



OPEN

## Deep learning-based image classification of turtles imported into Korea

Jong-Won Baek<sup>1,2</sup>, Jung-Il Kim<sup>1,2</sup> & Chang-Bae Kim<sup>1</sup>✉

Although turtles play a key role in maintaining healthy and balanced environments, these are endangered due to global trade to meet the high demand for food, medicine, and pets in Asia. In addition, imported non-native turtles have been controlled as alien invasive species in various countries, including Korea. Therefore, a rapid and accurate classification of imported turtles is needed to conserve and detect those in native ecosystems. In this study, eight Single Shot MultiBox Detector (SSD) models using different backbone networks were used to classify 36 imported turtles in Korea. The images of these species were collected from Google and were identified using morphological features. Then, these were divided into 70% for training, 15% for validation, and 15% for test sets. In addition, data augmentation was applied to the training set to prevent overfitting. Among the eight models, the Resnet18 model showed the highest mean Average Precision (mAP) at 88.1% and the fastest inference time at 0.024 s. The average correct classification rate of 36 turtles in this model was 82.8%. The results of this study could help in management of the turtle trade, specifically in improving detection of alien invasive species in the wild.

Turtles, order Testudines, are major biodiversity components of the ecosystems they inhabit, and often serve as keystone species beneficial to other animals and plants<sup>1,2</sup>. However, most turtles are endangered due to indiscriminate global trade<sup>2,3</sup>. Turtles have traditionally been used for food, medicine, and pets in East and Southeast Asia<sup>4–6</sup>. Consequently, Asia, including China, Hong Kong, Japan, Taiwan, and Korea, imports and exports large quantities of live turtles<sup>7,8</sup>. The annual trade of live turtles in Asia has exceeded 13,000 metric tons, the highest number expected to be collected in the wild<sup>8</sup>. According to the National Institute of Biological Resources (NIBR), the turtle was the second most commonly imported taxon among endangered species that were imported into Korea between 2009 and 2014<sup>9</sup>. In addition, the demand for turtles in Asia has increased in recent decades<sup>8</sup>. In order to satisfy the rising demand for food, medicines, and pets, the illegal trade of turtles has increased<sup>3</sup>. In particular, due to pet preference, the illegal trade of rare turtle species, such as *Cuora flavomarginata*, increases<sup>10,11</sup>. Various international conventions and conservation bodies, including the Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES) and the International Union for the Conservation of Nature and Natural Resources (IUCN), seek to protect turtles from extinction and illegal trade.

The wildlife trade has been known to be a source of alien invasive species<sup>12–14</sup>. These negatively affect native species through predation and competition, diffuse pathogens, and modify the functioning of the ecosystem and abiotic features of environments<sup>13,15,16</sup>. In particular, turtles have a greater impact on the native population than other animals as they are long-lived organisms which can survive for decades in suboptimal habitats where environmental characteristics are suitable for breeding<sup>17,18</sup>. The introduction of alien turtles into the wild has continued due to abandonment of breeding as pets, religious release, and careless management<sup>11,19,20</sup>. Alien invasive turtles have been reported in various countries. Since 1997, the European Union has prohibited the import of slider turtle, *Trachemys scripta elegans*, because it competes for food and basking spots with the threatened European pond turtle, *Emys orbicularis*<sup>21,22</sup>. Native to the Americas, *Chelydra* and *Macrochelys*, are found in the Yangtze River basin in China<sup>23</sup>. The snapping turtle, *Chelydra serpentina*, was imported to Japan from the American continent as a pet, but some have been and continue to be abandoned in the wild<sup>24</sup>. In Korea, invasive turtles have also been reported, where non-native turtles such as *Graptemys ouachitensis* and *Macrochelys temminckii* have been observed<sup>25,26</sup>. These are now included on the NIBR's list of imported turtle species<sup>9</sup>. In order to preserve native ecosystems, it is essential to detect and classify these non-native turtles coming no through the pet trade.

To identify which species are being traded, species identification based on morphological features has been used as a standard and effective method<sup>27–29</sup>. However, this requires skilled experts to identify targeted wildlife<sup>28</sup>,

<sup>1</sup>Department of Biotechnology, Sangmyung University, Seoul 03016, Korea. <sup>2</sup>These authors contributed equally: Jong-Won Baek and Jung-Il Kim. ✉email: evodevo@smu.ac.kr

whose numbers have declined in recent years<sup>30</sup>. In this regard, the identification of turtles, particularly the genera *Chelonoidis* and *Geochelone*, is not an easy task due to similar morphological features between species<sup>31,32</sup>. Additionally, the skeletal traits used to identify turtles can be altered by ecological factors, including food and behavior<sup>33,34</sup>. These challenges make the identification of turtles based on their morphological features difficult.

Early Detection and Rapid Response (EDRR) are recognized as a set of actions that increase the chances of invasive species containment and eradication before they enter irreversible stages of the invasion curve<sup>35</sup>. Currently, the detection of invasive species through the camera trap method has been widely conducted for EDRR<sup>36</sup>. However, this method has become more challenging due to the manual analysis used to identify species based on morphology using a large image dataset<sup>37</sup>. Therefore, rapid and accurate tools are needed to identify species among the large quantities imported via the turtle trade.

Molecular methods, such as DNA barcoding, have been suggested as tools to control the wildlife trade and manage invasive species<sup>38–41</sup>. Current approaches to DNA barcoding necessitate laboratory analysis and require significant time and expertise to examine samples<sup>36,42</sup>. Therefore, a rapid and accurate method is required to detect imported turtles based on morphological features.

Image classification based on deep learning is expected to improve the ability to identify traded wildlife and detect alien invasive species<sup>43,44</sup>. This method can reduce personnel and time spent identifying species in the wildlife trade<sup>44</sup>. Moreover, this can be used to verify the precision of species occurrence data, leading to improved cost-effectiveness in data administration and precise insights for decision support tools designed to detect alien invasive species<sup>36,37</sup>. A Convolutional Neural Network (CNN) is a deep learning model developed for the classification of image data. Object detection models are developed based on CNNs and consider classification and regression, allowing for the prediction of objects in images. These models have been widely applied to the automatic classification of species of various organisms<sup>45–47</sup>. The object detection model can be divided into two and one-stage detectors. A two-stage detector, such as Faster-R-CNN<sup>48</sup>, learns regression and classification independently and continuously, whereas a one-stage detector, such as Single Shot MultiBox Detector (SSD)<sup>49</sup>, learns regression and classification simultaneously. These characteristics imply that one-stage detectors are better suited for real-time species classification of large quantities of wildlife trade<sup>47</sup>.

In the present study, object detection models, SSD using eight different CNNs as backbone networks, were assessed to classify turtles imported into Korea. The performance of the eight models was evaluated by the mean Average Precision (mAP) and inference time it takes to process one image. The model test results, which showed the highest mAP, are presented as a confusion matrix.

