



## CROSS-SPECIES DEEP LEARNING FOR AUTOMATIC IDENTIFICATION AND BEHAVIOUR ANALYSIS OF ENDANGERED HAWKSBILL SEA TURTLES USING FACIAL SIDE SIMILARITY

**Eneasoba Nneka Charity<sup>1</sup>, Chinwe Sussan Oguejiofor<sup>2</sup>, Oyelade Zainab<sup>3</sup>, Karthick Kulandhaivelu<sup>4</sup>, Ogochukwu Gloria NNONYELU<sup>5</sup>, Egajivwie Frank ODIRI<sup>6</sup>**

<sup>1</sup>Department of Agricultural and Vocational Education (Business Education Programme)  
Michael Okpara University of Agriculture, Umudike Abia State – Nigeria

[nnekaemeasoba@yahoo.com](mailto:nnekaemeasoba@yahoo.com)

<sup>2</sup>Department of Vocational Education Chukwuemeka Odumegwu Ojukwu University Igbariam  
Campus [chysussyogu@gmail.com](mailto:chysussyogu@gmail.com)

<sup>3</sup>National Open University of Nigeria Enugu Study Centre, Enugu State  
[oyesladezainab@gmail.com](mailto:oyesladezainab@gmail.com)

<sup>4</sup>Assistant Professor Department of Computer Science and Engineering Kongunadu College of Engineering and Technology, Namakkal - Trichy Main Road, Tholurpatti(Po), Thottiyam(Tk), Trichy - 621 215. [karthivel.me@gmail.com](mailto:karthivel.me@gmail.com)

<sup>5</sup>Department of Technology and Vocational Education Faculty of Education Nnamdi Azikiwe University [og.nnonnyelu@unizik.edu.ng](mailto:og.nnonnyelu@unizik.edu.ng)

<sup>6</sup>Department of Business Education University of Delta Agbor [odiri.egajivwie@unidel.edu.ng](mailto:odiri.egajivwie@unidel.edu.ng)

### Abstract:

Onboard accelerometer allows for remote monitoring of animal movement and posture, enabling researchers to infer behaviours. Data scarcity is still a problem in ecology, even with deep learning's automated analytical capabilities. To categorise behaviours using acceleration data of critically endangered hawksbill sea turtles (*Eretmochelys imbricata*), we investigated transfer learning. In order to tackle a similar problem, transfer learning leverages a model that was trained on one task from a big dataset. We used this technique to detect hawksbill behaviours including swimming, resting, and eating by adapting a model trained on green turtles (*Chelonia mydas*). Additionally, we contrasted this with a model that was developed using data on human activities. According to the findings, transfer learning from the human and emerald turtle databases improved the F1-score by 4% and 8%, respectively. Through the application of knowledge transfer, researchers may use deep understanding to increase the usage of acceleration sensors for wildlife observation by tailoring current models for the sample animals.

**Keywords:** Accelerometer, Behavioral classification, Bio-logging, Neural networks, Machine Learning, Animal behavior



All the articles published by Chelonian Conservation and Biology are licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/) Based on a work at <https://www.acgpublishing.com/>

## 1. Introduction

Studying animals in their natural environments via behaviour monitoring is a non-invasive approach that yields important information about their ecology and health. By means of behavioural analysis, scientists may get a deeper comprehension of how animals engage with their surroundings and clarify important facets of their survival strategies, energy tactics, and reproductive success (McHuron et al., 2018; Sansom et al., 2009; Zhang et al., 2004).

A crucial method for assessing how well a species has adapted to harsh environments (Abernathy et al., 2019; Sergio et al., 2018) and rapid environmental change (Beever et al., 2017; Buchholz et al., 2019) is behaviour monitoring. Additionally, it makes it possible to quantify extraordinary physiological abilities like deep breath-hold diving and long-distance migration, which contributes to a more thorough comprehension of the mechanisms underlying these behaviours (Boel et al., 2014; Fossette et al., 2010; Pedersen et al., 2018).

Because it allows for ongoing observation of animals in their natural environments, bio-logging has become a useful technique for researching animal behaviour. It entails the installation of onboard sensors that gather detailed information on the physiology, behaviour, and environmental circumstances of animals. Among these sensors, the accelerometer works best for capturing the posture and motion of animals. It functions as a piezoelectric sensor, converting forces acting on a mass into voltage signals that resemble waves. These forces include the force of gravity and the force of inertia brought on by the animal's motion and posture, respectively (Brown et al., 2013). Researchers can detect behaviours like locomotion, eating, and inactivity that are typified by certain postures and motions by aligning three accelerometers orthogonally (Shepard et al., 2008). One of the main challenges in acceleration-based behaviour detection is automating data processing in order to make it easier to analyse big time-series datasets. In order to automatically analyse acceleration data and identify the behaviour of wild animals, deep learning has recently shown promise (Aulsebrook et al., 2024; Jeantet et al., 2021; Otsuka et al., 2024; Zhang et al., 2019a). Neural networks, another name for deep learning algorithms, are computer models that process input in numerous layers and carry out both linear and non-linear changes. Their ability to automatically identify intricate, highly discriminative patterns and characteristics in the data allows them to categorise acceleration signals into different behavioural groups. Specifically, deep learning has led the way in improving human activity recognition (HAR) and behaviour identification using accelerometers in livestock monitoring (Bao and Xie, 2022; Kumar et al., 2023). Many research have created deep learning models for livestock with the goal of tracking animal welfare and production. These models can identify lameness in cattle, sheep, and pigs as well as track behaviours like eating, drinking, walking, and resting (Kleanthous et al., 2022b; Mao et al., 2023). Accelerometers are widely used to monitor human physical activity, including walking, running, standing, and sitting. They are also used to detect falls, which is particularly crucial when caring for elderly individuals (Choi et al., 2022; Nunavath et al., 2021; Ramanujam et al., 2021).

In contrast, there has been little research done on creating new techniques to make deep learning easier to use or improve its effectiveness in the subject of ecology. The added difficulties in obtaining large, accountable, and precisely labelled datasets for environmental study may be the reason for the sparse use of deep learning algorithms in ecological as opposed to cattle tracking or HAR. In overall, deep learning algorithms are big models with several functional layers that follow one another. The model is tailored to the particular job by optimising, or fine-tuning, the weights—real-valued numbers—used in every operation throughout the development phase utilising labelled data. The capacity of the model to distinguish features improves with the number of layers, but this also requires a bigger labelled dataset for efficient training (He et al., 2016). Studying wild animals, on the other hand, poses difficulties for data collecting due to the need to reduce disruption, natural limitations, and restricted availability for threatened species, which leads to smaller datasets.

The procedure is further complicated by having to arrange for contemporaneous observations in order to create a labelled information set, which results in significantly fewer data sets for training. Overfitting is a major danger of utilising information that is that is insufficient for deep learning model development. When the body masses of the model are too closely matched to the particular instances in the training set, this is known as overfitting. As a result, the model does well on the experimental information but badly on the training information. Transfer methods of learning have been used in a variety of environmental tasks, such as bioacoustics, trapping with cameras, drone enquiries, and animals estimation of poses from video recordings (the last three falling according to computer vision) to address the lack of data in other ecological fields (Dufourq et al., 2022; Grey et al., 2019; Liao et al., 2023; Lu et al., 2021; Schneider et al., 2018). A model that has been trained on one task may be modified and improved for an alternate but comparable task using the transfer learnt approach. This approach reduces the possibility of overfitting on smaller, more constrained datasets while allowing the usage of deep learning models with several layers that have already been pre-trained on big datasets. In bioacoustics, Batist et al. (2024) used a pre-trained model trained on ImageNet to identify sounds from black-and-white ruffed lemurs (*Varecia variegata*) using spectrograms. The model was created for classifying objects from colour images. Although the goals and datasets are different (classifying spectroscopy images vs photos), neural networks that were previously trained on the data set for images may pick up abstract designs which may be applied to the spectrometer information. Although transfer learning has long been used to study human behaviour (Cook et al., 2013; Hernandez et al., 2020) and more recently in acceleration-based behaviour identification studies of livestock (Bloch et al., 2023; Kleanthous et al., 2022a), it has not yet been used to study wild animals.

