higher.

chronologically. Ideally, these curves should flatten out over time indicating a large fraction of the actual population has been seen. This trend is seen clearly for the GGR, but not as dramatically for the GZGC. While the slope of the GZGC curve does decline, its slower convergence suggests a lower relative percentage of resightings. This intuition is confirmed by Figure 3 (right) which explicitly shows a higher percentage of resights in the GGR.

Figure 6 plots a histogram of the number of photographs per animal. It shows that most frequently an animal was photographed only once during both rallies. The collection protocol encouraged volunteers to take three photographs of a sighted animal, which disagrees with this histogram. Encouragingly, the number of animals with single-sightings *decreased* between the GZGC and the GGR, even though the number of annotations more than tripled. This suggests that more a thorough sampling (i.e. more volunteers) and better training can help correct for this bias.

Algorithm Discussion

While the focus of this paper is not on the computer vision algorithms themselves, it is important to consider their role in the ability to scale population censusing. Working fully manually, with N photographs, $O(N^2)$ manual decisions are required to determine which photographs are of the same animals. Clearly, this process does not scale. Our current algorithms provide "suggestions" for pairs of photographs that show the same animal. Since at most two or three possible matches are shown per photograph, our current algorithms require $O(N)$ manual decisions. For 50,000 photographs and a small number of experts, this is still quite labor-intensive. Scaling beyond the current size requires either involvement of a much larger group of individuals (i.e. crowd-sourcing) in a distributed, citizen-science based decision-making process or automated decision-making whereby only a small subset of the decisions must be made manually. We are currently pursuing both directions.

Similarly, since our evidence shows that citizen science-based photograph collection has produced sufficient density, coverage, and adherence to the protocol, the question about

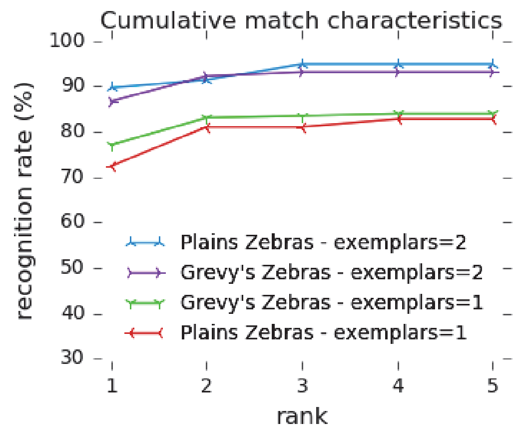

Figure 7: The accuracy of the identification algorithm as the percent of correct matches returned at or below each rank. We plot two curves for each species using 1 and 2 exemplars (other photographs of the same animals). This demonstrates the increase in accuracy due to having multiple sightings of an individual.

the accuracy of the count then depends on the accuracy of the identification process. Figure 7 shows that for a single run of the identification algorithm the correct match is found 80% of the time when there is only one other photo of the animal and over 90% when there are two — hence, the request for multiple photographs per animal. Since the matching process was repeated several times while factoring in previous matching decisions, and since the process was terminated only when further matches were no longer found deep in the ranked list, we have qualitative evidence that there are relatively few missed matches. The tightness of the Lincoln-Peterson estimate bounds shown in Table 2, especially for the GGR, supports this.

Conclusion

Our method has been shown to be a viable option for performing animal population censusing at scale. A photographic census can be an effective and less invasive method for producing an estimate of an animal population. Our estimates (Table 2) are consistent with previously known estimates for the resident population in the Nairobi National Park (Ogutu et al. 2013) and for Grevy's Zebra in Kenya (Ngene et al. 2013), but they provide tighter confidence bounds and a rich data source for further analysis about individual animals and their locations.

We have shown that citizen scientists can cover the needed area and take sufficient high-quality photographs from the required viewpoint to enable accurate, semi-automatic populations counts, driven by computer vision algorithms. Current limitations are that photographers don't quite gather enough redundant photographs and the counting methods should have a higher-degree of automation. Addressing these problems in the future will enable even higher volume censuses, faster processing, and greater accuracy.

References

Ancrenaz, M.; Hearn, A.; Ross, J.; Sollmann, R.; and Wilting, A. 2012. *Handbook for Wildlife Monitoring using Camera-traps*. Kota Kinabalu, Sabah, Malaysia: BBEC.

Berger-Wolf, T.; Crall, J.; Holberg, J.; Parham, J.; Stewart, C.; Mackey, B. L.; Kahumbu, P.; and Rubenstein, D. 2016. The Great Grevys Rally: The Need, Methods, Findings, Implications and Next Steps. Technical Report, Grevy's Zebra Trust, Nairobi, Kenya.