## Results

### Performance of the eight models

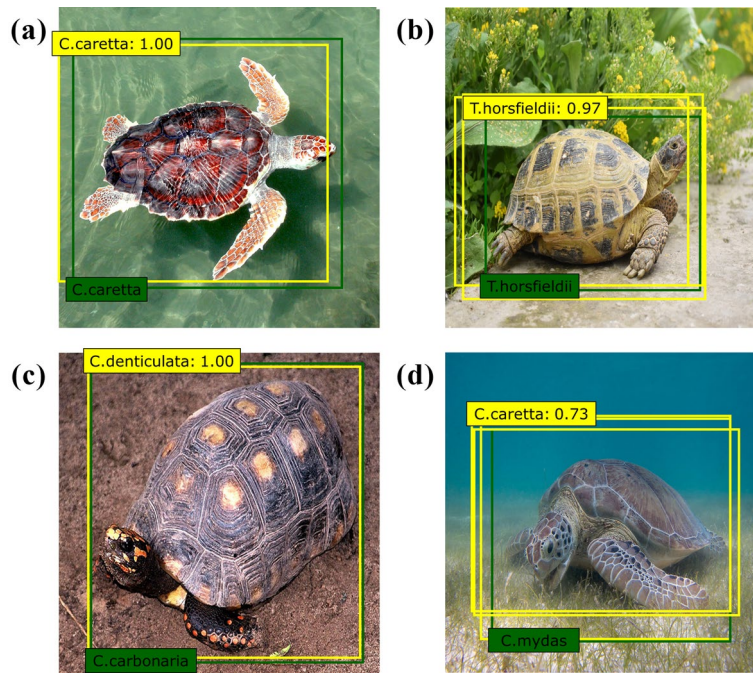
Figure 1 depicts representative prediction results from the eight models. There were four cases of prediction results of the eight models. First, one prediction bounding box was predicted and classified correctly (Fig. 1a). Second, multiple prediction bounding boxes were predicted, and the classification result with the highest confidence score provided by the model was determined to be correct (Fig. 1b). Third, a prediction bounding box was predicted and misclassified (Fig. 1c). Fourth, multiple prediction bounding boxes were expected, and it was determined that the classification result with the highest confidence score provided by the model was incorrect (Fig. 1d). Performances of the eight SSD models included in different CNN backbone networks are shown in Fig. 2 and Table S1. The mAP of the models varied based on the CNN used as the backbone network (Fig. 2). Among the three ResNet models, the ResNet18 model had the highest mAP, while the ResNet50 model had the lowest mAP (Fig. 2). The DenseNet30 model had the lowest mAP among the four DenseNet models, while the DenseNet18 model had the highest mAP (Fig. 2). Among all training results, the highest mAP was 88.1% of the ResNet18 model, and the lowest was 77.4% of the VGGNet16 model (Table S1). In the VGGNet16 model, which showed the lowest mAP, *Chelodina mccordi* presented the lowest AP at 55.0%, and *Clemmys guttata* presented the highest AP at 97.0% (Table S2). *Chelonia mydas* had the lowest AP (69.8%), while that of *Stigmochelys pardalis* was the highest (97.6%) in the ResNet18 model, which showed the highest mAP (Table S2). The inference time ranged from 0.024 s for the ResNet18 model to 0.052 s for the DenseNet121 model (Table S1).

### Classification results of the model which showed the highest mAP

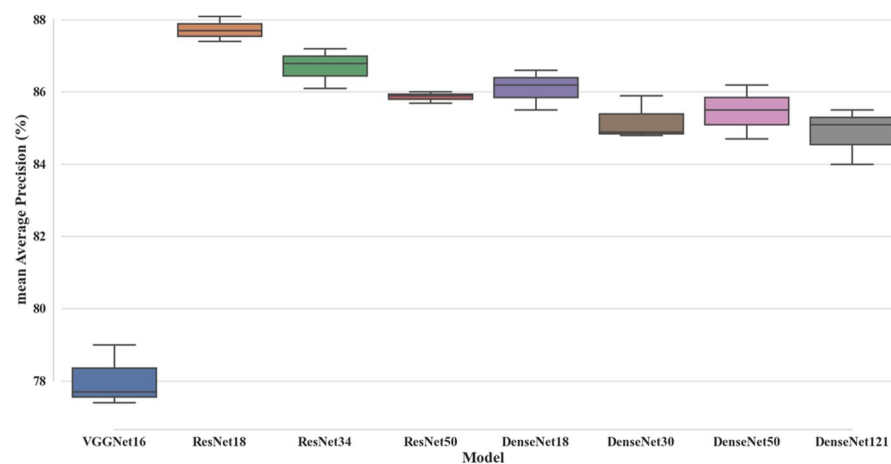
The confusion matrix displays the classification results of the ResNet18 model, which had the highest mAP (Table 1). This model correctly classified 36 turtles with an average accuracy of 82.8%, ranging from 60.7% for *Chelonia mydas* to 100.0% for *Stigmochelys pardalis* (Table 1). Nine cases showed a more than 10.0% misclassification rate (Table 1). At 26.0%, the classification of *Geochelone elagans* as *Geochelone platynota* was the model's most incorrect classification. Similarly, 25.0% of test images of *Geochelone platynota* were incorrectly classified as *Geochelone elagans*. The misclassification rate of *Chelonoidis chilensis* as *Geochelone sulcata* was 16.7%. The incorrect classification of *Chelonia mydas* as *Caretta caretta* and *Eretmochelys imbricata* was 14.3%, respectively. *Chelonoidis denticulata* and *Caretta caretta* were misclassified as *Chelonoidis carbonaria* and *Chelonia mydas* at 12.5%, respectively. 10.7% of *Cuora amboinensis* were misclassified as *Mauremys sinensis*. Additionally, *Chelonia mydas* had the lowest correct classification rate due to its misclassification as *Caretta caretta* and *Eretmochelys imbricata*.

## Discussion

The models using ResNet and DenseNet as backbone networks exhibited higher mAP than the using VGGNet models (Fig. 2). The difference may be attributable to the more complex architectures of the ResNet and DenseNet models compared to that of the VGGNet model. ResNet models enhanced network performance by resolving the degradation issue inherent to the VGGNet model used in previous layers by utilizing a skip connection that can



**Figure 1.** Representative images of four prediction result cases by the models. The yellow and green boxes on the images represent the prediction boxes and the ground-truth bounding boxes, respectively. The values in the yellow boxes of the images are confidence scores suggested by the models, indicating the probability that the prediction is correct. (a) Image of *Caretta caretta*, where one prediction bounding box was predicted and classified correctly, (b) Image of *Testudo horsfieldii*, where multiple prediction bounding boxes were predicted, and the classification result with the highest confidence score provided by the model was correct, (c) Image of *Chelonoidis carbonaria*, where one prediction bounding box was predicted and classified incorrectly, (d) Image of *Chelonia mydas*, where multiple prediction bounding boxes were predicted, and the classification result was incorrect. Photo credit: (b) Amirekul, (c) Dick Culbert, (d) P.Lindgren.



**Figure 2.** Comparison of the mean Average Precision (mAP) of the eight SSD models for classifying 36 turtles.

jump over layers and add features<sup>50</sup>. DenseNet models connect all layers directly and reuse all previous layers' features to maximize information delivery<sup>51</sup>. In addition, a direct correlation was discovered between inference time and model complexity, with the exception of the VGGNet16 model (Table S1). The more layers a model had, the longer it took to make inferences. The DenseNet models showed a slower inference time than those using ResNet as backbone networks.