Marine species are particularly affected by the lack of data needed to train deep learning algorithms in ecology. In this work, we investigated the application of transfer learning to the study of *Eretmochelys imbricata*, or hawksbill sea turtles. The Hawksbill and Kemp's ridley turtles (*Lepidochelys kempii*) are the two most endangered of the six sea turtle species, both of which are listed as critically endangered (IUCN, 2024). Hawksbill turtles and young green turtles (*Chelonia mydas*) are found in the Caribbean island of Martinique (Lelong et al., 2024; Nivierè et al., 2018). Understanding endangered species' behaviours is crucial for understanding their adaptation

to changing environments and anthropogenic pressures (e.g., coral bleaching, eutrophication; Li and Reidenbach, 2014; Hayes et al., 2017; Louis-Jean et al., 2009; Siegwalt et al., 2022). In order to better understand sea turtles' energy strategies, accelerometers are a useful tool for observing behaviours including swimming, resting, breathing, and eating (Fossette et al., 2012; Jeantet et al., 2020a). Significant monitoring is being done on Martinique's green turtle population (Bonola et al., 2019; Charrier et al., 2022; Siegwalt et al., 2022, 2020), and research has been done using accelerometers to automatically track their movements (Jeantet et al., 2020a, 2021). On the other hand, hawksbill turtle population research is more difficult since individuals are harder to locate than green turtles, making it challenging to outfit them with bio-loggers. Additionally, there aren't many research that utilise accelerometers to study hawksbill turtles; as far as we know, just two studies have looked at using this technology to examine the behaviour of these turtles (Jeantet et al., 2018; Okuyama et al., 2012). As a result, there is a significant dearth of information and research on this species, and the difficulties in gathering data make it challenging to use deep learning algorithms to examine their behaviours. This work sought to overcome the prevalent problem of limited datasets in ecological research when training deep learning algorithms by examining the use of transfer learning for the automated identification of behaviours in endangered species using accelerometer data. We concentrated on the hawksbill sea turtle, a species that is extremely difficult to monitor and for which there is little information and study. First, we investigated using a pre-trained model on the green turtle, a similarly related species whose postures and movement patterns are thought to resemble those of the hawksbill turtle while expressing behaviours of interest such breathing, swimming, eating, and resting. There is a trained model and dataset for this species that are accessible to the public. Second, we evaluated the viability of using transfer learning from a species that differs significantly from the hawksbill turtle in terms of morphology, since behavioural identification for this species should depend on distinct characteristics from those utilised for sea turtles. With little reliance between the species used in the pre-trained model and the species under study, this strategy allowed us to assess if the technique is species specific or applicable to a wider variety of species. Because data collecting is easier and there are many datasets and pre-trained models available, we concentrated on people for this. In order to provide a technique that can be used to a larger variety of species and enable the wider use of deep learning in the study of animal behaviour, this work is the first to investigate transfer learning for monitoring wild animals.

## 2. MATERIALS AND TECHNIQUES

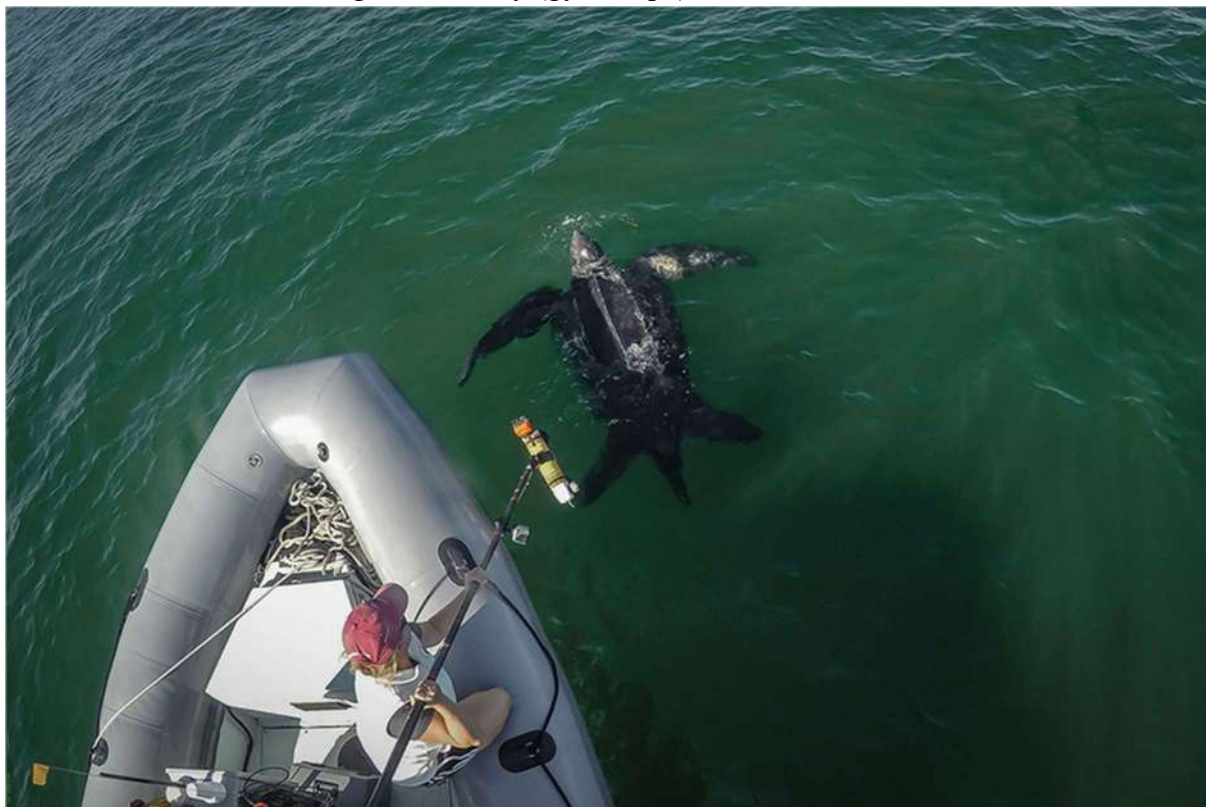
### 2.1 The Hawksbill dataset

This research complies with all institutional rules and legal requirements of the nations where it was conducted. The French Ministry for Ecology, Sustainable Development and Energy (permit no. 2013154-0037), which serves as an ethics committee in Martinique, and the "Conseil National de la Protection de la Nature" (<http://www.avis-biodiversite.developpement-durable.gouv.fr/bienvenue-sur-le-site-du-cnpn-et-du-cnb-a1.html>) both approved the protocol. To

minimise animal disturbance, the fieldwork was conducted strictly in compliance with the Prefecture of Martinique's guidelines (authorisation no. 201710-0005).

## 2.2 Information gathering

From November 2022 to May 2023, fieldwork was carried out on Martinique, a Caribbean island in France, to gather the hawksbill turtle dataset. As detailed in Niviere et al. (2018) and Jeantet et al. (2020a), six hawksbill turtles, *Eretmochelys imbricata* (Linnaeus 1766), that were free-ranging were manually captured, measured, and recognised. Over the course of two days, we fitted them with CATS (Customised Animal Tracking Solutions) devices, which included an automated release mechanism and four suction cups fastened to their carapace (Fig. 1; see Jeantet et al., 2020a, for details). Each CATS device had a pressure sensor, a tri-axial gyroscope, a tri-axial accelerometer, and a video recorder. The gadgets were set up to continually record pressure at 1 Hz and acceleration and angular velocity (gyroscope) at a rate of 20 Hz.



**Figure 1: TurtleCam: A “Smart” Autonomous Underwater Vehicle for Investigating Behaviors and Habitats of Sea Turtles**

The video recorders were set up to start recording at 18:00 and stop at 06:00, respectively, for nighttime. It was thought that the maximum battery capacity would allow for 18 hours of video recording and 48 hours for the other sensors.

## 2.3 Annotating videos and preparing data

Using specially designed software called TurtleCap (<https://github.com/Vadym-Hadetskyi/TurtleCap>), video recordings taken by the CATS devices were examined visually in order to find patterns. To the closest tenth of a second, the start and finish timings of every



behaviour that was seen were noted. The video recordings were labelled using six primary behavioural categories: breathing, feeding, gliding, resting, scratching, and swimming (Fig. 1). Any further behaviour that was seen was classified as Other. The behaviour of green turtles in earlier Martinique research was described using similar seven categories (for a detailed discussion of the behaviours, see Jeantet et al., 2020a). In order to visualise and verify the acceleration, gyroscope, and pressure data related to the relevant observed behaviours, R software (version 3.5.3; <http://www.R-project.org/>) and the rblt package (<https://CRAN.R-project.org/package=rblt>) were used.

The primary tasks of data preparation were eliminating the unlabelled night episodes for which there was no video footage and mapping the behaviour labels to the multi-sensor data. The gyroscopic and acceleration data were not pre-processed. The force information was analysed using a linear interpolation approach to boost the sampling frequency to 20 Hz. The variation among all of the data points was computed at 1 Hz.

Six distinct hawksbill turtles provided 69.7 hours of multisensor recordings in total, with an average of 11.6 hours per person, a maximum of 17.8 hours, a minimum of 6.3 hours, and a standard deviation of 3.6 hours. With a total of 38.6 hours, feeding was the most common behaviour shown in the movies. Resting and swimming came in second and third, with 19.1 and 7.9 hours, respectively (Table S1). Breathing took 2.2 hours, gliding took 1 hour, scratching took 0.8 hours, and other took 0.1 hours.

## **2.4 The pre-trained models' datasets**

### **2.4.1 Dataset of green turtles**

Similar to the hawksbill dataset, the green turtle dataset utilised in this research was gathered in Martinique between February 2018 and May 2019. Over the course of one or two days, the CATS devices were placed on thirteen green turtles, *Chelonia mydas* (Linnaeus 1758). The dataset has been thoroughly described in Jeantet et al. (2020a), is publicly accessible on the Dryad digital repository .

