

Optimising the automated recognition of individual animals to support population monitoring

Tijmen de Lorm¹, Catharine Horswill², Daniella Rabaiotti², Robert Ewers¹, Rosemary Groom², Jessica Watermeyer³, and Rosie Woodroffe²

¹Imperial College London - Silwood Park Campus

²Zoological Society of London Institute of Zoology

³African Wildlife Conservation Fund

March 10, 2023

Abstract

Reliable estimates of population size and demographic rates are central to assessing the status of threatened species. However, obtaining individual-based demographic rates requires long-term data, which is often costly and difficult to collect. Photographic data offer an inexpensive, non-invasive method for individual-based monitoring of species with unique markings, and could therefore increase available demographic data for many species. However, selecting suitable images and identifying individuals from photographic catalogues is prohibitively time-consuming. Automated identification software can significantly speed up this process. Nevertheless, automated methods for selecting suitable images are lacking, as are studies comparing the performance of the most prominent identification software packages. In this study, we develop a framework that automatically selects images suitable for individual identification, and compare the performance of three commonly used identification software packages; Hotspotter, I3S-Pattern, and WildID. As a case study, we consider the African wild dog *Lycaon pictus*, a species whose conservation is limited by a lack of cost-effective large-scale monitoring. To evaluate intra-specific variation in the performance of software packages, we compare identification accuracy between two populations (in Kenya and Zimbabwe) that have markedly different coat colouration patterns. The process of selecting suitable images was automated using Convolutional Neural Nets that crop individuals from images, filter out unsuitable images, separate left and right flanks, and remove image backgrounds. Hotspotter had the highest image-matching accuracy for both populations. However, the accuracy was significantly lower for the Kenyan population (62%), compared to the Zimbabwean population (88%). Our automated image pre-processing has immediate application for expanding monitoring based on image-matching. However, the difference in accuracy between populations highlights that population-specific detection rates are likely and may influence certainty in derived statistics. For species such as the African wild dog, where monitoring is both challenging and expensive, automated individual recognition could greatly expand and expedite conservation efforts.

Optimising the automated recognition of individual animals to support population monitoring

T.A. de Lorm¹, C. Horswill^{2,3,4}, D. Rabaiotti^{2,3}, R.M. Ewers¹, R. J. Groom^{2,5}, J. Watermeyer⁵, R. Woodroffe^{2,3}

Correspondence author: T.A. de Lorm, tad20@ic.ac.uk

Affiliations:

1: Department of Life Sciences, Imperial College London, Silwood Park, UK

2: Institute of Zoology, Zoological Society of London, London, UK

3: Division of Biosciences, Department of Genetics, Evolution and Environment, Centre for Biodiversity and Environment Research, University College London, London, UK

4: Department of Zoology, University of Cambridge, Downing St, Cambridge CB2 3EJ, UK

5: African Wildlife Conservation Fund, Chishakwe Ranch, Zimbabwe

Running head: Automating African wild dog recognition

Abstract

Reliable estimates of population size and demographic rates are central to assessing the status of threatened species. However, obtaining individual-based demographic rates requires long-term data, which is often costly and difficult to collect. Photographic data offer an inexpensive, non-invasive method for individual-based monitoring of species with unique markings, and could therefore increase available demographic data for many species. However, selecting suitable images and identifying individuals from photographic catalogues is prohibitively time-consuming. Automated identification software can significantly speed up this process. Nevertheless, automated methods for selecting suitable images are lacking, as are studies comparing the performance of the most prominent identification software packages.

In this study, we develop a framework that automatically selects images suitable for individual identification, and compare the performance of three commonly used identification software packages; Hotspotter, I³S-Pattern, and WildID. As a case study, we consider the African wild dog *Lycaon pictus*, a species whose conservation is limited by a lack of cost-effective large-scale monitoring. To evaluate intra-specific variation in the performance of software packages, we compare identification accuracy between two populations (in Kenya and Zimbabwe) that have markedly different coat colouration patterns.

The process of selecting suitable images was automated using Convolutional Neural Nets that crop individuals from images, filter out unsuitable images, separate left and right flanks, and remove image backgrounds. Hotspotter had the highest image-matching accuracy for both populations. However, the accuracy was significantly lower for the Kenyan population (62%), compared to the Zimbabwean population (88%).

Our automated image pre-processing has immediate application for expanding monitoring based on image-matching. However, the difference in accuracy between populations highlights that population-specific detection rates are likely and may influence certainty in derived statistics. For species such as the African wild dog, where monitoring is both challenging and expensive, automated individual recognition could greatly expand and expedite conservation efforts.

Key words: automated individual recognition, Hotspotter, I³S-pattern, *Lycaon pictus*, photographic identification, Wild-ID

1. Introduction

Reliable estimates of population size and demographic rates are central to monitoring the status of threatened species. However, obtaining individual-based demographic parameters requires long-term data, gathered through intensive monitoring that is often costly and difficult to conduct (Horswill, Humphreys & Robinson, 2018; Caughlan, 2001). Identification of individuals from photographic records could provide an inexpensive alternative, and open up the possibility of using camera traps and citizen scientists to expand the spatial coverage of monitoring (Seber, 1965; Marnewick et al., 2014; Wearn & Glover-Kapfer, 2019). This method can be used for species where individuals can be identified from individual markings, including many threatened species (Durant et al., 2014, Pierce & Norman, 2016).

Photographic records have already been used to estimate demographic parameters in several endangered species. For example, long term photographic data have been used to obtain survival and abundance esti-

mates of tigers *Panthera tigris* and cheetahs *Acinonyx jubatus* (Karanth & Nichols, 2011; Kelly et al., 1998), and tourist images have been used to estimate population sizes of whale sharks *Rhincodon typus* (Davies et al., 2013). In addition, photographs can provide data on individual movement, ranging behaviour, and social structure (Randić et al., 2012; Armstrong et al., 2019). Many species are photographed frequently as part of monitoring programs, and by members of the public, including tourists. Such image catalogues therefore represent a large, and potentially under-used, data resource that inform conservation action.

Nevertheless, visually identifying all individuals in large image databases is time-consuming. To partly automate this process, several software packages are available to match images based on an individual's unique body markings (e.g. APHIS and WildID, Óscar et al., 2015; Bolger et al., 2012). These image-matching software packages assist the user by ranking potential image matches using a similarity score. The algorithms underpinning the software packages find these potential matches by comparing images on either a pixel-by-pixel or feature basis. Pixel-based algorithms, such as APHIS, have been successfully applied to numerous species, including horseshoe whip snakes *Hemorrhhois hippocrepsis* and Balearic lizards *Podarcis lilfordi* (Óscar et al., 2015; Rotger, 2019). However, they are susceptible to differences in camera angle, scale, and cropping (Matthé et al., 2017), and are therefore unsuitable for animals that cannot be caught and photographed using a standardised methodology. By contrast, feature-based software packages, such as Wild-ID (Bolger, 2012), I³S-Pattern (Reijns, 2014) and Hotspotter (Crall et al., 2013), match images based on unique features including spots, stripes, blotches, or other marks. The algorithms that feature-based packages use vary, but all have a higher tolerance to differences in camera angle, scale, and lighting conditions than pixel-based algorithms. However, researchers are still required to select images that are suitable for identification, in that the distinctive marks must face the camera and must be clearly visible. Furthermore, when these suitable images are selected, the user has to crop the region of interest from the image, which is laborious, potentially preventing the application to large image catalogues (Miguel et al., 2019).

