

Comparison of two individual identification algorithms for snow leopards (*Panthera uncia*) after automated detection

Eve Bohnett^{a,g,*}, Jason Holmberg^b, Sorosh Poya Faryabi^c, Li An^{d,h}, Bilal Ahmad^e, Wajid Rashid^f, Stephane Ostrowski^c

^a Department of Biology, San Diego State University, San Diego, CA, USA

^b Wild Me, 1726 N Terry Street, Portland, OR, USA

^c Wildlife Conservation Society, 2300 Southern Boulevard Bronx, New York, NY, USA

^d Department of Geography, San Diego State University, San Diego, CA, USA

^e Institute of Agriculture Sciences and Forestry, University of Swat, Mingora, Pakistan

^f Department of Environmental and Conservation Sciences, University of Swat, Mingora, Pakistan

^g Department of Landscape Architecture, University of Florida, Gainesville, FL, USA

^h College of Forestry, Wildlife and Environment, Auburn University, Auburn, AL, USA

ARTICLE INFO

Keywords:

Background subtraction

Deep learning

Hotspotter

Individual identification

PIE v2

Snow leopards

ABSTRACT

Photo-identification of individual snow leopards (*Panthera uncia*) is the primary data source for density estimation via capture-recapture statistical methods. To identify individual snow leopards in camera trap imagery, it is necessary to match individuals from a large number of images from multiple cameras and historical catalogues, which is both time-consuming and costly. The camouflaged snow leopards also make it difficult for machine learning to classify photos, as they blend in so well with the surrounding mountain environment, rendering applicable software solutions unavailable for the species. To potentially make snow leopard individual identification available via an artificial intelligence (AI) software interface, we first trained and evaluated image classification techniques for a convolutional neural network, pose invariant embeddings (PIE) (a triplet loss network), and compared the accuracy of PIE to that of the HotSpotter algorithm (a SIFT-based algorithm). Data were acquired from a curated library of free-ranging snow leopards taken in Afghanistan between 2012 and 2019 and from captive animals in zoos in Finland, Sweden, Germany, and the United States. We discovered several flaws in the initial PIE model, such as a small amount of background matching, that was addressed, albeit likely not fixed, using background subtraction (BGS) and left-right mirroring (LR) techniques which demonstrated reasonable accuracy (Rank 1: 74% Rank-5: 92%) comparable to the Hotspotter results (Rank 1: 74% Rank 2: 84%) The PIE BGS LR model, in conjunction with Hotspotter, yielded the following results: Rank-1: 85%, Rank-5: 95%, Rank-20: 99%. In general, our findings indicate that PIE BGS LR, in conjunction with HotSpotter, can classify snow leopards more accurately than using either algorithm alone.

1. Introduction

Population density and size are crucial ecological variables, and recent research has demonstrated that they can be approximated using spatial capture-recapture models from data captures at different locations with closely spaced time intervals, which is dependent on individual-level animal data modeled across space and time (Borchers and Fewster, 2016; Jackson et al., 2006; McCarthy et al., 2016; Royle et al., 2009; Royle and Young, 2008). Camera trapping and genetic data analysis are the two main data sources for modeling population size and

density, where data preparation typically includes sorting the data at the outset to identify individuals by their unique characteristics (Crouse et al., 2017; Pereira et al., 2022; Strampelli et al., 2020; Valdez et al., 2015; Weckworth, 2021; Wegge et al., 2019; Zheng et al., 2016a). Due to the difficulty of identifying individual animals from large datasets for use in capture-recapture models, it has been challenging for ecologists to obtain precise population size and density estimates of wildlife (Alonso et al., 2015; Efford et al., 2009).

With the emergence of deep learning approaches in ecological informatics, ecologists have begun addressing conservation biology

* Corresponding author at: Department of Biology, San Diego State University, San Diego, CA, USA.

E-mail address: evebohnnett@yahoo.com (E. Bohnett).

<https://doi.org/10.1016/j.ecoinf.2023.102214>

Received 9 February 2023; Received in revised form 12 July 2023; Accepted 13 July 2023

Available online 18 July 2023

1574-9541/© 2023 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

concerns by integrating deep learning and artificial intelligence (AI) into their research workflows and designs (Miao et al., 2019). The availability of low-cost digital camera traps has necessitated the development of software solutions for organizing and analyzing photos of animals, as the need to manage massive photo sets has increased (Bohnett et al., 2022; Greenberg et al., 2019; Miao et al., 2019; Rashid et al., 2021; Schneider et al., 2020a; Sultan et al., 2022; Swanson et al., 2015; Tabak et al., 2019, 2020; Verschueren et al., 2023; Willi et al., 2019; Yu et al., 2013). As a means of reducing the time, effort, and sources of human error necessary to interpret camera trap data, scientists are increasingly turning to AI and computer vision to identify animals by species (Beery et al., 2019; Falzon et al., 2019; Nguyen et al., 2017a; Norouzzadeh et al., 2019; Parham et al., 2018), and pattern recognition software has been developed for distinguishing individually distinct patterns such as an individual's stripes or spots in camera trap photographs (Blount et al., 2018; Chehrsimin et al., 2018; Crouse et al., 2017; Dawson, 2021; Langley et al., 2021; Park et al., 2019; Shi et al., 2020; Wäldchen and Mäder, 2018; Weinstein, 2018). Machine learning software is becoming increasingly user-friendly, allowing individuals with limited computer science knowledge to independently apply the most recent algorithms to their issues and datasets through coding packages or software user interfaces (Wäldchen and Mäder, 2018).

Many software programs have been developed to incorporate an AI workflow into density and population size estimation via spatial-capture recapture modeling, for example, WildBook (Berger-Wolf et al., 2017), APHIS (Moya et al., 2015), I3S Pattern+ (Hartog and Reijns, 2013), AmphIdent (Matthé, 2015), and Wild-ID (Bolger et al., 2011, 2012), have all been tested for their ability to identify individual animals (Bardier et al., 2020; Dawson, 2021; Matthé et al., 2017). Various algorithms have been made available for individual identification (ID) classification for camera trapping imagery within this software, for example, HotSpotter, a Scale Invariant Feature Transform (SIFT) which casts points onto an image to then identify k-nearest neighbor clusters of hotspots within the image for which to match against (Crall et al., 2013; Lowe, 2004). Hotspotter has been one of the most popular methods for image classification for individual ID for various species and has been incorporated into software (Burgstaller et al., 2020; Crall et al., 2013; de Lorm et al., 2023; Khan et al., 2022; Miguel et al., 2019; Nipko et al., 2020; Park et al., 2019; Verschueren et al., 2023). Recent advances in deep learning have facilitated the growth of convolutional neural network (CNN)-based systems that can now extract layers of image characteristics for a wide variety of ecological and environmental image processing concerns, especially for camera trapping image analysis (Bogucki et al., 2019; El Abbadi and Alsaadi, 2020; Nguyen et al., 2017b; Norouzzadeh et al., 2018a; Shi et al., 2020). Due to the capacity to reduce images to numerous layers of distinct patterns, these advanced feature extraction and CNN-based deep learning approaches have demonstrated promising results for identifying individuals (Bhattacharya et al., 2022; Cheema and Anand, 2017; Miao et al., 2019).

Deep learning research has concentrated on the broad topic of re-identification using CNNs, which have been largely designed and widely implemented for human identification mainly through facial recognition (Hermans et al., 2017; Moskvayak et al., 2020a, 2021; Wang and Deng, 2021; Zheng et al., 2016b; Zheng et al., 2017). Despite the availability of effective techniques and numerous successful attempts to apply the method to non-human species, image ID and reidentification is a challenging task when applied to images of animals in the field (Chehrsimin et al., 2018; Crouse et al., 2017; de Silva et al., 2022; Dębicki et al., 2021; Gómez-Vargas et al., 2023; Guo et al., 2020; Hansen et al., 2018; He et al., 2019; Korschens and Denzler, 2019; Kuncheva et al., 2023; Liu et al., 2019; Loos, 2016; Moskvayak et al., 2019; Nepovinnikh et al., 2022a, 2022b; Park et al., 2019; Petso et al., 2022; Ravor and Sudarshan, 2020; Schneider et al., 2020b; Schofield et al., 2019; Shukla et al., 2019; Zuerl et al., 2023). Low accuracy, especially with low-quality imagery or imagery captured at different distances, angles, visibility, and lighting conditions, has been cited as a problem in previous studies

(Carter et al., 2014; Nepovinnikh et al., 2020; Papafitsoros et al., 2023). Wildlife animal re-identification is frequently hampered by the difficulty of accumulating a large enough dataset containing numerous photographs of many individuals, where the method must be able to identify an individual using a single or a small number of training examples (Nepovinnikh et al., 2022b; Vidal et al., 2021). In addition, the animal re-identification method must be able to ascertain whether or not a new individual is present and thus distinguished from a previously compiled dataset (Nepovinnikh et al., 2022b; Papafitsoros et al., 2023; Stewart et al., 2021).

The population size and density of snow leopards have primarily been estimated using camera trapping (Alexander et al., 2015, 2016; Bian et al., 2022; Zhang et al., 2020), whereas genetic analysis of scat samples has been used infrequently (Atzeni et al., 2021; Chetri et al., 2019; Hacker et al., 2023; Janečka et al., 2011; Laguardia et al., 2015). The inability to visually distinguish snow leopard scats from fox scats based on morphology, loss of genetic stability over time for older scats, and extensive fieldwork for repeated temporal samples over large areas of high altitude habitat all impede genetic analysis, necessitating molecular scatology of a large number of scat samples to differentiate species first and then to the individual snow leopard level, which is challenging to implement in the field and requires advanced lab equipment (Chetri et al., 2019; Laguardia et al., 2015). Using their unique genetic identifiers, however, genetics can be used to find specific individuals and their near genetic relatives comprehensively and reliably (Chetri et al., 2019). However, the most prominent and assessable technique for field-based data collection for individual-level data on snow leopards is the manual sorting of camera-trap photographs based on matching the snow leopard individual's distinctive rosette fur patterns (Alexander et al., 2015, 2016; Bian et al., 2022; Janečka et al., 2011; Oberosler et al., 2022; Sharma et al., 2020; Suryawanshi et al., 2019; Zhang et al., 2020). However, the misclassification of individuals in camera trap photos by manual observers has been demonstrated to introduce significant errors which result in overestimating snow leopard abundance by over one-third (Johansson et al., 2020). Using multiple observers to check large photo sets manually is only one example of the resource-intensive strategies now used to reduce the possibility of misclassifying images of snow leopards (Borchers and Fewster, 2016; Choo et al., 2020; Foster and Harmsen, 2012). If researchers can correctly identify individual snow leopards from photos captured by camera traps, they can monitor population trends better and focus conservation efforts (Johansson et al., 2020).

Emerging capabilities in AI may increase the accuracy of individual snow leopard identification and our ability to classify vast image collections. Field settings and animal morphology often make it challenging to train machine learning algorithms to distinguish patterns, colors, or forms in camera trap photos, as is typically the case with field-collected data. The ability to detect snow leopards from photos of other animals, such as blue sheep, has been effectively automated using AI-supported technology, with convolutional neural networks (CNN) achieving 91% accuracy for species discrimination (Tariq et al., 2018). For individual ID related tasks, snow leopard fur patterns are easily confounded in machine learning algorithms by their habitat's predominantly grey and white mountain background, making it difficult for algorithms to distinguish snow leopard individuals, requiring numerous data preparation steps to subset areas of interest with snow leopard patterns (Beery, 2016; Miguel et al., 2019). Several studies have piloted algorithms for ID of snow leopards, evaluating Groth Algorithm, which is a pattern recognition method for extracting raw rosette patterns and iterating through the patterns to verify matches (Beery, 2016), and HotSpotter (Miguel et al., 2019) has been piloted for use on snow leopards but not formally evaluated for accuracy. Considering the difficulty for snow leopards to be distinguished from snowy mountain terrain, studies have highlighted the image segmentation research necessary to remove the distracting background elements of a picture so that algorithms can better concentrate on the relevant patterns of the

snow leopards (Beery et al., 2019; Miguel et al., 2019). Even with these methodological advancements, field-based studies have raised concerns about machine learning's overall accuracy and dependability for this species, leading researchers not to implement it in data processing workflows (Rode et al., 2021). These publications emphasize the current limitations of early research on snow leopard pattern identification leading to a lack of field research programs incorporating these image classification algorithms or available software.

This study investigates how additional deep learning algorithms may facilitate semi-autonomous classification and detection strategies for individual snow leopards and improve the current effort for individual ID. This work is novel as the first attempt to test and comprehensively evaluate two computer vision algorithms to determine how well they match individual snow leopard sightings. Additionally, the project must match previously classified and curated snow leopard photo-ID catalogues with individuals from freshly gathered datasets, a problem known as "continual curation" (Stewart et al., 2021). To accomplish this, we must automate a more time-efficient pipeline, be less prone to misidentification errors, and be accessible to field-based researchers through a graphical user interface.