The green turtle dataset, which was labelled with the same seven behaviours as the hawksbill dataset, included 68.6 hours of recordings from 13 individuals (mean 5.29 hours per individual, maximum 14.67 hours, minimum 0.96 hours, standard deviation 3.39 hours). With a total of 34.3 hours, resting was the most common behaviour shown in the films. Swimming and breathing came in second and third, with 22.3 and 5.7 hours, respectively (Table S1). The other behaviors—gliding for 2.3 hours, feeding for 1.8 hours, scratching for 1.2 hours, and other for 1 hour—were shown in minority.

### **2.4.2 Dataset of humans**

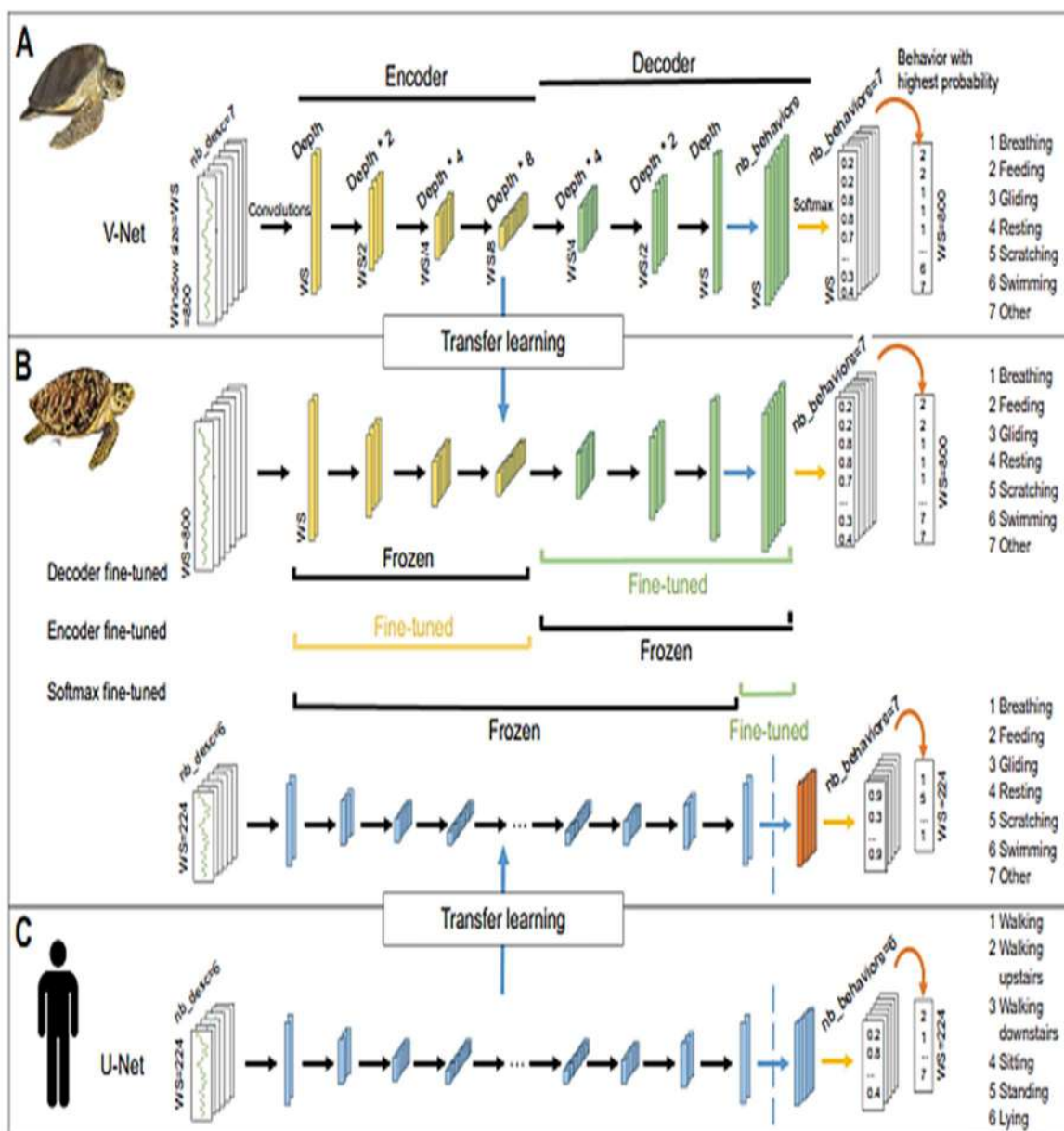
We also examined the possibility of transfer learning from other taxa, where data collecting is more practical, in order to gauge the importance of the species used for pre-training. Since it is far simpler to gather data from humans than from other animals, we concentrated on them, and as a consequence, there are many open-access datasets accessible. The Intensive Care Unit (ICU) HAR

dataset (Reyes-Ortiz et al., 2013), which is publically accessible and often utilised in the creation of deep learning models for HAR (Ramanujam et al., 2021; Zhang et al., 2022), was used in this instance.

Thirty participants in the ICU HAR wore smartphones around their waists to capture 3D acceleration and 3D angular speed, which were recorded at 50 Hz. A total of almost 220 hours of labelled sequences were completed by each participant in six different activities (walking, walking upstairs, walking downstairs, sitting, standing, and lying) with durations varying from one to two minutes each (see Table S1 for the particular time available per activity).

### 2.4.3 Deep learning model architecture

We evaluated V-Net (Jeantet et al., 2021), a model that was first modified to recognise the underwater behaviours of green turtles in Martinique, in this transfer learning research. One benefit of the V-Net model is that it produces an output that is the same length as the input, enabling predictions to be made from the recorded data at each data point, or time step. Without the need for preprocessing procedures like segmentation, filtering, or the computation of descriptive variables, it can predict behaviours straight from raw accelerometer data (Jeantet et al., 2020a). Additionally, the V-Net model was created especially to deal with unbalanced and tiny datasets (Milletari et al., 2016). Beyond its use with Martinian green turtles, the model has shown promising results in detecting prey capture events in Antarctica's chinstrap penguins (*Pygoscelis antarctica*) (Schoombie et al., 2024) and terrestrial behaviours in French Guiana's green turtles (Jeantet et al., 2022). This data points to its capacity to generalise across situations and species. We evaluated U-Net, a similarly comparable model with a similar architecture and extensive usage in human investigations (Meena and Sarawadekar, 2023; Zhang et al., 2019b), for the human dataset.



**Figure 2:** The V-Net and U-Net architectures are shown. The V-Net was pre-trained on the dataset of green turtles, while the U-Net was pre-trained on the dataset of humans. (B) The transfer learning applications evaluated on the Hawksbill dataset, displaying the various frozen layer configurations for the model pretrained on the Green Turtle dataset and the orange-colored changed layer for the model pretrained on the Human dataset. Two successive convolutional layers are shown by the black arrows, and one convolutional layer is represented by the blue arrows. The amount of features to be generated is represented by depth (depth=32), the quantity of input data by nb\_desc, and the number of behaviours by nb\_behaviors. The window size is represented by WS.



Because convolution is their main function, V-Net and U-Net are models sometimes referred to as fully convolutional neural networks. Convolution is a mathematical process used in machine learning, specifically in convolutional neural networks, to identify and extract hierarchical features from input data, such as edges, textures, and patterns in signals or pictures. It creates a new signal that emphasises certain characteristics depending on the filter by sliding a filter across an input, multiplying the filter values by the signal values, and adding them at each location. The weights of the model are represented by the values of the filter used in the summations and multiplications. Based on the labelled data supplied to the model, the training phase of a neural network optimises these weights to allow precise identification of the behaviours of the species under study.

A convolutional layer filters the incoming data using a predetermined number of filters. Every filter creates a feature map, which is a modified version of the input signal. Multiple convolutional layers arranged sequentially constitute a fully convolutional neural network (a schematic representation of the V-Net and U-Net architecture may be seen in Fig. 2). Behaviour categorisation is made possible by the sequential application of numerous filters, which allows for the extraction of different characteristics from the signal.

Two encoder-decoder designs are V-Net and U-Net (Fig. 2). These models initially decrease the input's dimensionality by mapping it to a lower-dimensional space using a sequence of convolutional layers. Depending on the number of layers and filters in the encoder, this procedure creates a significant number of discriminative features. The decoder, the second component of the model, then combines these information to forecast the result. An output matrix of the same length as the input is the consequence of this. A convolutional layer with seven channels, representing each of the behavioural types, makes up the last layer. A matrix with a value for each behaviour at each data point is the result at this moment. The probability of each behaviour for each data point is then determined by running these seven values through a softmax function (Eqn 1).

$$\sigma(z_i) = \frac{e^{z_i}}{\sum_{j=1}^n e^{z_j}} \quad (1)$$

where  $z_i$  is the  $i$ th element of the array of size  $n$ , where  $n$  is the number of classes, and  $\sigma$  is the softmax. The softmax function may be seen as a probability distribution as it generates an array of non-negative values that add up to 1. The behaviour with the greatest probability is then chosen as the anticipated behaviour for each data point. The Supplementary Materials and Methods include a thorough explanation of the V-Net and U-Net topologies.