The feature-based packages have been tested on a range of taxa (Table 1), and the reported proportion of true matches that the software detects, i.e., accuracy rate, varies markedly, ranging between 36% and 100%. This variation can be attributed to differences in species markings, image quality, size of database, how many potential matches were inspected per image, and the image-matching software used (Nipko, Holcombe & Kelly, 2020; Matthé et al., 2017; Crall et al., 2013). Studies directly comparing the accuracy of different feature-based packages are considerably more limited, even though the most accurate software differs between species. For example, studies on jaguars *Panthera onca*, ocelots *Leopardus pardalis*, and Saimaa ringed seals *Phoca hispida saimensis* found that Hotspotter outperformed Wild-ID (Nipko et al. 2020, Chehrsimin et al. 2018), whilst studies on amphibian species found that Wild-ID outperformed I³S-Pattern (Nipko et al., 2020, Matthé et al., 2017), and Hotspotter (Morrison et al., 2016). The only study that directly compared all three software packages found that Hotspotter was superior to I³S-pattern and Wild-ID for identifying individual green toads *Bufo viridis* (Burgstaller, Gollmann & Landler, 2021). To date, studies comparing image-matching accuracy across all three software packages for a mammal species are lacking.

African wild dogs *Lycaon pictus* (hereafter 'wild dogs') have unique coat markings, which vary between individuals (Figure 1, Maddock & Mills, 1994). Wild dogs are classified as globally endangered, and a lack of cost-effective large-scale monitoring has been highlighted as a major limitation in developing effective conservation strategies (Woodroffe & Sillero-Zubiri, 2020). Consequently, there is a pressing need to devise new approaches for monitoring wild dogs. Demographic processes of African wild dogs are typically studied by monitoring a subset of individuals fitted with tracking collars (Woodroffe et al., 2019; Rabaiotti et al. 2021; Jenkins et al., 2015). Such collar-based monitoring is labour-intensive and expensive, so upscaling is difficult. However, many wild dog packs have already been systematically photographed as part of monitoring programs, and many are also regularly photographed by tourists. Therefore, photographic identification of wild dogs potentially offers a non-invasive, cheaper approach for monitoring, and could reduce uncertainties in demographic rates and expand the spatial representation of monitoring (Maddock & Mills, 1994; Marnewick et al., 2014).

Wild dog coat patterns contain tan, white, and black patches that can vary considerably between populations. For example, wild dogs in eastern African populations tend to have coats consisting of predominantly black fur, whilst those in southern African populations have relatively more white and tan blotches (Figure 1, McIntosh, Woodroffe & Rabaiotti, 2016). In amphibian species with contrasting colour patterns of similar blotches, Wild-ID, Hotspotter, and I³S-Pattern have been shown to effectively match images of the same individual, reaching accuracy rates of up to 97% (Matthé et al., 2017; Burgstaller, Gollmann & Landler, 2021). Therefore, it is likely that feature-based image matching algorithms will effectively identify individual wild dogs from image catalogues. However, variation in the degree of contrast in the colour patterns amongst populations could affect the image-matching accuracy, and the best performing software package could therefore also vary between populations.

In this study, we develop a method to automatically isolate and crop images from catalogues that are suitable for automated image-matching. We then use these images to compare the efficacy of three feature-based software packages with different underlying image-matching algorithms (I³S-Pattern, Hotspotter, and Wild-ID; Reijns, 2014; Bolger, 2012; Crall *et al.*, 2013). Finally, we compare whether there is a difference in the accuracy of each software package between two populations with differing coat patterns.

2. Methods

2.1 Image datasets

To examine whether the performance of feature-based image matching software packages varies for different populations of African wild dogs, we considered image catalogues from two wild dog populations, one that spans Laikipia, Samburu, and Isiolo Counties in Kenya (37°2'E, 0deg6'N) and another from the Save Valley Conservancy in Zimbabwe (32deg4'E, 20deg3'S). The Kenyan dataset contained images taken between 2004 and 2017 (n = 9139), and the Zimbabwean dataset contained images taken between 2010 and 2013 (n = 2066). In Kenya, these images were taken with ten different cameras (Olympus(c) C765UZ, Canon(c) PowerShot A720IS, EOS Digital Rebel XT, 10D, 40D, and 60D, Fujifilm(c) FinePix S5500, Kodak(c) Easyshare Z1015IS, Nikon(c) D70s, and Nikon(c) Coolpix90). In Zimbabwe, they were taken with five different cameras (Canon(c) EOS 450D, 20D, and Digital Rebel XT, and Panasonic(c) DMC-FZ20, and Zoran(c) Coach). Both datasets were collected as part of long-term monitoring programmes, and contained images of both single wild dogs and groups of wild dogs, ranging in their posture from lying down to walking.

2.2 Pre-processing steps

To automate the selection of suitable images for image-matching we developed a five-step image pre-processing method (Figure 2).

2.2.1 Detecting and cropping individuals from images

The aim of the first step in the image pre-processing method was to automatically detect and crop wild dog individuals from the images. To do this, we used the Microsoft AI for Earth MegaDetector (hereafter 'MegaDetector', Beery, Morris & Yang, 2019) that automatically detects and crops animals in images. We assessed the efficacy of this method by visually recording the presence of wild dogs in a subset of 1060 images from the Kenyan dataset and 246 images from the Zimbabwean dataset, and comparing the results to the cropped images (hereafter 'crops') produced by the MegaDetector for the same subset of images. In this way, we obtained the MegaDetector's number of true positives (wild dogs that were successfully detected), false positives (detections which did not contain a wild dog), and false negatives (wild dogs which were found by visual inspection, but not by the MegaDetector). All images contained wild dogs, so there were no true negatives in the dataset.

2.2.2 Aspect-ratio filtering

The aim of the second step in the image pre-processing method was to filter out images that were unsuitable for identification due to the individual's body rotation in the image. We considered crops suitable for image-

matching if approximately [?]80% of the individual’s flank was visible, and the angle between the image axis and animal’s flank was less than approximately 30deg, i.e., the flank was facing the camera. Crops where the angle between the image axis and the animal’s flank was more than 30deg were expected to be narrower than crops suitable for image-matching and therefore demonstrate a relatively low aspect-ratio. By contrast, crops where the flank was concealed because the individual was lying down, or obscured by vegetation, were expected to be considerably wider and demonstrate a relatively high aspect-ratio. These criteria were visually assessed for the crops that the MegaDetector produced in the previous step. We then calculated the range of aspect-ratios for suitable crops, i.e., where an unobscured flank was facing the camera, using the “jpeg” package (Urbanek, 2021) in Program R (version 4.0.4, R Core Team, 2020). Images with an aspect-ratio outside of this range were removed from the dataset.

2.2.3 Selecting standing individuals

Not all sitting or lying individuals could be filtered out solely using image aspect-ratios. Therefore, the aim of the third step in the image pre-processing method was to filter out the remaining crops that were unsuitable for identification because the individual’s body position, i.e. sitting or lying, obscured the full coat pattern. To do this, we trained a Convolutional Neural Net (CNN) to classify crops as either a standing wild dog or a sitting wild dog. To obtain data to train this image classifier, we used the full image catalogues from both sites ($n = 11205$). The crops produced by steps 1 and 2 of the pre-processing ($n = 21745$) were then manually classified as either containing a standing wild dog ($n = 13500$) or sitting wild dog ($n = 6512$). We removed all crops depicting anything other than wild dogs (e.g., birds, rocks or logs), or wild dogs where it could not be confirmed whether they were standing or sitting, because only a small part of the animal was visible ($n = 1733$). We trained a CNN using the remaining 20012 pre-processed crops, to classify these as containing either a standing wild dog or not. The CNN was made using Tensorflow (Abadi et al., 2016) in Python (Version 3.6.10). The model was trained with 16012 crops, validated with 2000 crops, and tested with 2000 crops.