The [Whiskerbook.org](https://whiskerbook.org) online platform (Blount et al., 2021b) includes a web-based data management architecture and a computer vision pipeline (Parham et al., 2018) for the detection and individual ID of various species of large cats. The [Whiskerbook.org](https://whiskerbook.org) software implemented the Hotspotter algorithm (Crall et al., 2013), a SIFT-based analysis comparing important areas of visual texture between two or more photos, previously developed and implemented on jaguar (*Panthera onca*). In addition, recent advances in machine learning suggest that a new class of deep learning-based approaches, initially created for facial recognition via the extraction of complex patterned information and analyzed with a CNN, apply to wildlife individual ID (Moskvyak et al., 2019, 2020b). The deep learning method used in this study is a CNN called Pose Invariant Embeddings (PIE v2; Moskvyak et al., 2019), specifically an InceptionV3 optimized model that learns the extracted feature embeddings with a triplet loss network and is initially tested on manta ray bellies (*Mobula* spp.) and humpback whale (*Megaptera novaeangliae*) flukes for implementation in online software (Blount et al., 2018). The PIEv2 triplet loss function compares two images by computing the Euclidean distance between them using the embedding vector obtained from the pooled feature maps, which results from performing numerous dimensionality reductions to the images to extract visible patterns. PIEv2 is resilient to viewpoint changes of animal positions, lighting changes, and small occlusions. PIEv2 does not require retraining due to the addition of new pictures to the enlarged dataset because it will match both pictures that were in the database during training and pictures that were uploaded after training. (Moskvyak et al., 2019). The ability to quickly match individuals across photo ID catalogues (potentially saving a great deal of time and money on the path to more extensive and comprehensive catalogues and modeling efforts) is just one of the many ways that researchers can benefit from a better understanding of the performance of the HotSpotter and PIEv2 algorithms, which can inform the [Whiskerbook.org](https://whiskerbook.org) platform's usability and aid in snow leopard conservation.

2. Materials and methods

The required steps for implementing the workflow used in this project began with the snow leopard camera trap imagery, then training and testing the feature detection algorithms for accuracy in annotating snow leopards in the imagery, followed by data cleaning including graph matching steps to qualify individual snow leopard ID matches, subsetting the data according to min-3/max-10, and then testing the PIE and Hotspotter ID models. We also further subset the data according to individuals at the same location and different location to test the ID models for background matching. Finally, data augmentation for left/right mirroring and background subtraction were performed prior to

sending the data to another round of training and testing of the individual ID models (Fig. 1).

2.1. Camera trap imagery

The Wildlife Conservation Society (WCS) in Afghanistan curated images of snow leopard individuals taken between 2012 and 2019 using the [Whiskerbook.org](https://whiskerbook.org) interface; these were subjects of our ML tests and assessments (Blount et al., 2021b). Seven European zoos (Helsinki and Åtheri Zoos in Finland, Kolmården Zoo, Nordens Ark and Orsa Bear Park in Sweden, and Köln and Wuppertal Zoos in Germany) (Johansson et al., 2020), and two American zoos (WCS managed Bronx and Central Park zoos in New York City) supplied further data regarding captive snow leopards. Data in the sorted dataset used camera traps from both Reconyx and Bushnell. Three manual observers sorted the data initially according to individuals from the field-based dataset. The imagery collected in this study was 60% nighttime with black and white photos and 40% daytime illuminated color photographs. Previous studies have shown similar findings throughout the night and day activity (proportional movement activity of night [sunset to sunrise] versus day [sunrise to sunset] for males = 66% versus 34%, and for females = 60% versus 40%, respectively) (Johansson et al., 2022).

2.2. Feature detection

Feature detection is a technique used in computer vision and image processing pipelines to automatically extract pertinent features from unprocessed image data (Cheema and Anand, 2017). These features can then be implemented in various applications, including object identification and image classification. Feature detection methods are meant to learn and recognize complicated patterns in visual data, making them especially suitable for situations where traditional computer vision approaches may be ineffective, such as finding objects in crowded or variable backdrops (Christin et al., 2018; Duhart et al., 2019; Guo et al., 2017; LeCun et al., 2015; Miao et al., 2019; Mikolajczyk and Grochowski, 2018; Norouzzadeh et al., 2018b).

The snow leopard is the subject of this study, so we trained a detection algorithm to locate the snow leopard within an image. In the initial stage of the Wildbook Image Analysis (WBIA) pipeline (Parham et al., 2018), snow leopard annotations were generated by a machine learning detector, a modified PyTorch implementation of YOLO v2 (Redmon and Farhadi, 2016). The detector aims to localize animals in images, generating precise bounding boxes over ground-truth detections (made a priori by humans for a test set) while minimizing false positives and negatives (Fig. 2). Essentially, the bounding boxes restrict the photographs to only the areas that include snow leopards. Snow leopards are bilaterally asymmetric, with significantly different patterns on their left and right sides; therefore, these viewpoints are required to match photographs of the species, which are further labeled according to viewpoint classification to determine whether an image of the snow leopard is being viewed from the front or the back (such as left, right, back, front, front left, front right, back left, back right). We used a training dataset of 2000 images and 2078 annotated bounding boxes (2 empty images, 34 images with two boxes, and 23 images with two boxes) to train a classifier to predict the snow leopard viewpoint.

Our project had access to 359 sightings of snow leopards at hourly intervals, yielding 22,120 annotations (i.e., machine learning-detected bounding boxes around snow leopards in photos) (Fig. 1). The number of annotations dropped to 12,311, and the number of separate encounters dropped to 116 when the data was restricted to only individuals spotted three or more times (three annotations of a side are a minimum criterion for PIE model training).

2.3. Graph matching

These annotations were rigorously subset into individuals using a

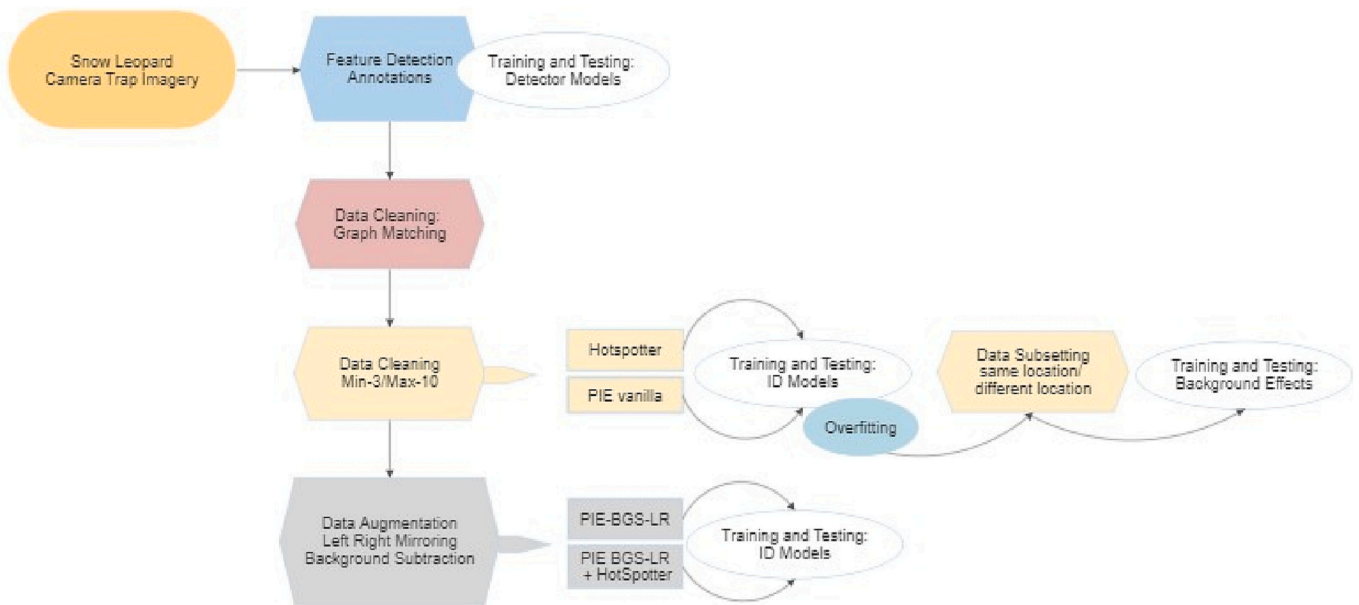


Fig. 1. Workflow diagram of methods.

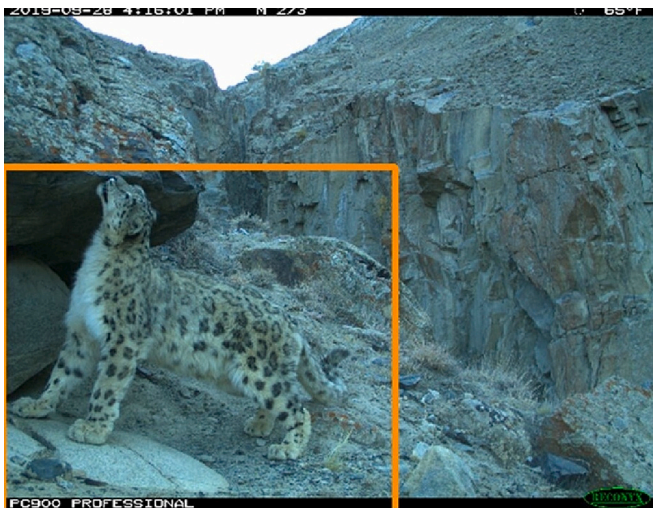


Fig. 2. An annotated snow leopard. Annotations were created using a computer vision model based on machine learning and were coupled with the human identification of known individuals. Annotations served as the core data from which ML learned and against which each algorithm was assessed. Photo courtesy of WCS Afghanistan.

comprehensive graph matching software within the Wild Me software suite for cleaning data at the standard suitable for deep learning algorithms. Two manual observers participated in exhaustive rounds of ID decisions in a process called graph matching of the 12,311 annotations, whereby two images are placed side by side, and the observer determines whether the images are the same individual or different individuals (same animal, different animal, cannot tell), and ranks their level of certainty (undeniable, certain, somewhat certain, mostly uncertain, completely uncertain). This level of comprehensive individual ID enabled the subsets of our dataset to have the highest certainty for imagery from a field-based dataset.

2.4. Min-3/Max-10

A Min-3/Max-10 data subset is frequently used for training machine learning algorithms such as PIE. This data subset comprises individuals

with at least three images of the same viewpoint (either left or right), limited to a maximum of ten images per individual/viewpoint. Three images are required for the training phase (two images for ML to learn from), and one image is required for testing (at least one for ML to test it against). For data set fairness and to avoid the ML system optimizing on highly sighted individuals and performing poorly on infrequently sighted individuals, a limit of 10 images is implemented. According to our research, setting a max-10 limit on individuals to be matched results in a more effective machine learning model for real-world matching (Blount et al., 2018, 2021a). Applying these criteria to the collected data on Whiskerbook.org yielded a total of 829 photos of 217 individual snow leopards. It was compiled using data from both the captive zoo and field camera-trapped snow leopards using the above described advanced filtering and verification techniques.

2.5. Individual identification algorithms

2.5.1. HotSpotter

The HotSpotter algorithm method casts a series of points onto the image, and compares high-importance texture patches using a scale-invariant feature transform (SIFT), then k-nearest neighbors scoring highlights clusters of multiple similar points of an image. (Crall et al., 2013). Highlighting the group of k neighbors in yellow draws attention to the potential similarities image. Each photo match receives a numerical score. A manual observer can then examine each image pair for matching clusters and determine if the individuals in the photos are the same. The algorithm may compare numerous photos of a known individual to a target image and scan the entire database for the best matches or similar images of the same individual (Crall et al., 2013).

2.5.2. Pose-invariant embeddings (PIE)

The Pose Invariant Embeddings (PIE) method uses the ResNeXt (Xie et al., 2017) convolutional neural network (CNN) trained on labeled and classed pictures. CNNs reduce images to their main features to discover patterns in dimensionality reduction, allowing multiple patterns to be extracted in layers. The algorithm “learns” from these extracted features to create a classification model for similar-shaped patterns. The computer “mines” the image for forms and features to build an “embedding” for a tagged photo database, exposing snow leopard fur patterns. After embedding, the k-nearest neighbor classifier may find images that match the focal image.

2.6. Network architecture, data pre-processing, training, and evaluation

With a machine learning detector developed and configured to extract snow leopard annotations, these annotations and accompanying information (in particular, the known identifications based on coat patterning) were then utilized in the WBIA pipeline (Parham et al., 2018) to first custom-train the Pose Invariant Embeddings (PIE) algorithm (Moskvyak et al., 2019). We used a Min-3/Max-10 data constraint for PIE ML training and divided the training and test data (Table 1).

Models were trained and tested by tuning the number of required epochs and then inspecting the model outputs and error outcomes for signs of overfitting. If the model focuses excessively on data fitting, it may imitate insignificant noise rather than the broad properties of interest, resulting in overfitting.

After performing the first round of PIE algorithm modeling, we observed PIE converging extremely quickly while training on this data. We believe this is partially due to snow leopard territoriality, such that images of an individual may come from only one camera trap location, and the training data was mainly from in-field-based camera traps (with a minority of captive zoo data). We theorized that these two factors resulted in “background matching” contributing to the matching strategy of the PIE model during training, with the model learning to identify individual snow leopards by identifying the environments in which those leopards were most frequently seen. Overfitting was further examined by modeling two datasets, one using a subset of individuals ($n = 14$) observed at multiple locations and a subset of individuals ($n = 29$) observed in the exact location many times.

