Assessing the performance of open-source, semi-automated pattern recognition software for harbour seal (*P. v. vitulina*) photo ID

Izzy Langley[1]*, Emily Hague[1,2] and Mònica Arso Civil[1]

[1]Sea Mammal Research Unit, Scottish Oceans Institute, University of St Andrews, St Andrews, Scotland, UK.

[2]Institute of Life and Earth Sciences, Heriot-Watt University, Edinburgh, Scotland, UK.

*il32@st-andrews.ac.uk (ORCID ID: 0000-0002-8957-1373)

Acknowledgements

Data collection was funded by the Scottish Government (grant number MMSS/002/15). The authors would like to thank the expert knowledge and assistance of collaborators at each of the study sites around Scotland, without whom data collection would not have been possible. The authors would also like to thank the publishers of all three pattern recognition software programmes for making them freely available; particularly Lex Hiby from Conservation Research Ltd. who has provided training and advice over the years. We would also like to thank the two anonymous reviewers whose comments helped to improve this manuscript.

This work is dedicated to Andy Law – a brilliant naturalist, photographer, colleague, and friend.

Abstract

Photographic identification (photo ID) is a well-established, non-invasive, and relatively cost-effective technique to collect longitudinal data from species that can be individually recognised based on natural markings. This method has been improved by computer-assisted pattern recognition software which speed up the processing of large numbers of images. Freely available algorithms exist for a wide range of species, but the choice of software can have significant effects on the accuracy of individual capture histories and derived demographic parameter estimates. We tested the performance of three open source, semi-automated pattern recognition software algorithms for harbour seal (*Phoca vitulina vitulina*) photo ID: ExtractCompare, I^{3}S Pattern and Wild-ID. Performance was measured as the ability of the software to successfully score matching images higher than non-matching images using the cumulative density function (CDF). The CDF for the top ranked potential match was highest for Wild-ID (CDF_1 = 0.34-0.58), followed by ExtractCompare (CDF_1 = 0.24-0.36) and I^{3}S Pattern (CDF_1 = 0.02-0.3). This trend emerged regardless of how many potential matches were inspected. The highest performing aspects in ExtractCompare were left heads, whereas in I^{3}S Pattern and Wild-ID these were front heads. Within each aspect, images collected using a camera and lens performed higher than images taken by a camera and scope. Data processing within ExtractCompare took >4x longer than Wild-ID, and >3x longer than I^{3}S Pattern. We found that overall, Wild-ID outperformed both ExtractCompare and I^{3}S Pattern under tested scenarios, and we therefore recommend its assistance in harbour seal photo ID.

Keywords

Pattern recognition; photo ID; software comparison; harbour seal; *Phoca vitulina vitulina*; Wild-ID

Introduction

Recognising individual animals is an important tool in the monitoring of wild populations (e.g. Wells and Scott 1990; Rotella et al. 2012; Letcher et al. 2015). For many species, individuals are artificially marked using a wide range of techniques, including bird ringing (e.g. spotted owl *Strix occidentalis*; Zimmerman et al. 2007), freeze-branding (e.g. *Chiroptera* spp; Sherwin et al. 2002), colour-marking (e.g. *Satyrinae* spp; Morton 1982) and tagging (e.g. pink abalone *Haliotis corrugate*; Button and Rogers-Bennet 2011). However, for some species individuals can be distinguished from one another from natural markings such as patterning and/or scarring (e.g. Asian elephant *Elephas maximus*; Goswami et al. 2007; whale shark *Rhincodon typus*; Bradshaw et al. 2007; wild horse *Equus ferus*; Vernes et al. 2009). These species can be photographed and, if the image is of sufficient quality, individuals can be identified. Photographic identification (photo ID) is a widely used, non-invasive and relatively cost-effective method to study the distribution and life-history parameters of wild populations (e.g. Thompson et al. 2008; Mackey et al. 2008; Gore et al. 2016; Langley et al. 2020).

A number of phocid seal species have individually unique pelage patterns which remain stable through adulthood, enabling populations to be monitored long-term through photo ID (e.g. grey seal *Halichoerus grypus*; Hiby et al. 2007; Saimaa ringed seal *Pusa hispida saimensis*; Koivuniemi et al. 2016; harbour seal *Phoca vitulina*; Yochem et al. 1990). While there is slight variation in the pelage colour and spot density among harbour seal sub-species, it is not consistent enough to confidently identify to sub-species level (Kelly 1981; Cunningham 2009; McCormack 2015). The repeated identification of individuals within species has been successful for three of these harbour seal sub-species: *P. v. richardii* in the northeast Pacific (Yochem et al. 1990), *P. v. concolor* in the northwest Atlantic (McCormack 2015), and *P. v. vitulina* in the northeast Atlantic (Cunningham 2009).

The matching efficiency and error rates of photo ID studies have been improved by the introduction of computer-assisted pattern recognition software (Arzoumanian et al. 2005; Caiafa et al. 2005; Morrison et al. 2011). Computer algorithms assist in the photo ID of species that have particularly fine-detailed patterning, and/or when dealing with large databases (e.g. Andrzejaczek et al. 2016; Germanov et al. 2019; Langley et al. 2020). Freely available algorithms exist for a wide range of species, but the choice of algorithm can have significant effects on the derived demographic parameter estimates. Misidentification of matches can introduce false positives (i.e. two different individuals matched to the same ID) and/or false negatives (i.e. one individual given two IDs). For example, a high false-acceptance rate results in an under-estimation of population size, whereas a high false rejection rate inflates estimates of population size (Hammond et al. 1990). The false-acceptance rate can be reduced to effectively zero by visually confirming potential matches, whereas the false-rejection rate is subject to multiple variables and so should be calculated and reported per analysis (Hastings et al. 2001; Cunningham 2009).

Here we focus on three freely available pattern recognition software programmes: ExtractCompare, I^{3}S Pattern and Wild-ID. ExtractCompare was originally developed for grey seals (Hiby and Lovell 1990) but has since been extended to other species (e.g. Eurasian lynx *Lynx lynx*; Gimenez et al. 2019; Amur leopard *Panthera pardus orientalis*; Jiang et al. 2015; Vitkalova and Shevtsova 2016) and is currently the only pattern recognition software which has a harbour seal specific model. The software builds a three-dimensional surface model from reference points in a manually annotated image. Pattern cells are then extracted from multiple aspects of the body (i.e. multibiometric identification; Jain 2007) to compare the patterning on non-planar surfaces (Hiby and Lovell 1990). Pairs of images are ranked by similarity scores and matches are manually confirmed. The software presents all potential matches, but a similarity score threshold can be assigned to streamline the processing of large datasets.

The Interactive Individual Identification System (I^3S) has multiple versions designed to extract and compare natural markings from a range of different species. I^3S *Pattern* was designed for species with hard to annotate markings such as lionfish (*Pterois volitans*; Chaves et al. 2016) and turtles (Calmanovici et al. 2018). It employs a SURF (speeded-up robust features) detector and descriptor, which first detects point correspondences between images, then describes the area of interest and detects matches between these areas (Bay et al. 2008). This is robust to noise, and variation in image scale and orientation, whilst computing faster than pre-existing alternatives (such as the SIFT operator described below). Similar to ExtractCompare, images are manually annotated with morphological reference points and an extractable area, although these are specified by the user at data entry, along with the number of potential matches presented.