#### 2.4.4 Assessment of the models

We evaluated the different models used in this work on a dataset different from the training set in order to assess their performance. With this method, we may evaluate how well the models predict data that is deemed unseen since it isn't utilised in model training. Training, validation, and testing sets were created from the three datasets previously mentioned. During the iterative training process, the training and validation sets are utilised concurrently, with the validation set being used to evaluate performance trends. The testing set, which replicates the circumstances in which the model will be used, is kept aside for the model's final analysis. According to Reyes-Ortiz et al.

(2013), the human dataset was first made available in two datasets, with 70% of the participants being randomly assigned to the training set and the remaining 30% to the test set. Thirty percent of the participants in the human training dataset were assigned at random to the validation dataset. Following the same distribution as in Jeantet et al. (2021), seven green turtles were utilised for training, three for validation, and three for testing. Three hawksbill turtles were chosen at random for the training dataset, two for testing, and one for validation. This method of data splitting was used consistently to every model that was evaluated.

We used the testing dataset to assess the generalisability of the model. In order to do this, we separated each test subject's whole recording into windows that overlapped by 10%. After running the trained model over all windows, we put the predictions back together to recreate the signal. We created the confusion matrix and determined the global accuracy (Eqn 2), precision (Eqn 3), recall (Eqn 4), and F1-score (Eqn 5) based on the predictions and the manually confirmed labels:

$$\text{global accuracy} = \frac{TP + TN}{TP + TN + FN + FP} \quad (2)$$

$$\text{Precision} = \frac{TP}{TP + FN} \quad (3)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (4)$$

$$\text{F1 - SCORE} = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \quad (5)$$

True positives (TP) and true negatives (TN) in these equations denote the total amount of behaviours that were properly detected, while false negatives (FN) and false positives (FP) denote the number of behaviours that were classified incorrectly. The F1-score is thought to be a more accurate indicator of the model's efficiency than global reliability for unbalanced datasets (Saito and Rehmsmeier, 2015).

#### 2.4.5 Learning transfer: The basics of transfer learning

In order to minimise the discrepancy among the outputs and the annotated numbers, a neural network's weights are repeatedly modified as the model learns on labelled information during training. These weights are first produced at random and then tailored to the particular classification job. According to Ying (2019), overfitting occurs when these weights are excessively customised for the labelled dataset, which impairs generalisation and model performance on fresh data. This problem is especially common when there are a lot of weights in the model and a little training dataset.

Using pre-trained models, transfer learning initialises model B using weights from model A. Usually, a greater amount of data is used to train a particular model on a task that is comparable with model B. Instead of using randomly initialised a weight, the weights from version A are used as the beginning point for design B in this procedure. The parameter values of the model A are then adjusted especially for the task B throughout the modelling stage using the smaller labelled dataset. In reality, the algorithm's early tiers are often "frozen," which means only the weights of the layers that are closest to the output are adjusted throughout training. There are computational

benefits to freezing certain layers as fewer weights need to be optimised. Early layers of a model should be frozen since they capture broad data characteristics, but deeper layers gradually pick up abstractions from the data (Gu et al., 2018; Yosinski et al., 2015 preprint). Since of this, the deeper layers need to be modified during transfers since they are more tailored to the particular task.

## 2.5 Pre-training using datasets of humans and green turtles

Model A is trained using the original dataset as the first stage in the transfer the learning process. We used the same settings as in Jeantet et al. (2021) to train the V-Net in order to get the trained model from the green turtle dataset. We employed the gyroscope (X, Y, Z), the raw acceleration (X, Y, Z), and the pressure difference as input. These were recorded at 20 Hz during a 40-second frame, producing a 7x800 matrix (Fig. 2). Utilising the Adam optimiser and a learning rate of 0.0001, we train the V-Net on thirty runs with a group size of 32, using the Generalised Dice loss (Sudre et al., 2017). Using the coloured turtle information set, we achieved a worldwide precision of 97.2% and an F1-score of 81.1%.

The green turtle's V-Net model has 15 convolutional layers, while Zhang et al.'s U-Net model (2019a,b) has 23 convolutional layers (Fig. 2, see Supplementary Materials and Methods for a thorough explanation of the V-Net and U-Net designs). The authors of the study used the ICU HAR dataset to test the model, segmenting the data using a fixed-length window of size 224. Using this method, we created a 6x224 matrix by training the U-Net on the ICU HAR dataset using raw acceleration (X, Y, Z) and gyroscope (X, Y, Z) data across a 224-length window. We trained the model for 50 epochs with 32 batches, using the Adam optimiser, cross-entropy loss, and 0.001 learning rate, as per Zhang et al. (2019a,b). On the human dataset, we achieved a global accuracy of 95.7% and an F1-score of 85.6%.

## 3. Using the Hawksbill dataset for transfer learning

We investigated a number of situations in order to evaluate the advantages of using transfer learning to identify hawksbill turtle behaviour from a short dataset. First, we used the V-Net trained on green turtles to predict the behaviours of hawksbill turtles in order to ascertain if a new model is required for recognising their behaviour. In this instance, no transfer learning was used, and the framework was just used on the six people to evaluate the forecasts rather than being refined on the Hawksbill datasets. Second, we trained the V-Net using randomly initialised parameters on the hawksbill information to create the starting point (Model-Hawksbill). In this instance, we trained the V-Net using the hawksbill database after arbitrarily initialising each layer instead of using the model that had already been trained on green turtles. Third, the V-Net pretrained on the green turtle database (Model-Green\_turtle) was subjected to transfer learning. Finally, we applied learning from transfers to the U-Net that had previously been trained on the Model-Human database. For our case investigation, we modified the pre-trained U-Net buildings by adding a seventh convolutional network layer with 7 channels (representing the seven behavioural categories in the hawksbill dataset versus six in the human database) and fine-tuning each layer (Fig. 2).

Additionally, given transfer learning includes freezing and refinement multiple layers, we evaluated four extra possibilities for the Model-Green\_turtle to find the ideal configuration (Fig. 2). We adjusted each layer of the model in the first case. In the second, we merely adjusted the decoder levels and frozen the encoder weights. By freezing the decoder and adjusting the encoder measurements, we explored the opposite strategy in the third situation. Lastly, we just adjusted the last convolutional layer's weights in the fourth situation.

We provided the method with a window size of 800 (40 s) to train or fine-tune the V-Net (Model-Hawksbill, ModelGreen\_turtle) on the hawksbill dataset. This included data from the three gyroscope and three accelerometer axes as well as the pressure differential (matrix size  $7 \times 800$ ). We replicated the input shape used with the human dataset for Model-Human by using a window of size 224 and only the three gyroscope and accelerometer axes (matrix size  $224 \times 6$ ). We created these windows by reusing the technique outlined in Jeantet et al. (2021), which enabled us to create a preset quantity of windows while encouraging certain behaviours to balance the dataset (for more details, see Jeantet et al., 2021). At each epoch, we created 6000 windows from the training dataset and 3000 from the validation dataset (the distribution of the behaviours based on this technique is shown in Fig. S1). With identical hyperparameters (epoch=20, learning rate=0.0001, batch size=32, Adam optimiser, the Generalised Dice loss function for the V-Net, and the cross-entropy loss function for the U-Net), we performed each execution 20 times. The model's performance was assessed on the validation dataset at each epoch, and when the optimal performance was attained throughout the course of the 20 epochs, the model weights were stored.

## 4. Results

### 4.1. Initial Trial: Assessing how comparable the left and right profiles are



Fig. 5 displays similarities scores for profiles assessments across all five configurations (A)-(E) of Fig. 3, using MegaDescriptor, ALIKED, and TORSOOI (three columns) for all four collections (four rows) of original (red) and flipped (blue) query photos. Higher scores indicate more similarity for all methods. Nevertheless, we see that the horizontally axis' similarities across the various approaches cannot be compared. MegaDescriptor learnt the individuals' inner descriptions regardless of the capturing profile since it did not distinguish among the initial and flipped photos. The parallels scores often fall into a particular order:

MegaDescriptor:  $A > B \sim C > D \sim E$

TORSOLL:  $B > \sim A \sim C > D \sim E$

ALIKED ORIGINAL QUERY:  $B > D > A \sim C \sim E$

ALIKED FLIPPED QUERY:  $A > C > E > B \sim D$

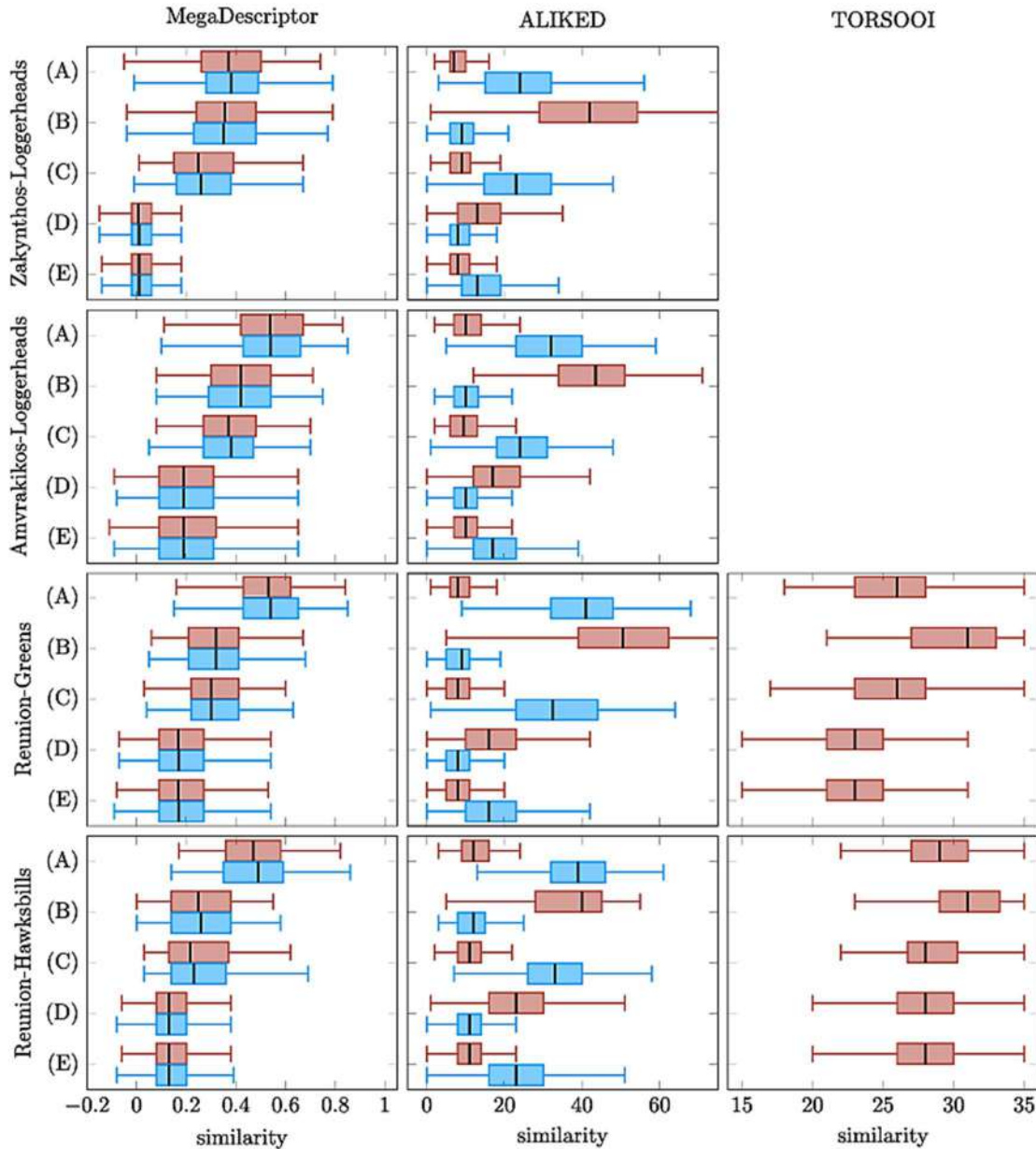
|             | Setting (A)   | Setting (B)   | Setting (C)   | Setting (D)  | Setting (E)   |
|-------------|---|---|---|--|---|
|             |  |  |  |  |  |
| Individual: | Same  | Same  | Same  | Different  | Different   |
| Profile:    | Opposite  | Same  | Opposite  | Same   | Opposite  |
| Year:       | Same  | Different   | Different   | Any  | Any   |
| Count:      | 1   | 1   | 1   | Many   | Many  |

**Figure 3: Experiment 1: Configuration for assessing sea turtle head feature similarity. "Individual," "Profile," and "Year" relate to the extent that identical people, characteristics, or captures seasons are shown in the photographs in pairwise contrasts with every circumstance. The number of photos against which each query picture is evaluated under every option is shown in the last row, "Count."**



**Figure 4: Inspiration for conducting experiments after reversing the query pictures across: ALIKED initially generates only 24 (mostly incorrect) matching key locations for opposing profiles; however, when the input picture is flipped vertically, 62 corresponding keypoints are produced, the majority of which correspond identical the head jobs.**



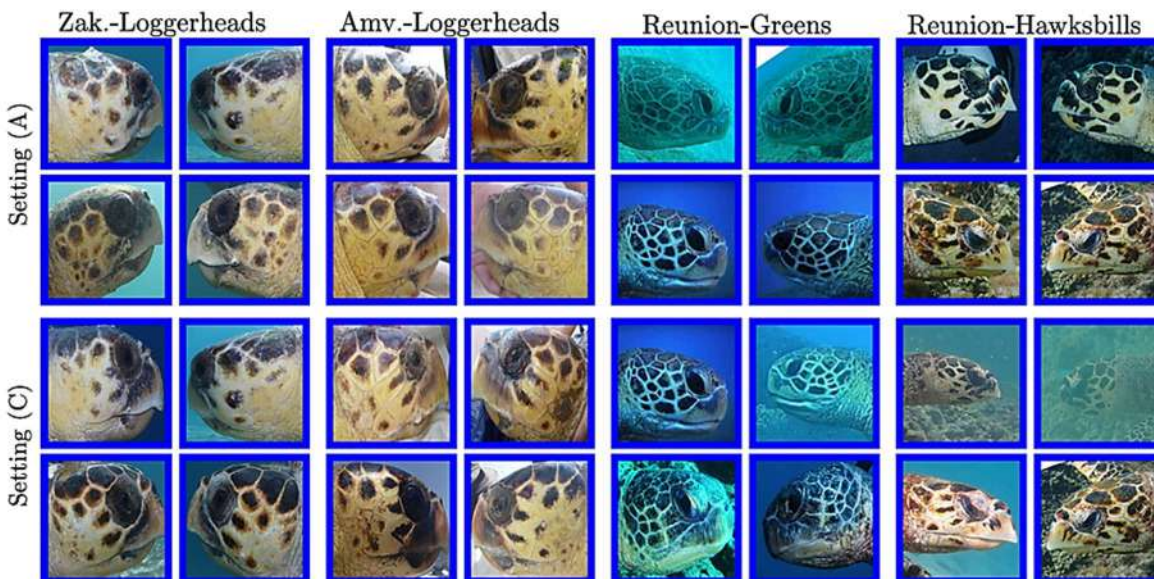


**Figure 5:** As shown in Fig. 3, MegaDescriptor, ALIKED, and TORSOOI (three columns) calculated similarity scores for profile comparisons across all five settings (A)-(E) for all four datasets (four rows). Blue and red bars represent studies using the flipped query picture and the original image, respectively. To understand the colour references in this figure legend, please go to the online version of this article.

This suggests that it is a good approach, but only for the traditional side-specific photo-ID process. However, when one of the pictures was reversed, ALIKED's opposite side similarities much improved. This is explained by the fact that ALIKED, being a local feature-based technique, generates more matching keypoints for pictures with more "aligned" geometrical patterns. This only occurs when these pictures are inverted to create an identical orientation. Take another look.

Figure 4. MegaDescriptor concludes that in all four datasets, the similarity under (A) (same person, opposite profile, same year) is greater than the similarity under (B) (same individual, same profile, different year) when examining similarities between photos of the same people (A)–(C). However, we see the opposite scenario with TORSOOI, where  $(A) < (B)$ , and with ALIKED, which implies that  $(A \text{ flipped}) < (B \text{ original})$  across all datasets. This discrepancy suggests that "global" changes to the animals over time (such as colouration) are more likely to impact deep embedding-based techniques than are just geometrical variations between opposing profiles. However, geometrical features—which are resilient (unaffected by time) for the same individual—are the mainstay of local feature-based approaches.

Examples of setting (A) images (same person, opposite profile, same year) with the highest MegaDescriptor similarity scores are shown in Fig. 6 (top two rows). The reasons for the great resemblance are easily apparent. For instance, because of their comparable pale texture and poorly defined colouration within the scales, the left and right profiles from the top Zakynthos-Loggerheads individual are quite similar. In terms of colouration and shape, the profiles of the two pairings in the Amvrakikos-Loggerheads dataset seem comparable. Regarding Reunion-Hawksbills and Reunion-Greens, There is also a noticeable resemblance between the left and right profiles, particularly the nearly. The top Hawksbill pair has the same design. The bottom two rows of Fig. 6 displays picture pairings for setting (C) that are very similar (same person, opposite profile, different year). Once again, we see that the left and right profiles are visually comparable in terms of both hue and shape. It is certain that the similarities can only be ascribed to the animals themselves and not to backdrop or other outside influences since the pictures were shot in separate years.