2.7. Data augmentation

Deep CNN-based identification methods require large training sets with numerous annotated images (Bansal et al., 2022; Masi et al., 2016). To generate enough data for a deep learning project, we collected and labeled additional zoo and field-based data after discovering only one available individually labeled snow leopard dataset (16 individuals) from a prior study on camera trapping captive snow leopards in zoos (Johansson et al., 2020). Furthermore, it was also necessary to expand the dataset with new data to achieve a larger sample size by artificially extending the dataset through data augmentation (Masi et al., 2016; Mikolajczyk and Grochowski, 2018). Examples of data augmentation include geometric transformations (Flusser and Suk, 1993), such as cropping (Chatfield et al., 2014), mirroring (Yang and Patras, 2015), and rotating (Xie and Tu, 2015), or photometric transformations (Krizhevsky et al., 2017), which are beneficial because they do not change the original base data features and can be used to enrich a data set, supplemented by as much as $10\times$ the amount of data while also reducing overfitting of the model (Bansal et al., 2022; Howard, 2013; Lv et al., 2017; Shorten and Khoshgoftaar, 2019).

During training, we took two approaches to reduce the likelihood of the PIE model overfitting. The first technique, aptly titled “background subtraction,” involves algorithmically erasing the background (Babae et al., 2017; Beery, 2016; Mahadevan and Vasconcelos, 2008; Sobral and Vacavant, 2014; Yousif et al., 2017). As part of detector training in the WBIA pipeline (Parham et al., 2018), Wild Me trained a background subtraction model for snow leopards (Fig. 3).

Our second data augmentation strategy involved mirroring left-side images so that each image was transformed to appear as an animal’s right-side viewpoint. This technique for augmenting data is typically



Fig. 3. A background-subtracted snow leopard photo used to train the final PIE model. Credit: Wild Me.

referred to as “mirroring or flip augmentation”, and they have been found to considerably increase the performance of deep learning algorithms (Chatfield et al., 2014). The left/right mirroring method helped us to expand the dataset. We anticipate that mirroring could double the backdrop texture pool used for background matching. This mirroring is a configuration setting that can be enabled on or off during PIE training. It has previously been used on complicated matching problems to achieve maximum image-level consistency: PIE can typically match multiple perspectives but standardizing them can improve accuracy even more.

2.8. Performance metrics

2.8.1. Detection algorithm

Precision-Recall performance curves are used for analyzing and optimizing detection algorithms to provide a comprehensive evaluation score that considers both false positives and false negatives. The recall is a fundamental measurement for false negatives and indicates a miss rate of 10%. The precision value indicates the percentage of correct detections (thereby measuring the number of false positives) and how many additional incorrect detections. Typically, researchers use Precision-Recall performance curves to evaluate detection algorithms, as they allow them to determine the optimal trade-off between precision and recall for a particular detection method, thereby facilitating the tuning of the model and optimizing its performance. This is particularly important for applications where the cost of missing a positive detection is high or false positives can have serious consequences.

In our detection scenario, a true-positive is defined as the proportion of intersection-over-union (IoU) between a prediction and a matched ground-truth bounding box. For all plots in the results section, we fix the acceptable IoU threshold to be 50% or greater. Non-maximum suppression (NMS) is a prevalent method for filtering duplicate detections by removing substantially overlapping and low-scoring predictions. A high NMS value will eliminate several bounding boxes based on their proportion of overlap area (leading to an increase in precision but a decrease in recall). No receiver operating characteristic (ROC) curve is shown in this research since true negatives are undefined, as that would imply that the algorithm did not identify the animal because there was none.

2.8.2. Image classification

To assess the efficacy of the various techniques, we compute the top-k accuracy on a test set, where $k = 1, 5,$ and 10 represents the position of the correct match (i.e., an annotation of the same individual represented by a query annotation) within a list of proposed matches. Consequently, a top-1 rank is an accurate result supplied by the algorithm as the most probable match for a candidate annotation. A top-5 rating indicates that the correct answer might be any of the top-5 values among the returned options, and so on.

Table 1

Initial data division for machine learning training with PIE.

Set	Individuals	Annotations
Train	91	745
Test	35	84
Total	126	829

3. Results

3.1. Detection algorithm

Precision-Recall performance curves were calculated on a 20% held-out test set (413 annotations) to evaluate the trained snow leopard recognition model's accuracy and report its comprehensive detection success (Fig. 4). The various colors of the curves represent different thresholds of non-maximum suppression (NMS) applied to the network's final bounding box predictions. The Average Precision (AP), as measured by the area under the curve, is a standard approach for summarizing accuracy. For example, the best-performing configuration with an NMS of 30% achieves an AP of 94.45%. If a detector is perfect, it will be represented by the colored dot at the place on the line that is closest to the upper right corner of the precision-recall coordinate system. Moreover, if a fixed recall of 80% is required, the yellow diamond specifies the best possible accuracy across all possible configurations.

The maximum recall values (intercept on the x-axis) on the Precision-Recall curves represent the maximum percentage of annotations that the detector configuration can "recover" or "recall" from the ground-truth detections. Therefore, a recall of 90% signifies that a given detection configuration identified 90% of the ground-truth annotations.

The confusion matrices give the accuracy for the best-colored point (left) and the yellow diamond (right) (Fig. 5). Notably, the 80% recall is an arbitrary number that can be changed to fit the project's performance targets better.

We can see that the best performing and our chosen configuration (highest AP at nearly 95%) has an NMS threshold of 40% and a score threshold of 44% (Fig. 4). Out of 413 total annotations, the overall detector makes 37 errors, with 22 false negatives (not detecting a leopard when one is present), and 15 false positives (a bounding box annotation is generated where there is no animal) (Fig. 5, left plot). If we relax the miss-rate criterion 10%, we make fewer false detections (a total of 9 down from 15), but we end up missing 41 animals (false negatives), and the overall accuracy decreases by over 3% (Fig. 5, right plot).

3.2. Investigation of overfitting

After running an initial set of PIE models, the neural network quickly

converged to its most accurate state, assessing each image only 10–30 times (in machine learning terms, after 10–30 "training epochs"), whereas the typical duration is between 70 and 250 epochs. Subsequent training had no effect or decreased the algorithm's performance on held-out test data, indicating that its long-term behavior resembled memorizing its training set rather than learning a generalized matching strategy (this is the machine learning definition of "overfitting").

Overfitting was studied by comparing the accuracy of matching 14 individuals seen in multiple locations (PIE results: Rank 1–68%, Rank-5 86%) to the initial model consisting of 29 individuals observed at the exact location with the same background (PIE results: Rank 1–77%, Rank-5 93%, and Rank-20 96%) (Fig. 6). The top-1 and top-5 accuracies of the algorithms for snow leopards observed at different locations were lower than the initial modeling attempt, which can be attributed to background matching aiding in the image categorization. The results indicate the importance of background matching in enhancing classification accuracy. After applying the filtering criterion, the number of individuals from multiple locations was nearly half compared to the number of individuals from the exact locations. This leads us to believe that the smaller subset should have had greater accuracy in distinguishing snow leopard individuals successfully.

3.3. Background subtraction

After the model was developed, it was trained on images that had been L/R mirrored and had their backgrounds subtracted. Over-fitting was reduced, and convergence was significantly slowed when the LR parameter was enabled during training. We have speculated that this mirroring could double the number of background textures available for background matching. It is also likely that randomness in the initial configurations of the neural network during each training cycle contributed to the slower convergence. We believe the PIE model with the mirrored and background-subtracted model has theoretical advantages and has shown higher accuracy, so we have chosen it for deployment in the Whiskerbook platform.

3.4. Min-2 accuracies

Thus far, we have only seen results on datasets containing at least

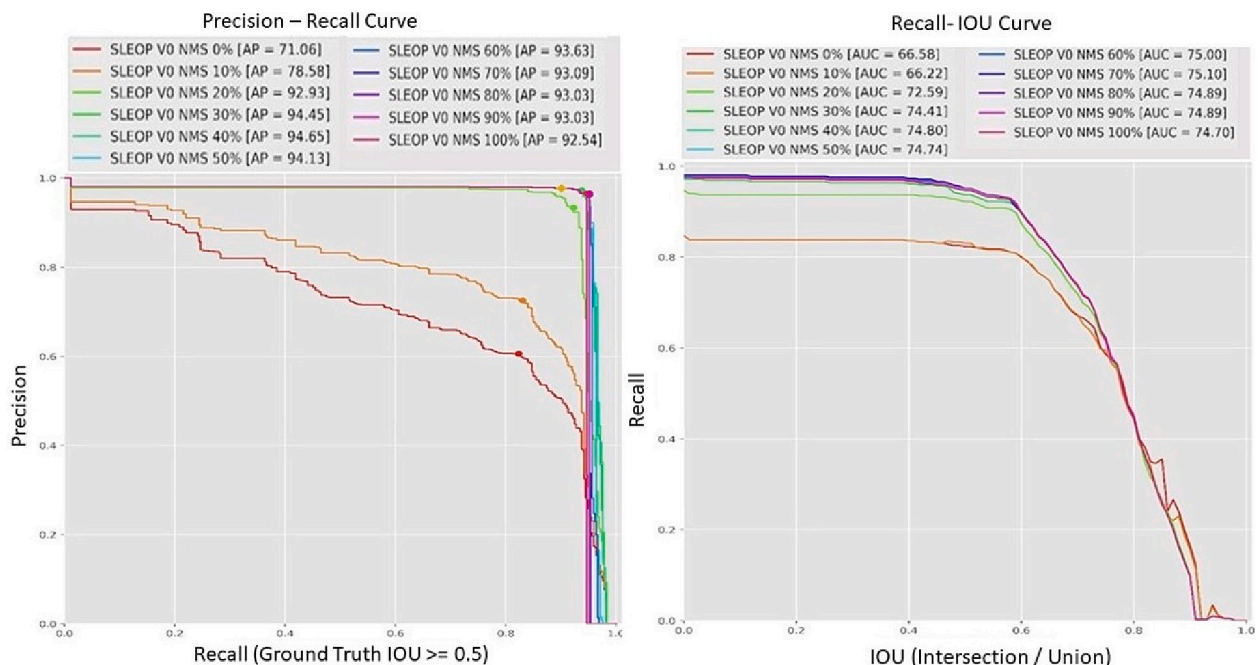


Fig. 4. The detector Precision-Recall curves for snow leopards.

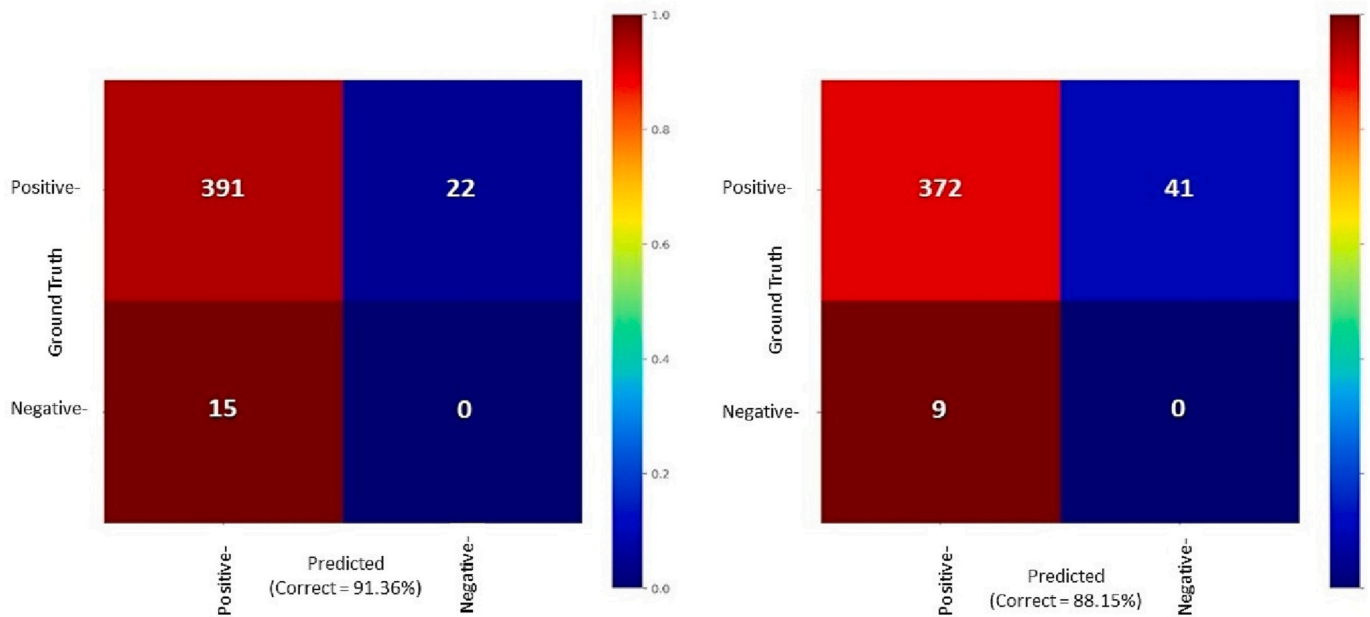


Fig. 5. Confusion matrices for the best-colored point (left) and the yellow diamond (right) from the Precision-Recall performance curves (Fig. 3). Spurious detections, or false positives, occur when a bounding box is formed when there is no snow leopard within it, while false negatives occur when a snow leopard is not detected when it is there in the image. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