Wild-ID was specifically designed to assist in the processing of large datasets generated by monitoring populations using camera traps. The software employs a SIFT (scale-invariant feature transform) operator which extracts distinctive image features whilst accounting for image scale and rotation (Lowe 2004). The images are cropped prior to data entry as the software does not distinguish the pattern of the subject from the pattern in the background (i.e. the noise; Bolger et al. 2012). The software pattern comparison function is not species-specific which enables its usability across a wide range of taxa, from amphibians (Bendik et al. 2013; Mettouris et al. 2016; Pereira and Maneyro 2016) to mammals (Bolger et al. 2012; Halloran et al. 2015). The standard version of the software then presents the top 20 potential matches which require visual confirmation or rejection (Bolger et al. 2012).

The aim of this study was to test the performance of these three freely available pattern recognition software programmes for the individual recognition of northeast Atlantic harbour seals (*P. v. vitulina*). Photo ID data were collected as part of an ongoing project investigating the regional decline in harbour seal numbers around Scotland (Arso Civil et al. 2016). Here, software performance was measured as its ability to successfully score matching images higher than non-matching images (Matthé et al. 2017). We investigated the effect of the data collection methods and the aspect of the body from which the pattern cell was compared. Data processing time was also compared between the three software programmes.

Methods

Data collection

Photo ID data were collected from harbour seal haulout sites in Kintyre, the Isle of Skye and Orkney (Scotland), during the breeding seasons (June and July) of 2016, 2017 and 2018. In Kintyre and Orkney, data were collected during dedicated land-based surveys from cliff tops and beaches, 50-150m from harbour seal haulout sites, using a digiscope system comprising of a DSLR camera attached to a scope (Swarovski ATS 80 with x20-60 eyepiece and TLS-APO 30mm). On the Isle of Skye, data were collected from small tourist boats that circumnavigate skerries where harbour seals haul out, 5-10m away from the boats, using a DSLR camera with an 80-400mm zoom lens.

Photographs were graded for quality on a scale of 1 (poor) to 4 (excellent), following a protocol adapted from Cunningham (2009), based on the focus of the image, the angle of the seal to the photographer and the clarity of the pelage markings (i.e. lighting, wet/dry, moult). Only images assigned a quality $\geq$3 were used in this analysis. Matches between pairs of images were initially found manually and confirmed by a trained expert. A catalogue of individual harbour seals with uniquely identifiable IDs was built and used to generate databases to test the performance of each software.

Database construction

Multiple databases consisting of pairs of images from individual harbour seals were constructed based on how the data were collected (scope, lens) and which aspect of the body the pattern cell was extracted from (front head, left head, left neck, left flank; Fig. 1). We excluded images from the right-hand side of the body as the algorithms should perform as well with these as the left. We ensured that each image from a single individual were collected on different sampling days, which avoided the likelihood of the backgrounds matching (seals return to the water on each tide). Front head aspects were images of seals facing the camera lens and included both eyes; left head aspects included the full side of the head including the nose, eye and ear; left neck aspects included the area between the ear and the fore-flipper; and left flank aspects included the area between the fore-flipper and the pelvis. Flank aspects were not available from the Isle of Skye data as the photographer was often too close to the seal to capture the entire body with a lens. Databases included pairs of images from all available individuals for each data collection method and aspect; this ranged from 65 to 178 individuals.

Data processing

We tested the performance of pattern recognition algorithms in detecting the one matching image in a set of non-matching images. In order to standardise the methodology across software (each has slightly different processing methods), data were entered in two batches and only the images with the top 20 similarity scores were manually inspected. Manual inspection in our study was of the image names which included the individual ID, but in a real-world scenario this would be manual inspection of the pelage. Batch 1 was entered first, containing a single image of each individual in that database. Batch 2 contained a second different image of each individual and was then entered systematically and compared to batch 1 (i.e. the library). Each database (n = 7) was run through each of the three software, except for databases containing front heads (front head aspects cannot be processed in the current ExtractCompare harbour seal model); this resulted in 19 trials. The process was timed for each trial, from the first stage of data preparation through to the final stage of match confirmation.

i) ExtractCompare

For ExtractCompare, images were reduced in size (i.e. cropped) prior to entry into the Microsoft Access database as in the authors experience, this speeds up the processing time. This software uses multibiometric identification and the pattern can be extracted from up to five aspects of the body. However, this is subject to data availability, and chest and abdomen aspects were underrepresented in our data. For this analysis, we focussed on left heads, necks, and flanks. The left head aspect covers an area behind the eye which includes the ear (Fig. 1a); the left neck aspect is the area between the ear and the fore flipper (Fig. 1b); and the left flank aspect is the area between the fore flipper and the pelvis

(Fig. 1c). Images were annotated with the outline of the body and morphological reference points which are specific to each aspect in question, but include the base of the skull, chin, nose, eyes, ears, post-orbital vibrissae, flippers, and pelvis (Fig. 1).

ii) I^3S Pattern

Cropping of images was not required for I^3S pattern, and as far as possible, reference points and extractable areas were specified so as to be as comparable across software as possible. For front head aspects, the reference points identified were the right eye, the left eye, and the nose, with the general identification area being a polygon from the eyes to the top of the head (Fig. 1d). Left head aspects were identified by the nose, the left eye and the left ear, and the area extended from the corner of the mouth to the back of the skull (Fig. 1e). Left neck aspects were identified by the nose, the post-orbital vibrissae, and the fore flipper, with the area extending from the corner of the mouth to the fore flipper (Fig. 1f). Finally, left flank aspects were identified by the nose, the fore flipper, and the pelvis, with the identifiable area extending from the fore flipper to the pelvis (Fig. 1g).

iii) Wild-ID

Wild-ID differs from the other two software programmes in that the pattern is not extracted from an aspect of the subject but is compared across the entire image. Images were therefore cropped to include only the desired aspect of the subject with as little of the background noise as possible. To make the analysis comparable across the three software programmes, we cropped images to the same aspects as with ExtractCompare and I^3S Pattern: front head (Fig. 1h), left head (Fig. 1i), left neck (Fig. 1j) and left flank (Fig. 1k).

Performance analysis

The pattern recognition software programmes used in this analysis are described as semi-automated, as all require a final manual confirmation stage where the user has to accept or reject each potential match. This reduces the overall likelihood of false acceptance (Sacchi et al. 2016). For the purpose of this study, we focused on the recognition rate, defined as the ability of the algorithm to successfully score matching images higher than non-matching images (Matthé et al. 2017). The image filenames (which included the individual ID) of the top 20 ranked similarity scores were visually inspected for each trial to manually confirm or reject the potential match. The cumulative density function (CDF) was calculated for each rank by dividing the cumulative sum of matches found by the number of matches available, and the corresponding two-sided 95% confidence intervals (based on the binomial distribution) were estimated using the *binom.test* function in R (R Core Team 2019). For a software to perform well, the CDF should reach 1 within the fewest ranks possible; i.e. if the match is not ranked high enough, the user could miss this (depending on any assigned similarity score threshold) and the false-rejection rate would increase. More generally, the lower down the potential matches a true match is ranked, the more time is required for the user to find the match.

