


The Future of Artificial Intelligence in Monitoring Animal Identification, Health, and Behaviour

Jenna V. Congdon ^{1,2,*} , Mina Hosseini ^{1,3}, Ezekiel F. Gading ¹, Mahdi Masousi ³, Maria Franke ⁴ and Suzanne E. MacDonald ¹

¹ Department of Psychology, Faculty of Health, York University, Toronto, ON M3J 1P3, Canada; minah91@yorku.ca (M.H.); efgading@yorku.ca (E.F.G.); suzmac@yorku.ca (S.E.M.)

² Toronto Zoo Wildlife Conservancy, Toronto Zoo, Toronto, ON M1B 5K7, Canada

³ EAIGLE, Markham, ON L3R 9Z7, Canada; marsousi@eaigle.com

⁴ Toronto Zoo, Toronto, ON M1B 5K7, Canada; mfranke@torontozoo.ca

* Correspondence: jcongdon@yorku.ca; Tel.: +1-587-873-9605

Simple Summary: Due to climate change and human interference, many species are now without habitats and on the brink of extinction. Zoos and other conservation spaces allow for non-human animal preservation and public education about endangered species and ecosystems. Monitoring the health and well-being of animals in care, while providing species-specific environments, is critical for zoo and conservation staff. In order to best provide such care, keepers and researchers need to gather as much information as possible about individual animals and species as a whole. This paper focuses on existing technology to monitor animals, providing a review on the history of technology, including recent technological advancements and current limitations. Subsequently, we provide a brief introduction to our proposed novel computer software: an artificial intelligence software capable of unobtrusively and non-invasively tracking individuals' location, estimating position, and analyzing behaviour. This innovative technology is currently being trained with orangutans at the Toronto Zoo and will allow for mass data collection, permitting keepers and researchers to closely monitor individual animal welfare, learn about the variables impacting behaviour and provide additional enrichment or interventions accordingly.

Abstract: With many advancements, technologies are now capable of recording non-human animals' location, heart rate, and movement, often using a device that is physically attached to the monitored animals. However, to our knowledge, there is currently no technology that is able to do this unobtrusively and non-invasively. Here, we review the history of technology for use with animals, recent technological advancements, current limitations, and a brief introduction to our proposed novel software. Canadian tech mogul EAIGLE Inc. has developed an artificial intelligence (AI) software solution capable of determining where people and assets are within public places or attractions for operational intelligence, security, and health and safety applications. The solution also monitors individual temperatures to reduce the potential spread of COVID-19. This technology has been adapted for use at the Toronto Zoo, initiated with a focus on Sumatran orangutans (*Pongo abelii*) given the close physical similarity between orangutans and humans as great ape species. This technology will be capable of mass data collection, individual identification, pose estimation, behaviour monitoring and tracking orangutans' locations, in real time on a 24/7 basis, benefitting both zookeepers and researchers looking to review this information.

Keywords: artificial intelligence; orangutan; *Pongo abelii*; animal behaviour; animal health; identification; monitoring; conservation



Citation: Congdon, J.V.; Hosseini, M.; Gading, E.F.; Masousi, M.; Franke, M.; MacDonald, S.E. The Future of Artificial Intelligence in Monitoring Animal Identification, Health, and Behaviour. *Animals* **2022**, *12*, 1711. <https://doi.org/10.3390/ani12131711>

Academic Editors: Ori Pomerantz and Kristine Coleman

Received: 21 May 2022

Accepted: 28 June 2022

Published: 1 July 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

1.1. The History of Technology

1.1.1. Chronological Background

Over time, technology for monitoring both human and non-human lives has evolved and advanced. Dating back to the 1970s, collars with electronic transponders were attached to cows to automatically record data [1,2]; since then, technology has become smaller and more affordable, with a focus on improving overall productivity in agriculture [3,4]. These types of devices are now less intrusive and commonplace in the home, allowing humans to monitor their pet companions (e.g., smart home cameras) [5]. Although contemporary technology has a wide range of uses, the focus of this paper is on zoo management to enhance non-human animal welfare. We provide a brief overview regarding the types of hard- and software currently available and implemented in both the wild and captivity, and address remaining gaps in zoo-focused technology.

Overall, technology has been critical to obtaining large amounts of data, greatly reducing the labour-intensive and invasive aspects of data collection. Not only has technology reduced time in conducting research, but it has allowed for more accurate collection compared to that of human observers. For example, when Desormeaux and colleagues [6] compared the information collected from motion triggered cameras located at migratory fence gaps to observers' recordings of mammal tracks (e.g., giraffe, zebra, elephant, and hyaena), a higher volume of crossing was reported with technology, suggesting that human observers may often miss important events. Methodologies that use manual collection of data are often time consuming, time limited, and potentially invasive and/or ill-suited depending on the research focus, supporting further investment in the advancement and refinement of technology-based alternatives.

1.1.2. Locations of Employed Technology

In addition to footprints used to collect an estimate of the number of animals in a particular geographical location, to further gather information from wild animals researchers have collected hair, urine, fecal matter, etc. (e.g., [7–9]). Although minimally invasive or noninvasive, these types of data collection can be tedious and limited, requiring consideration of both accuracy and potential confounding variables in interpreting results. As a result, a wide range of technology has been employed, often in wild settings. For example, accelerometers have been extensively employed because they are inexpensive, animal-attached tags used to remotely gain positional data and 3D reconstructions of behaviour of often “unwatchable” animals [10,11]. Visibility issues can also be combated with another form of technology: drones. The use of remote-controlled or software-controlled flying robots have revolutionized monitoring of many marine species, including sharks [12], rays [13], dolphins [14], and seabirds [15]. Additional forms of technology include radio-tracking and GPS collars, cameras, and artificial intelligence, used to detect lions (*Panthera leo*) [16], lynx (*Lynx lynx*) [17], giant pandas (*Ailuropoda melanoleuca*) [18], wild red fox (*Vulpes vulpes*) [19], and a range of primate species including blond capuchins (*Sapajus flavius*) [20], baboons (*Papio ursinus*) [11], and chimpanzees (*Pan troglodytes*) [21]. These projects have tracked physical locations of animals, with some capable of gathering information from individual animals, including behavioural data; i.e., face detection to record locomotor behaviour [16], camera traps to capture a range of daily activities [19], and tracking collars to quantify a range of daily activities [11].