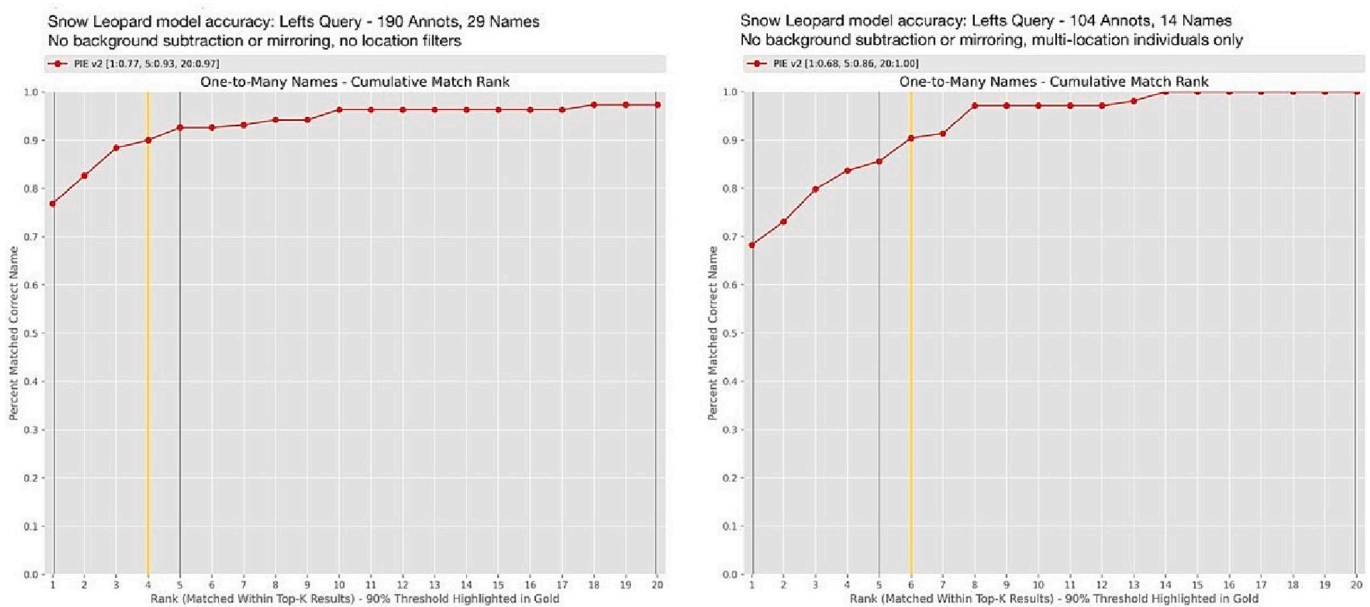


Fig. 6. Accuracy on a PIE model without background subtraction or L/R mirroring. Left is without location filtering; right is only multi-location individuals that moved between trapping stations.

three photographs per individual + viewpoint (e.g., individuals with at least three photos taken from the left). We also computed the accuracy of the more conservative “min-2” filter, which is the bare minimum required for a human reviewer to accomplish identification. This includes the scenario in which the algorithm matches a new animal to a single existing catalog image for the first time. The results sought to classify 40 individuals with PIE BGS-LR (Rank-1 72%, Rank-5 89%, Rank-20 94%), PIE without background subtraction (Rank-1 64%, Rank-5 86%, Rank-20 95%), as compared to Hotspotter (Rank-1 69%, Rank-5 82%, Rank-20 84%). Results showed that the PIE model with background subtraction performed the best on the min-2 side-by-side matching (Fig. 7). Overall, the BGS-LR model performs slightly better

when compared to other models.

3.5. PIE and Hotspotter model accuracy

Since Whiskerbook.org contains a multi-species, multi-feature, and multi-algorithm technical foundation (Blount et al., 2021b), more than one algorithm can be run concurrently when identifying the individual animal in an image. Consequently, we compared the performance of the best PIE model to that of the earlier HotSpotter algorithm (Fig. 8), as well as the accuracy of both algorithms combined (i.e., top-1 of PIE + HotSpotter refers to the proportion of instances in which at least one of the algorithms identified the correct match at the top rank). Combining

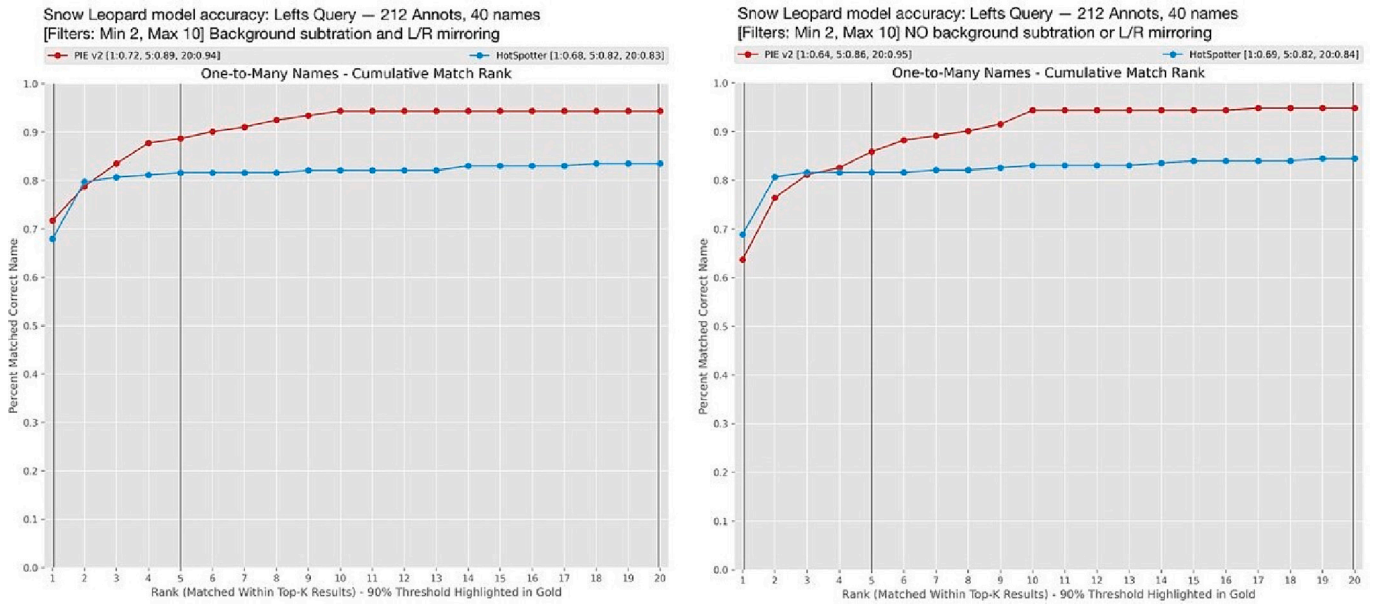


Fig. 7. Comparison of our two PIE models on data with a minimum of two left-side photos per individual. Model on the left used background subtraction and L/R mirroring; the model on the right used neither technique. HotSpotter (HS) shown in both cases; differences in HS accuracies are due to random noise as HotSpotter match scores are not strictly deterministic.

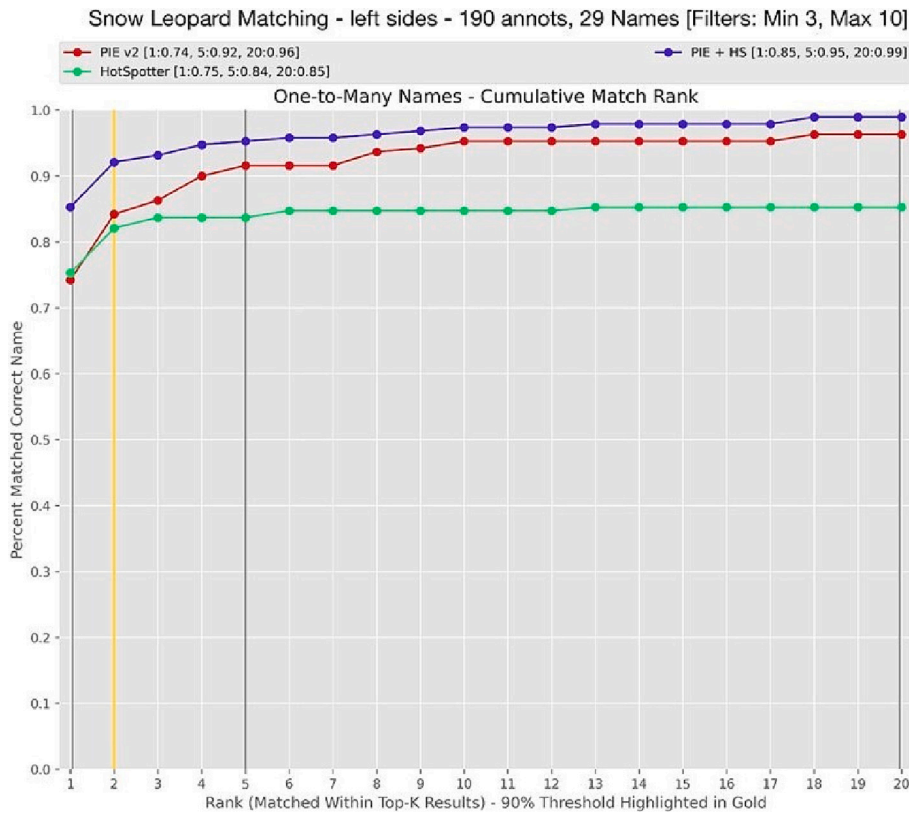


Fig. 8. Accuracy plot of the deployed matching algorithm. Left-side annotations of individuals contain at least three left-side photos. Right-side performance is roughly equivalent (this model mirrors left-side photos, so every animal it sees is from a right-side perspective, but queries are pre-filtered by viewpoint so as not to compare lefts and rights).

the two algorithms substantially enhances the overall accuracy of matchmaking:

- Top-1: 85%
- Top-5: 95%

- Top-20: 99%

There was a thorough examination of the accuracy plots for each model attempt and a comparison of the numerous trained PIE model variants (Table 2). Two things determine the accuracy of an ID

Table 2

Accuracy summary of algorithms on various data filters. BGS stands for “Background Subtraction.” “LR” stands for left-right mirroring. PIE-Vanilla indicates unmodified annotation used in a PIE model. *: NA accuracies are shown where the result would be a trivial 100% because fewer than 20 individuals meet the criteria.

Algorithm	Dataset	Number of annots.	Number of individuals	Top-1 ACC	Top-5 ACC	Top-20 ACC
PIE-BGS-LR	min-2	212	40	72%	89%	94%
PIE-vanilla	min-2	212	40	64%	86%	95%
HotSpotter	min-2	212	40	69%	82%	84%
PIE-BGS-LR	min-3	190	29	74%	92%	96%
PIE-vanilla	min-3	190	29	77%	93%	96%
HotSpotter	min-3	190	29	75%	84%	87%
PIE-BGS-LR + HotSpotter	min-3	190	29	85%	95%	99%
PIE-BGS-LR	min-3, multiloc.	104	14	67%	90%	NA*
PIE-vanilla	min-3, multiloc.	104	14	68%	86%	NA*
HotSpotter	min-3, multiloc.	104	14	72%	79%	85%

algorithm: the algorithm and the data. We tested two PIE models, one trained with background subtraction and L/R mirroring (the final, deployed model, referred to here as “PIE-BGS-LR”), and one without either option (“PIE-vanilla”), as well as HotSpotter. We have evaluated three subsets of the same data, consisting of photographs of snow leopards identified on [Whiskerbook.org](https://www.whiskerbook.org). The subsets are either min-3 or min-2, where the number represents the minimal number of left-side sightings required to include those photographs in the subset (right-side matching behavior is comparable but not shown here). Further, “multiloc” datasets imply that researchers filtered the data only to include species sighted in numerous locations while still meeting the min-X condition.

4. Discussion

Using camera traps, the population size and density of felid species can be assessed; however, processes must be developed to extract individual identities from camera trap imagery reliably. Even though snow leopards can have a diverse array of pelt pattern rosettes, they are famously difficult to reliably classify by human observers (Johansson et al., 2020), machine learning classifiers have shown to be challenging (Beery, 2016; Miguel et al., 2019), leading machine learning tools having been ignored or rejected by the snow leopard researchers for lack of reliability and accuracy (Rode et al., 2021). This project aimed to investigate and evaluate algorithms, on whether they would provide sufficient deep learning algorithmic support for big cat researchers, specifically snow leopard researchers with challenging image-matching issues.