ExtractCompare, I^3S Pattern and Wild-ID differ in data processing methodology and so processing was timed for all trials. The different stages were made up of both manual and automated steps. To run an image through ExtractCompare, there are five distinct stages: cropping, data input, pattern extraction, batch comparison and visual confirmation. In I^3S Pattern, the stages of data input (pattern extraction, comparison, and confirmation) are combined into a single step (combining manual and automated stages), and in Wild-ID, there are four distinct stages: cropping, input/extraction, comparison and confirmation. Each stage from data pre-processing to visual confirmation was timed separately and divided by the number of images to give the time in minutes and seconds required to process a single image (data processing rate).

Results

Across each tested scenario, Wild-ID outperformed both ExtractCompare and I^3S Pattern for harbour seal pattern recognition (Table 1, Fig 2). This trend was most pronounced when comparing the pelage pattern from the left head (CDF = 0.49-0.66) and neck regions (CDF = 0.45-0.64), regardless of data collection method, and for front head aspects taken using a camera and lens (CDF = 0.58-0.71). Data collected using a camera and lens had a higher proportion of the highest quality images (lens=0.62, scope=0.27) and in general, the highest performance for each software came from using data collected with a camera and lens (Fig. 2).

In Wild-ID, front head aspects performed highest; when only visually inspecting the top ranked potential match, the CDF was 0.58, translating to a false-rejection rate (FRR; 1-CDF) of 0.42. When the top 20 ranked potential matches were visually inspected, the CDF reached the highest recorded in this study: 0.71 (with an associated FRR of 0.29). Conversely, ExtractCompare performed best with left head aspects (CDF_1=0.36 with a FRR of 0.64; CDF_{20}=0.55 with a FRR of 0.45). Indeed, by rank 10, the uncertainty around the CDF for ExtractCompare overlapped with that of Wild-ID. I^3S Pattern performed poorly in most scenarios except for in trials which used front head aspects. As with Wild-ID, the highest CDF_1 for I^3S Pattern was recorded from front head aspects taken using a camera and lens (CDF_1=0.30 with a FRR of 0.70). The performance of front head aspects taken using a camera and scope however was much more comparable to that of Wild-ID.

With all processing stages combined, Wild-ID had the highest data processing rate (i.e. the least amount of time per image processed; mean ± sd mm:ss, 00:22 ± 00:04), followed by I^3S Pattern (00:31 ± 00:04) and ExtractCompare (01:36 ± 00:08; Table 2). For ExtractCompare, the vast proportion of time was spent in the pattern extraction stage (01:01 ± 00:06; 64% of total time) where images were annotated, and the three-dimensional model was applied. The remaining time was spread across cropping (00:09 ± 00:04; 9%), input (00:08 ± 00:02; 8%), comparison (00:05 ± 00:01; 5%) and confirmation stages (00:12 ± 00:02; 13%). The data processing in I^3S Pattern was shorter than ExtractCompare and cropping was not required prior to data entry. For Wild-ID, images were cropped prior to entry which took the greatest proportion of time (00:16 ± 00:04; 73%). Data input and pattern extraction stages were combined into one (<00:01 ± <00:01; 4%) and were followed by short comparison (00:01 ± <00:01; 5%) and confirmation stages (00:04 ± 00:01; 18%).

Discussion

The highest performing pattern recognition software tested for harbour seal photo ID was Wild-ID, followed by ExtractCompare and then I^3S Pattern. The strength of this trend varied with the data collection method and the aspect of the body that the pattern was compared from. Importantly, Wild-ID also required the least amount of time to run a single image through the stages from pre-processing to match confirmation. The highest recorded CDF, and therefore the lowest FRR, was recorded in Wild-ID for front head aspects collected using a camera and lens (CDF_{20}=0.71; FRR=0.29). This error is within a range deemed acceptable for the estimation of population parameters (Hiby et al. 2013).

In the present study, photo ID data were either collected from a platform 50-150m away from the seal (using a digiscope) or from a boat within 10m of the seal (using a lens). The data collection method was therefore used as a proxy for distance to haulout, which has been shown to influence image quality (Bendik et al. 2013). In this study, within each aspect, data collected using a lens performed marginally better than data collected using a scope. Previous photo ID studies have found that image quality has influenced the performance of pattern recognition algorithms. In ExtractCompare for harbour seals, the false-rejection rate has been shown to decrease from 73% to 2% by increasing image quality alone (Hastings et al. 2008). Similar trends have been reported for I^3S Pattern (Steinmetz et al. 2018) and Wild-ID (Bendik et al. 2013). Halloran et al. (2015) investigated the effect of image quality further and found that the only variable which affected the ability of Wild-ID to detect matches between images of Thornicroft's giraffe (*Giraffa camelopardalis thornicrofti*) was background complexity. This effect could therefore be reduced by cropping the images or by digitally removing the background entirely (Bolger et al. 2012; Chehrsimin et al. 2018).

The patterned surface of a seal's pelage is non-planar and can appear very different depending on the animal's orientation and torsion (Hiby and Lovell 1990). This is most pronounced on regions such as the neck and flank, whereas the region around the head is less susceptible to this distortion. Additionally, repeatability in the manual placement of the pattern cell is easier in the head region due to the proximity of obvious morphological features (i.e. eyes, ears, nose). In this study, ExtractCompare performed best with left head aspects. In previous studies, ExtractCompare has been shown to perform well for harbour seals using the shoulder/neck regions (Cunningham 2009) and ventral aspects (Hastings et al. 2008). The neck aspect is a larger region than the head and so contains more of an individual's unique "fingerprint", but it is also possibly more difficult to standardise across images. Ventral aspects were underrepresented in our dataset given the haulout behaviour of seals at the sites in this study, although it would be interesting to explore whether the performance of ExtractCompare, along with I^3S Pattern and Wild-ID, could be improved for northeast Atlantic harbour seal photo ID if images of the ventral side of the animals could be collected.

Conversely, we found that both Wild-ID and I^3S Pattern performed best for harbour seal photo ID using front head aspects. Previous studies have found I^3S Pattern to perform highly in the photo ID of green turtles (*Chelonia mydas*; Den Hartog and Reijns 2014), Hawksbill turtles (*Eretmochelys imbricate*; Steinmetz et al. 2018) and *Tarentola* geckos (Rocha et al. 2013); the natural patterning of all are found on rigid body parts (e.g. carapace scutes). The fore-head region of a harbour seal is also relatively rigid, and so best satisfies the assumption within I^3S Pattern that animals have linearity (i.e. their body parts do not move in respect to one another; Den Hartog and Reijns 2014).

When choosing a pattern recognition software to assist in the analysis of photo ID data, the ability of the software to detect a match is important, but often the amount of time required to process data is also crucial. Pattern matching algorithms have dramatically reduced the number of images which need to be visually inspected to find a match (Hastings et al. 2001; Morrison et al. 2011). This is important for long-term population studies that rely on detecting matches between thousands of images which would not be feasible though manual matching alone. In this study, the time required to process a single image using Wild-ID was on average 22s, compared with 31s in I^{3}S Pattern and 1m36s in ExtractCompare; the processing time of images in ExtractCompare was >4x greater than in Wild-ID. However, it is important to note that data processing included both manual and automated stages, and time can be saved by running automated stages overnight or alongside other tasks.