However, there are many situations in which animals live in captivity under human care. As mentioned, when considering the development of technologies across the globe, many of the advancements have been produced for agricultural applications, which have clear economic advantages for humans. In captive settings outside of agriculture, monitoring vulnerable individuals (e.g., endangered or injured) is not only critical to ensure the health and well-being of such animals, but it can provide information to inform greater animal management decisions. The type of data researchers collect from wild animals is quite different from data collected from animals in captivity as the locations of captive

animals are limited within relatively unchanging and predictable environments, and the individuals are known and remain in place across months, years, or decades. Thus, captive animal data are typically more welfare related, useful in informing animal husbandry and management decisions. Accredited zoos, aquariums, and conservation areas employ the best practices to monitor their animals and obtain necessary behavioural data however possible. For example, the low-cost “ZooMonitor” application has allowed for the tracking of pygmy hippos (*Choeropsis liberiensis*) and domestic chickens (*Gallus gallus*), providing data on physical appearance, habitat use, and behaviour within their enclosures [22]. Formerly expensive, intrusive, and inaccurate technologies have also recently been replaced by software and complex algorithms; for example, deep neural networks are computer technologies that have “learned” to detect and recognize individual giant pandas [23]. (Note: These neural networks are described in more detail in Section 1.2.2)

1.2. Recent Technological Advancements

1.2.1. Purposes for Technological Advances

Across the many countries that technology has been developed in order to monitor animals, the variety of purposes for implementation include monitoring health (e.g., [24,25]), differentiating between individual animals (e.g., [18,26]), and behaviour (e.g., [11,19,27]). A focus on health is the primary rationale for much of the technological advances in agriculture, given the focus on economic outcomes. Variables such as head motion, core body temperature, and heart rate [28], as well as surrounding temperature and humidity [24] are crucial for agricultural production. For example, in cattle, stress due to warm weather can influence vulnerability to disease, food consumption, weight gain, reproduction, and milk production [24,25]. Thus, in being able to track individual animals, technology allows for the monitoring of changes in livestock [3]. Monitoring individual animals is also integral to generating inferences about spatial ecology. For example, GPS tracking collars have been critical to mapping the urban movement patterns of raccoons [29], while camera traps have successfully tracked individual pandas amongst dense bamboo, eventually to be used to estimate population sizes [18]. Automated facial recognition has also been developed to identify individual wild gorillas, useful for evaluating spatial biodiversity in the wild [26]. In addition to monitoring health and differentiating between individuals, technology has advanced to be able record the actions of animals.

Only recently has technology been able to monitor animal behaviour as it is a complicated process; not only does the technology have to be able to recognize the individual from their environment, but must be capable of interpreting similar patterns of movements as distinct behaviours (e.g., agonistic behaviour vs. play). Traditional behaviour research involves many hours of training researchers and volunteers to correctly identify animal behaviour based on clear, detailed ethograms. Similarly it takes many hours and thousands of images to train artificial intelligence to “recognize” distinct behaviours. The payoff of this technology is that it allows for longer durations of monitoring, without the need for human observers, resulting in mass data collection [30]. Artificial intelligence can also be applied to large groups of animals (e.g., ant colonies and bee hives [31]), or implemented for monitoring animals that are often difficult to sample, namely cryptic species such as the nocturnal spotted-tailed quoll [32] or species that cover an extensive range such as seabirds [27]. Historically, collecting information from these species has proven difficult, but technology allows for the prediction of breeding biology and at-sea foraging behaviour, respectively [27,32]. In summary, much of the information that we learn from species would not be possible without these technological advancements.

1.2.2. Technology Types

Evidently, a variety of technologies are available, such as wearable sensors, described above for livestock, that can track individuals and monitor changes [33,34]. However, it is important to note that, as the title suggests, these devices are quite obtrusive and sometimes invasive. Research on wild animals instead often involves the deployment of

camera traps, which are triggered by wild animals passing by [18,20,35], or using thermal-sensing technology [32,35,36]. In the past two decades, facial recognition technology has developed, allowing for unobtrusive individual recognition and tracking [16,21,37,38]. With recent advancements in technology, computer software involving machine and deep learning, as well as artificial neural networks (ANNs), have been trained to provide surveillance of animals with regard to their location, species identity, individual identity, and/or behaviour (e.g., [19,21,27,30,31,39–44]). Machine learning focuses on algorithms and data to imitate the way humans learn; deep learning is a subfield of machine learning. Machine learning involves training computers to function to perform a task with minimal human intervention (i.e., explicit programming), whereas deep learning involves training computers to “think” using a human-like brain structure. The process and output of deep learning is more complex and requires larger datasets for training. ANNs mimic human brains through a set of algorithms. These ANNs consist of “layers” and require more than two layers to be classified as deep neural networks (DNNs) [45].

1.3. Current Limitations and Proposed Novel Technology

As outlined above, many technologies have been designed and developed, currently capable of recording animals’ identity, movement, heart rate, and temperature. Unfortunately, wearable devices can be intrusive and invasive when physically attached to the monitored animal, and evidence suggests that their presence may affect individuals’ behaviour (e.g., transmitters on snakes [46]; radio tags on birds [47]). Further, there are many species for which such devices are not compatible (e.g., some animals are too small or too large—GPS collars cannot be fitted on male polar bears because their necks are larger than their heads [48]), whereas camera technology is not capable of collecting all of the same information remotely (i.e., one technology capable of movement, identity, and temperature). While these technologies all provide critical advancements in monitoring animals, upon review of the literature at the initiation of this project, there appeared to be no single technology capable of providing all of the above information, both unobtrusively and non-invasively. Canadian tech mogul EAIGLE Inc. (hereafter “EAIGLE”) has developed and produced artificial intelligence (AI) software solution capable of determining where people and assets customers are within industry partners’ stores or public places or attractions for operational intelligence, security and health and safety applications. The solution also improves layout and monitors individual temperatures to reduce the potential spread of COVID-19 (Figure 1). Conservation of highly endangered species is a critically important issue in these days of climate and anthropogenic change. Modern zoos, such as the Toronto Zoo, are at the forefront of efforts to save species from extinction, and to ensure that animals in their care live enriched, healthy lives in species-typical environments. It is evident that mass collection of all of multiple sources of data points would provide zookeepers, conservation workers, and researchers with the necessary information to learn more about the species, and individuals under their supervision, in order to provide the best care attainable while exploring further reintroduction and conservation possibilities. Presently, research on animal behaviour is typically collected by observers collecting real-time data via paper, pen, and clipboard, or videos that are later manually annotated (i.e., labeled). Unfortunately, this type of data collection can be a lengthy and tedious process, and observations are limited to the hours in which observers are present.