CNN’s consist of convolutional layers (Albawi, Mohammed & Al-Zawi, 2017): filter layers which digitally ‘slide’ over the image and aim to recognise specific features. The convolutional layers pass a map of specific features to the next layer, a Max Pooling layer. The Max Pooling layer reduces the resolution of this feature map, thus reducing the importance of the position of features within this map. This step can help prevent the model from becoming too fine-tuned to the training data, which causes over-fitting and reduces the generalisability of the classifier. After this, a dropout layer is applied, which randomly removes 50% of connections made between layers. This benefits the model by teaching it to recognise robust features, again preventing over-fitting. The data are then passed on to a flattening layer, which turns the data into a 1-dimensional string, which is passed onto the final two layers. Firstly, the string goes through a layer which connects all the data from the previous layer and produces prediction scores from the inputs. Secondly, another layer turns these scores into a single prediction: standing, or not standing (for a more detailed description of CNNs, see O’Shea & Nash, 2015 and Albawi, Mohammed & Al-Zawi, 2017).

The number of convolutional layers and the size of the filters that they comprise was optimised using KerasTuner (O’Malley et al., 2019). KerasTuner runs CNNs with a range of values, and automatically selects the model with the highest validation accuracy, i.e., the proportion of correct classifications on the validation database. KerasTuner ran CNNs with between one and three convolutional layers, with 16, 32 and 64 filters per layer, and with a kernel size (the number of pixels in the filters) of 3x3 pixels. This was done for 20 different random combinations for the number of convolutional layers and number of filters per layer. Test runs showed that the maximum accuracy was reached before the 70th epoch, and therefore each combination was run for 70 epochs, meaning that the training data were passed through the CNN 70 times. The learning rate of the model, that is, the speed at which the model improved itself, was also optimised with KerasTuner, testing a rate of 10^{-3} , 10^{-4} , and 10^{-5} , with the optimal number of convolutional layers. The model with the highest test-accuracy was selected as the final model.

2.2.4 Separating left and right flanks

The aim of the fourth step in the image pre-processing method was to separate crops depicting left- and right-flanks of a wild dog, because image-matching software packages can only match images for one side of the animal. To do this, we made another CNN to automate the separation of left- and right-flanks. To obtain training data for this CNN, we visually classified all crops of standing dogs used for the CNN in step three ($n = 12357$) whose side was facing the camera, as showing the right ($n = 6140$) or left flank ($n = 6217$). We optimised this CNN's parameters as described in step three of the image pre-processing method, using KerasTuner to find the optimal number of convolutional layers and learning rate. Each CNN ran for 100 epochs, because test runs showed that this model took longer than the previous model to reach its maximum accuracy. The first layer of this CNN was an average pooling layer, a layer which reduced the resolution of the input images by a factor of four, which prevents overfitting. This layer was added to this CNN, because preliminary runs showed this CNN was more prone to overfitting than the CNN developed in step three of the image pre-processing method. We used 9857 crops as training data, 1246 as validation data, and 1246 as testing data. All other layers were equal to the previous CNN. For the full model conditions, see Table S1 in Supporting Information.

2.2.5 Image background removal

Lastly, we removed the image backgrounds of suitable images using the “rembg” package in Python (Gatis, 2020). We removed image backgrounds to remove the risk of the background confounding image-matching results, while eliminating the need to manually select an individual's flank.

2.3 Image-matching software packages

We compared the performance of three feature-based image-matching software packages that differ in the underlying algorithms used to match individuals: I³S-Pattern (Reijns, 2014), WildID (Bolger et al., 2012) and Hotspotter (Crall et al., 2013). All three assist the user by listing potential matches for each image, ranked by a similarity score. The user then confirms which of these potential matches are true matches.

2.3.1 I³S-Pattern

I³S-Pattern uses the Speeded-up Robust Features (SURF) algorithm (Reijns, 2014; Bay et al., 2008) that selects key points and compares each image-pair in a dataset based on the size and position of these key points. The software requires the user to select three reference points per image, as well as the outline of the animal. As reference points, we used the base of the tail, the withers (i.e., the ridge between the shoulder blades), and the base of the neck (Figure S1).

2.3.2 Wild-ID

Wild-ID uses the Scale Invariant Feature Transform (SIFT) algorithm (Bolger et al., 2012; Lowe, 2004). It requires the user to input crops of the region of interest: the part of the animal which bears unique marks. The SIFT algorithm detects salient features regardless of their scale and viewpoint. For each image pair in a database, these features are compared, both in how the features look, and how the different features are positioned relative to each other. Based on these two characteristics, a goodness-of-fit score is computed per image pair.

2.3.3 Hotspotter

Hotspotter also uses the SIFT algorithm to conduct pairwise comparisons (Lowe, 2004; Crall et al., 2013). Users enter either entire pictures of individuals and select the rectangular region of interest, or image crops containing the region of interest. Hotspotter supplements the pairwise comparisons used by Wild-ID with a one-vs-many approach that uses a Local Naive Bayes Nearest Neighbour method (McCann & Lowe, 2012) to take all of the images in the database into account when computing similarity scores.

2.4 Performance of the image-matching software

To test which image-matching software most accurately matched crops of the same individual, we created two separate datasets for the Kenyan and Zimbabwean populations. To select suitable crops, we used

the four-step image pre-processing method described above. We also visually inspected discarded crops to avoid missing suitable crops. We then manually identified individuals from the dataset of right flank crops, to provide a standard against which automated identifications could be compared, and randomly selected two crops per individual. To prevent similar lighting conditions and posture from creating a bias towards matching images of the same individual, we ensured selected images were taken on different days. The two generated datasets consisted of 104 individuals from the Kenyan population, and 48 individuals from the Zimbabwean population. To increase the dataset for the Zimbabwean population, we also included left-flank crops for 41 individuals and horizontally mirrored the crops to enable comparison with the right-flank crops. This increased the total number of unique flanks from the Zimbabwean population to 89. The coat pattern of wild dogs differs between right and left flanks, and we have no reason to expect that including left-flank crops would bias our results.

We analysed the Kenyan dataset with each of the three image-matching software packages: Hotspotter, WildID, and I³S-Pattern. We then analysed the Zimbabwean dataset with Hotspotter and WildID. I³S-Pattern was not tested with the Zimbabwean dataset because tests with the Kenyan dataset identified it to be considerably less accurate than the other software packages and considerably more time-consuming to input images and assign reference points.

We also examined whether image background removal increased the accuracy of WildID and Hotspotter. I³S-Pattern requires users to manually select the outline of the animal in the program and therefore was not included in this analysis, because it does not take the background into account in its default use. We compared the image-matching results obtained using images from which we manually cropped just the individuals' flanks, with those based on crops of complete individuals from which the background was automatically removed (see Figure S2). For three of the 178 images from the Zimbabwe site, the algorithm did not crop out the wild dog, instead cropping out vegetation in the foreground. For these images, a manually cropped flank of the wild dog was used.

To compare the image-matching performance of each software package, we examined the 10 crops identified as most similar to the sample individual. We used the first 10 ranked images, as the best performing software's accuracy started levelling off around this rank, indicating that inspecting the first 10 image matches could maximise recognition rates, while minimising the time spent visually inspecting and confirming potential matches. We used a mixed effects logistic regression to test for differences in the efficacy of the software packages. Here, the response variable was a binary variable describing whether or not an individual was successfully matched in the first 10 ranked images, and software package was the explanatory variable. Individual identity was included as a random effect to avoid pseudoreplication. Post-hoc pairwise comparisons were carried out using Tukey contrasts. This analysis was performed separately for the Zimbabwean and Kenyan datasets. Models were run using the "lme4" (v. 1.1-27.1, Bates et al., 2015) package in Program R (R Core Team, 2020, version 4.0.4).