This paper presents the first comprehensive analysis of the accuracy of two separate, production-ready pattern-matching algorithms for snow leopards. Between the two algorithms, PIEv2 (Rank-1 72%, Rank-5 89%) has demonstrated more accuracy in matching these data than HotSpotter, and it is currently the best-performing system we know for recognizing snow leopard individuals in camera trap images. HotSpotter’s findings are marginally less accurate (Rank-1 69%, Rank-5 82%), but their independence from PIE implies that they complement the matching from PIE. Using a combination of PIE and Hotspotter (Rank-1 85%, Rank-5 95%) achieved the highest accuracy of correctly re-identifying individuals in our dataset. The results indicate that the combined algorithms use can significantly enhance researchers’ ability to detect snow leopards in large catalogues rapidly. In a previous study, Hotspotter matching was performed on a series of 102 images of six snow leopards, yielding an accuracy of 70% (Miguel et al., 2019). In our study, combining HotSpotter + PIE algorithms enhances previous efforts at individual ID of snow leopards. Like the previous study, our research revealed a rank-1 accuracy of 69%, indicating that the Hotspotter algorithm alone does not perform well alone.

After applying our sophisticated identification and filtering techniques via graph matching and subset the dataset based on the Min 3/Max 10 strategy, we obtained a total of 829 photographs of 217 snow leopards. Data was collected in the field in Afghanistan and from zoos in

Europe and the United States. Previous studies on snow leopard individual ID used 102 images that contained 6 cats (Miguel et al., 2019); other reports on snow leopard image research mention enormous field datasets without describing the methods for human classification of individuals or for subsetting the data for the final analysis (Beery, 2016; Miguel et al., 2016). Although the dataset used in this work is relatively small, to the best of our knowledge, it is the largest dataset ever collected to test the efficacy of computer vision in snow leopard identification. To assure confidence in the matches being specific individuals, the field collected photos for individuals in this study were compiled utilizing a rigorous graph-based matching procedure of 12,311 side-by-side image inspections before processing. The images were 60% nighttime and 40% daytime illuminated, consistent with the natural activity patterns found for the species in previous studies (Johansson et al., 2022). Our research aimed to compile high-quality image sets of individuals from our field-based data for use in future research on snow leopard individual ID.

Prior research evaluating human manual observers without AI technologies demonstrated substantial error rates when classifying photos of captive snow leopards with known identities from zoos (Johansson et al., 2020). Johansson et al. (2020) revealed that manual observers accurately identified 87.5% of all capture events, demonstrating that misclassification errors would compound to inflate population abundance estimates by a third above the actual population size. An “artificially intelligent software assistant” can assist in the classification process, where our research reveals that the PIE + Hotspotter combined methods are capable of matching with 85% accuracy with the algorithms alone, similar to the abilities of the average manual human observer, 87.5%, reported in previous studies. In a separate study, human manual observers using the [Whiskerbook.org](https://www.whiskerbook.org) platform with AI algorithms demonstrated that all observers (experts and novices) underestimated population by 12% when assisted by the PIE and HotSpotter AI technology, although estimates were precise ~3% error for expert manual observers with experience in performing individual ID (Bohnett et al., 2023). Current research indicates that the performance of deep learning algorithms is still insufficient for solely autonomous matching (Bardier et al., 2020; Morrison et al., 2011; Nipko et al., 2020), necessitating skilled human observers in the loop to perform corrections and achieve a precise level of accuracy (Bohnett et al., 2023; Miao et al., 2021).

One shortcoming identified and addressed within the study showed a limited background matching for the PIE model (Fig. 7 helps visualize the impact), which we assessed by subsetting individuals according to their locations (different locations or same location). We discovered that the algorithm was less accurate for the image subset at different locations, but more accurate at the same location or for the entire dataset, suggesting that the algorithm was aided using the background to match the individuals. As a territorial species, the snow leopard is especially sensitive to context because multiple individuals may be seen in a single camera trap location. Observers have reported two territorial males, two territorial females, and several juveniles at a single camera-trapping location (Bohnett, 2021). Matching several individuals would be most

effective if the system made intelligent identification predictions based only on the animal's natural patterns. However, the model can filter the image and draws intelligent conclusions based on the landscape elements, and it is not necessarily bad if the list of probable individuals may be subset and ranked based on specific backgrounds and geographical information in the image. The study limitations for this small amount of background matching stem from using a relatively limited dataset compiled from zoo data worldwide and field camera trap observations from a set of camera stations in one national park in Afghanistan. Rapid convergence indicated that the "matching problem" was unusually straightforward on these data, typically the result of insufficient training data volume, lack of diversity, or both. To address these concerns, we evaluated a variant of the PIE algorithm with background subtraction and left-right mirroring data augmentation techniques, which reduced the issue of overfitting, and strengthened the algorithms capacity to match individuals based on pelt patterns rather than the background.

The snow leopard training pipeline described here may be repurposed rather quickly to train new models if more users utilize the system and additional data is submitted. The existing model may also help bootstrap that data-curation process. There is a considerable possibility for regional and global research collaborations with snow leopard research institutions to curate and independently identify data that would advance these existing models toward greater sophistication and improved performance. Deep learning systems have the ability to surpass humans in the re-identification of animals, even in the absence of obvious patterns and markings, if sufficient datasets and collaboration come together to develop new systems (Schneider et al., 2019).

While evaluating the detection algorithm to create a bounding box annotation, out of 413 total annotations, there was a 5% error in false positives and 4% error for false negatives. During industry use by human manual observers, the Whiskerbook.org software, running the detection algorithm is a precursor to running Hotspotter or PIE ID models. After the detection algorithm is run, human observers are expected to correct false positive and negative errors manually in the detection system after running the detection algorithm (Bohnett et al., 2023). The cleaning process can simultaneously handle errors produced by duplicate detections of the same individual, new encounters for a second or third snow leopard, and missing annotations (Bohnett et al., 2023). While our study did not formally analyze the impact of humans on data cleaning processes, prior research has shown that these errors can be manually mitigated using a hybrid combination of machine learning and humans in the feedback loop to correct falsely identified annotations (Miao et al., 2021). Users have noted that field camera trap data frequently contain many images of the same individual within an hour, suggesting that the consequences of false negatives are likely to be less significant on a larger dataset (Bohnett, 2021). For instance, the detection algorithm might erroneously not detect a snow leopard within several photographs within a 30-photo encounter spanning 5 min, but the annotated sample for that individual is often still sizable enough to classify the individual. False negatives in field based data result from low-quality catches, such as those that are too far away from the camera trap or are hazy or less defined (Bohnett, 2021). This research demonstrates the importance of human observers verifying annotations within the imagery uploaded to Whiskerbook.org before HotSpotter and PIEv2 algorithms are applied, as there will be a small percentage of both false positives and false negatives generated by the detection annotation algorithm.

We anticipate that the Whiskerbook.org PIE model will expedite snow leopard matching, especially when related Whiskerbook.org features are considered (Bohnett et al., 2023). Existing features already implemented in the Whiskerbook include a location-filtering option (where users can limit a match query to animals spotted in a specific area), one-to-one image matching or one-to-many image matching (for previously classified individuals), side-by-side comparison with HotSpotter or PIE results (Blount et al., 2021b). Whiskerbook's algorithm tools are supplemented by a "visual matcher" interface for manual classification by an observer, enabling easier side-by-side comparisons

of images against each camera trap station's photos taken on different dates (Blount et al., 2021b). These extra capabilities have somewhat compensated for the algorithm's error rate for the detection annotations as well as for individual-level classification, allowing users to manually search through datasets and identify matches precisely.

5. Conclusions

It is essential to accurately reidentify species to generate valid population estimates, which in turn inform numerous critical ecological issues. Whiskerbook.org is a platform for big cat ecologists that incorporates deep learning into an ecological research tool, removing a significant barrier and giving ecologists direct access to deep learning technology for photo reidentification (Berger-Wolf et al., 2017; Blount et al., 2018, 2021b; Bohnett et al., 2023). According to the results of this research, the combination of the PIE and Hotspotter algorithms offers an improvement to the methods on Whiskerbook.org for image reidentification of snow leopards, finding an 85% Rank-1 accuracy for matching individuals. The PIE algorithm performed the best with background subtraction and data augmentation with left-right mirroring, which significantly reduced overfitting. When it comes to animal reidentification studies, a lack of data is the greatest challenge and limitation researchers encounter (Schneider et al., 2019). With the data collection compiled in this work, we contribute the largest fully labeled snow leopard dataset for deep learning research. In addition, we anticipate that our novel deep learning methods will encourage other academics to contribute more datasets to Whiskerbook.org for the purpose of continuously curating and enhancing the data used to improve the deep learning pipeline. Comparing these results to those of an earlier study by Johansson et al. (2020) that depended on human manual classification, without AI or software, determined that observers significantly overestimate the true abundance. AI-based individual ID within Whiskerbook.org has demonstrated the potential to enhance the precision and efficiency of manual observers, approaching more accurate estimations of population size (Bohnett et al., 2023), and may be transferable to the monitoring of populations of other big cat species.

Code availability

All software used in this analysis is available in the Wild Me open-source repository at: <https://github.com/wildmeorg>

The base application for algorithm analysis as defined in Parham et al., 2018 is: <https://github.com/WildMeOrg/wildbook-ia>

Specific algorithm plugins for the three algorithms evaluated here can be found at: <https://github.com/WildMeOrg/wbia-plugin-pie-v2>

Funding

This study was supported by the Global Environment Fund (GEF) and the United Nations Development Programme (UNDP)- grant AA/Pj/PIMS: 00105859/00106885/5844; project 'Conservation of snow leopards and their critical ecosystems in Afghanistan' and the European Union project "Improving participatory management and efficiency of rangeland and watershed focusing on Wakhan, Yakawlang, Kahmard and Sayghan Districts (Contract ACA/2018/399-742)" executed by the Wildlife Conservation Society (WCS) in Afghanistan. The paper's contents are the sole responsibility of the authors and do not necessarily reflect the views of the European Union. The National Science Foundation partially funded this research under the Dynamics of Coupled Natural and Human Systems program [BCS-1826839]. This research also received financial and research support from San Diego State University. Machine learning advancements used in this study were partially funded by a grant from the Gordon and Betty Moore Foundation. A Microsoft Sponsorship supported Azure-based development and model deployment in Whiskerbook.org.

CRediT authorship contribution statement

Eve Bohnett: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Visualization, Writing - original draft, Writing - review & editing. **Jason Holmberg:** Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Resources, Software, Supervision, Validation, Visualization, Writing - original draft. **Sorosh Poya Faryabi:** Conceptualization, Data curation, Project administration, Investigation, Supervision, Writing - review & editing. **Li An:** Funding acquisition, Project administration. **Bilal Ahmad:** Writing - review & editing. **Wajid Rashid:** Writing - review & editing. **Stephane Ostrowski:** Conceptualization, Data curation, Funding acquisition, Project administration, Resources, Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Research-related requests for annotations and data used for ML training in this paper can be requested in COCO format (Lin et al., 2020) via the corresponding author and must be expressly and independently permitted by author Eve Bohnett or through an established collaboration on [Whiskerbook.org](https://whiskerbook.org). Data can also be reviewed and shared via a collaboration request to user Eve Bohnett inside the [Whiskerbook.org](https://whiskerbook.org) system.

Acknowledgements

We thank Wild Me for all their initiative and support for this project: Drew Blount performed the PIE algorithm training and multi-algorithm analyses, and Jason Parham trained the machine learning detector before PIE training. We thank Patrick Thomas, WCS, and Craig Piper, WCS, for sharing photographs of captive snow leopards from Bronx and Central Park zoos, NY, USA. Thank you, Örjan Johansson for supplying images of captive leopards from European Zoos.

References

- Alexander, J.S., Gopalaswamy, A.M., Shi, K., Riordan, P., 2015. Face value: towards robust estimates of snow leopard densities. *PLoS One* 10, e0134815. <https://doi.org/10.1371/journal.pone.0134815>.
- Alexander, J., Zhang, C., Shi, K., Riordan, P., 2016. A granular view of a snow leopard population using camera traps in Central China. *Biol. Conserv.* 197, 27–31. <https://doi.org/10.1016/j.biocon.2016.02.023>.
- Alonso, R.S., McClintock, B.T., Lyren, L.M., Boydston, E.E., Crooks, K.R., 2015. Mark-recapture and mark-resight methods for estimating abundance with remote cameras: a carnivore case study. *PLoS One* 10, e0123032. <https://doi.org/10.1371/journal.pone.0123032>.
- Atzeni, L., Cushman, S.A., Wang, J., Riordan, P., Shi, K., Bauman, D., 2021. Evidence of spatial genetic structure in a snow leopard population from Gansu, China. *Heredity* 127, 522–534. <https://doi.org/10.1038/s41437-021-00483-0>.
- Babae, M., Dinh, D.T., Rigoll, G., 2017. A Deep Convolutional Neural Network for Background Subtraction.
- Bansal, Ms.A., Sharma, Dr.R., Kathuria, Dr.M., 2022. A systematic review on data scarcity problem in deep learning: solution and applications. *ACM Comput. Surv.* 54, 1–29. <https://doi.org/10.1145/3502287>.
- Bardier, C., Székely, D., Augusto-Alves, G., Matínez-Latorraca, N., Schmidt, B.R., Cruickshank, S.S., 2020. Performance of visual vs. software-assisted photo-identification in mark-recapture studies: a case study examining different life stages of the Pacific horned frog (*Ceratophrys stolzmanni*). *Amphib.-Reptilia* 1–12. <https://doi.org/10.1163/15685381-bja10025>.
- Beery, S., 2016. Orientation Invariant Autonomous Recognition of Individual Snow Leopards, p. 9.
- Beery, S., Morris, D., Yang, S., Simon, M., Norouzzadeh, A., Joshi, N., 2019. Efficient pipeline for automating species ID in new camera trap projects. *BISS* 3, e37222. <https://doi.org/10.3897/biss.3.37222>.
- Berger-Wolf, T.Y., Rubenstein, D.L., Stewart, C.V., Holmberg, J.A., Parham, J., Menon, S., Crall, J., Van Oast, J., Kiciman, E., Joppa, L., 2017. Wildbook: Crowdsourcing, Computer Vision, and Data Science for Conservation arXiv:1710.08880 [cs].