With consideration of EAIGLE’s AI software, a novel opportunity arose for an industry partner (EAIGLE), an academic institution (York University), and a not-for-profit institution (Toronto Zoo Wildlife Conservancy) to become partners in developing an AI that fills the gap in current technology: tracking orangutans’ location in real time, identifying individuals, pose estimation (i.e., specifying body landmarks, e.g., shoulder and head), and behaviour monitoring to a high level of specificity. Given that the AI was originally developed for use with humans, we initiated this project with a focus on our close relative, the critically endangered Sumatran orangutan (*Pongo abelii*). The research team members are collecting, preparing, and analyzing the images from the orangutans in their zoo

enclosure, and EAIGLE is training the AI model. The purpose of the AI model is to identify individuals and to analyze their behaviours, as well as their body temperature, and limb/joint position.

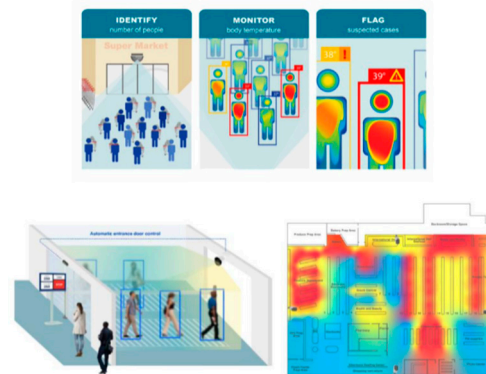


Figure 1. An image of EAIGLE’s current artificial intelligence that can identify the number of people in a room (top left), monitor individual body temperature (top center), and flag suspected cases (top right) of COVID-19. This allows the technology to monitor crowds of people and count them (bottom left), while mapping their real-time density on the facility layout (bottom right) as deployed in high-foot traffic facilities and public places.

Orangutans are members of the same great ape taxon as humans, diverging from us in the evolutionary tree only 12–16 million years ago (e.g., [49,50]). Studying orangutans can provide a strong comparative model for the evolution of complex cognitive abilities, including perception, memory, and sociality. Orangutans have a unique ecological and behavioural repertoire, resulting in some important differences between these two species. Orangutans are: (1) non-vocal learners (i.e., do not learn species-specific vocalizations from a parent or teacher), (2) arboreal (i.e., tree dwelling), and (3) semi-solitary (i.e., social bonds are more easily formed between adult females and weaned offspring). Sumatran orangutans are critically endangered (~7,500 remain) [51] and the future of this species consists of living in zoos and conservation sites. Because these animals spend their lives in human care, it is important to understand how the animals interact with their artificial environments, including where they spend their time and which behaviours they complete within a day. For example, where are animals active vs. resting within their habitat? Are there changes in foraging, urination, and defecation indicating a health issue? Are there any increases in agitated, agonistic, or stereotypical behaviours indicating an increase in stress? Understanding how animals interact with their environment is useful for making changes in the layout of their habitat, ensuring a species-typical lifestyle, and improving both group and individual enrichment. Changes in behaviour from baseline, however, can indicate potential health or welfare issues with an individual or provide general data that could assist with husbandry improvements. Therefore, this artificial intelligence will greatly improve animal welfare outcomes. The ability for location detection and individual recognition, for instance, will be critical to monitoring pregnant or sick animals, and track development from infancy to old age. An ability that many technologies lack is pose estimation; the capacity to track body position is not only necessary to flag any potential limb issues, but behaviours are defined by distinct patterns of movements determined by the position of the limbs and joints. All of these capacities combined will ensure for artificial intelligence that addresses the current gap in available technology.

2. Materials and Methods

2.1. Dataset

For Phase 1 of this project, a dataset of 20,000 images was collected in which 5000 images were collected from online resources and 15,000 images were collected from fixed cameras that are installed inside the orangutan habitat at the Toronto Zoo to ensure that the trained

model would be able to detect orangutans (see Figure 2 for examples). Five 12 MP cameras with variable focus lenses were installed inside the orangutans' habitat in order to provide nearly complete coverage of the habitat, minimizing blind spots (see Figures 3 and 4 for panoramic photo and blueprint of the enclosure and camera locations, respectively). These stationary cameras are capable of collecting image stills (Phases 1 and 2) and video clips (Phase 3). Images included multiple orangutans in a single scene which allowed for training of the AI in order to recognize the potential for multiple orangutans visible on camera.



Figure 2. Examples of images collected from five cameras installed in the orangutans' habitat in Toronto Zoo.



Figure 3. A panoramic photo (Google Pixel 5) taken of the indoor orangutan exhibit. Blue boxes indicate the location of each of the five (5) cameras around the perimeter of the enclosure. Orangutan (Sekali) can be seen brachiating between bars.

Since there was no publicly available pre-trained deep learning model for orangutan detection, we used generic object detection and recognition deep neural network models, including YOLOv4, EfficientNet, and Efficient Pose, which are used for orangutan/zookeeper/visitor detection, orangutan recognition, and behavioural analysis based on pose alignments and movements [52–54]. However, to train these models, we first needed to collect a dataset. Therefore, we started with using a method for object detection based on change detection in the background image. A change in the scene corresponds to an object movement. Any changes in the background image with a minimum of 50x50 connected pixels area in the image was taken as a non-noise change in the image; images were taken from each of the five cameras at the frequency of 30 frames per seconds. Once a change was detected, the images were selected and stored in a 2 TB hard drive. The software ensured that at least a 30 s delay existed between two consecutive frames, taken from each camera. Once the images were collected, we refined them by removing redundant images or images without objects of interest. This measure was taken to ensure that collected images were distinct enough from each other to avoid any bias in the dataset.



Figure 4. A blueprint of the indoor orangutan exhibit study area. The location of each of the five (5) cameras is indicated above in red. Beige (middle) indicates the dirt ground, light grey (center) indicates the cement ground, dark grey (bottom) indicates to the raised walkway and platform, and blue (upper right) indicates the water of the moat. Additional grey shapes (throughout) signifies the location of enrichment objects, with green dots referring to the base of the jungle gym bars. Note for scale: Each section of the dark grey walkway is 6 ft wide.