Bird, T. J.; Bates, A. E.; Lefcheck, J. S.; Hill, N. A.; Thomson, R. J.; Edgar, G. J.; Stuart-Smith, R. D.; Wotherspoon, S.; Krkosek, M.; Stuart-Smith, J. F.; Pecl, G. T.; Barrett, N.; and Frusher, S. 2014. Statistical solutions for error and bias in global citizen science datasets. *Biological Conservation* 173(1):144 – 154.

Bolger, D. T.; Morrison, T. A.; Vance, B.; Lee, D.; and Farid, H. 2012. A computer-assisted system for photographic markrecapture analysis. *Methods in Ecology and Evolution* 3(5):813–822.

Chapman, D., and Chapman, N. 1975. *Fallow deer: their history, distribution, and biology*. Ithaca, NY: Dalton.

Chase, M. J.; Schlossberg, S.; Griffin, C. R.; Bouch, P. J.; Djene, S. W.; Elkan, P. W.; Ferreira, S.; Grossman, F.; Kohi, E. M.; Landen, K.; Omondi, P.; Peltier, A.; Selier, S. J.; and Sutcliffe, R. 2016. Continent-wide survey reveals massive decline in African savannah elephants. *PeerJ* 4(1):e2354.

Cohn, J. P. 2008. Citizen science: Can volunteers do real research? *BioScience* 58(3):192–197.

Dickinson, J. L.; Zuckerberg, B.; and Bonter, D. N. 2010. Citizen Science as an Ecological Research Tool: Challenges and Benefits. *Annual Review of Ecology, Evolution, and Systematics* 41(1):149–172.

Forrester, T.; McShea, W. J.; Keys, R. W.; Costello, R.; Baker, M.; and Parsons, A. 2014. eMammalcitizen science camera trapping as a solution for broad-scale, long-term monitoring of wildlife populations. In *North America Congress for Conservation Biology*.

Foster, R. J., and Harmsen, B. J. 2012. A critique of density estimation from camera-trap data. *The Journal of Wildlife Management* 76(2):224–236.

Hiby, L.; Paterson, W. D.; Redman, P.; Watkins, J.; Twiss, S. D.; and Pomeroy, P. 2013. Analysis of photo-id data allowing for missed matches and individuals identified from opposite sides. *Methods in Ecology and Evolution* 4(3):252–259.

Hodgson, A.; Kelly, N.; and Peel, D. 2013. Unmanned Aerial Vehicles (UAVs) for Surveying Marine Fauna: A Dugong Case Study. *PLoS ONE* 8(11):e79556.

Hombal, V.; Sanderson, A.; and Blidberg, D. R. 2010. Multiscale adaptive sampling in environmental robotics. In *In Proceedings of the 2010 IEEE Conference on Multisensor Fusion and Integration for Intelligent Systems*, 80–87.

Irwin, A. 1995. *Citizen Science: A Study of People, Expertise and Sustainable Development*. Environment and Society. New York, NY: Routledge.

Jolly, G. M. 1965. Explicit estimates from capture-recapture data with both death and immigration-stochastic model. *Biometrika* 52(1/2):225–247.

Lowe, D. G. 2004. Distinctive image features from scale-invariant keypoints. *2004 International Conference on Computer Vision (ICCV)* 60(2):91–110.

Maputla, N. W.; Chimimba, C. T.; and Ferreira, S. M. 2013. Calibrating a camera trapbased biased markrecapture sampling design to survey the leopard population in the N'wanetsi concession, Kruger National Park, South Africa. *African Journal of Ecology* 51(3):422–430.

Marsh, D. M., and Hanlon, T. J. 2004. Observer gender and observation bias in animal behaviour research: experimental tests with red-backed salamanders. *Animal Behaviour* 68(6):1425 – 1433.

Marsh, D. M., and Hanlon, T. J. 2007. Seeing What We Want to See: Confirmation Bias in Animal Behavior Research. *Ethology* 113(11):1089–1098.

McCann, S., and Lowe, D. G. 2012. Local naive bayes nearest neighbor for image classification. In *In Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition*, 3650–3656.

Mikolajczyk, K.; Tuytelaars, T.; Schmid, C.; Zisserman, A.; Matas, J.; Schaffalitzky, F.; Kadir, T.; and Gool, L. V. 2005. A Comparison of Affine Region Detectors. *2005 International Conference on Computer Vision (ICCV)* 65(1-2):43–72.

Muja, M., and Lowe, D. G. 2009. Fast Approximate Nearest Neighbors with Automatic Algorithm Configuration. In *In Proceedings of the 12th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*, 331–340.