Previous studies have shown that the image-matching performance of different software packages is affected by database size (Matthe et al., 2017). Therefore, to compare software performance on wild dogs from Kenyan and Zimbabwean populations, we randomly selected a subset of the Kenyan individuals to equal the number of identified individuals in the Zimbabwean dataset ($n = 89$). We then used the best performing software package identified in the previous step of the analysis to rerun the image-matching analysis for both datasets. Differences in software performance between the two populations were then assessed using a mixed effects logistic regression with a binomial link function. The response variable in the model was whether or not a match was detected in the first 10 ranked images, and study site (Kenya or Zimbabwe) was the explanatory variable. To correct for possible differences in image quality, two proxies for image quality were included in the model. Firstly, we included image size (total number of pixels of the crop) as a continuous predictor. Secondly, all images were visually scored on a scale of 1 to 3, based on how well their distinct marks could be recognised. This approach followed Nipko, Holcombe & Kelly (2020), where score 1 was given to images that were out of focus, of a moving animal, or badly lit, score 2 was given to images of intermediate quality, and score 3 was given to images where all features were clearly visible (for examples,

see Figure S3). Score was included as a fixed effect, and individual identity was included as a random effect. Furthermore, a Wilcoxon Rank Sum test was performed to test for differences between the quality score of crops from Kenya and Zimbabwe. The model was fit using the “lme4” package (v 1.1-27.1, Bates et al., 2015) in R (version 4.0.4, R Core Team, 2020).

3. Results

3.1 Automatic detection and aspect-ratio filtering

The Microsoft AI for Earth MegaDetector, which is designed to automatically detect and crop animals from images, produced 2652 crops from the test dataset of 1306 images, meaning that on average, approximately two detections were made per image. Of these, 2523 crops contained a wild dog (true positive rate = 0.951), 129 were false detections, such as rocks or vegetation (false positive rate = 0.049), while 531 wild dogs were not successfully detected (false negative rate = 0.174). However, only five of these false negatives were images suitable for image-matching. By contrast, the flank of the wild dog was not visible in the other 526 false negatives. In total, 722 crops were suitable for identification using image matching software, of which five were not detected by the automated processing (false negative rate = 0.007). For the 2652 crops that were produced by the MegaDetector, all crops considered suitable for identification on visual inspection had an aspect-ratio between 0.65 and 2.25. Applying the MegaDetector to the entire image dataset ($n = 11205$), as opposed to the test dataset, resulted in 21745 crops. Of these, 5788 (21%) fell outside the suitable range of aspect ratios and were therefore removed from the dataset.

3.2 Using Convolutional Neural Nets to filter out unsuitable images

The optimal conditions for the Convolutional Neural Net (CNN) trained to recognise wild dogs standing up were two convolutional layers, with kernel sizes of 32 and 64, respectively, and a learning rate of 10^{-5} . This model achieved a training accuracy of 100%, a validation accuracy of 91% (95% C.I. 90 – 92), and a testing accuracy of 90% (95% C.I. 88 – 91, Table 2). For the CNN designed to separate images of the left and right flanks, the optimal conditions were three convolutional layers, one with a kernel size of 64 and two with kernel sizes of 32, with a learning rate of 10^{-4} . Its training, validation and testing accuracy were 100%, 96% (95% C.I. 95 – 97), and 95% (95% C.I. 94 – 96), respectively.

3.3 Performance of the image-matching software packages

For both the Kenyan and Zimbabwean datasets, Hotspotter achieved the highest image-matching accuracy (Figure 3). For the Kenyan dataset, using Hotspotter with crops of the full individual from which the background was removed, was most effective. This method detected 62% of the matches in the 10 highest ranked crops (Figure 3B). This was significantly higher than using manually cropped flanks of the individual in both WildID ($z = 5.0$, $p < 0.01$) and Hotspotter ($z = 2.8$, $p = 0.046$), as well as crops of the full individual in WildID ($z = 4.7$, $p < 0.01$) and I³S-Pattern ($z = 5.0$, $p < 0.01$). For the Zimbabwean dataset, Hotspotter detected 88% of matches within the first 10 ranked images when the background was removed from a crops of the full individual (Figure 3A). The matching performance was significantly lower when crops of just the flank were used ($z = 2.7$, $p = 0.03$). Hotspotter with background removal performed significantly better than WildID with background removal, ($z = 4.7$, $p < 0.01$), as well as WildID with crops of the flanks ($z = 5.1$, $p < 0.01$).

The probability of accurate image-matching occurring within the first 10 ranked images was significantly higher for wild dogs from Zimbabwe than wild dogs from Kenya (OR = 9.64, 95% CI 3.65 – 15.63, Figure 4). The proportion of matched individuals identified in this analysis was not significantly associated with image size ($X^2_1 = 0.16$, $p = 0.69$) or image quality (OR_{Quality Score 2 / Quality Score 1} = 0.89, 95% C.I. -2.26 – 4.04, OR_{Quality Score 3 / Quality Score 1} = 1.82, 95% C.I. -2.20 – 5.83). In addition, the image quality score did not differ between the populations ($W = 15008$, $p = 0.33$).

4. Discussion

This study presents a novel framework for automating the individual recognition of species with distinct marks. The framework includes an automated pre-processing method for identifying images suitable for image-matching, and then using image-matching software for individual recognition. The automated pre-processing method consists of five steps that (1) crop all images containing animals from a large database, (2) filter out a portion of the unsuitable images based on image aspect ratio, (3) use convolutional neural nets to select images of standing individuals (accuracy of 90%), (4) separate images into left and right flanks (accuracy of 95%), and (5) remove image backgrounds. As a case study, we applied the described methods to an image catalogue of African wild dogs and found that Hotspotter (Crall *et al.*, 2013) was the most efficient software package for matching images. Image-matching performance was also significantly improved by using the full image of an individual from which the background was removed, as opposed to just the cropped flank. Finally, we found that image-matching performance differed between populations of wild dogs with different coat coloration patterns. This work showed that image-matching software could become a powerful method for monitoring populations of African wild dogs. However, caution is needed as detection rates are likely to vary between – and even within – populations. This could affect the certainty of derived population-specific demographic parameters, such that careful consideration is needed to account for individual heterogeneity in detection when large variation in coat colouration occurs within a population.

The automated pre-processing method presented in this study could eliminate the need to manually select suitable images for image-matching and crop individuals from original photographs. This method thus enables processing of large image catalogues where selection using visual inspection would be extremely time consuming. We found that the method does discard a small number of suitable images, and therefore in situations where it is important to include all suitable images, the pre-processing method outlined here could also be used as a pre-sorting approach. The user could then visually review images that were classified as not suitable, to prevent usable images from being discarded.

The described method of pre-processing is particularly useful for wild dogs, since an individuals' posture varies substantially between images. Images taken by tourists provide an opportunity to bolster and spatially extend image catalogues. However, these images are also likely to contain many images unsuitable for identification, as they are not taken for the purpose of identification. Accordingly, filtering unsuitable images from these datasets using an automated approach could be especially timesaving. The described pre-processing method is therefore highly suitable to species targeted by wildlife watching excursions, that have distinctive marks and where individual posture influences image suitability, for example cheetahs, leopards *Panthera pardus*, and tigers.