Two annotation programs were developed: one of the annotation programs was designed to draw bounding boxes for each subject (i.e., orangutan, zookeeper, or zoo visitors), whereas the second one was used to segment each part of the orangutans' body and specify body landmarks (i.e., header, neck, shoulder, elbow, hand, hip, knee, etc.). In Phase 2, 15,000 images were annotated and the specific names of each of the six (6) Toronto Zoo orangutans (4 female/2 male, range 15–54 years old) were recorded within every image's annotation (EfficientNet model used for Phase 2 recognition; see Table 1 for orangutan names and defining characteristics used to identify each individual; see Figure 5 for an annotated example image during training). Thus, following Phase 1's image collection and orangutan detection (i.e., distinct from their habitat), the technology was further trained to recognize individual orangutans, distinct from zookeepers, visitors, and other orangutans. Phase 2 was essential as individual identification will be critical to provide personally adapted care for each individual based on the information collected; see Figure 6 for a screen capture of the current technology, successfully identifying the location of individual 'Puppe'.

2.2. Implementation

Each phase further trains the AI model to be able to recognize orangutans' location, identity, physical stature, and all executed behaviours. Presently, we are in the process of completing Phase 3 (see Figure 7 for a flowchart of the project phases). This third phase includes pose estimation and annotating the behaviours of orangutans including foraging, brachiation, locomotion, object play, object manipulation, fiddling, scanning, patrolling, hiding, inactivity, urination, defecation, agitated movement, affiliative vs. agonistic behaviours, and keeper-, guest-, self-, baby-, and tech-directed behaviours (see Table A1 for a full ethogram of orangutan behaviours; note: baby-directed behaviours have recently been added due to the April 2022 birth of an infant male, mothered by Sekali and fathered

by Budi). This level of data collection will be timely, including annotating series of brief video clips and specifying body landmarks using our second annotation tool. Due to the importance and time-consuming nature of this process, this phase will be the most critical piece of this project.

Table 1. Toronto Zoo Sumatran orangutan information, including name, birth date, current age (range 15–54), biological sex (F = female; M = male), birth place, and defining features.

Orangutan Name	Birth Date	Current Age (June 2022)	Biological Sex	Birth Place	Defining Features
Puppe	1967/09/07	54	F	Sumatra	Dark face with dimpled cheeks; yellow hair around ears and on chin; reddish-orange hair, matted hair on back/legs; medium body size; always haunched with elderly gait/shuffle; curled/stumpy feet
Ramai	1985/10/04	36	F	Toronto Zoo	Oval-shaped face with flesh colour on eyelids (almost white); red hair that falls on forehead; pronounced nipples; medium body size
Sekali	1992/08/18	29	F	Toronto Zoo	Uniformly dark face (cheeks/mouth), with flesh coloured upper eyelids and dots on upper lip; horizontal lines/wrinkles under eyes; long, hanging, smooth orange hair with “bowl-cut”, and lighter orange hair to the sides of the mouth; mixed light and dark orange hair on back (light spot at neck); medium-sized body; brachiates throughout the enclosure
Budi	2006/01/18	16	M	Toronto Zoo	Wide, dark face with pronounced flanges (i.e., cheek pads); dark brown, thick/shaggy hair on body and arms, wavy hair on front of shoulders; large/thick body; stance with rolled shoulders; often climbing throughout enclosure
Kembali	2006/07/24	15	M	Toronto Zoo	Oval/long face with flesh colour around eyes and below nose; dark skin on nose and forehead; small flange bulges; hanging throat; reddish-orange shaggy hair with skin breaks on shoulders and inner arm joints; thicker hair falls down cheeks; large/lanky body; strong, upright gait; often brachiating throughout the enclosure
Jingga	2006/12/15	15	F	Toronto Zoo	Oval/long face with flesh colour around eyes and mouth; full lips; reddish-orange hair with skin breaks on shoulders, inner arm joints, and buttocks; thicker hair falls on forehead and cheeks; medium-small body size

The developed algorithm is capable of merging repetitive objects in the event of overlap between cameras. The developed algorithm is able to identify the same object detected by multiple cameras. This is performed by a similarity check of the objects detected by multiple cameras. The similarity check is performed using both the location of the object and features extracted using deep convolutional network from detected object. It provides a feature vector including 64 values of the object which compresses the appearance information of the animal. Using the similarity check, the algorithm merges multiple detections of each single animal in the habitat.

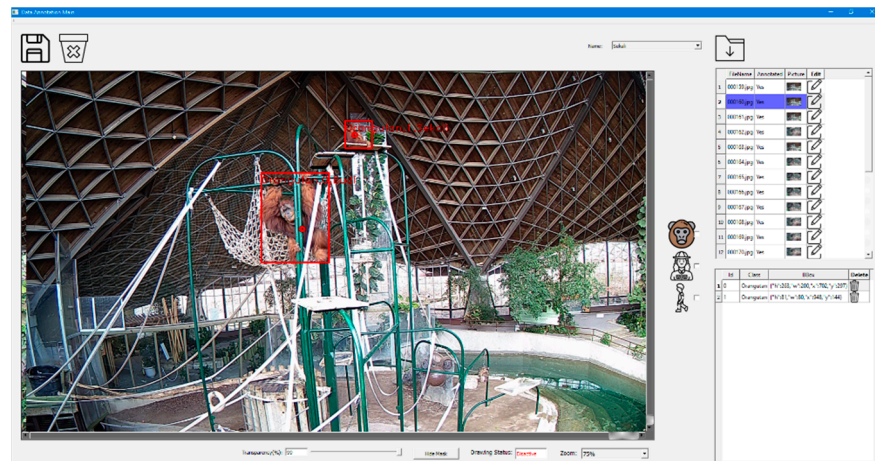


Figure 5. An example of an annotated image for training the artificial intelligence model to identify Sumatran orangutans (“Budi”, lower; “Sekali”, higher). The displayed tool allows for a choice between annotating orangutan vs. zookeeper vs. guest, and identification between the six orangutans.

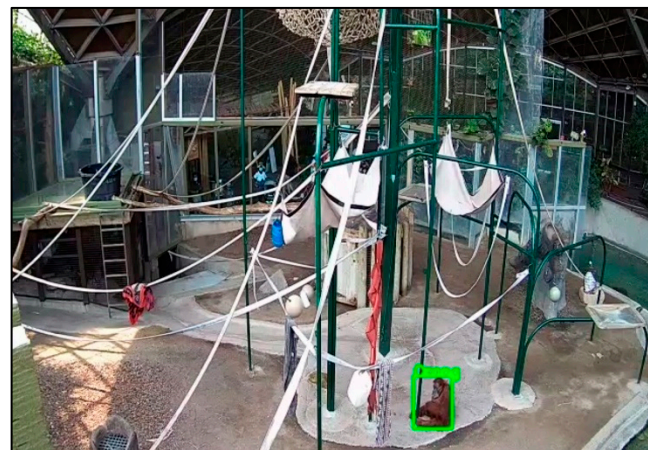


Figure 6. Screen capture of the artificial intelligence detecting an orangutan (“Puppe”). Characteristics include facial details (eyes, mouth, colouration), hair colour and consistency, etc., as compiled from thousands of images. Refer to Table 1 for a list of distinct, individual characteristics that raters used to annotate these images.