Ngene, S.; Mukeka, J.; Ihwagi, F.; Mathenge, J.; Wandera, A.; Anyona, G.; Tobias, N.; Kawira, L.; Muthuku, I.; Kathiwa, J.; and others. 2013. Total aerial count of elephants, Grevyszebra and other large mammals in Laikipia-Samburu-Marsabit Ecosystem in (November 2012). Technical Report, Kenya Wildlife Service, Nairobi, Kenya.

Ogutu, J. O.; Owen-Smith, N.; Piepho, H.-P.; Said, M. Y.; Kifugo, S. C.; Reid, R. S.; Gichohi, H.; Kahumbu, P.; and Andanje, S. 2013. Changing wildlife populations in Nairobi National Park and adjoining Athi-Kaputiei Plains: collapse of the migratory wildebeest. *Open Conservation Biology Journal* 7(1):11–26.

Pacala, S. W., and Roughgarden, J. 1985. Population experiments with the Anolis lizards of St. Maarten and St. Eustatius. *Ecology* 66(1):129–141.

Parham, J., and Stewart, C. 2016. Detecting Plains and Grevy's Zebras in th Real World. In *2016 IEEE Winter Applications of Computer Vision Workshops*, 1–9.

Parham, J. R. 2015. *Photographic censusing of zebra and giraffe in the Nairobi National Park*. M.S. thesis, Department of Computer Science, Rensselaer Polytechnic Institute, Troy, NY.

Perdoch, M.; Chum, O.; and Matas, J. 2009. Efficient representation of local geometry for large scale object retrieval. In

In Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition, 9–16.

Philbin, J.; Chum, O.; Isard, M.; Sivic, J.; and Zisserman, A. 2007. Object retrieval with large vocabularies and fast spatial matching. In *In Proceedings of the 2007 IEEE Conference on Computer Vision and Pattern Recognition*, 1–8.

Pradel, R. 1996. Utilization of capture-mark-recapture for the study of recruitment and population growth rate. *Biometrics* 52(2):703–709.

Redmon, J.; Divvala, S.; Girshick, R.; and Farhadi, A. 2015. You only look once: Unified, real-time object detection. *CoRR* abs/1506.02640:1–10.

Robson, D. S., and Regier, H. A. 1964. Sample size in petersen markrecapture experiments. *Transactions of the American Fisheries Society* 93(3):215–226.

Rowcliffe, J. M.; Kays, R.; Carbone, C.; and Jansen, P. A. 2013. Clarifying assumptions behind the estimation of animal density from camera trap rates. *The Journal of Wildlife Management* 77(5):876–876.

Rubenstein, D. I.; Stewart, C. V.; Berger-Wolf, T. Y.; Parham, J.; Crall, J.; Machogu, C.; Kahumbu, P.; and Maingi, N. 2015. The Great Zebra and Giraffe Count: The Power and Rewards of Citizen Science. Technical Report, Kenya Wildlife Service, Nairobi, Kenya.

Seber, G. A. 1965. A note on the multiple-recapture census. *Biometrika* 52(1/2):249–259.

Seber, G. 2002. *The Estimation of Animal Abundance and Related Parameters*. Caldwell, NJ: Blackburn, 2 edition.

Sermanet, P.; Eigen, D.; Zhang, X.; Mathieu, M.; Fergus, R.; and LeCun, Y. 2013. OverFeat: Integrated recognition, localization and detection using convolutional networks. *CoRR* abs/1312.6229:1–16.

Simpson, R.; Page, K. R.; and De Roure, D. 2014. Zooniverse: Observing the World's Largest Citizen Science Platform. In *In Proceedings of the 23rd International Conference on World Wide Web*, WWW '14 Companion, 1049–1054. New York, NY: ACM.

Swanson, A.; Kosmala, M.; Lintott, C.; Simpson, R.; Smith, A.; and Packer, C. 2015. Snapshot Serengeti, high-frequency annotated camera trap images of 40 mammalian species in an African savanna. *Scientific Data* 2(150026):1–14.

van Strien, A. J.; van Swaay, C. A.; and Termaat, T. 2013. Opportunistic citizen science data of animal species produce reliable estimates of distribution trends if analysed with occupancy models. *Journal of Applied Ecology* 50(6):1450–1458.

Wiggins, A.; Newman, G.; Stevenson, R. D.; and Crowston, K. 2011. Mechanisms for Data Quality and Validation in Citizen Science. In *2011 IEEE Seventh International Conference on e-Science Workshops*, 14–19.

Xue, Y.; Davies, I.; Fink, D.; Wood, C.; and Gomes, C. P. 2016. Avicaching: A two stage game for bias reduction in citizen science. In *In Proceedings of the 2016 International Conference on Autonomous Agents & Multiagent Systems*, 776–785.