Hotspotter outperformed I³S-Pattern and Wild-ID at matching images of individual wild dogs. This finding agrees with studies on green toads that compared Hotspotter and I³S-pattern (Burgstaller, Gollmann & Landler 2021), as well as studies comparing Hotspotter and WildID (Nipko, Holcombe & Kelly, 2020; Burgstaller, Gollmann & Landler, 2021; Chehrsimin *et al.*, 2018). Nevertheless, this result is not ubiquitous. Wild-ID was superior to Hotspotter at matching images for a blotched amphibian species, the Wyoming toad *Anaxyrus baxteri* (Morrison *et al.*, 2016). This indicates that the identification performance of different software packages is dependent on species, even when two species' patterns show similarities. Consequently, we recommend that all three software packages are tested on new species before deciding on which one to use.

Using crops of full individuals from which the background was removed significantly increased the image-matching accuracy of Hotspotter, compared to using crops of just individuals' flanks. This method also speeds up image pre-processing by eliminating the need to manually crop the region of interest. The improved accuracy is likely caused by two factors. Firstly, removing the background prevents images being matched based on similar backgrounds, as the flanks are not perfect rectangles, meaning that crops of the flank also contain some background (see Figure S2). Secondly, using complete individuals allows images to be matched based on unique features on the legs, in addition to the flanks. This result is in line with studies on Saimaa ringed seals *Pusa hispida* and Thornicroft's giraffes *Giraffa camelopardalis thornicrofti*, which found evidence

that using a full individual from which the background is removed, could result in a higher accuracy (Chermin *et al.*, 2018; Halloran, Murdoch & Becker 2015). However, neither of these previous studies statistically tested whether background removal increased identification accuracy. Our study therefore provides the first statistical evidence that background removal can increase the performance of image-matching software. This also indicates that the common usage of Hotspotter, in which a rectangular region of interest is manually cropped (e.g. Dunbar *et al.*, 2021; Nipko, Holcombe & Kelly, 2020), could be improved by removing the image background.

Hotspotter was significantly better at matching images from Zimbabwean wild dogs, compared to Kenyan individuals. The higher image-matching accuracy found for the Zimbabwean population is likely to reflect the regional difference in wild dog coat colouration patterns. The Kenyan population has darker, more uniform coats, consisting of large black patches, often with few white or tan areas (McIntosh, Woodroffe & Rabaiotti, 2016, Daniels, Woodroffe & Rabaiotti, 2022). By contrast, the proportion of tan fur is ~ 1.5 times higher, and the proportion of white fur is almost 7 times higher for the Zimbabwean population (Figure 1, Daniels, Woodroffe & Rabaiotti, 2022). Therefore, the higher contrast within the patterns of the Zimbabwean wild dogs could make it easier for the software to match images of these individuals. The identified relationship between image-matching performance and software package remained unaltered when image quality and image size were included in analyses, and there was no significant difference between the image quality scores between the Zimbabwean and Kenyan populations. The image quality score approach was modelled after Nipko, Holcombe & Kelly (2020), who found that it significantly affected the probability of matching ocelot and jaguar individuals. As a result, we are confident that the differences in coat colouration patterns between wild dogs from Zimbabwe and Kenya reflect variation in identification performance between populations.

Inter-population variation in image-matching performance indicates that detection probabilities derived from using this approach will not be directly comparable between populations. Since the probability of finding an accurate image-match depends on individual coat pattern, this finding highlights that individual heterogeneity in detection may also occur if large variation in coat colouration occurs within a population. Capture-mark-recapture techniques assume individuals experience equal detection probability across a population (White and Burnham, 2009). Therefore, individual coat pattern may also need accounting for when deriving survival estimates using such analysis. This also applies to other species whose coat pattern varies regionally, such as Asian golden cats *Catopuma temminckii* and ocelots (Allen *et al.*, 2011; Khan, Ali & Mohammed, 2017). Furthermore, the coat patterns of other wild dog populations can differ considerably from the two populations included in this study (McIntosh, Woodroffe & Rabaiotti, 2016, Daniels, Woodroffe & Rabaiotti, 2022). Consequently, we advocate that estimating a population-specific image-matching accuracy score becomes an essential pre-requisite step for applying these techniques in different locations.

Automatically pre-processing wild dog image datasets and using image-matching software facilitates the use of archived and citizen science image catalogues where visually identifying all individuals would be extremely time-consuming. Although the best performing image-matching software did not detect all matches, it could be used to identify a large proportion of the individuals in a dataset. Afterwards, individuals that were not matched to any other images could be visually identified, to prevent missing actual matches. Using image-matching software in this way still saves time by rapidly identifying a large portion of the matches, without compromising on accuracy. Furthermore, it is plausible that the likelihood of correctly detecting matching images increases if more than two images per individual are included, for example if multiple viewpoints per individual are present in a dataset, the probability of matching these is expected to increase (Crall *et al.*, 2013). Our accuracy values therefore represent a conservative estimate of Hotspotter's true accuracy.

Our study indicates that image-matching could provide a valuable new approach for monitoring wild dogs. A combination of citizen science and image-matching has already been successfully employed to monitor other species, such as Blanding's turtles *Emydoidea blandingii* and whale sharks (Araujo *et al.*, 2017; Cross *et al.*, 2021). Similarly, previous studies have used tourist images to estimate the population size of wild dogs in Kruger National Park, South Africa (Marnewick *et al.*, 2014). Combining citizen science, image-

matching software, and capture-recapture methods therefore has the potential to improve understanding of wild dog demography. However, more research is needed to investigate whether photographic data could improve our understanding of wild dog demography beyond population size, by estimating parameters such as pack structure, dispersal rates, and death and birth rates. This can be achieved by applying image-matching software to existing image datasets, to assess whether they generate enough data to estimate key demographic parameters, or whether more intensive monitoring - for example using long term camera trap surveys - would be necessary.

In conclusion, we have developed a new automated method for pre-processing image datasets, by automatically cropping animals from images, removing images in which the individuals' posture hinders identification, separating left and right flanks, and removing the image background. This framework will enable large image datasets to be analysed rapidly, thereby expanding monitoring efforts and expediting conservation action. Furthermore, we have shown how well different image-matching software packages perform on African wild dogs. Hotspotter outperformed the other software packages, while its performance differed between two populations which exhibit intra-specific variation in their coat patterns. Our pre-processing method, in combination with Hotspotter, has immediate application in research and monitoring efforts for wild dogs and other species. Data obtained in this way could provide cost-effective large-scale monitoring for endangered species, therefore supporting the implementation of effective conservation.

References

- Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., et al. (2016) TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems. *arXiv:1603.04467 [cs]* . <http://arxiv.org/abs/1603.04467>.
- Agarap, A.F. (2019) Deep Learning using Rectified Linear Units (ReLU). *arXiv:1803.08375 [cs, stat]* . <http://arxiv.org/abs/1803.08375>.
- Albawi, S., Mohammed, T.A. & Al-Zawi, S. (2017) Understanding of a convolutional neural network. In: *2017 International Conference on Engineering and Technology (ICET)* . August 2017 Antalya, IEEE. pp. 1–6. doi:10.1109/ICEngTechnol.2017.8308186.
- Allen, W.L., Cuthill, I.C., Scott-Samuel, N.E. & Baddeley, R. (2011) Why the leopard got its spots: relating pattern development to ecology in felids. *Proceedings of the Royal Society B: Biological Sciences* . 278 (1710), 1373–1380. doi:10.1098/rspb.2010.1734.
- Araujo, G., Snow, S., So, C.L., Labaja, J., Murray, R., Colucci, A. & Ponzo, A. (2017) Population structure, residency patterns and movements of whale sharks in Southern Leyte, Philippines: results from dedicated image-ID and citizen science: Whale sharks of Southern Leyte. *Aquatic Conservation: Marine and Freshwater Ecosystems* . 27 (1), 237–252. doi:10.1002/aqc.2636.
- Armstrong, A.O., Armstrong, A.J., Bennett, M.B., Richardson, A.J., Townsend, K.A. & Dudgeon, C.L. (2019) Photographic identification and citizen science combine to reveal long distance movements of individual reef manta rays *Mobula alfredi* along Australia's east coast. *Marine Biodiversity Records* . 12 (1), 14. doi:10.1186/s41200-019-0173-6.
- Bates, D., Machler, M., Bolker, B. & Walker, S. (2015) Fitting Linear Mixed-Effects Models Using **lme4** . *Journal of Statistical Software* . 67 (1). doi:10.18637/jss.v067.i01.
- Bay, H., Ess, A., Tuytelaars, T. & Van Gool, L. (2008) Speeded-Up Robust Features (SURF). *Computer Vision and Image Understanding* . 110 (3), 346–359. doi:10.1016/j.cviu.2007.09.014.
- Beery, S., Morris, D. & Yang, S. (2019) Efficient Pipeline for Camera Trap Image Review. *arXiv:1907.06772 [cs]* . <http://arxiv.org/abs/1907.06772>.
- Bolger, D., Morrison, T., Vance, B., Lee, D. & Farid, H. (2012) A computer-assisted system for photographic mark-recapture analysis. *Methods in Ecology and Evolution* . 3, 813–822. doi:10.1111/j.2041-