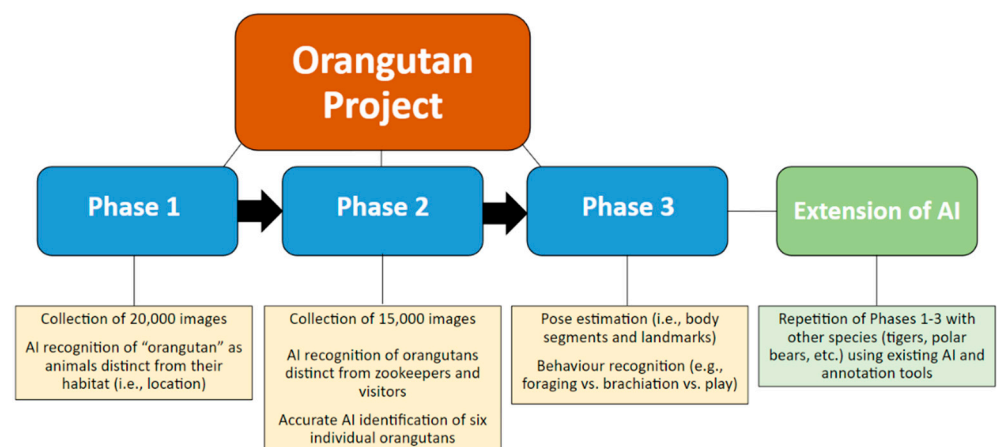


Figure 7. A flowchart of the orangutan project and extension of the resulting artificial intelligence (AI). The details of each of the three (3) phases of the orangutan project are provided, along with a brief explanation of the extension.

3. Results and Discussion

At this time, EAIGLE technology is still in the process of being trained to the highest accuracy. The AI is already successful in detecting orangutans from the background, distinguishing between orangutans vs. zookeepers vs. zoo visitors, and correctly labeling each of the six orangutans (refer to Table 1). Current rates of accuracy have been assessed and are high; specifically, the accuracy of detection is 94% when the animal is at a minimum size of 50×50 pixels, whereas the accuracy of recognition on an individual basis is 80% per single image and increases to 92% with majority voting over 50 frames.

Currently, the AI is being trained to monitor pose estimation and classify animal behaviour (refer to Table A1 for ethogram), a critical and lengthy phase of calibrating this technology. This product will be capable of mass data collection, which will not only benefit the animals at the Toronto Zoo, but which will, we hope, become the 'gold standard' for zoos worldwide.

Our proposed AI addresses a current gap in technology for unobtrusive, 24/7 monitoring non-human animals in collecting a wide-range of information. In understanding where animals spend their time, zoo staff can rearrange their enclosures to be more suitable, comfortable, and cognitively stimulating, improving enrichment and decreasing any detrimental impacts on these captive animals (e.g., [55]). In large conservation areas, gamekeepers could monitor the location and health of the animals (e.g., alarmingly high or low temperatures). Therefore, this computer technology will ensure accurate, unbiased, reliable, and rapid data collection to address a multitude of research questions, from applied welfare to comparative evolution and cognition. Data from captive-based studies such as these will also be useful to provide baselines for species that have yet to be studied extensively in the wild.

4. Conclusions

When the AI is fully validated, we propose extending this project to record and analyze the behaviours of other non-human animals, including (quadrupedal and mammalian) tigers and polar bears, which are also critically endangered. We expect that this technology will be adaptable upon extension to other species. This can be approached in two (2) ways: (1) training a separate model for each animal, or (2) training the current model to detect animals and their species. Considering the variation of characteristics between animals, we recommend grouping animals which are usually placed in the same enclosure and training a model based on the data collected to reduce the complexity of training deep learning models. This helps to keep the size of deep learning models small which allows them to run on small on-edge processors. Additionally, this helps to increase the accuracy of classifying species in a habitat. In addition, both of these species (tigers and polar bears) are also typically solitary, making the identification process easier. This technology is capable of locating orangutans within their indoor habitat (The Indomalayan Pavilion), distinct from foliage and background movements (e.g., birds, visitors, and keepers). In addition, this AI can be accommodated for the tigers and polar bears to distinguish them from their surroundings. Lastly, orangutan behaviour is considered to be quite complex due to the dexterity of their limbs; tiger and polar bear behaviour could be classified as more simplistic, but they make many fine motor movements that may be difficult to capture on camera. Eventually, this AI will be applicable for large mammals throughout the Toronto Zoo, can be implemented in other major zoos in North America and beyond, and eventually modified to implement in conservation areas for monitoring endangered wildlife.

Author Contributions: J.V.C. conceived this paper, contributed Section 1.1, Section 1.2, Section 1.3, Section 3, and Section 4, and edited all sections of this communication. M.H. contributed Sections 2.1 and 2.2. E.F.G. contributed to the behavioural ethogram for AI behaviour training and the ethogram. M.M. developed the software and annotation tools. The entire manuscript was reviewed and received substantial input by S.E.M., with further review by M.F. All authors have read and agreed to the published version of the manuscript.

Funding: This project was funded by two (2) Mitacs Accelerate postdoctoral fellowships. J.V.C.'s postdoctoral funding (Mitacs Reference #IT23450) is in collaboration with Toronto Zoo Wildlife Conservancy (Toronto, ON, Canada). M.H.'s postdoctoral funding (Mitacs Reference #IT25818) is in collaboration with EAIGLE Inc. (Markham, ON, Canada) J.V.C. also received additional postdoctoral funding from an NSERC Postdoctoral Fellowship (Reference #456426328; Natural Science and Engineering Research Council of Canada, Ottawa, ON, Canada).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. Ethogram of behaviours for *Pongo abelii*.