- Bhattacharya, S., Sultana, M., Das, B., Roy, B., 2022. A deep neural network framework for detection and identification of bengal tigers. *Innov. Syst. Softw. Eng.* <https://doi.org/10.1007/s11334-021-00431-5>.
- Bian, X., Liang, X., Weckworth, B., Jyal, D., Hull, V., Yang, L., 2022. Spatial density estimate of the snow leopard, *Panthera uncia*, in the Central Tibetan Plateau, China. *Integr. Zool.* <https://doi.org/10.1111/1749-4877.12672>, 1749-4877.12672.
- Blount, D., Minton, G., Khan, C., Levenson, J., Dulau, V., Gero, S., Parham, J., Holmberg, J., 2018. Flukebook – Continuing Growth and Technical Advancement for Cetacean Photo Identification and Data Archiving, Including Automated Fin, Fluke, and Body Matching, p. 13.
- Blount, D., Holmberg, J., Parham, J., Gero, S., Gordon, J., Levenson, J.J., 2021a. Comparison of three individual identification algorithms for sperm whales (*Physeter macrocephalus*) after automated detection (preprint). *Ecology.* <https://doi.org/10.1101/2021.12.22.473895>.
- Blount, D., Parham, J., Holmberg, J., 2021b. Whiskerbook.org.
- Bogucki, R., Cygan, M., Khan, C.B., Klimek, M., Milczek, J.K., Mucha, M., 2019. Applying deep learning to right whale photo identification. *Conserv. Biol.* 33, 676–684. <https://doi.org/10.1111/cobi.13226>.
- Bohnett, E., 2021. Snow Leopard Numbers in the Hindukush Mountain Range of the Wakhan National Park Camera Trap Survey Assessment in Sept 2017–June 2018, and July 2019–October 2019. Wildlife Conservation Society. WCS Afghanistan Program Academic Publications.
- Bohnett, E., Goossens, B., Bakar, M.S.A., Abidin, T.R., Lim, H.-Y., Hulse, D., Ahmad, B., Hocht, T., Gardner, P., 2022. Examining diversity of terrestrial mammal communities across forest reserves in Sabah, Borneo. *Biodivers. Conserv.* 31, 1709–1734. <https://doi.org/10.1007/s10531-022-02423-8>.
- Bohnett, E., Poya Faryabi, S., Lewison, R., An, L., Bian, X., Rajabi, A.M., Jahed, N., Rooyesh, H., Mills, E., Ramos, S., Mesnildrey, N., Santoro Perez, C.M., Taylor, J., Terentyev, V., Ostrowski, S., 2023. Human expertise combined with artificial intelligence improves performance of snow leopard camera trap studies. *Glob. Ecol. Conserv.* 41, e02350 <https://doi.org/10.1016/j.gecco.2022.e02350>.
- Bolger, D., Vance, B., Morrison, T., Farid, H., 2011. WildID User Guide: WildID Camera Trap Image Processing [WWW Document]. URL <https://userguide.wildid.app/>.
- Bolger, D.T., Morrison, T.A., Vance, B., Lee, D., Farid, H., 2012. A computer-assisted system for photographic mark-recapture analysis: software for photographic mark-recapture. *Methods Ecol. Evol.* 3, 813–822. <https://doi.org/10.1111/j.2041-210X.2012.00212.x>.
- Borchers, D., Fewster, R., 2016. Spatial capture–recapture models. *Stat. Sci.* 31, 219–232. <https://doi.org/10.1214/16-STS557>.
- Burgstaller, S., Gollmann, G., Landler, L., 2020. The Green Toad Example: A Comparison of Pattern Recognition Software (Preprint). In Review. <https://doi.org/10.21203/rs.3.rs-76239/v1>.
- Carter, S.J.B., Bell, I.P., Miller, J.J., Gash, P.P., 2014. Automated marine turtle photograph identification using artificial neural networks, with application to green turtles. *J. Exp. Mar. Biol. Ecol.* 452, 105–110. <https://doi.org/10.1016/j.jembe.2013.12.010>.
- Chatfield, K., Simonyan, K., Vedaldi, A., Zisserman, A., 2014. Return of the devil in the details: delving deep into convolutional nets. In: Proceedings of BMVC. <https://doi.org/10.48550/arXiv.1405.3531>.
- Cheema, G.S., Anand, S., 2017. Automatic detection and recognition of individuals in patterned species. In: Altun, Y., Das, K., Mielikäinen, T., Malerba, D., Stefanowski, J., Read, J., Zitnik, M., Ceci, M., Džeroski, S. (Eds.), Machine Learning and Knowledge Discovery in Databases, Lecture Notes in Computer Science. Springer International Publishing, Cham, pp. 27–38. https://doi.org/10.1007/978-3-319-71273-4_3.
- Chehresim, T., Eerola, T., Koivuniemi, M., Auttila, M., Levänen, R., Niemi, M., Kunnasranta, M., Kälviäinen, H., 2018. Automatic individual identification of Saimaa ringed seals. *IET Comput. Vis.* 12, 146–152. <https://doi.org/10.1049/iet-cvi.2017.0082>.
- Chetri, M., Odden, M., Sharma, K., Flagstad, Ø., Wegge, P., 2019. Estimating snow leopard density using fecal DNA in a large landscape in north-Central Nepal. *Glob. Ecol. Conserv.* 17, e00548 <https://doi.org/10.1016/j.gecco.2019.e00548>.
- Choo, Y.R., Kudavidanage, E.P., Amarasinghe, T.R., Nimalrathna, T., Chua, M.A.H., Webb, E.L., 2020. Best practices for reporting individual identification using camera trap photographs. *Glob. Ecol. Conserv.* 24, e01294 <https://doi.org/10.1016/j.gecco.2020.e01294>.
- Christin, S., Hervet, É., Lecomte, N., 2018. Applications for deep learning in ecology (preprint). *Ecology.* <https://doi.org/10.1101/334854>.
- Crall, J.P., Stewart, C.V., Berger-Wolf, T.Y., Rubenstein, D.L., Sundareshan, S.R., 2013. HotSpotter; patterned species instance recognition. In: 2013 IEEE Workshop on Applications of Computer Vision (WACV). Presented at the 2013 IEEE Workshop on Applications of Computer Vision (WACV). IEEE, Clearwater Beach, FL, USA, pp. 230–237. <https://doi.org/10.1109/WACV.2013.6475023>.
- Crouse, D., Jacobs, R.L., Richardson, Z., Klum, S., Jain, A., Baden, A.L., Tecot, S.R., 2017. LemurFaceID: a face recognition system to facilitate individual identification of lemurs. *BMC Zool.* 2, 2. <https://doi.org/10.1186/s40850-016-0011-9>.
- Dawson, J., 2021. Comparisons of image-matching software when identifying pool frog (*Pelophylax lessonae*) individuals from a reintroduced population. *Herpetol. J.* 55–59. <https://doi.org/10.33256/31.1.5559>.
- de Lorm, T., Horswill, C., Rabaiotti, D., Ewers, R., Groom, R., Watermeyer, J., Woodroffe, R., 2023. Optimising the Automated Recognition of Individual Animals to Support Population Monitoring (Preprint). Preprints. <https://doi.org/10.22541/au.167845477.77416758/v1>.
- de Silva, E.M.K., Kumarasinghe, P., Indrajith, K.K.D.A.K., Pushpakumara, T.V., Vimukthi, R.D.Y., de Zoysa, K., Gunawardana, K., de Silva, S., 2022. Feasibility of using convolutional neural networks for individual-identification of wild Asian