The ResNet18 model showed the highest mAP at 88.1% among the eight models examined in this study (Fig. 2), while nine of the misclassification results from the ResNet18 model showed a rate of more than 10.0% (Table 2). Four of these nine misclassification cases could be due to the morphological similarity between species. Representative images of these misclassification cases are presented in Fig. 3. Morphological identification

		Predicted results																																							
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36				
True results	1	91.2	-	2.9	-	-	-	-	-	-	-	-	5.9	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-			
	2	8.7	69.6	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	4.4	8.7	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	4.4			
	3	4.2	-	79.2	12.5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	4.2		
	4	-	-	14.3	60.7	14.3	-	-	-	-	3.6	-	3.6	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	3.6		
	5	-	-	4.2	8.3	83.3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		
	6	-	-	-	2.9	2.9	77.1	-	-	-	-	-	-	-	-	-	-	-	5.7	2.9	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2.9		
	7	-	-	-	-	-	-	90.9	-	-	-	3.0	-	-	-	-	3.0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
	8	-	-	-	-	-	-	-	84.0	-	-	-	-	-	-	-	-	-	-	12.0	4.0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
	9	-	-	-	-	-	-	-	-	88.9	-	2.8	-	-	-	-	-	2.8	2.8	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2.8	
	10	-	-	-	-	-	-	-	-	-	94.7	2.6	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
	11	-	-	-	-	-	-	-	-	-	5.7	94.3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
	12	-	-	-	-	-	-	-	-	-	-	82.1	3.6	-	-	-	-	-	10.7	3.6	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
	13	-	-	-	-	-	-	-	-	-	-	4.0	76.0	4.0	-	-	-	4.0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
	14	-	-	-	-	-	-	-	-	-	-	-	4.2	87.5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
	15	-	-	-	-	-	-	-	-	-	-	-	-	-	91.7	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
	16	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	83.3	-	4.2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	4.2	
	17	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2.6	89.5	2.6	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2.6	
	18	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	3.5	3.5	3.5	75.9	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	3.5	
	19	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	91.5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2.1	
	20	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	96.4	-	3.6	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
	21	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	75.0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	4.2	
	22	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	81.3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	3.1
	23	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	3.7	-	96.3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	24	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	73.3	26.7	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	25	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	25.0	67.9	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	26	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	65.2	8.7	-	-	-	-	-	-	-	-	-	-	-	-	4.4
	27	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	91.3	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	28	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	75.0	8.3	-	-	-	-	-	-	-	-	-	-	4.2
	29	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
	30	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
	31	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
	32	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
	33	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	34	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	35	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	36	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

**Table 1.** The confusion matrix of the ResNet18 model, showing the highest mean average precision (mAP). The numbers 1 to 36 indicate the 36 turtles, as shown in Table 1. The rows contain the actual species, while the columns contain the species predicted by the models. The prediction results for the models are shown as percentage values, while the zero values are presented as a Endash. The correct predictions are shaded in blue, and the incorrect predictions are shaded in red; the deeper the blue, the higher the value of the correct prediction, and the deeper the red, the higher the value of the incorrect prediction.

between *Geochelone elegans* and *Geochelone platynota*, which had the highest misclassification rate by the model, has been known to be challenging (Figs. 3a and b)<sup>32</sup>. These species have star-like patterns on their shells, and their carapace is dark brown with light yellow radiating markings<sup>32,52</sup>. However, the color of the top of the head makes *Geochelone elegans* and *Geochelone platynota* distinguishable. Although *Geochelone elegans* has a yellow and black top of the head with small scales, *Geochelone platynota* has a primarily yellow top of the head with large scales<sup>32,52</sup>. In order to increase the classification accuracy between these species, images that showed morphological differences between the two species should be added to train the models. Additionally, although *Chelonoidis chilensis* and *Geochelone sulcata* belong to different genera, the misclassification of these could be due to morphological similarity (Fig. 3c and d). Furthermore, these species are similar due to their black annuli with brown–yellow lines on the carapace<sup>52,53</sup>. The convex scutes on the carapace and the serrated anterior and posterior margins of the shell of *Geochelone sulcata* make this species distinguishable from *Chelonoidis chilensis*<sup>52,53</sup>. Therefore, the collection of images of *Geochelone sulcata* that showed the shape of scutes and shell’s margins should increase this species’ correct classification rate. *Chelonoidis carbonaria* is difficult to classify because its morphological characteristics are very similar to those of *Chelonoidis denticulata* (Fig. 3e and f)<sup>31</sup>. Both species lack a cervical scute on their carapace. Their carapaces are black with yellow or orange patches. However, these species can be distinguished based on head characteristics<sup>52,53</sup>. In contrast, while the head of *Chelonoidis carbonaria* is yellow to red with one large frontal scale, the head of *Chelonoidis denticulata* has multiple frontal scales and large yellow or orange scales at the top of the head<sup>52,53</sup>.

The classifications between three sea turtles, *Chelonia mydas*, *Caretta caretta*, and *Eretmochelys imbricata*, were not clear, despite the distinguishable morphological features of these species (Table 1). A pair of prefrontal scales and one claw per limb make *Chelonia mydas* distinguishable from the other two sea turtles<sup>52,53</sup>. The misclassifications among these species could be due to the underwater images used to train the models. The use of underwater images presents challenges such as color bias, unclear image quality, and image quality degradation<sup>54</sup>. Methods for enhancing underwater image have been developed, including the histogram-based method, which can improve the visual appearance by manipulating pixels directly without regard to the imaging principle<sup>55,56</sup>. In future studies, methods for enhancing underwater image quality should be applied. In order to improve the classification accuracy of these sea turtles, images depicting their morphological characteristics should be added to the training data for the classification models. The misclassification of *Graptemys ouachitensis* as *Mauremys sinensis* and *Cuora amboinensis* as *Mauremys sinensis* could be due to the lack of images to train the models. *Graptemys ouachitensis* and *Mauremys sinensis* can be easily distinguished by small yellow spots on the upper and lower jaws of *Graptemys ouachitensis*<sup>52</sup>. *Cuora amboinensis* can be distinguished from *Mauremys sinensis* by