Category	Code	Description
Foraging	F	Consumption of water or plant matter (e.g., leaves, soft vine barks, soft stalks, and round fleshy parts). Marked by insertion of plant matter into the mouth with the use of the hands. It starts with the use of the hands to pick plant matter from a bunch or pile, to pick apart plant matter and to break plant matter into small pieces. The hands are then used to bring plant matter into the mouth. This is followed by chewing (i.e., open and close movement of the jaw whilst the plant matter is either partially or fully in the mouth). This is culminated by swallowing; that is, the plant matter is no longer in the mouth nor outside the animal and the animal moves to get more. The bout stops when there is a pause in the behaviour >3 s or another behaviour is performed.
Brachiation	BR	Arm-over-arm movement along bars or climbing structures. The orangutan is suspended fully from only one or two hands (Full BR) or with the support of a third limb (Partial BR). The movement is slow and there is no air time between alternating grasps on the bars (i.e., at no point will the orangutan be mid-air without support).
Locomotion	L	The orangutan moves with the use of limbs from one point in the exhibit to the next point at least within a metre away from the origin. The orangutan may end up in the same location as the origin, but along the path should have gone at least a metre away from the origin. The orangutan may be locomoting bipedally or quadrupedally on plane surfaces such as platforms and the ground. If the orangutan is on climbing structures but is supported by all four limbs, the movement is classified as locomotion.
Object Play	OP	Repetitive manipulation and inspection (visual and/or tactile) of inedible objects which are not part of another individual's body. The individual is visibly engaged (i.e., the facial/head orientation is on the object being manipulated). Inspection or manipulation is done by mouth, hands or feet. Movement may appear like other behavioural categories but the size/speed of movements of limbs are exaggerated.
Fiddling	FD	Slow and repetitive manipulation of an object with <i>no apparent purpose or engagement</i> (i.e., The orangutan may appear like staring in space and not paying attention to the movement). The orientation of the head must not be facing the object being manipulated. Manipulation may be subtle repetitive finger movements along the object being manipulated.
Inactive	I	The animal stays in the same spot or turns around but does not go beyond a metre from origin. The animal is not engaged in self-directed behaviours, foraging, hiding, defecation, urination, scanning behaviours, or social interaction. The animal may be lying prone, supine, sideways, upright sitting, or quadrupedal, but remains stationary.
Affiliative	AF	The animal engages in social interactions with another individual such as allogrooming, begging for food, food sharing, hugging, tolerance. Behaviours would appear to maintain bond as seen by maintenance of close proximity. These behaviours do not have audible vocalizations or vigorous movements.

Table A1. Cont.

Category	Code	Description
Agonistic	AG	Social interactions with individuals where distance from each other is the outcome unless there is a physical confrontation or fight. The animal may be engaged in, rejection of begging, or avoidance, or vigorously grabbing food from the grasp of the receiver of the interaction. Characterized by vigorous movements towards or away from the other individual.
Keeper Directed	KD	Staring, following, locomoting towards the keeper, or obtaining food from the keeper. Attention/head orientation must be placed on the keeper. The keeper should be visible around the perimeter of the exhibit or in the keeper's cage.
Guest Directed	GD	Staring, following, or moving towards the guests. Attention must be placed on the guest. Volunteers (humans in white shirts and beige trousers are considered guests)
Self Directed Baby Directed	SD BD	Inspection of hair, body, or mouth with hands, feet, mouth or with the use of water or objects such as sticks or non-food enrichment. The body part being inspected is prod repeatedly by any of the abovementioned implements. The animal may scratch, squeeze, poke, or pinch the body part being inspected. Attention does not have to be on the body part. Physical inspection of infant offspring, including limbs, face, and hair of baby, using mouth and/or hands. Attention does not have to on the infant or their body part(s).
Tech Directed	TD	The animal uses or waits at a computer touchscreen enrichment.
Hiding	H	The animal covers itself with a blanket, a leaf, or goes in the bucket such that only a portion of the head is visible. The animal remains stationary.
Urination	U	Marked by the presence of darker wet spot on the ground. Urine flows from the hind of the orangutan. The orangutan may be hanging on climbing structures using any combination of limbs or may be sitting at the edge of the moat, platform, or on a bar with the hind facing where the urine would land.
Object Manipulation	OM	Moving objects with limbs or the mouth from one point in the enclosure to the other point. There is a very clear purpose that usually stops once the purpose has been achieved (e.g., filling a water bottle).
Scanning	SC	The animal makes a short sweeping movement of the head and the eyes stay forward following the gaze. The attention has to be on anything outside the exhibit. The animal may be sitting on the floor or bipedally/quadrupedally locomoting towards a window or the edge of the exhibit.
Patrolling	PT	The animal follows a repeated path around a portion or the entirety of the perimeter of the exhibit. The animal seems vigilant with repeated scans as movement happens.
Defecation	D	Marked by the presence of fecal matter on the floor. Feces drops from the hind of the orangutan. The orangutan may be hanging on climbing structures using any combination of limbs or may be sitting at the edge of the moat, platform, or on a bar with the hind facing where the feces would land. The orangutan may also reach around such that the feces would land on the palm and the orangutan would drop the collected feces on the floor. The orangutan may also gradually orient the upper body from an upright sitting position to a more acute prone posture.
Agitated Movement	AM	Locomotion that is fast, with fast scanning of surroundings, may or may not stop at a destination. Usually follows after a loud noise. Brachiation along the bars is hasty and may involve short air time. Scans towards the keeper's kitchen or the entrance to the exhibit may be possible.

References

- Rossing, W. Cow identification for individual feeding in or outside the milking parlor. In Proceedings of the Symposium on Animal Identification Systems and their Applications, Wageningen, The Netherlands, 8–9 April 1976.
- Rossing, W.; Maatje, K. Automatic data recording for dairy herd management. In Proceedings of the International Milking Machine Symposium, Louisville, KY, USA, 21–23 February 1978.
- Eradus, W.J.; Jansen, M.B. Animal identification and monitoring. *Comput. Electron. Agric.* **1999**, *24*, 91–98. [[CrossRef](#)]
- Helwatkar, A.; Riordan, D.; Walsh, J. Sensor technology for animal health monitoring. *Int. J. Smart Sens. Intell. Syst.* **2014**, *7*, 1–6. [[CrossRef](#)]
- Tan, N.H.; Wong, R.Y.; Desjardins, A.; Munson, S.A.; Pierce, J. Monitoring pets, deterring intruders, and casually spying on neighbors: Everyday uses of smart home cameras. *CHI Conf. Hum. Factors Comput. Syst.* **2022**, *617*, 1–25. [[CrossRef](#)]