We tested the ability of the software algorithms to not only detect a positive match but also to rank it higher than non-matching images. The time required to manually inspect each potential match can be substantial and so often thresholds are assigned, below which potential matches are rejected without inspection. In ExtractCompare, previous studies have assigned thresholds on similarity scores of 0.95 (Hiby et al. 2013) and 0.75 (Langley et al. 2020) for grey seal photo ID, and 0.45 for cheetah photo ID (Kelly 2001). In I^{3}S Pattern, it has been more common to assign a threshold on the number of potential matches that are visually inspected (e.g. 50; Rocha et al. 2013; Steinmetz et al. 2018). Previous studies which use Wild-ID have also assigned thresholds on the similarity scores generated, and for species that can be easily manipulated, cleaned and posed against white backgrounds (e.g. Amphibians; Bardier et al. 2017), similarity scores are consistently predictive of positive matches (Bendik et al. 2013). However, with other taxa there is evidence that the similarity scores in Wild-ID can be affected by allometric variation; i.e. when individuals are still growing (Bardier et al. 2017), and in these cases the time between photographs can reduce similarity scores (Bendik et al. 2013). In this study our data were limited to adult harbour seals, but it would be useful to test the performance of pattern recognition software in detecting matches between pups, juveniles and adults; as has been successful using ExtractCompare for grey seals (Paterson et al. 2013).

Setting thresholds for manual review can significantly increase the efficiency of data processing but comes with associated false-rejection rates (Hiby et al. 2013). These errors are not consistent across studies and/or sub-species, with previous harbour seal photo ID analyses using ExtractCompare reporting error rates of 6.2% (Hastings et al. 2001) and 21.4% (McCormack 2015). False-rejection rates for I^{3}S Pattern and Wild-ID are not available for harbour seals but are low for the species that the algorithms were initially designed for. For example, the false rejection rate (using only the top ranked potential match) for the photo ID of green turtles (*Chelonia mydas*) in I^{3}S Pattern was 14% (Den Hartog and Reijns 2014), and for Masai giraffe (*Giraffa camelopardalis tippelskirchi*) photo ID in Wild-ID was 0.7% (Bolger et al. 2012). False-rejection rates are therefore variable and can be influenced by the experience of software users (Bolger et al. 2012) and the number of images from the same individuals (Hiby et al. 2013); along with variables tested in this study.

In this study we compared the performance of three freely available software, but there are additional software algorithms available; these include, but are not limited to, ICEIS/Hotspotter (Crall et al. 2013), Discovery (Gailey and Karczmarski 2012) and StripeSpotter (Lahiri et al. 2011). Investigation into the performance of other software algorithms for harbour seal photo ID, and their comparison to Wild-ID, would be a valuable next step. Beyond that, as ecological research becomes increasingly data-heavy, methods such as photo ID lend themselves to automation. Existing photo ID databases are required to train algorithms to automatically locate a seal within an image (i.e. segmentation; Chehrsimin et al. 2018), extract pelage pattern, describe this pattern and then compare it to a library of known individuals. As it stands, artificial intelligence for pattern recognition requires manually annotated databases. There is also a strong argument for manual confirmation of detected matches, at least until the error rates are below an accepted threshold. However, at the very least, automating the data pre-processing and input stages will help to improve the efficiency of pattern recognition software further.

Compliance with ethical standards

This research was approved by the University of St Andrews Animal Welfare and Ethics Committee (AWEC) and data collection was funded by the Scottish Government (grant number MMSS/002/15). The authors declare that they have no conflicts of interest.

Author contributions

IL & MAC conceived and designed the analysis, IL, EH & MAC collected the data, IL performed the analysis and wrote the paper, and EH & MAC read and commented on multiple drafts.

References

Andrzejaczek S, Meeuwig J, Rowat D, Pierce S, Davies T, Fisher R, Meekan M (2016) The ecological connectivity of whale shark aggregations in the Indian Ocean: a photo-identification approach. Royal Society Open Science 3(11):160455. https://doi.org/10.1098/rsos.160455

Arso Civil M, Smout S, Onoufriou J, Thompson D, Brownlow A, Davison N, Cummings C, Pomeroy P, McConnell B, Hall A (2016) Harbour Seal Decline - vital rates and drivers (Report to Scottish Government HSD2). Retrieved from Sea Mammal Research Unit, University of St Andrews: http://www.smru.st-andrews.ac.uk/files/2016/10/HSD-2-annual-report-year-1.pdf

Arzoumanian Z, Holmberg J, Norman B (2005) An astronomical pattern-matching algorithm for computer-aided identification of whale sharks Rhincodon typus. Journal of Applied Ecology 42(6):999-1011. https://doi.org/10.1111/j.1365-2664.2005.01117.x

Bay H, Ess A, Tuytelaars T, Van Gool L (2008) Speeded-up robust features (SURF). Computer Vision and Image Understanding 110(3):346-359. https://doi.org/10.1016/j.cviu.2007.09.014

Bendik NF, Morrison TA, Gluesenkamp AG, Sanders MS, O'Donnell LJ (2013) Computer-assisted photo identification outperforms visible implant elastomers in an endangered salamander, *Eurycea tonkawae*. PloS one 8(3):e59424. https://doi.org/10.1371/journal.pone.0059424

Bolger DT, Morrison TA, Vance B, Lee D, Farid H (2012) A computer-assisted system for photographic mark-recapture analysis. Methods in Ecology and Evolution 3(5):813-822. https://doi.org/10.1111/j.2041-210X.2012.00212.x

Bradshaw CJ, Mollet HF, Meekan MG (2007) Inferring population trends for the world's largest fish from mark-recapture estimates of survival. Journal of Animal Ecology 76(3):480-489. https://doi.org/10.1111/j.1365-2656.2006.01201.x

Button CA, Rogers-Bennett L (2011) Vital rates of pink abalone *Haliotis corrugata* estimated from mark-recapture data to inform recovery. Marine Ecology Progress Series 431:151-161. https://doi.org/10.3354/meps09094

Caiafa CF, Proto AN, Vergani D, Stanganelli Z (2005) Development of individual recognition of female southern elephant seals, *Mirounga leonina*, from Punta Norte Península Valdés, applying principal components analysis. Journal of Biogeography 32(7):1257-1266. https://doi.org/10.1111/j.1365-2699.2004.01215.x

Calmanovici B, Waayers D, Reisser J, Clifton J, Proietti M (2018) I^3S Pattern as a mark-recapture tool to identify captured and free-swimming sea turtles: an assessment. Marine Ecology Progress Series 589:263-8. https://doi.org/10.3354/meps12483

Chaves LCT, Hall J, Feitosa JLL, Côté IM (2016) Photo-identification as a simple tool for studying invasive lionfish *Pterois volitans* populations. Journal of Fish Biology 88(2):800-804. https://doi.org/10.1111/jfb.12857