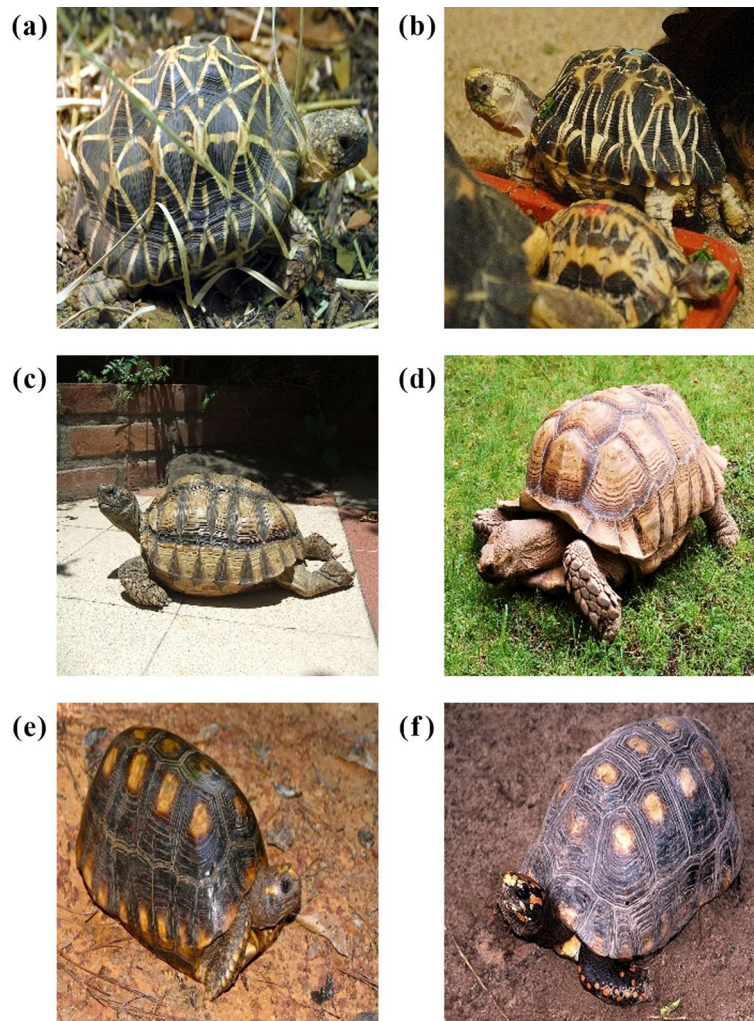
Number	Family	Species	Training set	Validation set	Test set
1	Carettochelyidae	<i>Carettochelys insculpta</i>	156	33	34
2	Chelidae	<i>Chelodina mccordi</i>	105	22	23
3	Cheloniidae	<i>Caretta caretta</i>	109	23	24
4	Cheloniidae	<i>Chelonia mydas</i>	126	27	28
5	Cheloniidae	<i>Eretmochelys imbricata</i>	106	22	24
6	Cheloniidae	<i>Macrochelys temminckii</i>	158	33	35
7	Emydidae	<i>Clemmys guttata</i>	150	32	33
8	Emydidae	<i>Graptemys ouachitensis</i>	109	23	25
9	Emydidae	<i>Malaclemys terrapin</i>	164	35	36
10	Emydidae	<i>Terrapene carolina</i>	173	37	38
11	Emydidae	<i>Terrapene ornata</i>	157	33	35
12	Geoemydidae	<i>Cuora amboinensis</i>	124	26	28
13	Geoemydidae	<i>Cuora flavomarginata</i>	110	23	25
14	Geoemydidae	<i>Cuora galbinifrons</i>	108	23	24
15	Geoemydidae	<i>Geoclemys hamiltonii</i>	106	22	24
16	Geoemydidae	<i>Heosemys spinosa</i>	107	23	24
17	Geoemydidae	<i>Mauremys sinensis</i>	173	37	38
18	Podocnemididae	<i>Podocnemis unifilis</i>	128	27	29
19	Testudinidae	<i>Aldabrachelys gigantea</i>	214	45	47
20	Testudinidae	<i>Chelonoidis carbonaria</i>	252	54	55
21	Testudinidae	<i>Chelonoidis chilensis</i>	112	24	24
22	Testudinidae	<i>Chelonoidis denticulata</i>	143	30	32
23	Testudinidae	<i>Chersina angulata</i>	121	26	27
24	Testudinidae	<i>Geochelone elegans</i>	135	29	30
25	Testudinidae	<i>Geochelone platynota</i>	125	26	28
26	Testudinidae	<i>Geochelone sulcata</i>	105	22	23
27	Testudinidae	<i>Gopherus berlandieri</i>	105	22	23
28	Testudinidae	<i>Indotestudo elongata</i>	107	22	24
29	Testudinidae	<i>Malacochersus tornieri</i>	147	31	32
30	Testudinidae	<i>Manouria emys</i>	105	22	23
31	Testudinidae	<i>Manouria impressa</i>	109	23	24
32	Testudinidae	<i>Stigmochelys pardalis</i>	192	41	42
33	Testudinidae	<i>Testudo horsfieldii</i>	166	22	23
34	Testudinidae	<i>Testudo hermanni</i>	86	31	33
35	Testudinidae	<i>Testudo marginata</i>	142	30	31
36	Trionychidae	<i>Lissemys punctata</i>	105	22	23
Total			4840	1023	1071

**Table 2.** The dataset of turtle species examined in this study.

three bright yellow stripes on the sides of the head of *Cuora amboinensis*<sup>52,53</sup>. More images should also be included to train the models to increase the classification accuracy of *Graptemys ouachitensis* and *Cuora amboinensis*.

Despite the importance of classifying globally traded turtles to conserve and detect invasive turtle species, there are few studies on automatic and accurate classification of those<sup>57–59</sup>. These previous studies applied deep learning models to detect or classify two to five species of turtles. Furthermore, most studies have been conducted to detect sea turtles rather than classify those<sup>58,59</sup>. Unlike previous studies, deep learning models were evaluated to classify 36 turtles, including sea turtles, freshwater turtles, and tortoises, imported into Korea in the present study. Moreover, this is the first study to classify turtles, including endangered species, imported into Asian countries. Among the turtles examined in this study, eight species are categorized as Critically Endangered, nine are Endangered, and 11 are Vulnerable, according to the IUCN Red List (<https://www.iucnredlist.org>) (Table S3). Additionally, according to the CITES Checklist (<https://www.checklist.cites.org>), eight species are listed in Appendix I, where their trade is prohibited, and 26 are in Appendix II, where their trade is closely monitored through permits (Table S3). Furthermore, two species, *Macrochelys temminckii*, and *Graptemys ouachitensis*, examined in this study are reported as alien invasive turtles in Korea<sup>25,26</sup>. Therefore, the novelty of this study is that the models can be used to control the global trade of endangered turtles for conservation purposes and to detect alien invasive turtles in the native ecosystem for EDRR.

Although this study applied the object detection model to classify turtles, it has limitations. Typically, object detection models focus on detecting and classifying multiple objects in images. However, this study's dataset primarily consists of a single object in the images. Therefore, it is necessary to collect images of multiple objects



**Figure 3.** Representative images of the incorrect classification results of the ResNet18 model caused by morphological similarities between species. (a) *Geochelone elegans*, (b) *Geochelone platynotan*, (c) *Chelonoidis chilensis*, (d) *Geochelone sulcata*, (e) *Chelonoidis denticulate*, (f) *Chelonoidis carbonaria*. Photo credit: (a) Tatarinov A. C., (b) Jacob.jose, (c) Arteivanna, (d) MartinThoma, (e) Bernard DUPONT, (f) Dick Culbert.

at various scales, and test models using these images in further research. Given that the images for this study were obtained from the Internet also presents a challenge, which made it difficult to collect a comprehensive set of images of turtles in various poses, which could reduce the classification accuracy.