6. Dupuis-Desormeaux, M.; Davidson, Z.; Mwololo, M.; Kisio, E.; MacDonald, S.E. Comparing motion capture cameras versus human observer monitoring of mammal movement through fence gaps: A case study from Kenya. *Afr. J. Ecol.* **2016**, *54*, 154–161. [[CrossRef](#)]
7. Heimbürge, S.; Kanitz, E.; Otten, W. The use of hair cortisol for the assessment of stress in animals. *Gen. Comp. Endocrinol.* **2019**, *270*, 10–17. [[CrossRef](#)]
8. Inoue, E.; Inoue-Murayama, M.; Takenaka, O.; Nishida, T. Wild chimpanzee infant urine and saliva sampled noninvasively usable for DNA analyses. *Primates* **2007**, *48*, 156–159. [[CrossRef](#)]
9. Touma, C.; Palme, R. Measuring fecal glucocorticoid metabolites in mammals and birds: The importance of validation. *Ann. N. Y. Acad. Sci.* **2005**, *1046*, 54–74. [[CrossRef](#)]
10. Brown, D.D.; Kays, R.; Wikelski, M.; Wilson, R.; Klimley, A.P. Observing the unwatchable through acceleration logging of animal behavior. *Anim. Biotelemetry* **2013**, *1*, 20. [[CrossRef](#)]
11. Fehlmann, G.; O’Riain, M.J.; Hopkins, P.W.; O’Sullivan, J.; Holton, M.D.; Shepard, E.L.; King, A.J. Identification of behaviours from accelerometer data in a wild social primate. *Anim. Biotelemetry* **2017**, *5*, 47. [[CrossRef](#)]
12. Butcher, P.A.; Colefax, A.P.; Gorkin, R.A.; Kajiura, S.M.; López, N.A.; Mourier, J.; Raoult, V. The drone revolution of shark science: A review. *Drones* **2021**, *5*, 8. [[CrossRef](#)]
13. Oleksyn, S.; Tosetto, L.; Raoult, V.; Joyce, K.E.; Williamson, J.E. Going batty: The challenges and opportunities of using drones to monitor the behaviour and habitat use of rays. *Drones* **2021**, *5*, 12. [[CrossRef](#)]
14. Giles, A.B.; Butcher, P.A.; Colefax, A.P.; Pagendam, D.E.; Mayjor, M.; Kelaher, B.P. Responses of bottlenose dolphins (*Tursiops spp.*) to small drones. *Aquat. Conserv. Mar. Freshw. Ecosyst.* **2021**, *31*, 677–684. [[CrossRef](#)]
15. Brisson-Curadeau, É.; Bird, D.; Burke, C.; Fifield, D.A.; Pace, P.; Sherley, R.B.; Elliott, K.H. Seabird species vary in behavioural response to drone census. *Sci. Rep.* **2017**, *7*, 17884. [[CrossRef](#)] [[PubMed](#)]
16. Burghardt, T.; Čalić, J. Analysing animal behaviour in wildlife videos using face detection and tracking. *IEE Proc. Vis. Image Signal Processing* **2006**, *153*, 305–312. [[CrossRef](#)]
17. Zviedris, R.; Elsts, A.; Strazdins, G.; Mednis, A.; Selavo, L. Lynxnet: Wild animal monitoring using sensor networks. In *International Workshop on Real-World Wireless Sensor Networks*; Springer: Berlin/Heidelberg, Germany, 2010; pp. 170–173. [[CrossRef](#)]
18. Zheng, X.; Owen, M.A.; Nie, Y.; Hu, Y.; Swaisgood, R.R.; Yan, L.; Wei, F. Individual identification of wild giant pandas from camera trap photos—A systematic and hierarchical approach. *J. Zool.* **2016**, *300*, 247–256. [[CrossRef](#)]
19. Rast, W.; Kimmig, S.E.; Giese, L.; Berger, A. Machine learning goes wild: Using data from captive individuals to infer wildlife behaviours. *PLoS ONE* **2020**, *15*, e0227317. [[CrossRef](#)]
20. Bezerra, B.M.; Bastos, M.; Souto, A.; Keasey, M.P.; Eason, P.; Schiel, N.; Jones, G. Camera trap observations of nonhabituated critically endangered wild blonde capuchins, *Sapajus flavius* (formerly *Cebus flavius*). *Int. J. Primatol.* **2014**, *35*, 895–907. [[CrossRef](#)]
21. Schofield, D.; Nagrani, A.; Zisserman, A.; Hayashi, M.; Matsuzawa, T.; Biro, D.; Carvalho, S. Chimpanzee face recognition from videos in the wild using deep learning. *Sci. Adv.* **2019**, *5*, eaaw0736. [[CrossRef](#)]
22. Wark, J.D.; Cronin, K.A.; Niemann, T.; Shender, M.A.; Horrigan, A.; Kao, A.; Ross, M.R. Monitoring the behavior and habitat use of animals to enhance welfare using the ZooMonitor app. *Anim. Behav. Cogn.* **2019**, *6*, 158–167. [[CrossRef](#)]
23. Chen, P.; Swarup, P.; Matkowski, W.M.; Kong, A.W.K.; Han, S.; Zhang, Z.; Rong, H. A study on giant panda recognition based on images of a large proportion of captive pandas. *Ecol. Evol.* **2020**, *10*, 3561–3573. [[CrossRef](#)]
24. Kumar, A.; Hancke, G.P. A zigbee-based animal health monitoring system. *IEEE Sens. J.* **2014**, *15*, 610–617. [[CrossRef](#)]
25. Salman, M.D. (Ed.) *Animal Disease Surveillance and Survey Systems: Methods and Applications*; John Wiley & Sons: Hoboken, NJ, USA, 2008.
26. Brust, C.A.; Burghardt, T.; Groenenberg, M.; Kading, C.; Kuhl, H.S.; Manguette, M.L.; Denzler, J. Towards Automated Visual Monitoring of Individual Gorillas in the Wild. In *Proceedings of the IEEE International Conference on Computer Vision Workshops, Venice, Italy, 22–29 October 2017*; pp. 2820–2830.
27. Browning, E.; Bolton, M.; Owen, E.; Shoji, A.; Guilford, T.; Freeman, R. Predicting animal behaviour using deep learning: GPS data alone accurately predict diving in seabirds. *Methods Ecol. Evol.* **2018**, *9*, 681–692. [[CrossRef](#)]
28. Smith, K.; Martinez, A.; Craddolph, R.; Erickson, H.; Andresen, D.; Warren, S. An Integrated Cattle Health Monitoring System. In *Proceedings of the 2006 International Conference of the IEEE Engineering in Medicine and Biology Society, New York, NY, USA, 30 August–3 September 2006*; pp. 4659–4662. [[CrossRef](#)]
29. MacDonald, S.E.; Ritvo, S. Comparative cognition outside the laboratory. *Comp. Cogn. Behav. Rev.* **2016**, *11*, 49–61. [[CrossRef](#)]
30. Valletta, J.J.; Torney, C.; Kings, M.; Thornton, A.; Madden, J. Applications of machine learning in animal behaviour studies. *Anim. Behav.* **2017**, *124*, 203–220. [[CrossRef](#)]
31. Gernat, T.; Jagla, T.; Jones, B.M.; Middendorf, M.; Robinson, G.E. Automated monitoring of animal behaviour with barcodes and convolutional neural networks. *BioRxiv* **2020**, 1–30. [[CrossRef](#)]
32. Claridge, A.W.; Mifsud, G.; Dawson, J.; Saxon, M.J. Use of infrared digital cameras to investigate the behaviour of cryptic species. *Wildl. Res.* **2005**, *31*, 645–650. [[CrossRef](#)]
33. Gelardi, V.; Godard, J.; Paleressompouille, D.; Claidière, N.; Barrat, A. Measuring social networks in primates: Wearable sensors vs. direct observations. *BioRxiv* **2020**, 1–20. [[CrossRef](#)]
34. Neethirajan, S. Recent advances in wearable sensors for animal health management. *Sens. Bio-Sens. Res.* **2017**, *12*, 15–29. [[CrossRef](#)]