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

Crall JP, Stewart CV, Berger-Wolf TY, Rubenstein DI, Sundaresan SR (2013) Hotspotter – Patterned species instance recognition. IEEE workshop on applications of computer vision (WACV) 230-237. https://doi.org/10.1109/WACV.2013.6475023

Cunningham L (2009) Using computer-assisted photo-identification and capture-recapture techniques to monitor the conservation status of harbour seals (*Phoca vitulina*). Aquatic Mammals 35(3):319-329. https://doi.org/10.1578/AM.35.3.2009.319

Den Hartog J, Reijns R (2014) I^3S Pattern Manual. Interactive Individual Identification System. Version 4.0.2. http://www.reijns.com/i3s/download/I3S%20Pattern.pdf. Accessed 25th September 2020

Gailey G, Karczmarski L (2012) Discovery: Photo-identification data-management system for individually recognizable animals. Available from: http://www.biosch.hku.hk/ecology/staffhp/lk/Discovery/

Germanov ES, Bejder L, Chabanne DB, Dharmadi D, Hendrawan IG, Marshall AD, Pierce SJ, van Keulen M, Loneragan NR (2019) Contrasting habitat use and population dynamics of reef manta rays within the Nusa Penida Marine Protected Area, Indonesia. Frontiers in Marine Science 6(215). https://doi.org/10.3389/fmars.2019.00215

Gimenez O, Gatti S, Duchamp C, Germain E, Laurent A, Zimmermann F, Marboutin E (2019) Spatial density estimates of Eurasian lynx (*Lynx lynx*) in the French Jura and Vosges Mountains. Ecology and Evolution 9(20):11707-11715. https://doi.org/10.1002/ece3.5668

Gore MA, Frey PH, Ormond RF, Allan H, Gilkes G (2016) Use of photo-identification and mark-recapture methodology to assess basking shark (*Cetorhinus maximus*) populations. PLoS One 11(3):e0150160. https://doi.org/10.1371/journal.pone.0150160

Goswami VR, Madhusudan MD, Karanth KU (2007) Application of photographic capture-recapture modelling to estimate demographic parameters for male Asian elephants. Animal Conservation 10(3):391-399. https://doi.org/10.1111/j.1469-1795.2007.00124.x

Halloran KM, Murdoch JD, Becker MS (2015) Applying computer-aided photo-identification to messy datasets: a case study of Thornicroft's giraffe (*Giraffa camelopardalis thornicrofti*). African Journal of Ecology 53(2):147-155. https://doi.org/10.1111/aje.12145

Hammond PS, Mizroch SA, Donovan GP (1990) Individual recognition of cetaceans: Use of photo-identification and other techniques to estimate population parameters. Report of the International Whaling Commission 12.

Hastings KK, Small RJ, Hiby L (2001) Use of computer-assisted matching of photographs to examine population parameters of Alaskan harbor seals. In: Harbor Seal Investigations in Alaska. Annual Report for NOAA Award NA87FX0300. Alaska Department of Fish and Game, Division of Wildlife Conservation, Anchorage, AK, 146-160.

Hastings KK, Hiby LA, Small RJ (2008) Evaluation of a computer-assisted photograph-matching system to monitor naturally marked harbor seals at Tugidak Island, Alaska. Journal of Mammalogy 89(5):1201-1211. https://doi.org/10.1644/07-MAMM-A-151.1

Hiby L, Lovell P (1990) Computer aided matching of natural markings: a prototype system for grey seals. Report of the International Whaling Commission 12:57-61.

Hiby L, Lundberg T, Karlsson O, Watkins J, Jüssi M, Jüssi I, Helander B (2007) Estimates of the size of the Baltic grey seal population based on photo-identification data. NAMMCO scientific publications 6:163-175. https://doi.org/10.7557/3.2731

Hiby L, Paterson WD, Redman P, Watkins J, Twiss SD, 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):252-9. https://doi.org/10.1111/2041-210x.12008

Jain AK (2007) Biometric recognition. Nature 449(7158):38-40. https://doi.org/10.1038/449038a

Jiang G, Qi J, Wang G, Shi Q, Darman Y, Hebblewhite M, Miquelle DG, Li Z, Zhang X, Gu J, Chang Y (2015) New hope for the survival of the Amur leopard in China. Scientific Reports 5:15475. https://doi.org/10.1038/srep15475

Kelly BP (1981) Pelage polymorphism in Pacific harbor seals. Canadian Journal of Zoology 59(7):1212-1219. https://doi.org/10.1139/z81-173

Kelly MJ (2001) Computer-aided photograph matching in studies using individual identification: an example from Serengeti cheetahs. Journal of Mammalogy 82:440-449. https://doi.org/10.1644/1545-1542(2001)082<0440:CAPMIS>2.0.CO;2

Koivuniemi M, Auttila M, Niemi M, Levänen R, Kunnasranta M (2016) Photo-ID as a tool for studying and monitoring the endangered Saimaa ringed seal. Endangered Species Research 30:29-36. https://doi.org/10.3354/esr00723

Lahiri M, Tantipathananandh C, Warungu R, Rubenstein DI, Berger-Wolf TY (2011) Biometric animal databases from field photographs: Identification of individual zebra in the wild. Proceedings of the ACM International Conference on Multimedia Retrieval (ICMR) 1-8. https://doi.org/10.1145/1991996.1992002

Langley I, Rosas Da Costa Oliver TV, Hiby L, Stringell T, Morris C, O'Cahdla O, Morgan L, Lock K, Perry S, Westcott S, Boyle D, Büche BI, Stubbings EM, Boys RM, Self H, Lindenbaum C, Strong P, Baines M, Pomeroy PP (2020) Site use and connectivity of female grey seals (*Halichoerus grypus*) around Wales. Marine Biology 167:86. https://doi.org/10.1007/s00227-020-03697-8

Letcher BH, Schueller P, Bassar RD, Nislow KH, Coombs JA, Sakrejda K, Morrissey M, Sigourney DB, Whiteley AR, O'Donnell MJ, Dubreuil TL (2015) Robust estimates of environmental effects on population vital rates: An integrated capture-recapture model of seasonal brook trout growth, survival and movement in a stream network. Journal of Animal Ecology 84(2):337-352. https://doi.org/10.1111/1365-2656.12308

Lowe DG (2004) Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision 60(2):91-110. https://doi.org/10.1023/B:VISI.0000029664.99615.94

Mackey BL, Durban JW, Middlemas SJ, Thompson PM (2008) A Bayesian estimate of harbour seal survival using sparse photo-identification data. Journal of Zoology 274(1):18-27. https://doi.org/10.1111/j.1469-7998.2007.00352.x

Matthé M, Sannolo M, Winiarski K, Spitzen-van der Sluijs A, Goedbloed D, Steinfartz S, Stachow U (2017) Comparison of photo-matching algorithms commonly used for photographic capture-recapture studies. Ecology and Evolution 7(15):5861-5872. https://doi.org/10.1002/ece3.3140

McCormack M (2015) Assessing the applicability of computer aided photo-identification for Pinniped studies through the determination of site fidelity in Long Island, NY harbor seals (*Phoca vitulina concolor*). Dissertation, University of Miami.