In further studies, a comprehensive collection of turtle images should be conducted using various methods, such as taking a photo directly at customs. Furthermore, future studies should include more turtles that are frequently traded worldwide. Object detection models are still being developed, for example, EfficientDet<sup>60</sup> and YOLOv5<sup>61</sup>. Thus, it is necessary to apply the other models to find a more suitable model to classify turtles. Recently, transfer learning has been applied to increase the classification accuracy of various organisms<sup>62,63</sup>. This can transfer the knowledge from the pre-trained model using vast data, such as the ImageNet benchmark dataset, to the new task model<sup>64</sup>. The accuracy of the model developed in this study will be increased using transfer learning in future studies. Finally, the model will be supplied as a mobile application for use in customs control of the turtle trade and in the wild for the detection of invasive turtles.

### Conclusion

In this study, the eight SSD models with different CNN backbone networks were evaluated to classify 36 turtle species. Among the eight models, the ResNet18 model showed the highest mAP and the fastest inference time. This model classified the 36 turtles correctly with an average accuracy of 82.8%, ranging from 60.7% for *Chelonia mydas* to 100.0% for *Stigmochelys pardalis*. The models developed in this study have the potential to automatically and accurately classify turtle species among the vast quantities of traded turtles. Application of this model can aid in regulating turtle trade for conservation purposes and detecting the alien invasive turtles in native ecosystems for EDRR. In order to improve the classification accuracy of the models in future studies, a comprehensive collection of turtle images should be considered despite the novelty of this study. Databases providing images of

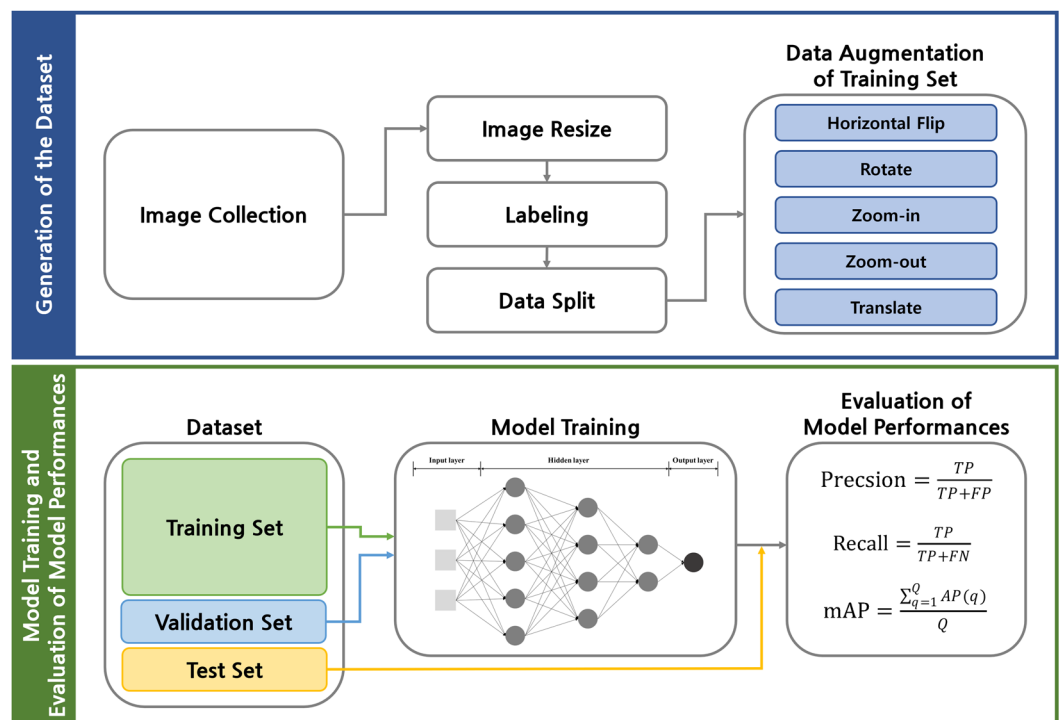
various species taken by citizen scientists, such as iNaturalist (<https://www.inaturalist.org/>), might be helpful in comprehensive image collection. Future research should also employ other recently developed object detection models to identify a suitable classification scheme for turtles. Furthermore, transfer learning will be used to increase the accuracy of classification. Finally, the model will be supplied as a mobile application.

## Methods

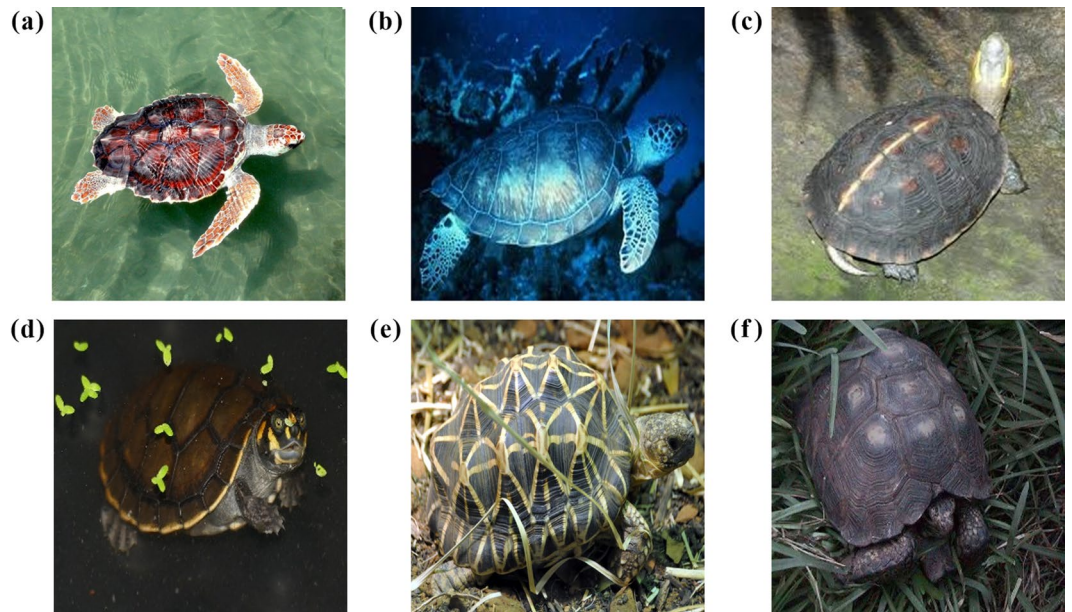
The entire methodology of this study is presented as a flowchart in Fig. 4. Images of sea turtles imported into Korea were collected. These were integrated into the size required by the SSD model used in the study, and the object in the images was labeled. The images were split into the training, validation, and test sets, and data augmentation was applied to the training set. The models were trained using the final dataset established in this study, and the performances of the models were evaluated.