35. McShea, W.J.; Forrester, T.; Costello, R.; He, Z.; Kays, R. Volunteer-run cameras as distributed sensors for macrosystem mammal research. *Landsc. Ecol.* **2016**, *31*, 55–66. [[CrossRef](#)]
36. Dong, R.; Carter, M.; Smith, W.; Joukhadar, Z.; Sherwen, S.; Smith, A. Supporting Animal Welfare with Automatic Tracking of Giraffes with Thermal Cameras. In Proceedings of the 29th Australian Conference on Computer-Human Interaction, Brisbane, Australia, 28 November–1 December 2017; pp. 386–391. [[CrossRef](#)]
37. Loos, A.; Ernst, A. An automated chimpanzee identification system using face detection and recognition. *EURASIP J. Image Video Process.* **2013**, *2013*, 49. [[CrossRef](#)]
38. Witham, C.L. Automated face recognition of rhesus macaques. *J. Neurosci. Methods* **2018**, *300*, 157–165. [[CrossRef](#)] [[PubMed](#)]
39. Duhart, C.; Dublon, G.; Mayton, B.; Davenport, G.; Paradiso, J.A. Deep Learning for Wildlife Conservation and Restoration Efforts. In Proceedings of the 36th International Conference on Machine Learning, Long Beach, CA, USA, 10–15 June 2019; pp. 1–4.
40. Norouzzadeh, M.S.; Nguyen, A.; Kosmala, M.; Swanson, A.; Packer, C.; Clune, J. Automatically identifying wild animals in camera trap images with deep learning. *Proc. Natl. Acad. Sci. USA* **2017**, *115*, E5716–E5725. [[CrossRef](#)]
41. Patil, H.; Ansari, N. Smart surveillance and animal care system using IOT and deep learning. *SSRN* **2020**, 1–6. [[CrossRef](#)]
42. Pons, P.; Jaen, J.; Catala, A. Assessing machine learning classifiers for the detection of animals' behavior using depth-based tracking. *Expert Syst. Appl.* **2017**, *86*, 235–246. [[CrossRef](#)]
43. Schindler, F.; Steinhage, V. Identification of animals and recognition of their actions in wildlife videos using deep learning techniques. *Ecol. Inform.* **2021**, *61*, 101215. [[CrossRef](#)]
44. Willi, M.; Pitman, R.T.; Cardoso, A.W.; Locke, C.; Swanson, A.; Boyer, A.; Fortson, L. Identifying animal species in camera trap images using deep learning and citizen science. *Methods Ecol. Evol.* **2019**, *10*, 80–91. [[CrossRef](#)]
45. Janiesch, C.; Zschech, P.; Heinrich, K. Machine learning and deep learning. *Electron. Mark.* **2021**, *31*, 685–695. [[CrossRef](#)]
46. Lentini, A.M.; Crawshaw, G.J.; Licht, L.E.; McLelland, D.J. Pathologic and hematologic responses to surgically implanted transmitters in eastern massasauga rattlesnakes (*Sistrurus catenatus catenatus*). *J. Wildl. Dis.* **2011**, *47*, 107–125. [[CrossRef](#)]
47. Snijders, L.; Weme, L.E.N.; de Goede, P.; Savage, J.L.; van Oers, K.; Naguib, M. Context-dependent effects of radio transmitter attachment on a small passerine. *J. Avian Biol.* **2017**, *48*, 650–659. [[CrossRef](#)]
48. Pagano, A.M.; Durner, G.M.; Amstrup, S.C.; Simac, K.S.; York, G.S. Long-distance swimming by polar bears (*Ursus maritimus*) of the southern Beaufort Sea during years of extensive open water. *Can. J. Zool.* **2012**, *90*, 663–676. [[CrossRef](#)]
49. Goodman, M.; Porter, C.A.; Czelusniak, J.; Page, S.L.; Schneider, H.; Shoshani, J.; Groves, C.P. Toward a phylogenetic classification of primates based on DNA evidence complemented by fossil evidence. *Mol. Phylogenetics Evol.* **1998**, *9*, 585–598. [[CrossRef](#)] [[PubMed](#)]
50. Chen, F.C.; Li, W.H. Genomic divergences between humans and other hominoids and the effective population size of the common ancestor of humans and chimpanzees. *Am. J. Hum. Genet.* **2001**, *68*, 444–456. [[CrossRef](#)] [[PubMed](#)]
51. World Wildlife Fund (n.d.). Available online: <https://www.worldwildlife.org/species/orangutan> (accessed on 17 May 2022).
52. Gkioxari, G.; Girshick, R.; Dollár, P.; He, K. Detecting and Recognizing Human-Object Interactions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 8359–8367. [[CrossRef](#)]
53. Tan, M.; Le, Q. Efficientnet: Rethinking Model Scaling for Convolutional Neural Networks. In Proceedings of the International Conference on Machine Learning, Long Beach, CA, USA, 10–15 June 2019; pp. 6105–6114. [[CrossRef](#)]
54. Bochkovskiy, A.; Wang, C.Y.; Liao, H.Y.M. Yolov4: Optimal speed and accuracy of object detection. *arXiv* **2020**, arXiv:2004.10934. [[CrossRef](#)]
55. Ward, S.J.; Sherwen, S.; Clark, F.E. Advances in applied zoo animal welfare science. *J. Appl. Anim. Welf. Sci.* **2018**, *21*, 23–33. [[CrossRef](#)] [[PubMed](#)]