Mettouris O, Megremis G, Giokas S (2016) A newt does not change its spots: using pattern mapping for the identification of individuals in large populations of newt species. Ecological Research 31(3):483-489. https://doi.org/10.1007/s11284-016-1346-y

Morrison TA, Yoshizaki J, Nichols JD, Bolger DT (2011) Estimating survival in photographic capture-recapture studies: Overcoming misidentification error. Methods in Ecology and Evolution 2(5):454-463. https://doi.org/10.1111/j.2041-210X.2011.00106.x

Morton AC (1982) The effects of marking and capture on recapture frequencies of butterflies. Oecologia 53(1):105-110. https://doi.org/10.1007/BF00377143

Paterson WD, Redman P, Hiby LA, Moss SE, Hall AJ, Pomeroy P (2013) Pup to adult photo-ID: Evidence of pelage stability in gray seals. Marine Mammal Science 29(4):537-541. https://doi.org/10.1111/mms.12043

Pereira G, Maneyro R (2016) Movement patterns in a Uruguayan population of *Melanophryniscus montevidensis* (Philippi, 1902) (Anura: Bufonidae) using photo-identification for individual recognition. South American Journal of Herpetology 11(2):119-126. https://doi.org/10.2994/SAJH-D-15-00020.1

R Core Team (2019) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Rocha R, Carrilho T, Rebelo R (2013) Iris photo-identification: A new methodology for the individual recognition of Tarentola geckos. Amphibia-Reptilia 34:590-596. https://doi.org/10.1163/15685381-00002918

Rotella JJ, Link WA, Chambert T, Stauffer GE, Garrott RA (2012) Evaluating the demographic buffering hypothesis with vital rates estimated for Weddell seals from 30 years of mark–recapture data. Journal of Animal Ecology 81(1):162-173. https://doi.org/10.1111/j.1365-2656.2011.01902.x

Sacchi R, Scali S, Mangiacotti M, Sannolo M, Zuffi MA (2016) Digital identification and analysis. In: Dodd CK (ed) Reptile Ecology and Conservation A Handbook of Techniques. Oxford University Press, pp 59-72.

Sherwin RE, Haymond S, Stricklan D, Olsen R (2002) Freeze-branding to permanently mark bats. Wildlife Society Bulletin 97-100. https://www.jstor.org/stable/3784641

Steinmetz K, Webster I, Rowat D, Bluemel JK (2018) Evaluating the Software I3S Pattern for Photo-Identification of Nesting Hawksbill Turtles (*Eretmochelys imbricata*). Marine Turtle Newsletter 155:15-19.

Thompson PM, Wheeler H (2008) Photo-ID-based estimates of reproductive patterns in female harbor seals. Marine Mammal Science 24(1):138-146. https://doi.org/10.1111/j.1748-7692.2007.00179.x

Vernes K, Freeman M, Nesbitt B (2009) Estimating the density of free-ranging wild horses in rugged gorges using a photographic mark-recapture technique. Wildlife Research 36(5):361-367. https://doi.org/10.1071/WR07126

Vitkalova AV, Shevtsova EI (2016) A complex approach to study the Amur leopard using camera traps in protected areas in the southwest of Primorsky Krai (Russian Far East). Nature Conservation Research Заповедная наука 1(3).

Wells RS, Scott MD (1990) Estimating bottlenose dolphin population parameters from individual identification and capture-release techniques. Report of the International Whaling Commission 12:407-415.

Yochem PK, Stewart BS, Mina M, Zorin A, Sadovov V, Yablokov A (1990) Non-metrical analyses of pelage patterns in demographic studies of harbor seals. Report of the International Whaling Commission 12:87-90.

Zimmerman GS, Gutierrez RJ, Lahaye WS (2007) Finite study areas and vital rates: sampling effects on estimates of spotted owl survival and population trends. Journal of Applied Ecology 44(5):963-971. https://doi.org/10.1111/j.1365-2664.2007.01343.x