### Collection of image data

Images of 51 turtle species that are imported into Korea were collected from the Internet because there was no standard dataset for these species. Scientific and common names were used as keywords for comprehensive image collection. The image collection from the Internet has been used to establish a dataset for deep learning to obtain images of various individuals with diverse backgrounds in the absence of a standard dataset<sup>47,65–67</sup>. Turtle images were collected at the species level due to a lack of data at the subspecies level. The images were identified using morphological features described in scholarly books, such as the color and shape of carapaces, the absence or presence of scute, the scale's relative size, and the number of claws per limb<sup>32,52,53</sup>. While sex dimorphism in turtles is mainly due to differences in size<sup>68</sup>, *Graptemys ouachitensis* and *Podocnemis unifilis* have morphological differences depending on sex<sup>69,70</sup>. Therefore, images of males and females of these two species were included in the dataset. Images in which species could not be identified or had low quality were excluded from the dataset. This study included all species with more than 150 initial images. In the end, 36 species from 8 families and 26 genera were chosen for the study after meeting the outlined criteria (Table 2). Table S3 displays the native distribution, the IUCN Red List category, and the CITES Appendix of the turtles examined in the present study. Except for *Caretta caretta*, *Chelonia mydas*, and *Eretmochelys imbricata*, none of the examined species are indigenous to Korea (Table S3). Figure 5 depicts representative images of examined species with various backgrounds and angles. Marine (Fig. 5a and b), freshwater (Fig. 5c and d), and grassland (Fig. 5e and f) images depict the habitat characteristics of turtles. In addition, the images from various angles, such as dorsal (Fig. 5a), dorsal lateral (Fig. 5b and e), dorsal posterior (Fig. 5c), and dorsal frontal (Fig. 5d and f), are included. The images were integrated into a 300 × 300-pixel image which was the size required by the object detection model used in the study. Additionally, the horizontal and vertical resolutions of the images were converted to 96 DPI. The ground-truth bounding boxes have designated the entire body using DarkLabel (<https://darkpgmr.tistory.com/16>) since the



**Figure 4.** The flowchart of the overall methodology for classifying 36 turtles. (1) Generation of turtle image dataset, and applying data augmentation on the training set, (2) Model training using the training and validation sets, and evaluation of trained model performances using the test set.

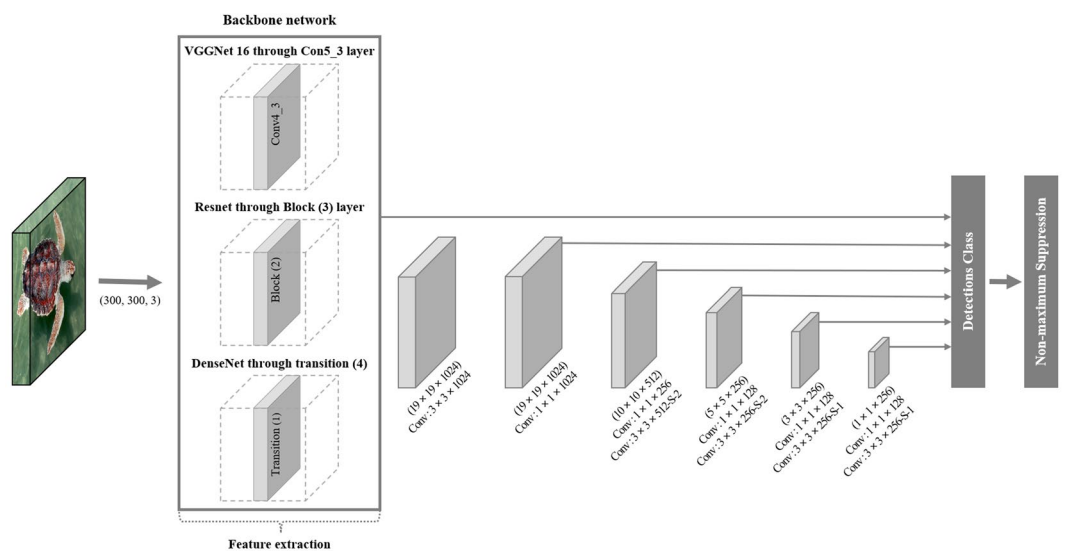


**Figure 5.** Representative images of examined species with various backgrounds. (a) *Caretta caretta*, (b) *Chelonia mydas*, (c) *Cuora flavomarginata*, (d) *Podocnemis unifilis*, (e) *Geochelone elegans*, (f) *Gopherus berlandieri*. Photo credit: (d) Bernad DUPONT, (e) Jacob.jose, (f) Dawson.

morphological features that classify the turtles are located all over the bodies. The dataset was divided into a 70% training set, a 15% validation set, and a 15% test set. Subsequently, data augmentation was applied to the training set, which included horizontal flip, random horizontal and vertical translation between  $-30$  and  $30$  pixels, random rotation between  $-10^\circ$  and  $10^\circ$ , and random zoom-in and zoom-out between  $100$  and  $200\%$ . At least  $10,000$  images of the examined species were included in the training set through data augmentation (Table S4).

### Model training

In this study, the SSD model was used to classify 36 turtle species (Fig. 6). Regression and classification were applied using a convolution layer on a multi-scale feature map. As backbone networks, eight CNN models were used in total. There is one VGGNet model<sup>71</sup>, three ResNet models<sup>50</sup>, and four DenseNet models<sup>51</sup> among the eight CNN backbone networks. These models' structures are presented in Tables S5–S7. The same environment and equipment were used to train eight SSD models. These models' experimental platform is based on the Ubuntu



**Figure 6.** Architecture of the Single Shot MultiBox Detector (SSD) model with different Convolutional Neural Networks (CNN) as the backbone network to classify the 36 turtles.

20.04 operating system and includes RTX 2080 Ti Graphics with 11G video memory, two Intel Xeon Silver 4110 CPUs, and 16 GB of REG.ECC DDR4 SDRAM. The experimental program is written in Python 3.9.7 and runs on the software PyCharm 2021.1 in Keras-TensorFlow environments. The early stop function of Keras was employed to prevent overfitting. In order to ensure consistency, each of the eight models was trained three times.

### Evaluation of model performances for detection and classification of object

The performance of the model was assessed using mean Average Precision (mAP) and inference time. Average Precision (AP) values for each class were determined using precision-recall curves obtained with precision (true positive/true positive + false positive) and recall (true positive/true positive + false negative) measurements. The Intersection over Union (IoU) was used to determine the overlap ratio between the hand-labeled ground-truth bounding boxes and the predicted bounding boxes suggested by the model. When a model made a prediction with an IoU threshold value greater than the value determined by the researcher, the prediction was deemed to be a true positive. The IoU threshold value was determined as 0.5 in this study. Finally, the mAP value was calculated using formula (1), where  $Q$  is the number of queries of the dataset, and  $AP(q)$  is the AP for the given query  $q$ .

$$\text{mean Average Precision (mAP)} = \frac{\sum_{q=1}^Q AP(q)}{Q} \quad (1)$$

The inference time was calculated as the time taken to process one image. Finally, the classification results of the model showing the highest mAP are represented in a confusion matrix. The classification result with the highest confidence value was chosen when models predicted multiple classification results.

### Data availability

The datasets examined in this study are available from the corresponding author upon a reasonable request.