- elephants. *Mamm. Biol.* 102, 931–941. <https://doi.org/10.1007/s42991-021-00206-2>.
- Dębicki, I.T., Mittell, E.A., Kristjánsson, B.K., Leblanc, C.A., Morrissey, M.B., Terzić, K., 2021. Re-identification of individuals from images using spot constellations: a case study in Arctic charr (*Salvelinus alpinus*). *R. Soc. Open Sci.* 8, 201768. <https://doi.org/10.1098/rsos.201768>.
- Duhart, C., Dublon, G., Mayton, B., Davenport, G., Paradiso, J.A., 2019. *Deep Learning for Wildlife Conservation and Restoration Efforts*, p. 5.
- Efford, M.G., Borchers, D.L., Byrom, A.E., 2009. Density estimation by spatially explicit capture–recapture: likelihood-based methods. In: Thomson, D.L., Cooch, E.G., Conroy, M.J. (Eds.), *Modeling Demographic Processes in Marked Populations*. Springer US, Boston, MA, pp. 255–269. https://doi.org/10.1007/978-0-387-78151-8_11.
- El Abbadi, N.K., Alsaadi, E.M.T.A., 2020. An automated vertebrate animals classification using deep convolution neural networks. In: 2020 International Conference on Computer Science and Software Engineering (CSASE). Presented at the 2020 International Conference on Computer Science and Software Engineering (CSASE). IEEE, Duhok, Iraq, pp. 72–77. <https://doi.org/10.1109/CSASE48920.2020.9142070>.
- Falzon, G., Lawson, C., Cheung, K.-W., Vernes, K., Ballard, G.A., Fleming, P.J.S., Glen, A. S., Milne, H., Mather-Zardain, A., Meek, P.D., 2019. ClassifyMe: a field-scouting software for the identification of wildlife in camera trap images. *Animals* 10, 58. <https://doi.org/10.3390/ani10010058>.
- Flusser, J., Suk, T., 1993. Pattern recognition by affine moment invariants. *Pattern Recogn.* 26, 167–174. [https://doi.org/10.1016/0031-3203\(93\)90098-H](https://doi.org/10.1016/0031-3203(93)90098-H).
- Foster, R.J., Harmsen, B.J., 2012. A critique of density estimation from camera-trap data: density estimation from camera-trap data. *J. Wildl. Manag.* 76, 224–236. <https://doi.org/10.1002/jwmg.275>.
- Gómez-Vargas, N., Alonso-Fernández, A., Blanquero, R., Antelo, L.T., 2023. Re-identification of fish individuals of undulate skate via deep learning within a few-shot context. *Ecol. Inform.* 102036. <https://doi.org/10.1016/j.ecoinf.2023.102036>.
- Greenberg, S., Godin, T., Whittington, J., 2019. Design patterns for wildlife-related camera trap image analysis. *Ecol. Evol.* 9, 13706–13730. <https://doi.org/10.1002/ece3.5767>.
- Guo, T., Dong, J., Li, H., Gao, Y., 2017. Simple convolutional neural network on image classification. In: 2017 IEEE 27th International Conference on Big Data Analysis (ICBDA). Presented at the 2017 IEEE 27th International Conference on Big Data Analysis (ICBDA). IEEE, Beijing, China, pp. 721–724. <https://doi.org/10.1109/ICBDA.2017.8078730>.
- Guo, S., Xu, P., Miao, Q., Shao, G., Chapman, C.A., Chen, X., He, G., Fang, D., Zhang, H., Sun, Y., Shi, Z., Li, B., 2020. Automatic identification of individual primates with deep learning techniques. *iScience* 23, 101412. <https://doi.org/10.1016/j.isci.2020.101412>.
- Hacker, C., Atzeni, L., Munkhtsog, Bariushaa, Munkhtsog, Bayaraa, Galsandorj, N., Zhang, Yuguang, Liu, Y., Buyanaa, C., Bayandonoi, G., Ochirjav, M., Farrington, J.D., Jevit, M., Zhang, Yu, Wu, L., Cong, W., Li, D., Gavette, C., Jackson, R., Janecka, J.E., 2023. Genetic diversity and spatial structures of snow leopards (*Panthera uncia*) reveal proxies of connectivity across Mongolia and northwestern China. *Landscape Ecol.* 38, 1013–1031. <https://doi.org/10.1007/s10980-022-01573-y>.
- Hansen, M.F., Smith, M.L., Smith, L.N., Salter, M.G., Baxter, E.M., Farish, M., Grieve, B., 2018. Towards on-farm pig face recognition using convolutional neural networks. *Comput. Ind.* 98, 145–152. <https://doi.org/10.1016/j.compind.2018.02.016>.
- Hartog, J., Reijns, R., 2013. About I3S.
- He, Q., Zhao, Q., Liu, N., Chen, P., Zhang, Z., Hou, R., 2019. *Distinguishing Individual Red Pandas from their Faces*.
- Hermans, A., Beyer, L., Leibe, B., 2017. In Defense of the Triplet Loss for Person Re-Identification.
- Howard, A.G., 2013. Some Improvements on Deep Convolutional Neural Network Based Image Classification. *arXiv*. <https://doi.org/10.48550/arXiv.1312.5402>.
- Jackson, R.M., Roe, J.D., Wangchuk, R., Hunter, D.O., 2006. Estimating snow leopard population abundance using photography and capture–recapture techniques. *Wildl. Soc. Bull.* 34, 772–781. [https://doi.org/10.2193/0091-7648\(2006\)34\[772:ESLPAU\]2.0.CO;2](https://doi.org/10.2193/0091-7648(2006)34[772:ESLPAU]2.0.CO;2).
- Janečka, J.E., Munkhtsog, B., Jackson, R.M., Naranbaatar, G., Mallon, D.P., Murphy, W. J., 2011. Comparison of noninvasive genetic and camera-trapping techniques for surveying snow leopards. *J. Mammal.* 92, 771–783. <https://doi.org/10.1644/10-MAMM-A-036.1>.
- Johansson, Ö., Samelius, G., Wikberg, E., Chapron, G., Mishra, C., Low, M., 2020. Identification errors in camera-trap studies result in systematic population overestimation. *Sci. Rep.* 10, 6393. <https://doi.org/10.1038/s41598-020-63367-z>.
- Johansson, Ö., Mishra, C., Chapron, G., Samelius, G., Lkhagvajav, P., McCarthy, T., Low, M., 2022. Seasonal variation in daily activity patterns of snow leopards and their prey. *Sci. Rep.* 12, 21681. <https://doi.org/10.1038/s41598-022-26358-w>.
- Khan, C., Blount, D., Parham, J., Holmberg, J., Hamilton, P., Charlton, C., Christiansen, F., Johnston, D., Rayment, W., Dawson, S., Vermeulen, E., Rowntree, V., Groch, K., Levenson, J.J., Bogucki, R., 2022. Artificial intelligence for right whale photo identification: from data science competition to worldwide collaboration. *Mamm. Biol.* 102, 1025–1042. <https://doi.org/10.1007/s42991-022-00253-3>.
- Korschens, M., Denzler, J., 2019. ELPePants: a fine-grained dataset for elephant re-identification. In: 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW). Presented at the 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW). IEEE, Seoul, Korea (South), pp. 263–270. <https://doi.org/10.1109/ICCVW.2019.00035>.
- Krizhevsky, A., Sutskever, I., Hinton, G.E., 2017. ImageNet classification with deep convolutional neural networks. *Commun. ACM* 60, 84–90. <https://doi.org/10.1145/3065386>.
- Kuncheva, L.L., Garrido-Labrador, J.L., Ramos-Pérez, I., Hennessey, S.L., Rodríguez, J.J., 2023. An experiment on animal re-identification from video. *Ecol. Inform.* 74, 101994. <https://doi.org/10.1016/j.ecoinf.2023.101994>.
- Laguardia, A., Wang, J., Shi, F.-L., Shi, K., Riordan, P., 2015. *Species Identification Refined by Molecular Scatology in a Community of Sympatric Carnivores in Xinjiang, China*. *Zoological Research*, p. 11.
- Langley, I., Hague, E., Civil, M.A., 2021. Assessing the performance of open-source, semi-automated pattern recognition software for harbour seal (*P. v. vitulina*) photo ID. *Mamm. Biol.* <https://doi.org/10.1007/s42991-021-00165-8>.
- LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. *Nature* 521, 436–444. <https://doi.org/10.1038/nature14539>.
- Liu, N., Zhao, Q., Zhang, N., Cheng, X., Zhu, J., 2019. Pose-guided complementary features learning for Amur Tiger re-identification. In: 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW). Presented at the 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW). IEEE, Seoul, Korea (South), pp. 286–293. <https://doi.org/10.1109/ICCVW.2019.00038>.
- Loos, A., 2016. *Face Recognition for Great Apes: Identification of Primates in Real-World Environments*, p. 291.
- Lowe, D.G., 2004. Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vis.* 60, 91–110. <https://doi.org/10.1023/B:VISI.0000029664.99615.94>.
- Lv, J.-J., Shao, X.-H., Huang, J.-S., Zhou, X.-D., Zhou, X., 2017. Data augmentation for face recognition. *Neurocomputing* 230, 184–196. <https://doi.org/10.1016/j.neucom.2016.12.025>.
- Mahadevan, V., Vasconcelos, N., 2008. Background subtraction in highly dynamic scenes. In: 2008 IEEE Conference on Computer Vision and Pattern Recognition. Presented at the 2008 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, Anchorage, AK, USA, pp. 1–6. <https://doi.org/10.1109/CVPR.2008.4587576>.
- Masi, I., Trän, A.T., Hassner, T., Leksut, J.T., Medioni, G., 2016. Do we really need to collect millions of faces for effective face recognition? In: Leibe, B., Matas, J., Sebe, N., Welling, M. (Eds.), *Computer Vision – ECCV 2016, Lecture Notes in Computer Science*. Springer International Publishing, Cham, pp. 579–596. https://doi.org/10.1007/978-3-319-46454-1_35.
- Matthé, M., 2015. About AmphibiIdent [WWW Document]. URL. <http://www.amphibident.de/en/pages/about.html>.
- 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. *Ecol. Evol.* 7, 5861–5872. <https://doi.org/10.1002/ece3.3140>.
- McCarthy, T., Mallon, D., Nyhus, P.J. (Eds.), 2016. *Snow Leopards, Biodiversity of the World: Conservation from Genes to Landscapes*. Academic Press is an imprint of Elsevier, London; San Diego, CA.
- Miao, Z., Gaynor, K.M., Wang, J., Liu, Z., Muellerklein, O., Norouzzadeh, M.S., McInturff, A., Bowie, R.C.K., Nathan, R., Yu, S.X., Getz, W.M., 2019. Insights and approaches using deep learning to classify wildlife. *Sci. Rep.* 9, 8137. <https://doi.org/10.1038/s41598-019-44565-w>.
- Miao, Z., Liu, Z., Gaynor, K.M., Palmer, M.S., Yu, S.X., Getz, W.M., 2021. Iterative human and automated identification of wildlife images. *Nat. Mach. Intell.* 3, 885–895. <https://doi.org/10.1038/s42256-021-00393-0>.
- Miguel, A., Beery, S., Flores, E., Klemesrud, L., Bayrakcismith, R., 2016. Finding areas of motion in camera trap images. In: 2016 IEEE International Conference on Image Processing (ICIP). Presented at the 2016 IEEE International Conference on Image Processing (ICIP). IEEE, Phoenix, AZ, USA, pp. 1334–1338. <https://doi.org/10.1109/ICIP.2016.7532575>.
- 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: Mao, K., Jiang, X. (Eds.), *Tenth International Conference on Signal Processing Systems*. Presented at the Tenth International Conference on Signal Processing Systems. SPIE, Singapore, Singapore, p. 36. <https://doi.org/10.1117/12.2521856>.
- Mikolajczyk, A., Grochowski, M., 2018. Data augmentation for improving deep learning in image classification problem. In: 2018 International Interdisciplinary PhD Workshop (IIPHDW). Presented at the 2018 International Interdisciplinary PhD Workshop (IIPHDW). IEEE, Swinoujście, pp. 117–122. <https://doi.org/10.1109/IIPHDW.2018.8388338>.
- Morrison, T.A., Yoshizaki, J., Nichols, J.D., Bolger, D.T., 2011. Estimating survival in photographic capture–recapture studies: overcoming misidentification error: unbiased survival estimation in photograph-ID. *Methods Ecol. Evol.* 2, 454–463. <https://doi.org/10.1111/j.2041-210X.2011.00106.x>.
- Moskvyak, O., Maire, F., Armstrong, A.O., Dayoub, F., Baktashmotlagh, M., 2019. *Robust Re-Identification of Manta Rays from Natural Markings by Learning Pose Invariant Embeddings* arXiv:1902.10847 [cs].
- Moskvyak, O., Maire, F., Dayoub, F., Baktashmotlagh, M., 2020a. *Keypoint-Aligned Embeddings for Image Retrieval and Re-Identification*.
- Moskvyak, O., Maire, F., Dayoub, F., Baktashmotlagh, M., 2020b. *Learning Landmark Guided Embeddings for Animal Re-Identification*.
- Moskvyak, O., Maire, F., Dayoub, F., Baktashmotlagh, M., 2021. *Going Deeper into Semi-Supervised Person Re-Identification*.
- Moya, Ó., Mansilla, P.-L., Madrazo, S., Igual, J.-M., Rotger, A., Romano, A., Tavecchia, G., 2015. APHIS: a new software for photo-matching in ecological studies. *Ecol. Inform.* 27, 64–70. <https://doi.org/10.1016/j.ecoinf.2015.03.003>.
- Nepovinykh, E., Eerola, T., Kalviainen, H., 2020. Siamese network based pelage pattern matching for ringed seal re-identification. In: 2020 IEEE Winter Applications of