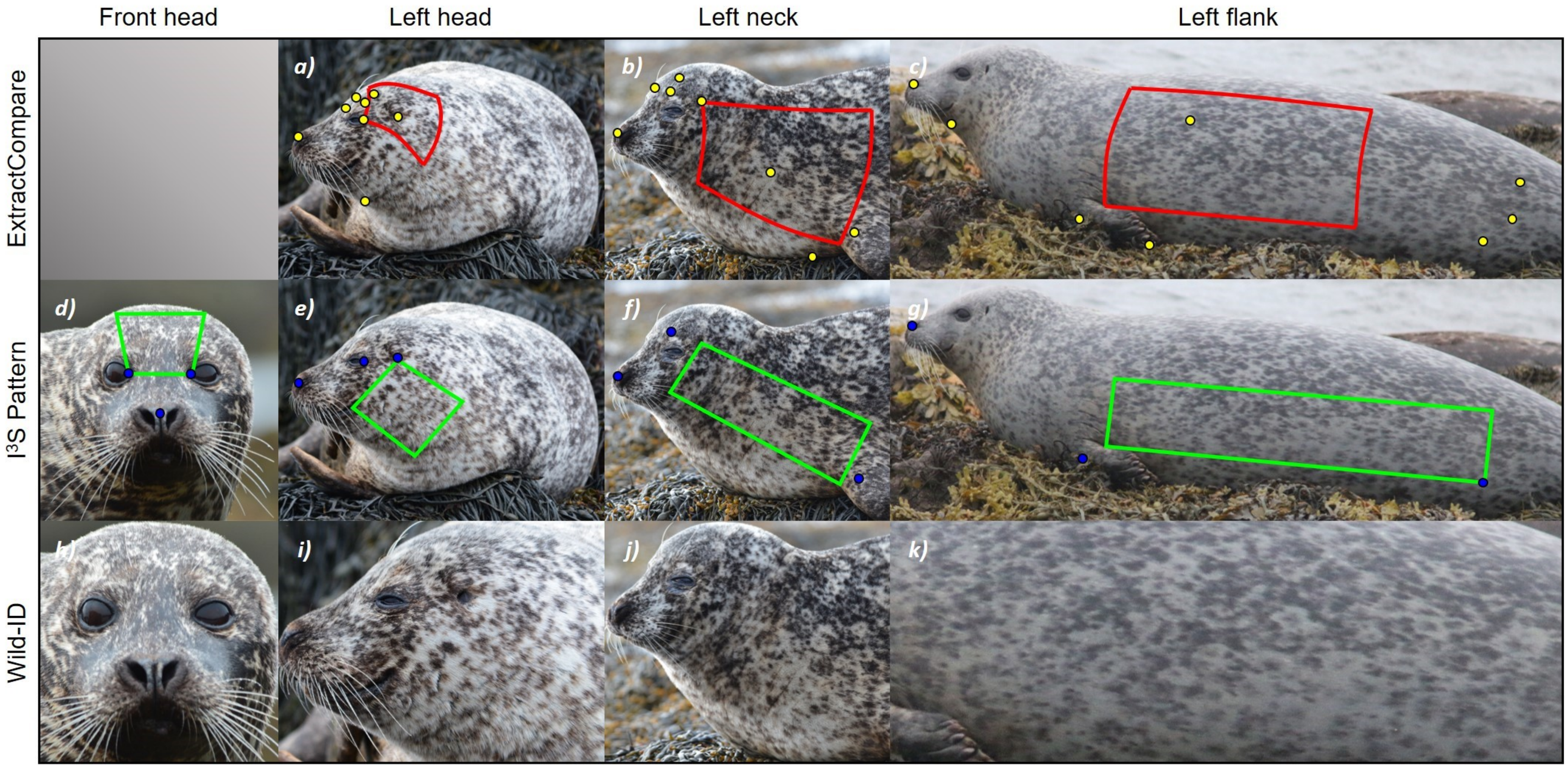

Fig. 1 Aspect specific reference points and extractable areas used to compare the pelage pattern of harbour seals using ExtractCompare, I^{3}S Pattern and Wild-ID. Top row (ExtractCompare): reference points (yellow dots) and extractable area (red box). Middle row (I^{3}S Pattern): reference points (blue dots) and extractable area (green box). Bottom row (Wild-ID): Wild-ID does not use reference points and so the extractable area is the cropped aspect of the subject.

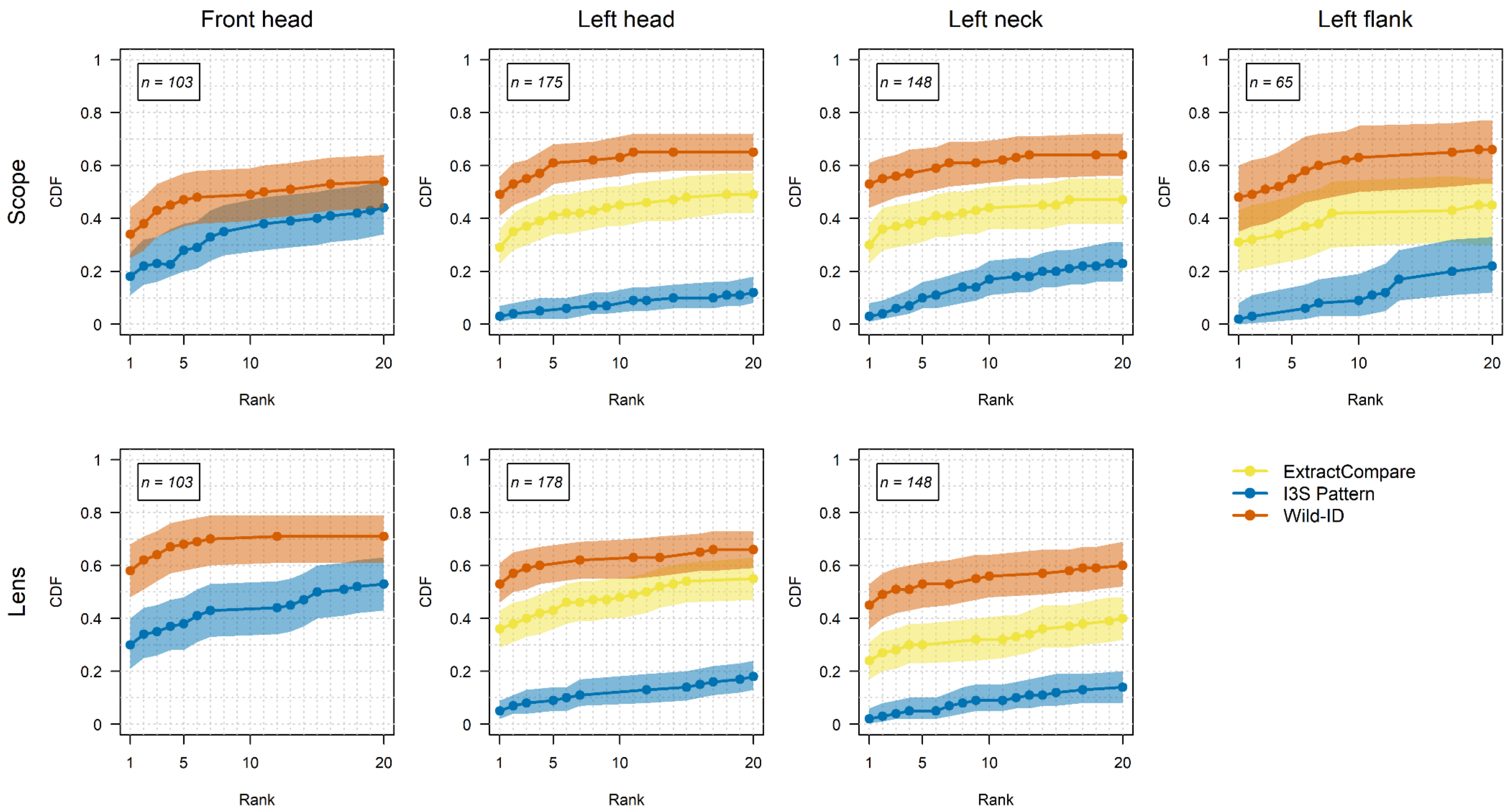

Fig. 2 The cumulative density function (CDF) of the matches detected by ranked similarity score. Trials were run for each pattern recognition software: ExtractCompare (yellow), I[3]S Pattern (blue) and Wild-ID (red), by data collection equipment (scope, lens) and seal aspect (front head, left head, left neck, left flank). Shaded areas represent 95% confidence intervals based on the binomial distribution.

Table 1. The cumulative density function and 95% confidence intervals for potential matches ranked in first position (CDF$_1$), and within the top 5, 10 and 20 ranks (CDF$_5$, CDF$_{10}$, and CDF$_{20}$ respectively); *n* is the number of individuals in each database, Db (number of images = *2n*). Flank aspects were not available from the lens data (as often the photographer was too close to the seal to capture the entire body) and front head aspects cannot be processed in the current ExtractCompare harbour seal model.