Received: 30 June 2023; Accepted: 3 December 2023

Published online: 07 December 2023

### References

- Stanford, C. B. *et al.* Turtles and tortoises are in trouble. *Curr. Biol.* **30**, R721–R735 (2020).
- Turtle Conservation Fund, *A global action plan for conservation of tortoises and freshwater turtles* (Washington, DC, 2002).
- Luiselli, L., Starita, A., Carpaneto, G. M., Segniagbeto, G. H. & Amori, G. A short review of the international trade of wild tortoises and freshwater turtles across the world and throughout two decades. *Chelonian Conserv. Biol.* **15**, 167–172 (2016).
- Gibbons, J. W. *et al.* The global decline of reptiles, Déjà Vu Amphibians: Reptile species are declining on a global scale. Six significant threats to reptile populations are habitat loss and degradation, introduced invasive species, environmental pollution, disease, unsustainable use, and global climate change. *Bioscience* **50**, 653–666 (2000).
- Myers, N., Mittermeier, R. A., Mittermeier, C. G., Da Fonseca, G. A. & Kent, J. Biodiversity hotspots for conservation priorities. *Nature* **403**, 853–858 (2000).
- Moll, D. & Moll, E. O. Syntax of referencing. in *The Ecology, Exploitation and Conservation of River Turtles* (Oxford University Press on Demand, 2004).
- Shepherd, C. R. Export of live freshwater turtles and tortoises from North Sumatra and Riau, Indonesia: A case study. *Chelonian Res. Monogr.* **2**, 112–119 (2000).
- Van Dijk, P. P., Stuart, B. L. & Rhodin, A. G. Syntax of referencing. In *Asian Turtle Trade: Proceedings of a Workshop on Conservation and Trade of Freshwater Turtles and Tortoises in Asia--Phnom Penh 1–4* (Cambodia, 2000).
- Seo, M. H., & Kim, H. J. *The Guideline for Import and Export Review of CITES Species* 76–132 (Korea, 2016).
- Courchamp, F. *et al.* Rarity value and species extinction: The anthropogenic Allee effect. *PLoS Biol.* **4**, e415 (2006).
- Sung, Y. & Fong, J. J. Assessing consumer trends and illegal activity by monitoring the online wildlife trade. *Biol. Conserv.* **227**, 219–225 (2018).
- García-Díaz, P., Ross, J. V., Woolnough, A. P. & Cassey, P. The illegal wildlife trade is a likely source of alien species. *Conserv. Lett.* **10**, 690–698 (2017).
- Smith, K. F. *et al.* Reducing the risks of the wildlife trade. *Science* **324**, 594–595 (2009).
- Souviron-Priego, L., Muñoz, A. R., Olivero, J., Vargas, J. M. & Fa, J. E. The legal international wildlife trade favours invasive species establishment: The monk and ring-necked parakeets in Spain. *Ardeola* **65**, 233–246 (2018).
- Ricciardi, A. Are modern biological invasions an unprecedented form of global change?. *Conserv. Biol.* **21**, 329–336 (2007).
- Strayer, D. L., Eviner, V. T., Jeschke, J. M. & Pace, M. L. Understanding the long-term effects of species invasions. *Trends Ecol. Evol.* **21**, 645–651 (2006).
- Gariboldi, A. & Zuffi, M. A. Notes on the population reinforcement project for *Emys orbicularis* (Linnaeus, 1758) in a natural park of northwestern Italy. *Herpetozoa* **7**, 83–89 (1994).
- Spinks, P. Q., Pauly, G. B., Crayon, J. J. & Shaffer, H. B. Survival of the western pond turtle (*Emys marmorata*) in an urban California environment. *Conserv. Biol.* **13**, 257–267 (2003).
- Oh, H. & Hong, C. Current conditions of habitat for *Rana catesbeiana* and *Trachemys scripta elegans* imported to Jeju-do, including proposed management plans. *Korean J. Environ.* **21**, 311–317 (2007).
- Thomas, M. & Hartnell, P. An occurrence of a red-eared turtle (*Trachemys scripta elegans*) in the Waikato River at Hamilton. *Herpetofauna* **30**, 15–17 (2000).
- Ficetola, G. F., Thuiller, W. & Padoa-Schioppa, E. From introduction to the establishment of alien species: Bioclimatic differences between presence and reproduction localities in the slider turtle. *Divers. Distrib.* **15**, 108–116 (2009).
- Cadi, A. & Joly, P. Impact of the introduction of the red-eared slider (*Trachemys scripta elegans*) on survival rates of the European pond turtle (*Emys orbicularis*). *Biodivers. Conserv.* **13**, 2511–2518 (2004).
- Liu, S. *et al.* E-commerce promotes trade in invasive turtles in China. *Oryx* **55**, 352–355 (2021).
- Mito, T. & Uesugi, T. Invasive alien species in Japan: The status quo and the new regulation for prevention of their adverse effects. *Glob. Environ. Res.* **8**, 171–193 (2004).
- Koo, K. S., Kwon, S., Do, M. S. & Kim, S. Distribution characteristics of exotic turtles in Korean wild—Based on Gangwon-do and Gyeong-sangnam-do. *Korean J. Ecol. Environ.* **50**, 286–294 (2017).
- Koo, K. S., Park, S., Choi, J. H. & Sung, H. New report of an alligator snapping turtle (*Macrochelys temminckii* Troost, 1835) introduced into the wild in the Republic of Korea. *Bioinvasions Rec.* **10**, 188–199 (2021).