210X.2012.00212.x.

Bolger, D.T. (2012) *Wild.ID* . <https://faculty-directory.dartmouth.edu//douglas-thomas-bolger>.

Burgstaller, S., Gollmann, G. & Landler, L. (2021) The Green Toad example: a comparison of pattern recognition software. *North-Western Journal of Zoology* . 17 ((1)), 96–99.

Caughlan, L. (2001) Cost considerations for long-term ecological monitoring. *Ecological Indicators* . 1 (2), 123–134. doi:10.1016/S1470-160X(01)00015-2.

Chehrsimin, T., Eerola, T., Koivuniemi, M., Auttila, M., Levanen, R., Niemi, M., Kunnasranta, M. & Kalviainen, H. (2018) Automatic individual identification of Saimaa ringed seals. *IET Computer Vision* . 12 (2), 146–152. doi:10.1049/iet-cvi.2017.0082.

Crall, J.P., Stewart, C.V., Berger-Wolf, T.Y., Rubenstein, D.I. & Sundaresan, S.R. (2013) HotSpotter — Patterned species instance recognition. In: *2013 IEEE Workshop on Applications of Computer Vision (WACV)* . January 2013 pp. 230–237. doi:10.1109/WACV.2013.6475023.

Cross, M.D., Mayer, J., Breymaier, T., Chiotti, J.A. & Bekker, K. (2021) Estimating Population Size of a Threatened Turtle Using Community and Citizen Science. *Chelonian Conservation and Biology* . 20 (1). doi:10.2744/CCB-1384.1.

Davies, T.K., Stevens, G., Meekan, M.G., Struve, J. & Rowcliffe, J.M. (2013) Can citizen science monitor whale-shark aggregations? Investigating bias in mark-recapture modelling using identification photographic sourced from the public. *Wildlife Research* . 39 (8), 696–704. doi:10.1071/WR12092.

Daniels, L., Rabaiotti, D., Woodroffe, R. (2022): Correlates of Coat Colour in the African wild dog. University College London

Dunbar, S.G., Anger, E.C., Parham, J.R., Kingen, C., Wright, M.K., Hayes, C.T., Safi, S., Holmberg, J., Salinas, L. & Baumbach, D.S. (2021) HotSpotter: Using a computer-driven image-id application to identify sea turtles. *Journal of Experimental Marine Biology and Ecology* . 535, 151490. doi:10.1016/j.jembe.2020.151490.

Durant, S., Mitchell, N., Ipavec, A. & Groom, R. (2014) *Acinonyx jubatus*: IUCN Red List of Threatened Species 2015 . doi:10.2305/IUCN.UK.2015-4.RLTS.T219A50649567.en.

Gatis, D. (2020) *Rembg* . <https://github.com/danielgatis/rembg>.

Glorot, X., Bordes, A. & Bengio, Y. (2011) Deep Sparse Rectifier Neural Networks. *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics* . 15:315-323. <https://proceedings.mlr.press/v15/glorot11a.html>.

Halloran, K.M., Murdoch, J.D. & Becker, M.S. (2015) Applying computer-aided image-identification to messy datasets: a case study of Thornicroft’s giraffe (*Giraffa camelopardalis thornicrofti*). *African Journal of Ecology* . 53 (2), 147–155. doi:<https://doi.org/10.1111/aje.12145>.

Horswill, C., Humphreys, E.M. & Robinson, R.A. (2018) When is enough ... enough? Effective sampling protocols for estimating the survival rates of seabirds with mark-recapture techniques. *Bird Study* . 65 (3), 290–298. doi:10.1080/00063657.2018.1516191.

Jenkins, E., Silva-Opps, M., Opps, S.B. & Perrin, M.R. (2015) Home Range and Habitat Selection of a Reintroduced African Wild Dog (*Lycaon pictus*) Pack in a Small South African Game Reserve. *African Journal of Wildlife Research* . 45 (2), 233–246. doi:10.3957/056.045.0233.

Karanth, K.U. & Nichols, J.D. (2011) Estimation of Demographic Parameters in a Tiger Population from Long-term Camera Trap Data. In: A.F. O’Connell, J.D. Nichols, & K.U. Karanth (eds.). *Camera Traps in Animal Ecology* . Tokyo, Springer Japan. pp. 145–161. doi:10.1007/978-4-431-99495-4_9.