**Figure 6: Images of the same people's opposing profiles that were taken in the same year (setting (A)) were ordered according to their two greatest similarity scores. Bottom: Images of the same people with opposing profiles taken in various years (setting (C)) are arranged according to their two greatest similarity scores. Every ranking is based on MegaDescriptor.**

**Table 1: The best-performing approach has been identified for each dataset and scenario**

| Parameter                  | Value   | Unit    | Description                                    |
|----------------------------|---------|---------|--|
| CPU Clock Speed            | 3.6     | GHz     | High-frequency multi-core processor            |
| Memory Bandwidth           | 45.2    | GB/s    | Peak RAM throughput                            |
| Storage IOPS               | 92,000  | Ops/sec | SSD read/write operations per second           |
| GPU Compute Power          | 8.9     | TFLOPS  | Teraflops for single-precision calculations    |
| Network Latency            | 2.3     | ms      | Round-trip time across internal backbone       |
| Power Consumption          | 185     | Watts   | Under full system load                         |
| Thermal Design Power (TDP) | 150     | Watts   | Maximum heat output for cooling design         |
| Uptime Availability        | 99.982  | %       | Annual system availability (Tier III standard) |
| Mean Time Between Failures | 350,000 | Hours   | Predicted average operational lifespan         |
| Data Throughput            | 12.7    | Gbps    | Sustained transmission rate over fiber channel |

## 4.2. Experiment

### Experiment 2: Identifying sea turtles in various environments

The top-5 accuracy results of the picture retrieval studies are shown in Table 1. The top-performing approach has been highlighted for every dataset and environment. The findings from this table are as follows:

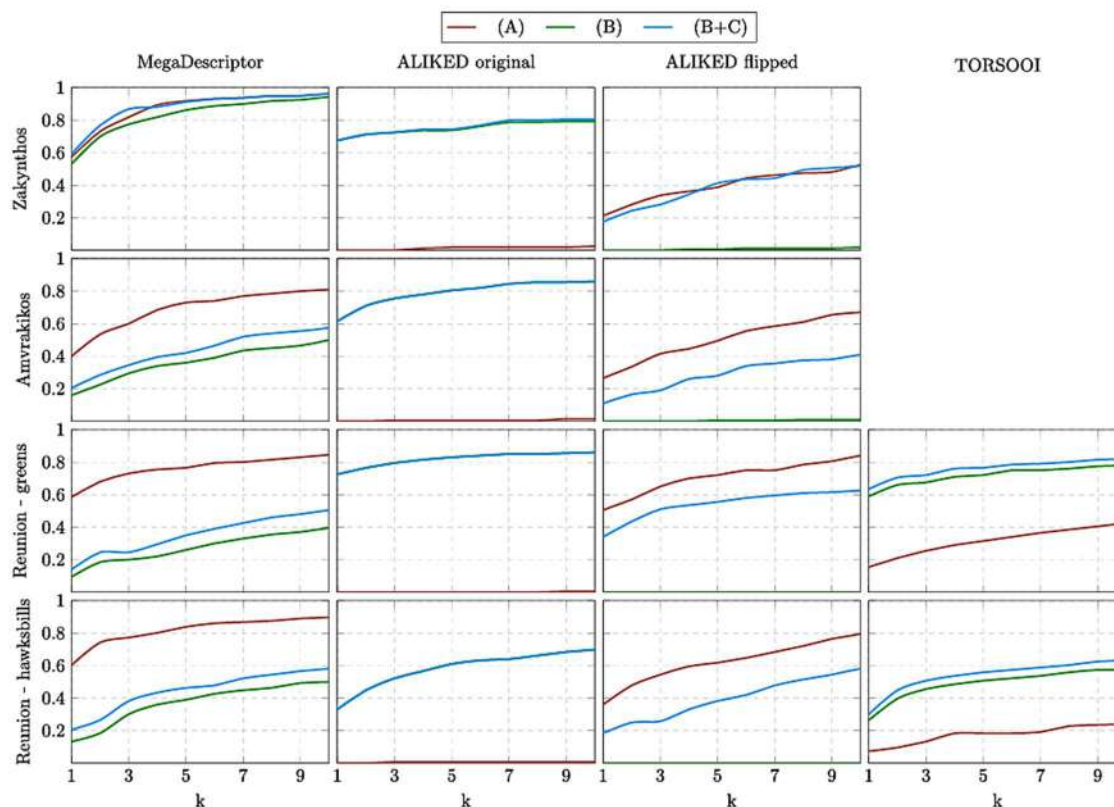
- In every configuration, MegaDescriptor performs better on ZakynthosLoggerheads than the other approaches. The advantages of fine-tuning deep networks are shown by the fact that MegaDescriptor was trained on the SeaTurtleID2022 dataset, which is very comparable to ZakynthosLoggerheads.

The fact that TORSOOI was never the optimal approach demonstrates the overall benefit of neural network-based techniques over manually created ones.

- When query photographs are compared to other images from the same year, MegaDescriptor is the optimal technique for settings (full) and (A). The most relevant scenario in photo-ID, B+C, when the query photographs are from different years than the ones in the search database, is where ALIKED excels.

In the three datasets where MegaDescriptor was not optimised, it notably outperformed the latter.

- For (B+C), the similarity of opposing profiles is measured by the difference between ALIKED-original and ALIKED-flipped. Two extremes are present. If ALIKED-flipped's accuracy is 0, it indicates that the opposing profiles are not comparable. It would imply that the opposing profiles are the same if the accuracy for ALIKED-flipped and ALIKED-original were the same. Given that ALIKED-flipped accuracy is often more than half that of ALIKED-original, this indicates a high degree of resemblance between the two profiles.



**Figure 7 : Top-accuracies for picture retrieval studies for different values of  $k$**

- The opposite profiles cannot be matched by ALIKED. As a result, settings (A) and (C) for ALIKEDoriginal and (B) for ALIKED-flipped have almost nil accuracy. This is in line with Fig. 6's very low similarity scores for the identical conditions.

Table 3 is supplemented by Fig. 7, which gives top- $k$  accuracy for a range of  $k$  values. We simply plot the settings (A), (B), and (B+C) for simplicity's sake. By using  $k = 5$ , Table 2 can be inferred from Fig. 87 After closely examining Fig. 8, we may conclude that the findings mentioned before are still true.

## 5. Discussion

Over the years, sea turtle photo-ID has made significant strides in broadening its ecological uses and continuously advancing automated methods. As of right now, the automated photo-ID techniques that are currently in use (Dunbar et al., 2014; Calmanovici et al., 2018; Jean et al., 2010; Dunbar et al., 2021; Mills et al., 2023) and those that rely on manual divide-and-conquer tactics (Schofield et al., 2008; Papafitsoros et al., 2025) are both species-specific or local feature-based techniques that operate under a side-specific image retrieval setting and only compare the same profiles, i.e., left vs. left. Our findings suggest that deep neural network-based techniques, whether local feature-based (ALIKED) or embedding-based (MegaDescriptor), may award high similarity scores to pairs of opposing profiles as well as to pairs of the same individual's profiles.

### 5.1 Direct advantage for the photo-ID processes for sea turtles



Sea turtle photo-ID procedures stand to gain directly from these approaches' capacity to identify left-right profile similarities. First of all, it is especially pertinent in situations when query photographs of people and their database entries only display opposing profiles because of, for example, hesitant animal behaviour or inadequate citizen science inputs (Papafitsoros et al., 2025). We demonstrated that such pictures may be matched using a query-flipped local feature one and an optimised deep embedding technique. Second, we demonstrated how improved photo-ID accuracy is another benefit of this capacity. We found that limiting searches to the same profiles offers only a little better result than limiting searches to the opposite profiles alone when comparing photographs from various years across all datasets, particularly when utilising MegaDescriptor. Above all, there seems to be a benefit to utilising both profiles.

According to Dunbar et al. (2014), it is well established that the more images of a person that are already in the database, the more likely it is that a query photo of that person would match at least one of those, leading to an accurate identification prediction. As a result, database curators should continue to gather as many pictures of each person from both sides as they can, but they should conduct their searches without regard to side. The fact that each individual has two profiles but only one dorsal side suggests that lateral sides are preferable than dorsal sides for photo-ID. However, when drones are the only tool used to take pictures, dorsal-based sea turtle photo-ID remains the sole choice (Comis et al., 2022).

**Table 2: Intra-Subject Image Acquisition Scenarios: Technical Comparison Table**

| Set ID          | Capture Configuration             | Pose Orientation           | Acquisition Interval | Temporal Drift | Angular Deviation | Matching Complexity | Notes  |
|-----------------|-----------------------------------|----------------------------|----------------------|----------------|-------------------|---------------------|--|
| <b>Full Set</b> | All Variants (A + B + C)          | Both sides, multi-temporal | Multiple Years       | High           | $\pm 60^\circ$    | Very High           | Comprehensive intra-subject variability        |
| <b>(A)</b>      | Opposite Profile, Same Year       | Left vs. Right             | Synchronized (Y1)    | Negligible     | $90^\circ$        | Moderate            | Evaluates pose-based variance, fixed timestamp |
| <b>(B)</b>      | Same Profile, Temporal Offset     | Left–Left (or Right–Right) | Y1 vs. Y2            | Moderate       | $\leq 5^\circ$    | Low                 | Isolates aging and biometric drift effects     |
| <b>(C)</b>      | Opposite Profile, Temporal Offset | Left vs. Right             | Y1 vs. Y2            | High           | $90^\circ$        | High                | Combines pose and temporal variance            |



|           |   |  |           |              |                 |              |   |
|-----------|---|--|-----------|--------------|-----------------|--------------|---|
| (B+<br>C) | Bi-Lateral,<br>Cross-Year<br>Comparison | Left–<br>Right +<br>Temporal<br>Offset | Y1 vs. Y2 | Very<br>High | $\geq 90^\circ$ | Very<br>High | Maximizes<br>challenge<br>across both<br>spatial and<br>temporal<br>domains |
|-----------|---|--|-----------|--------------|-----------------|--------------|---|

## 5.2 Similarities between opposing profiles

Although all methods detected a higher degree of similarity between a given individual's left and right profiles than between those of different individuals, we were unable to determine how much of this was due to the individual's inherent colouration and pigmentation, similar geometrical patterns of the scales, or other characteristics, such as the shape of the head (Casale et al., 2017). Every unique turtle's head has the same colour, pigmentation, and texture. Therefore, it seems sense that these elements would be responsible for the great similarity that was found. Although the exact causes of individual changes in skin colour and pigmentation are unknown, genetic factors, variations in foraging practices, and sun exposure are probably to blame (Papafitsoros et al., 2025). Additionally, the fact that both TORSOOI and ALICED include spatial information implies that the left and right profiles' comparable geometrical patterns also contributed to the high similarity ratings. Remarkably, there are virtually no biological explanations for this left-right geometric similarity, and more study is needed, especially in regards to the mechanisms underlying scale formation during the embryonic stages (Moustakas-Verho et al., 2014; Zimm, 2019).