27. Scheffers, B. R., Oliveira, B. F., Lamb, I. & Edwards, D. P. Global wildlife trade across the tree of life. *Science* **366**, 71–76 (2019).
28. Trail, P. W. Morphological species identification of wildlife forensic evidence based on digital images. *Forensic Sci. Int.* **1**, 100021 (2021).
29. Mahendiran, M., Parthiban, M., Azeez, P. A. & Nagarajan, R. In situ measurements of animal morphological features: A non-invasive method. *Methods Ecol. Evol.* **9**, 613–623 (2018).
30. Hopkins, G. W. & Freckleton, R. P. Declines in the numbers of amateur and professional taxonomists: Implications for conservation. *Anim. Conserv.* **5**, 245–249 (2002).
31. Farias, I. P., Jerzolimski, A., Melo, A., das Neves Viana, M., Martins, M. & dos Santos Monjeló, L. A. Population genetics of the Amazonian tortoises, *Chelonoidis denticulata* and *C. carbonaria*, (Cryptodira: Testudinidae) in an area of sympatry. *Amphib-Reptil.* **28**, 357–365 (2007).
32. McCloud, K. A photographic identification Guide to star-patterned tortoises. *Identif. Guides Wildl. Law Enforc.* **12** (2008)
33. Bramble, D. M. Scaptochelys: Generic revision and evolution of gopher tortoises. *Copeia* **1982**, 852–867 (1982).
34. Snell, H. L. & Fritts, T. H. The significance of diurnal terrestrial emergence of green turtles (*Chelonia mydas*) in the Galapagos Archipelago. *Biotropica* **15**, 285–291 (1983).
35. National Invasive Species Council (NISC) 2016–2018 NISC Management Plan. NISC Secretariat (Washington, DC 2016)
36. Martinez, B. *et al.* Technology innovation: Advancing capacities for the early detection of and rapid response to invasive species. *Biol. Invas.* **22**, 75–100 (2020).
37. Martinez, B. *et al.* Syntax of referencing. In *Advancing Federal Capacities for the Early Detection of and Rapid Response to Invasive Species Through Technology Innovation* (Washington, DC, 2017).
38. Abe, H., Hayano, A. & Inoue-Murayama, M. Forensic species identification of large macaws using DNA barcodes and microsatellite profiles. *Mol. Biol. Rep.* **39**, 693–699 (2012).
39. Bilodeau, P. *et al.* Biosurveillance of forest insects: Part II—Adoption of genomic tools by end user communities and barriers to integration. *J. Pest Sci.* **92**, 71–82 (2018).
40. Kim, J. I., Do, T. D., Choi, T. J., Yeo, Y. & Kim, C. B. Gender determination in parrots from Korean zoos using chromo-helicase-DNA binding protein 1 (CHD1) gene fragments. *Korean J. Environ. Biol.* **38**, 350–354 (2020).
41. Roe, A. D. *et al.* Biosurveillance of forest insects: Part I—Integration and application of genomic tools to the surveillance of non-native forest insects. *J. Pest Sci.* **92**, 51–70 (2018).
42. Brandis, K. J. *et al.* Novel detection of provenance in the illegal wildlife trade using elemental data. *Sci. Rep.* **8**, 15380 (2018).
43. Elias, N. Deep learning methodology for early detection and outbreak prediction of invasive species growth. *IEEE Winter Conf. Appl. Comput. Vis.* 6335–6343 (2023).
44. Esipova, O. *et al.* Syntax of Referencing in *Wildlife Trafficking Detection Tools: Best Practices and Application to the Illegal Rhino Horn Trade* 30–33 (USA, 2021).
45. Bergler, C. *et al.* FIN-PRINT a fully-automated multi-stage deep-learning-based framework for the individual recognition of killer whales. *Sci. Rep.* **11**, 23480 (2021).
46. Chang, C. *et al.* Using deep learning to identify maturity and 3D distance in pineapple fields. *Sci. Rep.* **12**, 8749 (2022).
47. Kim, J. I., Baek, J. W. & Kim, C. B. Image classification of amazon parrots by deep learning: A potentially useful tool for wildlife conservation. *Biology* **11**, 1303 (2022).
48. Ren, S., He, K., Girshick, R. & Sun, J. Faster r-cnn: Towards real-time object detection with region proposal networks. *Adv. Neural. Inf. Process. Syst.* **28** (2015).
49. Liu, W. *et al.* Ssd: Single shot multibox detector. In *European Conference on Computer Vision* (Springer, 2016).
50. He, K., Zhang, X., Ren, S. & Sun, J. Deep residual learning for image recognition. *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.* (2016).
51. Huang, G., Liu, Z., Van Der Maaten, L. & Weinberger, K. Q. Densely connected convolutional networks. *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.* (2017).
52. Shi, H. Syntax of Referencing. In *Identification Manual for the Conservation of Turtles in China* (China, 2013).
53. Charette, R. Syntax of referencing in *CITES Identification Guide—Turtles & Tortoises*. Blue Section “Introduction”-8, Yellow section “Introduction”-60 (Canada, 1999).
54. Hu, K., Weng, C., Zhang, Y., Jin, J. & Xia, Q. An overview of underwater vision enhancement: From traditional methods to recent deep learning. *J. Mar. Sci.* **10**, 241 (2022).
55. Li, X., Hou, G., Tan, L. & Liu, W. A hybrid framework for underwater image enhancement. *IEEE Access* **8**, 197448–197462 (2020).
56. Plutino, A., Barricelli, B. R., Casiraghi, E. & Rizzi, A. Scoping review on automatic color equalization algorithm. *J. Electron. Imaging* **30**, 020901 (2021).
57. Liu, J. *et al.* Classification and recognition of turtle images based on convolutional neural network. *Mater. Sci. Eng.* **782**, 052044 (2020).
58. Badawy, M., & Direkoglu, C. Sea turtle detection using faster r-cnn for conservation purpose. IN: *10th International Conference on Theory and Application of Soft Computing, Computing with Words and Perceptions-ICSCCW-2019*. 535–541 (2019).
59. Gray, P. C. *et al.* A convolutional neural network for detecting sea turtles in drone imagery. *Methods Ecol. Evol.* **10**, 345–355 (2019).
60. Tan, M., Pang, R. & Le, Q. V. Efficientdet: Scalable and efficient object detection. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 10781–10790 (2020).
61. Jocher, G. *et al.* Ultralytics/yolov5: v6. 1-tensorrt, tensorflow edge tpu and openvino export and inference. *Zenodo* (2022).
62. Elhamod, M. *et al.* Hierarchy-guided neural network for species classification. *Methods Ecol. Evol.* **13**, 642–652 (2022).
63. Suzuki-Ohno, Y. *et al.* Deep learning increases the availability of organism photographs taken by citizens in citizen science programs. *Sci. Rep.* **12**, 1210 (2022).
64. Tan, C. *et al.* A survey on deep transfer learning. In *International Conference on Artificial Neural Networks*, 270–279 (Springer, 2018).
65. Chabot, D., Stapleton, S. & Francis, C. M. Using Web images to train a deep neural network to detect sparsely distributed wildlife in large volumes of remotely sensed imagery: A case study of polar bears on sea ice. *Ecol. Inform.* **68**, 101547 (2022).
66. Jang, W. & Lee, E. C. Multi-class parrot image classification including subspecies with similar appearance. *Biology* **10**, 1140 (2021).
67. Patel, A. *et al.* Revealing the unknown: Real-time recognition of Galápagos snake species using deep learning. *Animals* **10**, 806 (2020).
68. Berry, J. F. & Shine, R. Sexual size dimorphism and sexual selection in turtles (Order Testudines). *Oecologia* **44**, 185–191 (1980).
69. Schneider, L., Iverson, J. B. & Vogt, R. C. *Podocnemis unifilis* Troschel 1848 yellow-spotted river turtle, Tracajá. *Cat. Am. Amphibians Reptiles (CAAR)* (2012).
70. Vogt, R. C. *Graptemys ouachitensis* Cagle 1953—Ouachita Map Turtle. In: Rhodin, A. G. J. *et al.* Conservation Biology of Freshwater Turtles and Tortoises: A Compilation Project of the IUCN/SSC Tortoise and Freshwater Turtle Specialist Group. Chelonian Research Monographs **5**(11), 1–13. <https://doi.org/10.3854/crm.5.103.ouachitensis.v1.2018> (2018)
71. Simonyan, K. & Zisserman, A. Very deep convolutional networks for large-scale image recognition. *ArXiv.org* (2014).

## Acknowledgements

This work was supported by a grant from the National Institute of Biological Resources (NIBR), funded by the Ministry of Environment (MOE) of the Republic of Korea (NIBR202333202).

## Author contributions

J.W.B., J.I.K. and C.B.K. conceptualized the study and wrote the first draft of the manuscript, J.W.B. and J.I.K. performed research and analyzed data, C.B.K. revised the final draft of the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-49022-3>.

**Correspondence** and requests for materials should be addressed to C.-B.K.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023