- Computer Vision Workshops (WACVW). Presented at the 2020 IEEE Winter Applications of Computer Vision Workshops (WACVW). IEEE, Snowmass Village, CO, USA, pp. 25–34. <https://doi.org/10.1109/WACVW50321.2020.9096935>.
- Nepovinnyykh, E., Chelak, I., Eerola, T., Kälviäinen, H., 2022a. NORPPA: NOvel Ringed Seal Re-Identification by Pelage Pattern Aggregation.
- Nepovinnyykh, E., Eerola, T., Biard, V., Mutka, P., Niemi, M., Kunnasranta, M., Kälviäinen, H., 2022b. SealID: Saimaa ringed seal re-identification dataset. *Sensors* 22, 7602. <https://doi.org/10.3390/s22197602>.
- Nguyen, H., Maclagan, S.J., Nguyen, T.D., Nguyen, T., Flemons, P., Andrews, K., Ritchie, E.G., Phung, D., 2017a. Animal recognition and identification with deep convolutional neural networks for automated wildlife monitoring. In: 2017 IEEE International Conference on Data Science and Advanced Analytics (DSAA). Presented at the 2017 IEEE International Conference on Data Science and Advanced Analytics (DSAA). IEEE, Tokyo, pp. 40–49. <https://doi.org/10.1109/DSAA.2017.31>.
- Nguyen, H., Maclagan, S.J., Nguyen, T.D., Nguyen, T., Flemons, P., Andrews, K., Ritchie, E.G., Phung, D., 2017b. Animal recognition and identification with deep convolutional neural networks for automated wildlife monitoring. In: 2017 IEEE International Conference on Data Science and Advanced Analytics (DSAA). Presented at the 2017 IEEE International Conference on Data Science and Advanced Analytics (DSAA). IEEE, Tokyo, pp. 40–49. <https://doi.org/10.1109/DSAA.2017.31>.
- Nipko, R.B., Holcombe, B.E., Kelly, M.J., 2020. Identifying individual jaguars and ocelots via pattern-recognition software: comparing HotSpotter and wild-ID. *Wildl. Soc. Bull.* 44, 424–433. <https://doi.org/10.1002/wsb.1086>.
- Norouzzadeh, M.S., Nguyen, A., Kosmala, M., Swanson, A., Palmer, M.S., Packer, C., Clune, J., 2018a. Automatically identifying, counting, and describing wild animals in camera-trap images with deep learning. *Proc. Natl. Acad. Sci. U. S. A.* 115 <https://doi.org/10.1073/pnas.1719367115>.
- Norouzzadeh, M.S., Nguyen, A., Kosmala, M., Swanson, A., Palmer, M.S., Packer, C., Clune, J., 2018b. Automatically identifying, counting, and describing wild animals in camera-trap images with deep learning. *Proc. Natl. Acad. Sci. U. S. A.* 115 <https://doi.org/10.1073/pnas.1719367115>.
- Norouzzadeh, M.S., Morris, D., Beery, S., Joshi, N., Jovic, N., Clune, J., 2019. A Deep Active Learning System for Species Identification and Counting in Camera Trap Images. *arXiv:1910.09716* [cs, eess, stat].
- Oberosler, V., Tenan, S., Groff, C., Krofel, M., Augugliaro, C., Munkhtsog, B., Rovero, F., 2022. First spatially-explicit density estimate for a snow leopard population in the Altai Mountains. *Biodivers. Conserv.* 31, 261–275. <https://doi.org/10.1007/s10531-021-02333-1>.
- Papafitsoros, K., Adam, L., Čermák, V., Píček, L., 2023. SeaTurtleID: A Novel Long-Span Dataset Highlighting the Importance of Timestamps in Wildlife Re-Identification.
- Parham, J., Stewart, C., Crall, J., Rubenstein, D., Holmberg, J., Berger-Wolf, T., 2018. An animal detection pipeline for identification. In: 2018 IEEE Winter Conference on Applications of Computer Vision (WACV). Presented at the 2018 IEEE Winter Conference on Applications of Computer Vision (WACV). IEEE, Lake Tahoe, NV, pp. 1075–1083. <https://doi.org/10.1109/WACV.2018.00123>.
- 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. *J. Ecol. Environ.* 43, 39. <https://doi.org/10.1186/s41610-019-0138-z>.
- Pereira, K.S., Gibson, L., Biggs, D., Samarasinghe, D., Braczkowski, A.R., 2022. Individual identification of large felids in field studies: common methods, challenges, and implications for conservation science. *Front. Ecol. Evol.* 10, 866403 <https://doi.org/10.3389/fevo.2022.866403>.
- Petso, T., Jamisola, R.S., Mpoeleng, D., 2022. Review on methods used for wildlife species and individual identification. *Eur. J. Wildl. Res.* 68, 3. <https://doi.org/10.1007/s10344-021-01549-4>.
- Rashid, W., Shi, J., Rahim, I. ur, Qasim, M., Baloch, M.N., Bohnett, E., Yang, F., Khan, I., Ahmad, B., 2021. Modelling potential distribution of snow leopards in Pamir, Northern Pakistan: implications for human–snow leopard conflicts. *Sustainability* 13, 13229. <https://doi.org/10.3390/su132313229>.
- Ravoor, P.C., Sudarshan, T.S.B., 2020. Deep learning methods for multi-species animal re-identification and tracking – a survey. *Comput. Sci. Rev.* 38, 100289 <https://doi.org/10.1016/j.cosrev.2020.100289>.
- Redmon, J., Farhadi, A., 2016. YOLO9000: Better, Faster, Stronger *arXiv:1612.08242* [cs].
- Rode, J., Lambert, C., Marescot, L., Chaix, B., Beesau, J., Bastian, S., Kyrbashev, J., Cabanat, A.-L., 2021. Population monitoring of snow leopards using camera trapping in Naryn state reserve, Kyrgyzstan, between 2016 and 2019 (preprint). *Ecology*. <https://doi.org/10.1101/2021.02.24.432722>.
- Royle, J.A., Young, K.V., 2008. A hierarchical model for spatial capture-recapture data. *Ecology* 89, 2281–2289. <https://doi.org/10.1890/07-0601.1>.
- Royle, J.A., Nichols, J.D., Karanth, K.U., Gopalaswamy, A.M., 2009. A hierarchical model for estimating density in camera-trap studies. *J. Appl. Ecol.* 46, 118–127. <https://doi.org/10.1111/j.1365-2664.2008.01578.x>.
- Schneider, S., Taylor, G.W., Linnquist, S., Kremer, S.C., 2019. Past, present and future approaches using computer vision for animal re-identification from camera trap data. *Methods Ecol. Evol.* 10, 461–470. <https://doi.org/10.1111/2041-210X.13133>.
- Schneider, S., Greenberg, S., Taylor, G.W., Kremer, S.C., 2020a. Three critical factors affecting automated image species recognition performance for camera traps. *Ecol. Evol.* 10, 3503–3517. <https://doi.org/10.1002/ece3.6147>.
- Schneider, S., Taylor, G.W., Kremer, S.C., 2020b. Similarity learning networks for animal individual re-identification - beyond the capabilities of a human observer. In: 2020 IEEE Winter Applications of Computer Vision Workshops (WACVW). Presented at the 2020 IEEE Winter Applications of Computer Vision Workshops (WACVW). IEEE, Snowmass, CO, USA, pp. 44–52. <https://doi.org/10.1109/WACVW50321.2020.9096925>.
- Schofield, D., Nagrani, A., Zisserman, A., Hayashi, M., Matsuzawa, T., Biro, D., Carvalho, S., 2019. Chimpanzee face recognition from videos in the wild using deep learning. *Sci. Adv.* 5 <https://doi.org/10.1126/sciadv.aaw0736> eaaw0736.
- Sharma, R.K., Sharma, K., Borchers, D., Bhatnagar, Y.V., Suryawanshi, K.S., Mishra, C., 2020. Spatial variation in population-density, movement and detectability of snow leopards in a multiple use landscape in Spiti Valley, Trans-Himalaya (preprint). *Ecology*. <https://doi.org/10.1101/2020.09.09.289181>.
- Shi, C., Liu, D., Cui, Y., Xie, J., Roberts, N.J., Jiang, G., 2020. Amur Tiger stripes: individual identification based on deep convolutional neural network. *Integr. Zool.* <https://doi.org/10.1111/1749-4877.12453>.
- Shorten, C., Khoshgoftaar, T.M., 2019. A survey on image data augmentation for deep learning. *J. Big Data* 6, 60. <https://doi.org/10.1186/s40537-019-0197-0>.
- Shukla, A., Anderson, C., Sigh Cheema, G., Gao, P., Onda, S., Anshuman, D., Anand, S., Farrell, R., 2019. A hybrid approach to Tiger re-identification. In: 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW). Presented at the 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW). IEEE, Seoul, Korea (South), pp. 294–301. <https://doi.org/10.1109/ICCVW.2019.00039>.
- Sobral, H., Vacavant, A., 2014. A comprehensive review of background subtraction algorithms evaluated with synthetic and real videos. *Comput. Vis. Image Underst.* 122, 4–21. <https://doi.org/10.1016/j.cviu.2013.12.005>.
- Stewart, C.V., Parham, J.R., Holmberg, J., Berger-Wolf, T.Y., 2021. The Animal ID Problem: Continual Curation.
- Strampelli, P., Andresen, L., Everatt, K.T., Somers, M.J., Rowcliffe, J.M., 2020. Leopard *Panthera pardus* density in southern Mozambique: evidence from spatially explicit capture–recapture in Xonghile game reserve. *Oryx* 54, 405–411. <https://doi.org/10.1017/S0030605318000121>.
- Sultan, H., Rashid, W., Shi, J., Rahim, I. ur, Nafees, M., Bohnett, E., Rashid, S., Khan, M. T., Shah, I.A., Han, H., Ariza-Montes, A., 2022. Horizon scan of transboundary concerns impacting Snow leopard landscapes in Asia. *Land* 11, 248. <https://doi.org/10.3390/land11020248>.
- Suryawanshi, K.R., Khanyari, M., Sharma, K., Lkhagvajav, P., Mishra, C., 2019. Sampling bias in snow leopard population estimation studies. *Popul. Ecol.* 61, 268–276. <https://doi.org/10.1002/1438-390X.1027>.
- Swanson, A., Kosmala, M., Lintott, C., Simpson, R., Smith, A., Packer, C., 2015. Snapshot Serengeti, high-frequency annotated camera trap images of 40 mammalian species in an African savanna. *Sci. Data* 2, 150026. <https://doi.org/10.1038/sdata.2015.26>.
- Tabak, M.A., Norouzzadeh, M.S., Wolfson, D.W., Sweeney, S.J., Vercauteren, K.C., Snow, N.P., Halseth, J.M., Di Salvo, P.A., Lewis, J.S., White, M.D., Teton, B., Beasley, J.C., Schlichting, P.E., Boughton, R.K., Wight, B., Newkirk, E.S., Ivan, J.S., Odell, E.A., Brook, R.K., Lukacs, P.M., Moeller, A.K., Mandeville, E.G., Clune, J., Miller, R.S., 2019. Machine learning to classify animal species in camera trap images: applications in ecology. *Methods Ecol. Evol.* 10, 585–590. <https://doi.org/10.1111/2041-210X.13120>.
- Tabak, M.A., Norouzzadeh, M.S., Wolfson, D.W., Newton, E.J., Boughton, R.K., Ivan, J.S., Odell, E.A., Newkirk, E.S., Conrey, R.Y., Stenglein, J., Iannarilli, F., Erb, J., Brook, R. K., Davis, A.J., Lewis, J., Walsh, D.P., Beasley, J.C., Vercauteren, K.C., Clune, J., Miller, R.S., 2020. Improving the accessibility and transferability of machine learning algorithms for identification of animals in camera trap images: MLWIC2. *Ecol. Evol.* 10, 10374–10383. <https://doi.org/10.1002/ece3.6692>.
- Tariq, N., Saleem, K., Mushtaq, M., Nawaz, M.A., 2018. Snow leopard recognition using deep convolution neural network. In: Proceedings of the 2nd International Conference on Information System and Data Mining - ICISDM '18. Presented at the 2nd International Conference. ACM Press, Lakeland, FL, USA, pp. 29–33. <https://doi.org/10.1145/3206098.3206114>.
- Valdez, F.P., Haag, T., Azevedo, F.C.C., Silveira, L., Cavalcanti, S.M.C., Salzano, F.M., Eizirik, E., 2015. Population genetics of jaguars (*Panthera onca*) in the Brazilian Pantanal: molecular evidence for demographic connectivity on a regional scale. *J. Hered.* 106, 503–511. <https://doi.org/10.1093/jhered/evs046>.
- Verschueren, S., Fabiano, E.C., Kakove, M., Cristescu, B., Marker, L., 2023. Reducing identification errors of African carnivores from photographs through computer-assisted workflow. *Mamm. Res.* 68, 121–125. <https://doi.org/10.1007/s13364-022-00657-z>.
- Vidal, M., Wolf, N., Rosenberg, B., Harris, B.P., Mathis, A., 2021. Perspectives on individual animal identification from biology and computer vision. *Integr. Comp. Biol.* 61, 900–916. <https://doi.org/10.1093/icb/icab107>.
- Wäldchen, J., Mäder, P., 2018. Machine learning for image based species identification. *Methods Ecol. Evol.* 9, 2216–2225. <https://doi.org/10.1111/2041-210X.13075>.
- Wang, M., Deng, W., 2021. Deep face recognition: a survey. *Neurocomputing* 429, 215–244. <https://doi.org/10.1016/j.neucom.2020.10.081>.
- Weckworth, B., 2021. Snow leopard (*Panthera uncia*) genetics: the knowledge gaps, needs, and implications for conservation. *J. Indian Inst. Sci.* 101, 279–290. <https://doi.org/10.1007/s41745-021-00236-2>.
- Wegge, P., Bakke, B.B., Odden, M., Rolstad, J., 2019. DNA from scats combined with capture–recapture modeling: a promising tool for estimating the density of red foxes—a pilot study in a boreal forest in Southeast Norway. *Mamm. Res.* 64, 147–154. <https://doi.org/10.1007/s13364-018-0408-7>.
- Weinstein, B.G., 2018. A computer vision for animal ecology. *J. Anim. Ecol.* 87, 533–545. <https://doi.org/10.1111/1365-2656.12780>.
- Willi, M., Pitman, R.T., Cardoso, A.W., Locke, C., Swanson, A., Boyer, A., Veldthuis, M., Fortson, L., 2019. Identifying animal species in camera trap images using deep learning and citizen science. *Methods Ecol. Evol.* 10, 80–91. <https://doi.org/10.1111/2041-210X.13099>.
- Xie, S., Tu, Z., 2015. Holistically-Nested Edge Detection. *arXiv:1504.06375*.

- Xie, S., Girshick, R., Dollár, P., Tu, Z., He, K., 2017. Aggregated Residual Transformations for Deep Neural Networks.
- Yang, H., Patras, I., 2015. Mirror, Mirror on the Wall, Tell Me, is the Error Small? arXiv. <https://doi.org/10.48550/arXiv.1501.05152>.
- Yousif, H., Yuan, J., Kays, R., He, Z., 2017. Fast human-animal detection from highly cluttered camera-trap images using joint background modeling and deep learning classification. In: 2017 IEEE International Symposium on Circuits and Systems (ISCAS). Presented at the 2017 IEEE International Symposium on Circuits and Systems (ISCAS). IEEE, Baltimore, MD, USA, pp. 1–4. <https://doi.org/10.1109/ISCAS.2017.8050762>.
- Yu, X., Wang, J., Kays, R., Jansen, P.A., Wang, T., Huang, T., 2013. Automated identification of animal species in camera trap images. J. Image Video Process. 2013, 52. <https://doi.org/10.1186/1687-5281-2013-52>.
- Zhang, L., Lian, X., Yang, X., 2020. Population density of snow leopards (*Panthera uncia*) in the Yage Valley region of the Sanjiangyuan National Park: conservation implications and future directions. Arct. Antarct. Alp. Res. 52, 541–550. <https://doi.org/10.1080/15230430.2020.1816341>.
- Zheng, X., Owen, M.A., Nie, Y., Hu, Y., Swaisgood, R.R., Yan, L., Wei, F., 2016a. Individual identification of wild giant pandas from camera trap photos - a systematic and hierarchical approach. J. Zool. 300, 247–256. <https://doi.org/10.1111/jzo.12377>.
- Zheng, L., Yang, Y., Hauptmann, A.G., 2016b. Person Re-Identification: Past, Present and Future.
- Zheng, L., Huang, Y., Lu, H., Yang, Y., 2017. Pose Invariant Embedding for Deep Person Re-Identification.
- Zuerl, M., Dirauf, R., Koeferl, F., Steinlein, N., Sueskind, J., Zanca, D., Brehm, I., von Fersen, L., Eskofier, B., 2023. PolarBearVidID: a video-based re-identification benchmark dataset for polar bears. Animals 13, 801. <https://doi.org/10.3390/ani13050801>.