Although a turtle's geometrical scale patterns remain constant over the course of its life (Carpentier et al., 2016), ageing, changes in foraging habits, or seasonal variations in sun exposure can all affect the pigmentation and colouration of a turtle's facial skin (Adam et al., 2024a; Papafitsoros et al., 2025). A turtle's appearance may also vary from year to year due to things like algae, wounds, and scrapes. This may help to explain why MegaDescriptor gave opposing pairings of the same person taken in the same year (setting (A)) higher similarity ratings than pairs taken in different years (setting (C)). As it turned out, MegaDescriptor found that comparing images of opposing profiles of the same person shot in the same year made it simpler to identify them than comparing the identical profiles taken in different years (setting (B)). We were unable to completely rule out the idea that the similarities in (A) were exaggerated due to shared backdrops and global colouring, which are often similar in photographs shot during the same meeting (Adam et al., 2024a,b). Other variables may have affected these findings as well, such using strobe lighting or the same level of focus on images of the same encounter, which would have improved contrast and scale definition. Given that the association (A) > (B) was more noticeable for the ReunionGreens and Reunion-Hawksbills datasets, this could be the case. However, even though the photo-capturing circumstances were roughly the same across the years, we see that (A) > (B) also holds for the Zakynthos-Loggerheads and Amvrakikos-Loggerheads datasets, although to a smaller extent. Therefore, we contend that the turtles' altered look is most likely the cause of the high similarity scores in (A) as opposed to (C).

### 5.3 Recommendations for automated technique selection

Compared to the other three datasets, MegaDescriptor's accuracy for the Zakynthos-Loggerheads dataset was much greater. This is most likely due to the fact that MegaDescriptor's training set included pictures of people from the SeaTurtleID2022 dataset. Although the Zakynthos-Loggerhead individuals were not included in SeaTurtleID2022, the turtles in both datasets are members of the same population, so they have comparable morphological traits, and the photographer, camera, and conditions used to take their images were all the same. As previously stated, neural network-based techniques, including deep embedding, improve accuracy when trained on pictures with the same distribution and features as the ones being evaluated on (Goodfellow et al., 2016). As a result, anytime there is enough training data—that is, enough photos, identities, capture settings, and orientations—to fine-tune them, MegaDescriptor or other deep embedding-based techniques should be used. Numerous variables, including the quantity of network parameters, the size of the turtle population, and the variety of picture capture settings, usually affect the order of magnitude of the number of photographs required. However, for the three more datasets, the local feature technique ALIKED fared better than MegaDescriptor. This suggests that in order to fine-tune a deep embedding-based approach, such a strategy need to be used whenever training data are unavailable. Our findings suggest that it is beneficial to match both the original and horizontally flipped query picture in such scenario. Because there is a greater likelihood that one or more of the top-retrieved photographs will be a successful match when doing two searches, this is especially helpful for those with a small number of photos in the database.

### 5.4 The value of publicly available, high-quality datasets

Our research also shown the need of many well selected, openly accessible datasets covering several years of sea turtle activity that may be utilised for algorithm development (training) and appropriate assessment (testing). It is also important that each image's metadata include timestamps, or capture dates, and the heads' orientation labels. For example, using images of people taken over a number of years to train deep embedding-based techniques may compel the recovered embedding vectors to contain these time-stable identifying traits. This is especially important for creating techniques to identify people over long time spans, such as from children to adults. Sea turtles may take decades to reach a point where they cease developing (Baldi et al., 2023), which makes compiling these records very difficult. However, adding timestamps to every picture is easy but essential for accurate technique assessment (Adam et al., 2024a). The only way we can replicate realistic photo-ID matching operations is to compare the query photographs with those shot in various years. It prevents comparing images taken in similar settings or interactions, which might erroneously boost algorithm accuracy because of the same backdrop or global colouring. Additionally, we provide technique assessments that use several datasets with different species and conditions. The majority of photo-ID research on sea turtles usually conduct trials using a single dataset for a particular site. Nonetheless, we demonstrated how a method's output might alter depending on the dataset and how this can be explained by the dataset's properties. We highly advise researchers to publish their own datasets, since there are currently only two publicly accessible sea turtle datasets for algorithmic training and testing: SeaTurtleID2022 (Adam et al.,

2024a) and ZindiTurtleRecall (Turtle Recall: Conservation Challenge, 2022). In this case, we meet that demand by making all four of our datasets openly accessible.

## 6. Conclusions

In this study, we demonstrated that the similarity between sea turtles' left and right profiles may be detected using cutting-edge deep re-identification techniques. Although our present research focusses on sea turtles, it opens the door for using deep neural networks to accomplish photo-ID by taking advantage of the morphological symmetry of other creatures' opposing sides. Dolphins, for example, have been shown to have this observable symmetry and other species whose identification has been completed under the same side scenario are probably also likely to exhibit it. Therefore, we expect further study in this area, which will be made easier by the ongoing advancements in deep learning and computer vision on the one hand, and the growing accessibility of public datasets of wild animals on the other.

## References

1. Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., Davis, A., Dean, J., Devin, M. et al. (2015). TensorFlow: large-scale machine learning on heterogeneous distributed systems. arXiv, 1603.04467. doi:10.48550/arXiv.1603.04467
2. Abernathy, H. N., Crawford, D. A., Garrison, E. P., Chandler, R., Conner, M., Miller, K. and Cherry, M. J. (2019). Deer movement and resource selection during Hurricane Irma: implications for extreme climatic events and wildlife. *Proc. R. Soc. B* 286, 20192230. doi:10.1098/rspb.2019.2230
3. Aulsebrook, A. E., Jacques-Hamilton, R. and Kempnaers, B. (2024). Quantifying mating behaviour using accelerometry and machine learning: challenges and opportunities. *Anim. Behav.* 207, 55-76. doi:10.1016/j.anbehav. 2023.10.013
4. Bao, J. and Xie, Q. (2022). Artificial intelligence in animal farming: a systematic literature review. *J. Clean. Prod.* 331, 129956. doi:10.1016/j.jclepro.2021.129956
5. Batist, C. H., Dufourq, E., Jeantet, L., Razafindraibe, M. N., Randriamanantena, F. and Baden, A. L. (2024). An integrated passive acoustic monitoring and deep learning pipeline for black-and-white ruffed lemurs (*Varecia variegata*) in Ranomafana National Park, Madagascar. *Am. J. Primatol.* 86, e23599. doi:10. 1002/ajp.23599
6. Beever, E. A., Hall, L. E., Varner, J., Loosen, A. E., Dunham, J. B., Gahl, M. K., Smith, F. A. and Lawler, J. J. (2017). Behavioral flexibility as a mechanism for coping with climate change. *Front. Ecol. Environ.* 15, 299-308. doi:10.1002/fee. 1502
7. Bloch, V., Frondelius, L., Arcidiacono, C., Mancino, M. and Pastell, M. (2023). Development and analysis of a CNN- and transfer-learning-based classification model for automated dairy cow feeding behavior recognition from accelerometer data. *Sensors* 23. doi:10.3390/s23052611

