Using neural networks to identify individual animals from photographs *

Emmanuel Kabuga[1,2], Ian Durbach[1,2], Bubacarr Bah[2,3], and Allan Clark[1]

[1] University of Cape Town (https://www.uct.ac.za/)
[2] AIMS South Africa (https://aims.ac.za/)
[3] Stellenbosch University (http://www.sun.ac.za/english)

Wildlife conservation relies on sound knowledge of population information. This information is crucial for addressing questions related to community, ecosystem function, population dynamics, and behavioural ecology. It is obtained via studies that recognise individuals. A commonly used technique to achieve individual species recognition is the use of invasive methods that apply a tag to the animal's body. These methods have been applied to both marine and terrestrial species to assess both theoretical and applied questions [5, 6]. However, invasive procedures are expensive to implement and potentially reduce the animal's natural behaviour and performance, disturb its activities and relationship to others. In addition to being impracticable with large populations, invasive methods cause ethical and welfare conflicts due to temporary or permanent application of tags. Alternatively, many species have body marks such as spots and fins that are individual-specific and hence can be utilised for individual recognition. These methods are cost-effective and not harmful to the animal's life. These approaches have evolved as a reliable alternative to invasive methods and have been applied to a range of animals including mammals, amphibians, reptiles and fishes [8, 9]. This paper aims at developing a machine learning algorithm which exploits individual-specific marks to automate the individual identification task and compares the model results with some of the existing computer-aided software used by the ecology community.

The developed model is tested on two case studies, a humpback whale (HBW) dataset and a western leopard toad (WLT) dataset. The HBW dataset consists of 25 631 images from 14 668 individuals. They are originally collected by various institutions across the globe and uploaded to the Happywhale platform [4]. HBWs can be identified by their fins and special marks. The WLT dataset consists of 1 770 images collected by citizen scientists in South Africa. They were either uploaded to iSpot [7], a citizen science project which collects images or sent to the WLT project, a conservation project staffed by volunteers. WLTs can be identified by their unique spots. One part of this dataset consists of 164 labelled individuals comprising 430 images and an unlabelled proportion comprising 1 340 images.

The model developed in this paper consists of two main components, an object detection model and a matching classifier model. In some images, the

* Acknowledgement to AIMS South Africa for the Research Masters scholarship, CHPC for computational resources, Dr. John Measey and Mr. Alex Rebelo for collecting and providing the toad photographs.

animal only occupies a small region of the image. This makes the individual-specific marks – spots for WLTs, and tail fins for HBWs to not be clearly visible while they are the key feature of this study. As a result, the goal of the object detection model is to – detect the region of the image containing the animal, localise it using a bounding box, and extract the animal, which is then taken to the next level of photo-matching. The object detection model is a custom convolutional neural network (CNN) originally inspired from VGG16 [1] which takes an image as input and outputs the coordinates of the region containing the animal within the image. The matching classifier model is a special kind of CNN called a Siamese network. The Siamese network is a custom ResNet [2] model which uses a pair of CNNs that share weights to summarise the images, followed by some dense layers which combine the summaries into measures of similarity which can be used to predict a match. This model takes a pair of two regions containing animals extracted by the detection model and outputs their matching probability. A threshold probability set by the user is used to decide if a pair of two animal images originates from the same individual or not. The individual IDs are extracted from the obtained matches. One of the computer-aided photo-matching algorithms used by the ecology community is WildID [3]. It utilises the scaled invariant feature transform (SIFT) to extract distinctive features from the images. It compares SIFT features of a new image with ones of the existing images in the catalogue and ranks the top 20 potential matches. The true match can appear anywhere between 1 and 20 or not. For a fair comparison with the developed Siamese network, we checked if WildID ranked the true match in the first position or not (top-1 accuracy).

The detection model achieved reliable results on both datasets for the task of localising the region of the image containing the animal on both datasets. The model achieved the intersection over union (IoU) of 0.90 and the coefficient of determination R^2 of 0.91 for HBWs and the IoU of 0.86 and R^2 of 0.85 for WLTs. The Siamese network model results are good for both HBWs and WLTs. The model correctly identified if a pair of images is from the same individual or not respectively in 95% of cases for HBWs and 87% of cases for WLTs. The main difference in the performance is due to the different amount of data used to train the model. In this study, the semi-supervised approach on WLT unlabelled dataset has been partially successful. The model was able to identify 47 new matches from 26 individuals comprising 63 images. These identified matches seem to be relatively few in numbers. Without an exhaustive check of the data, it is not clear whether this is due to the failure of the semi-supervised approach, or because there are not many matches in the data. After adding the newly identified and labelled individuals to the WLT labelled dataset, the model slightly improved its performance and correctly identified 89% of WLT pairs. WildID achieved good results on WLTs compared to HBWs. It ranked the true match in the first position in 64% of cases for WLTs and 36% of the cases for HBWs.

The Siamese network model achieved good results on the individual identification task for species dotted with individual-specific marks. Its performance was very competitive compared with WildID.

References

1. K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556, 2014
2. K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770-778, 2016.
3. D. T. Bolger, T. A. Morrison, B. Vance, D. Lee, and H. Farid. A computer-assisted system for photo- graphic markrecapture analysis. Methods in Ecology and Evolution, 3(5):813822, 2012.
4. https://happywhale.com/home
5. J. N. Auckland, D. M. Debinski, and W. R. Clark. Survival, movement, and resource use of the butterfly parnassius clodius. Ecological Entomology, 29(2):139149, 2004.
6. D. J. Booth. Synergistic effects of conspecifics and food on growth and energy allocation of a damselfish. Ecology, 85(10):28812887, 2004.
7. https://www.ispotnature.org/node/137767
8. L. Gamble, S. Ravela, and K. McGarigal. Multi-scale features for identifying individuals in large biological databases: an application of pattern recognition technology to the marbled salamander ambystoma opacum. Journal of Applied Ecology, 45(1):170180, 2008.
9. C. W. Speed, M. G. Meekan, and C. J. Bradshaw. Spot the matchwildlife photo-identification using information theory. Frontiers in zoology, 4(1):2, 2007.