- Kelly, M.J., Laurenson, M.K., FitzGibbon, C.D., Collins, D.A., Durant, S.M., Frame, G.W., Bertram, B.C. & Caro, T.M. (1998) Demography of the Serengeti cheetah (*Acinonyx jubatus*) population: the first 25 years. *Journal of Zoology* . 244 (4), 473–488. doi:10.1111/j.1469-7998.1998.tb00053.x.
- Khan, K., Ali, S. & Mohammed, R.S. (2017) Colours and Spots: Do they Tell the Story of an Ocelot's Origin? *Living World, Journal of the Trinidad and Tobago Field Naturalists' Club* . 38–42.
- Lowe, D.G. (2004) Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision* . 60 (2), 91–110. doi:10.1023/B:VISI.0000029664.99615.94.
- Maddock, A.H. & Mills, M.G.L. (1994) Population characteristics of African wild dogs *Lycaon pictus* in the Eastern transvaal lowveld, South Africa, as revealed through photographic records. *Biological Conservation* . 67 (1), 57–62. doi:10.1016/0006-3207(94)90009-4.
- Marnewick, K., Ferreira, S.M., Grange, S., Watermeyer, J., Maputla, N. & Davies-Mostert, H.T. (2014) Evaluating the Status of and African Wild Dogs *Lycaon pictus* and Cheetahs *Acinonyx jubatus* through Tourist-based photographic Surveys in the Kruger National Park R.M. Brigham (ed.). *PLoS ONE* . 9 (1), e86265. doi:10.1371/journal.pone.0086265.
- Matthe, M., Sannolo, M., Winiarski, K., Spitzen - van der Sluijs, A., Goedbloed, D., Steinfartz, S. & Stachow, U. (2017) Comparison of image-matching algorithms commonly used for photographic capture-recapture studies. *Ecology and Evolution* . 7 (15), 5861–5872. doi:10.1002/ece3.3140.
- McCann, S. & Lowe, D.G. (2012) Local Naive Bayes Nearest Neighbor for image classification. In: *2012 IEEE Conference on Computer Vision and Pattern Recognition* . June 2012 Providence, RI, IEEE. pp. 3650–3656. doi:10.1109/CVPR.2012.6248111.
- McIntosh, A., Woodroffe, R. & Rabaiotti, D. (2016) *Coat colour with temperature in the African wild dog, Lycaon pictus* . University College London.
- Miguel, A.C., Bayrakcismith, R., Ferre, E., Bales-Heisterkamp, C., Beard, J., Dioso, M., Grob, D., Hartley, R., Nguyen, T. & Weller, N. (2019) Identifying individual snow leopards from camera trap images. In: K. Mao & X. Jiang (eds.). *Tenth International Conference on Signal Processing Systems* . 17 April 2019 Singapore, Singapore, SPIE. p. 36. doi:10.1117/12.2521856.
- Morrison, T.A., Keinath, D., Estes-Zumpf, W., Crall, J.P. & Stewart, C.V. (2016) Individual Identification of the Endangered Wyoming Toad *Anaxyrus baxteri* and Implications for Monitoring Species Recovery. *Journal of Herpetology* . 50 (1), 44–49. doi:10.1670/14-155.
- Nipko, R.B., Holcombe, B.E. & Kelly, M.J. (2020) Identifying Individual Jaguars and Ocelots via Pattern-Recognition Software: Comparing HotSpotter and Wild-ID. *Wildlife Society Bulletin* . 44 (2), 424–433. doi:https://doi.org/10.1002/wsb.1086.
- O'Malley, T., Long, J., Chollet, F., Haifeng, J. & Invernizzi, L. (2019) *Keras Tuner* . <https://github.com/keras-team/keras-tuner>.
- Oscar, M., Pep-Luis, M., Sergio, M., Jose-Manuel, I., Andreu, R., Antonio, R. & Giacomo, T. (2015) APHIS: A new software for image-matching in ecological studies. *Ecological Informatics* . 7.
- O'Shea, K. & Nash, R. (2015) An Introduction to Convolutional Neural Networks. *arXiv:1511.08458 [cs]* . <http://arxiv.org/abs/1511.08458>.
- Park, H., Lim, A., Choi, T.-Y., Baek, S.-Y., Song, E.-G. & Park, Y.C. (2019) Where to spot: individual identification of leopard cats (*Prionailurus bengalensis euptilurus*) in South Korea. *Journal of Ecology and Environment* . 43 (1), 39. doi:10.1186/s41610-019-0138-z.
- Pierce, S.J. & Norman, B. (2016) *Rhincodon typus: The IUCN Red List of Threatened Species 2016* . doi:10.2305/IUCN.UK.2016-1.RLTS.T19488A2365291.en.

R Core Team (2020) *R: A language and environment for statistical computing*. Vienna, Austria., R Foundation for Statistical Computing. <https://www.R-project.org/>.

Randić, S., Connor, R.C., Sherwin, W.B. & Krützen, M. (2012) A novel mammalian social structure in Indo-Pacific bottlenose dolphins (*Tursiops* sp.): complex male alliances in an open social network. *Proceedings of the Royal Society B: Biological Sciences* . 279 (1740), 3083–3090. doi:10.1098/rspb.2012.0264.

Reijns (2014) *I3S Pattern* . <https://reijns.com/i3s/i3s-pattern/>.

Rotger, A. (2019) Image-identification of horseshoe whip snakes (*Hemorrhois hippocrepis*, Linnaeus, 1758) by a semi-automatic procedure applied to wildlife management. *Herpetological Journal* . 29 (4), 304–307. doi:10.33256/hj29.4.304307.

Sannolo, M., Gatti, F., Mangiacotti, M., Scali, S. & Sacchi, R. (2016) Image-identification in amphibian studies: a test of I3S Pattern. *Acta Herpetologica* . 63-68 Pages. doi:10.13128/ACTA_HERPETOL-17198.

Seber, G.A.F. (1965) A Note on the Multiple-Recapture Census. *Biometrika* . 52 (1/2), 249–259. doi:10.2307/2333827.

Suriyamongkol, T. & Mali, I. (2018) Feasibility of Using Computer-Assisted Software for Recognizing Individual Rio Grande Cooter (*Pseudemys gorzugi*). *Copeia* . 106 (4), 646–651. doi:10.1643/CH-18-101.

Tieleman, T. & Hinton, E. (2012) *Lecture 6.5 - rmsprop, COURSERA: Neural networks for machine learning* .

Urbanek, S. (2021) *Package ‘jpeg’* . <http://www.rforge.net/jpeg/>.

Wearn, O.R. & Glover-Kapfer, P. (2019) Snap happy: camera traps are an effective sampling tool when compared with alternative methods. *Royal Society Open Science* . 6 (3), 181748. doi:10.1098/rsos.181748.

White, G. C., & Burnham, K. P. (2009). Program MARK: survival estimation from populations of marked animals. 3657. doi: 10.1080/00063659909477239

Woodroffe, R., Rabaiotti, D., Ngatia, D.K., Smallwood, T.R.C., Strebel, S. & O’Neill, H.M.K. (2019) Dispersal behaviour of African wild dogs in Kenya. *African Journal of Ecology* . 58 (1), 46–57. doi:10.1111/aje.12689.

Woodroffe, R. & Sillero-Zubiri, C. (2020) *Lycaon pictus*: The IUCN Red List of Threatened Species 2020 (amended version of the 2012 assessment). Available from: doi:10.2305/IUCN.UK.2020-1.RLTS.T12436A166502262.en [Accessed: 16 November 2020].

Table 1. A selection of studies that tested the identification accuracy achieved by different image-matching software packages.

Species	Species	Software package	Software package	Accuracy (%)	Accuracy (%)	Number of inspected ranks per image	Number of inspected ranks per image	Reference	Re
Jaguar (<i>Panthera onca</i>)	Jaguar (<i>Panthera onca</i>)	Hotspotter	Hotspotter	77	77	1	1	Nipko, Holcombe & Kelly, 2020	Ni Ho co & Ke 20
Ocelot (<i>Leopardus pardalis</i>)	Ocelot (<i>Leopardus pardalis</i>)			76	76				

Species	Species	Software package	Software package	Accuracy (%)	Accuracy (%)	Number of inspected ranks per image	Number of inspected ranks per image	Reference	Re
Jaguar	Jaguar	Wild-ID	Wild-ID	68	68				
Ocelot	Ocelot			63	63				
Thornicroft's giraffe (<i>Giraffa camelopardalis thornicrofti</i>)	Thornicroft's giraffe (<i>Giraffa camelopardalis thornicrofti</i>)	Wild-ID	Wild-ID	71.6	71.6	20	20	Halloran <i>et al.</i> , 2014	Ha <i>et al.</i> , 2014
Thornicroft's giraffe	Thornicroft's giraffe	Wild-ID	Wild-ID	100	100	1	1	Bolger <i>et al.</i> , 2012	Bo <i>et al.</i> , 2012
Leopard cat (<i>Prionailurus bengalensis</i>)	Leopard cat (<i>Prionailurus bengalensis</i>)	Hotspotter	Hotspotter	100	100	1	1	Park <i>et al.</i> , 2019	Pa <i>et al.</i> , 2019
Jaguar	Jaguar	Hotspotter	Hotspotter	100	100	1	1	Crall <i>et al.</i> , 2013	Cr <i>et al.</i> , 2013
Giraffe (<i>Giraffa giraffa</i>)	Giraffe (<i>Giraffa giraffa</i>)			100	100				
<i>Ambystoma opachum</i>	<i>Ambystoma opachum</i>	I ³ S-Pattern	I ³ S-Pattern	30.8–48.4	30.8–48.4	10	10	Matthé <i>et al.</i> , 2017	Ma <i>et al.</i> , 2017
Yellow-bellied toad (<i>Bombina variegata</i>)	Yellow-bellied toad (<i>Bombina variegata</i>)	Wild-ID I ³ S-Pattern	Wild-ID I ³ S-Pattern	65.9–82.3 88.6–92.0	65.9–82.3 88.6–92.0				
Rio grande cooter (<i>Pseudemys gorzugi</i>)	Rio grande cooter (<i>Pseudemys gorzugi</i>)	Wild-ID I ³ S-Pattern	Wild-ID I ³ S-Pattern	96.4–97.3 41.94	96.4–97.3 41.94	20	20	Suriyamongkol & Mali, 2018	Su & Ma, 2018
		Wild-ID	Wild-ID	66.13	66.13	20	20		