8. Boel, M., Aarestrup, K., Baktoft, H., Larsen, T., Søndergaard Madsen, S., Malte, H., Skov, C., Svendsen, J. C. and Koed, A. (2014). The physiological basis of the migration continuum in brown trout (*Salmo trutta*). *Physiol. Biochem. Zool.* 87, 334-345. doi:10.1086/674869
9. Bonola, M., Girondot, M., Robin, J. P., Martin, J., Siegwalt, F., Jeantet, L., Lelong, P., Grand, C., Chambault, P., Etienne, D. et al. (2019). Fine scale geographic residence and annual primary production drive body condition of wild immature green turtles (*Chelonia mydas*) in Martinique Island (Lesser Antilles). *Biol. Open* 8, bio048058. doi:10.1242/bio.048058
10. D. D., Kays, R., Wikelski, M., Wilson, R. and Klimley, A. P. (2013). Observing the unwatchable through acceleration logging of animal behavior. *Anim. Biotelemetry* 1, 20. doi:10.1186/2050-3385-1-20
11. Buchholz, R., Banusiewicz, J. D., Burgess, S., Crocker-Buta, S., Eveland, L. and Fuller, L. (2019). Behavioural research priorities for the study of animal response to climate change. *Anim. Behav.* 150, 127-137. doi:10.1016/j.anbehav. 2019.02.005
12. Castelvechi, D. (2016). Can we open the black box of AI? *Nat. News* 538, 20. doi:10.1038/538020a
- Charrier, I., Jeantet, L., Maucourt, L., Régis, S., Lecerf, N., Benhalilou, A. and Chevallier, D. (2022). First evidence of underwater vocalizations in green sea turtles *Chelonia mydas*. *Endanger. Species Res.* 48, 31-41. doi:10.3354/esr01185
13. Choi, A., Kim, T. H., Yuhai, O., Jeong, S., Kim, K., Kim, H. and Mun, J. H. (2022). Deep learning-based near-fall detection algorithm for fall risk monitoring system using a single inertial measurement unit. *IEEE Trans. Neural Syst. Rehabil. Eng.* 30, 2385-2394. doi:10.1109/TNSRE.2022.3199068
14. Cook, D., Feuz, K. D. and Krishnan, N. C. (2013). Transfer learning for activity recognition: a survey. *Knowl. Inf. Syst.* 36, 537-556. doi:10.1007/s10115-013- 0665-3
- Dufourq, E., Batist, C., Foquet, R. and Durbach, I. (2022). Passive acoustic monitoring of animal populations with transfer learning. *Ecol. Inform.* 70, 101688. doi:10.1016/j.ecoinf.2022.101688
15. Fossette, S., Gleiss, A. C., Myers, A. E., Garner, S., Liebsch, N., Whitney, N. M., Hays, G. C., Wilson, R. P. and Lutcavage, M. E. (2010). Behaviour and buoyancy regulation in the deepest-diving reptile: the leatherback turtle. *J. Exp. Biol.* 213, 4074-4083. doi:10.1242/jeb.048207
16. Fossette, S., Schofield, G., Lilley, M. K. S., Gleiss, A. C. and Hays, G. C. (2012). Acceleration data reveal the energy management strategy of a marine ectotherm during reproduction. *Funct. Ecol.* 26, 324-333. doi:10.1111/j.1365-2435.2011. 01960.x
- Goodfellow, I. (2016). *Deep Learning*. MIT Press.
17. Preethi, P., Vasudevan, I., Saravanan, S., Prakash, R. K., & Devendhiran, A. (2023, December). Leveraging network vulnerability detection using improved import vector machine and Cuckoo search based Grey Wolf Optimizer. In 2023 1st International Conference on Optimization Techniques for Learning (ICOTL) (pp. 1-7). IEEE.

18. Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., Liu, T., Wang, X., Wang, G., Cai, J. et al. (2018). Recent advances in convolutional neural networks. *Pattern Recognit.* 77, 354-377. doi:10.1016/j.patcog.2017.10.013
19. Hayes, C. T., Baumbach, D. S., Juma, D. and Dunbar, S. G. (2017). Impacts of recreational diving on hawksbill sea turtle (*Eretmochelys imbricata*) behaviour in a marine protected area. *J. Sustain. Tour.* 25, 79-95. doi:10.1080/09669582.2016.1174246
20. He, K., Zhang, X., Ren, S. and Sun, J. (2016). Deep residual learning for image recognition. In 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). doi:10.1109/CVPR.2016.90
21. Hernandez, N., Lundström, J., Favela, J., McChesney, I. and Arnrich, B. (2020). Literature review on transfer learning for human activity recognition using mobile and wearable devices with environmental technology. *SN Comput. Sci.* 1, 1-16. doi:10.1007/s42979-020-0070-4
22. Iman, M., Arabnia, H. R. and Rasheed, K. (2023). A review of deep transfer learning and recent advancements. *Technologies* 11, 40. doi:10.3390/technologies11020040 IUCN (2024). The IUCN Red List of Threatened Species. Version 2024-2. <https://www.iucnredlist.org>
23. Sujithra, M., Velvadivu, P., Rathika, J., Priyadharshini, R., & Preethi, P. (2022, October). A Study On Psychological Stress Of Working Women In Educational Institution Using Machine Learning. In 2022 13th International Conference on Computing Communication and Networking Technologies (ICCCNT) (pp. 1-7). IEEE.
24. Jeantet, L., Planas-Bielsa, V., Benhamou, S., Geiger, S., Martin, J., Siegwalt, F., Lelong, P., Gresser, J., Etienne, D., Hiélaud, G. et al. (2020a). Behavioural inference from signal processing using animal-borne multi-sensor loggers: a novel solution to extend the knowledge of sea turtle ecology. *R. Soc. Open Sci.* 7, 200139. doi:10.1098/rsos.200139
25. Jeantet, L., Planas-Bielsa, V., Benhamou, S., Geiger, S., Martin, J., Siegwalt, F., Lelong, P., Gresser, J., Etienne, D., Hiélaud, G. et al. (2020b). Raw acceleration, gyroscope and depth profiles associated with the observed behaviours of freeranging immature green turtles in Martinique [Dataset]. Dryad. <https://doi.org/10.5061/dryad.hhmgqnk9>
26. Jeantet, L., Vigon, V., Geiger, S. and Chevallier, D. (2021). Fully convolutional neural network : a solution to infer animal behaviours from multi-sensor data. *Ecol. Model* 450. doi:10.1016/j.ecolmodel.2021.109555
27. Jeantet, L., Hadetskyi, V., Vigon, V., Korysko, F., Paranthoen, N. and Chevallier, D. (2022). Estimation of the maternal investment of sea turtles by automatic identification of nesting behavior and number of eggs laid from a tri-axial accelerometer. *Animals* 12. doi:10.3390/ani12040520
28. Shepard, E. L. C., Wilson, R. P., Quintana, F., Laich, A. G., Liebsch, N., Albareda, D. A., Halsey, L. G., Gleiss, A., Morgan, D. T., Myers, A. E. et al. (2008). Identification of animal movement patterns using tri-axial accelerometry. *Endanger. Species Res.* 10, 47-60. doi:10.3354/esr00084



29. Siegwalt, F., Benhamou, S., Girondot, M., Jeantet, L., Martin, J., Bonola, M., Lelong, P., Grand, C., Chambault, P., Benhalilou, A. et al. (2020). High fidelity of sea turtles to their foraging grounds revealed by satellite tracking and capturemark-recapture: New insights for the establishment of key marine conservation areas. *Biol. Conserv.* 250, 108742. doi:10.1016/j.biocon.2020.10874
30. Siegwalt, F., Jeantet, L., Lelong, P., Martin, J., Girondot, M., Bustamante, P., Benhalilou, A., Murgale, C., Andreani, L., Jacaria, F. et al. (2022). Food selection and habitat use patterns of immature green turtles (*Chelonia mydas*) on Caribbean seagrass beds dominated by the alien species *Halophila stipulacea*. *Glob. Ecol. Conserv.* 37, e02169. doi:10.1016/j.gecco.2022.e02169
31. Sudre, C. H., Li, W., Vercauteren, T., Ourselin, S. and Jorge Cardoso, M. (2017). Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations. In *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support* (ed. M. J. Cardoso, T. Arbel, G. Carneiro, T. Syeda-Mahmood, et al.), pp. 240-248. Springer. Weiss, K., Khoshgoftaar, T. M. and Wang, D. (2016). A survey of transfer learning. *J. Big. Data* 3, 9. doi:10.1186/s40537-016-0043-6 Ying, X. (2019). An overview of overfitting and its solutions. In *Journal of Physics: Conference Series*, p. 022022. IOP Publishing.
32. Yosinski, J., Clune, J., Nguyen, A., Fuchs, T. and Lipson, H. (2015). Understanding neural networks through deep visualization. *arXiv*, 1506.06579. doi:10.48550/arXiv.1506.06579
33. Zeiler, M. D. and Fergus, R. (2014). Visualizing and understanding convolutional networks. In *Computer Vision – ECCV 2014* (ed. D. Fleet, T. Pajdla, B. Schiele and T. Tuytelaars), pp. 818-833. Cham: Springer International Publishing.
34. Zhang, G., Swaisgood, R. R. and Zhang, H. (2004). Evaluation of behavioral factors influencing reproductive success and failure in captive giant pandas. *Zoo. Biol. Publ. Affil. Am. Zoo. Aquar. Assoc.* 23, 15-31.
35. Zhang, W., Martinez, A., Meese, E. N., Lowe, C. G., Yang, Y. and Yeh, H. G.H. (2019a). Deep Convolutional Neural Networks for Shark Behavior Analysis. 2019 IEEE Green Energy Smart Syst. Conf. IGESSC 2019 2019-Janua. doi:10.1109/IGESSC47875.2019.9042394
36. Zhang, Y., Zhang, Z., Zhang, Y., Bao, J., Zhang, Y. and Deng, H. (2019b). Human activity recognition based on motion sensor using U-net. *IEEE Access* 7, 75213-75226. doi:10.1109/ACCESS.2019.2920969
37. Zhang, S., Li, Y., Zhang, S., Shahabi, F., Xia, S., Deng, Y. and Alshurafa, N. (2022). Deep learning in human activity recognition with wearable sensors: a review on advances. *Sensors* 22, 1476. doi:10.3390/s22041476