Trial	Db	Method	Aspect	*n*	Software	CDF$_1$	CDF$_5$	CDF$_{10}$	CDF$_{20}$
1	A	Scope	Front head	103	I^3S Pattern	0.18 (0.11, 0.27)	0.28 (0.20, 0.38)	0.35 (0.26, 0.45)	0.44 (0.34, 0.54)
2	A	Scope	Front head	103	Wild-ID	0.34 (0.25, 0.44)	0.47 (0.37, 0.57)	0.49 (0.39, 0.59)	0.54 (0.44, 0.64)
3	B	Lens	Front head	103	I^3S Pattern	0.30 (0.21, 0.40)	0.38 (0.28, 0.48)	0.43 (0.33, 0.53)	0.53 (0.43, 0.63)
4	B	Lens	Front head	103	Wild-ID	0.58 (0.48, 0.68)	0.68 (0.58, 0.77)	0.70 (0.60, 0.79)	0.71 (0.61, 0.79)
5	C	Scope	Left head	175	ExtractCompare	0.29 (0.23, 0.36)	0.41 (0.34, 0.49)	0.45 (0.37, 0.52)	0.49 (0.42, 0.57)
6	C	Scope	Left head	175	I^3S Pattern	0.03 (0.01, 0.07)	0.05 (0.02, 0.10)	0.07 (0.04, 0.12)	0.12 (0.08, 0.18)
7	C	Scope	Left head	175	Wild-ID	0.49 (0.41, 0.56)	0.61 (0.53, 0.68)	0.63 (0.56, 0.71)	0.65 (0.58, 0.72)
8	D	Lens	Left head	178	ExtractCompare	0.36 (0.29, 0.43)	0.43 (0.36, 0.51)	0.48 (0.40, 0.55)	0.55 (0.47, 0.63)
9	D	Lens	Left head	178	I^3S Pattern	0.05 (0.02, 0.09)	0.09 (0.05, 0.14)	0.11 (0.07, 0.17)	0.18 (0.13, 0.24)
10	D	Lens	Left head	178	Wild-ID	0.53 (0.46, 0.61)	0.60 (0.53, 0.67)	0.62 (0.55, 0.69)	0.66 (0.59, 0.73)
11	E	Scope	Left neck	148	ExtractCompare	0.30 (0.23, 0.38)	0.39 (0.31, 0.47)	0.44 (0.36, 0.52)	0.47 (0.38, 0.55)
12	E	Scope	Left neck	148	I^3S Pattern	0.03 (0.01, 0.08)	0.10 (0.06, 0.16)	0.17 (0.11, 0.24)	0.23 (0.16, 0.31)
13	E	Scope	Left neck	148	Wild-ID	0.53 (0.44, 0.61)	0.57 (0.49, 0.66)	0.62 (0.54, 0.70)	0.64 (0.56, 0.72)
14	F	Lens	Left neck	148	ExtractCompare	0.24 (0.17, 0.31)	0.30 (0.23, 0.38)	0.32 (0.24, 0.40)	0.40 (0.32, 0.48)
15	F	Lens	Left neck	148	I^3S Pattern	0.02 (0.004, 0.06)	0.05 (0.02, 0.10)	0.09 (0.05, 0.15)	0.14 (0.08, 0.20)
16	F	Lens	Left neck	148	Wild-ID	0.45 (0.36, 0.53)	0.53 (0.44, 0.61)	0.56 (0.48, 0.64)	0.60 (0.52, 0.69)
17	G	Scope	Left flank	65	ExtractCompare	0.31 (0.20, 0.43)	0.34 (0.23, 0.47)	0.42 (0.29, 0.54)	0.45 (0.30, 0.55)
18	G	Scope	Left flank	65	I^3S Pattern	0.02 (0.0004, 0.08)	0.03 (0.004, 0.11)	0.09 (0.03, 0.19)	0.22 (0.12, 0.33)
19	G	Scope	Left flank	65	Wild-ID	0.48 (0.35, 0.60)	0.55 (0.43, 0.68)	0.63 (0.50, 0.75)	0.66 (0.53, 0.77)

Table 2. Data processing rate (time in minutes:seconds for a single image to be processed, from pre-processing to visual confirmation). Trials correspond to Table 1; *n* is the number of individuals within each database (number of images = *2n*); timed stages were crop (image cropping), input (data input), extract (pattern extract), compare (pattern comparison), and confirm (visual confirmation). In I^{3}S Pattern, the stages from data input to visual confirmation were combined into a single step, represented below by merged cells.

<table>
<tr><th>Software</th><th>Trial</th><th>Method</th><th>Aspect</th><th>n</th><th>Crop</th><th>Input</th><th>Extract</th><th>Compare</th><th>Confirm</th><th>Overall</th></tr>
<tr><td rowspan="5">ExtractCompare</td><td>5</td><td>Scope</td><td>Left head</td><td>175</td><td>00:10</td><td>00:07</td><td>00:53</td><td>00:05</td><td>00:10</td><td>01:25</td></tr>
<tr><td>8</td><td>Lens</td><td>Left head</td><td>178</td><td>00:10</td><td>00:11</td><td>00:59</td><td>00:06</td><td>00:13</td><td>01:39</td></tr>
<tr><td>11</td><td>Scope</td><td>Left neck</td><td>148</td><td>00:10</td><td>00:07</td><td>01:10</td><td>00:06</td><td>00:12</td><td>01:45</td></tr>
<tr><td>14</td><td>Lens</td><td>Left neck</td><td>148</td><td>00:09</td><td>00:09</td><td>01:03</td><td>00:05</td><td>00:15</td><td>01:40</td></tr>
<tr><td>17</td><td>Scope</td><td>Left flank</td><td>65</td><td>00:10</td><td>00:08</td><td>00:59</td><td>00:03</td><td>00:12</td><td>01:32</td></tr>
<tr><td rowspan="7">I^{3}S Pattern</td><td>1</td><td>Scope</td><td>Front head</td><td>103</td><td>NA</td><td colspan="4">00:33</td><td>00:33</td></tr>
<tr><td>3</td><td>Lens</td><td>Front head</td><td>103</td><td>NA</td><td colspan="4">00:31</td><td>00:31</td></tr>
<tr><td>6</td><td>Scope</td><td>Left head</td><td>175</td><td>NA</td><td colspan="4">00:35</td><td>00:35</td></tr>
<tr><td>9</td><td>Lens</td><td>Left head</td><td>178</td><td>NA</td><td colspan="4">00:31</td><td>00:31</td></tr>
<tr><td>12</td><td>Scope</td><td>Left neck</td><td>148</td><td>NA</td><td colspan="4">00:22</td><td>00:22</td></tr>
<tr><td>15</td><td>Lens</td><td>Left neck</td><td>148</td><td>NA</td><td colspan="4">00:31</td><td>00:31</td></tr>
<tr><td>18</td><td>Scope</td><td>Left flank</td><td>65</td><td>NA</td><td colspan="4">00:33</td><td>00:33</td></tr>
<tr><td rowspan="7">Wild-ID</td><td>2</td><td>Scope</td><td>Front head</td><td>103</td><td>00:16</td><td><00:01</td><td>00:01</td><td>00:06</td><td><00:01</td><td>00:22</td></tr>
<tr><td>4</td><td>Lens</td><td>Front head</td><td>103</td><td>00:15</td><td><00:01</td><td>00:01</td><td>00:03</td><td><00:01</td><td>00:19</td></tr>
<tr><td>7</td><td>Scope</td><td>Left head</td><td>175</td><td>00:14</td><td><00:01</td><td>00:01</td><td>00:04</td><td><00:01</td><td>00:20</td></tr>
<tr><td>10</td><td>Lens</td><td>Left head</td><td>178</td><td>00:14</td><td>00:01</td><td>00:01</td><td>00:06</td><td>00:01</td><td>00:21</td></tr>
<tr><td>13</td><td>Scope</td><td>Left neck</td><td>148</td><td>00:15</td><td>00:01</td><td>00:01</td><td>00:05</td><td>00:01</td><td>00:22</td></tr>
<tr><td>16</td><td>Lens</td><td>Left neck</td><td>148</td><td>00:14</td><td>00:01</td><td>00:02</td><td>00:05</td><td>00:01</td><td>00:21</td></tr>
<tr><td>19</td><td>Scope</td><td>Left flank</td><td>65</td><td>00:25</td><td><00:01</td><td>00:01</td><td>00:05</td><td><00:01</td><td>00:31</td></tr>
</table>