Species	Species	Software package	Software package	Accuracy (%)	Accuracy (%)	Number of inspected ranks per image	Number of inspected ranks per image	Reference	Re
Green toad (<i>Bufotes viridis</i>)	Green toad (<i>Bufotes viridis</i>)	I ³ S- Pattern Wild-ID Hotspotter	I ³ S- Pattern Wild-ID Hotspotter	~100 ~60 ~60	~100 ~60 ~60	1	1	Burgstaller, Gollmann & Landler, 2021	Bu Go & La 20
Wyoming toad (<i>Anaxyrus baxteri</i>)	Wyoming toad (<i>Anaxyrus baxteri</i>)	Wild-ID Hotspotter	Wild-ID Hotspotter	53 36	53 36	20	20	Morrison <i>et al.</i> , 2016	Mo <i>et</i> 20
Italian crested newt (<i>Triturus carnifex</i>)	I ³ S- Pattern	I ³ S- Pattern	100	100	5	5	Sannolo <i>et al.</i> , 2016	Sannolo <i>et al.</i> , 2016	

Table 2 : The accuracy and 95% confidence intervals of the best performing Convolution Neural Nets aiming to classify images of African wild dog into (1) those depicting an individual standing up and not standing up and (2) those depicting left or right flanks.

Model	Training Accuracy (%)	Validation Accuracy (%)	Testing Accuracy (%)
Standing/not standing classifier	100	91 (90 - 92)	90 (88 - 91)
Left/right flank classifier	100	96 (94 - 97)	95 (94 - 96)



Hosted file

image3.emf available at <https://authorea.com/users/594521/articles/628946-optimising-the-automated-recognition-of-individual-animals-to-support-population-monitoring>

Figure 2: A flowchart describing the image pre-processing steps applied to the combined Zimbabwean and Kenyan image catalogue, illustrated using two example images. Images outlined in red are filtered out of the dataset.

Hosted file

image4.emf available at <https://authorea.com/users/594521/articles/628946-optimising-the-automated-recognition-of-individual-animals-to-support-population-monitoring>

Figure 3 : The proportion of true matches in the dataset that were detected within the top 20 ranked images using different software packages for A) 89 Zimbabwean and B) 104 Kenyan African wild dogs, that were visually matched as a reference.

Hosted file

image6.emf available at <https://authorea.com/users/594521/articles/628946-optimising-the-automated-recognition-of-individual-animals-to-support-population-monitoring>

Figure 4 : The proportion of image matches detected within the 10 highest ranked images by Hotspotter for the Zimbabwean and Kenyan populations ($n = 89$ individuals for both populations). Error bars denote 95% confidence intervals.

Supplementary Information for

Optimising the automated recognition of individual animals to support population monitoring

T.A. de Lorm, C. Horswill, D. Rabaiotti, R.M. Ewers, R. J. Groom, J. Watermeyer, R. Woodroffe

Table S1 The structure of a Convolutional Neural Net that classifies images of African Wild Dogs into “standing” or “not standing” (S1.A), and a Convolutional Neural Net that classifies images of African Wild Dogs into “left flank” or “right flank” (S1.B). The layers are given in the order at which they occur in the model. The models were optimised using RMSprop, an algorithm which guides how the model improves itself (Tieleman & Hinton, 2012). The activation column refers to which activation function was used in each layer, which determines how the nodes within layers convert its input to an output-value. ReLu was used as activator function, as this has been found to improve multi-layer networks (Glorot, Bordes & Bengio, 2011). The last layer is activated with a Sigmoid function, which turns the input into a single, binary prediction (“standing” or “not standing”, “left” or “right”).

S1.A

Layer	Kernel size	Activation
Convolutional	32	Rectified Linear Activator (ReLU) (Agarap, 2019)
Convolutional	64	ReLU
Max Pooling 2D	NA	NA
Dropout	NA	NA
Flatten	NA	NA
Dense	128	ReLU
Dense	1	Sigmoid

S1.B

Layer	Kernel size	Activation
Average Pooling 2D	NA	NA
Convolutional	32	ReLU
Convolutional	64	ReLU
Convolutional	64	ReLU
Max Pooling 2D	NA	NA
Dropout	NA	NA
Flatten	NA	NA
Dense	128	ReLU
Dense	1	Sigmoid

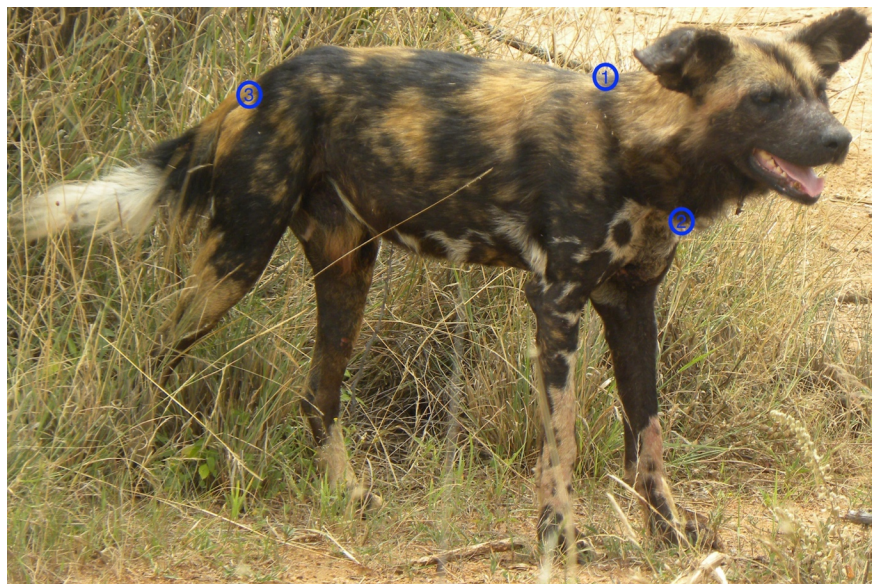


Figure S1. An image with the reference points used for I³S-Pattern: (1) the withers, (2) the base of the neck, and (3) the base of the tail.



Figure S2. The two ways in which WildID and Hotspotter were tested: with crops of the full individual from which the background was removed (left), and manually cropped flanks of the individual (right).

Hosted file

image10.emf available at <https://authorea.com/users/594521/articles/628946-optimising-the-automated-recognition-of-individual-animals-to-support-population-monitoring>

Figure S3. Three images representative of the image quality scores that were assigned to images included in a Mixed Effect Logistic Regression. The regression tested whether there was a difference in the accuracy of Hotspotter between the Kenyan and Zimbabwean population. The image quality score was included to correct for any potential influence of image quality.

References

Glorot, X., Bordes, A. & Bengio, Y. (2011) Deep Sparse Rectifier Neural Networks. In: Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics. 14 June 2011 JMLR Workshop and Conference Proceedings. pp. 315–323. <https://proceedings.mlr.press/v15/glorot11a.html>.

Tieleman, T. & Hinton, E. (2012) *Lecture 6.5 - rmsprop, COURSERA: Neural networks for machine learning* .