

Detection of Wildlife Species in the Peruvian Amazon Using Transfer Learning

Luis Alberto Holgado-Apaza^{1,*}, Ruth Nataly Aragon-Navarrete²,
Coren Luhana Ancco-Calloapaza³, Edgar E. Carpio-Vargas⁴,
Marleny Quispe-Layme⁵, José Miguel Barrón-Adame⁶,
Rafael Guzman-Cabrera⁷, Wilian Quispe-Layme⁸

¹ Universidad Nacional Amazónica de Madre de Dios,
Departamento Académico de Ingeniería de Sistemas e Informática,
Peru

² Universidad Nacional Amazónica de Madre de Dios,
Departamento Académico de Ecoturismo,
Peru

³ Universidad Nacional San Agustín de Arequipa,
Peru

⁴ Universidad Nacional del Altiplano,
Departamento Académico de Estadística e Informática,
Peru

⁵ Universidad Nacional Amazónica de Madre de Dios,
Departamento Académico de Contabilidad y Administración,
Peru

⁶ Universidad Tecnológica del Suroeste de Guanajuato,
Cuerpo Académico de Desarrollo Tecnológico Multidisciplinario,
Mexico

⁷ Universidad de Guanajuato,
División de Ingenierías del Campus Irapuato-Salamanca,
Mexico

⁸ Universidad Nacional Amazónica de Madre de Dios,
Departamento Académico de Educación y Humanidades,
Peru

{lholgado, raragon, maquispe, wquispe}@unamad.edu.pe, canccoca@unsa.edu.pe,
ecarpio@unap.edu.pe, guzmanc@ugto.mx, mbarrona@utsoe.edu.mx

Abstract. Wildlife holds an important role within the Amazon biome. However, wildlife identification and documentation methods in the Amazonian wilderness pose considerable challenges for fauna biology and ecology professionals. This complexity arises from the demand for specialized expertise and the substantial investment of time required. This challenge is compounded by the remarkable resemblance between various animal species. In this study, we delve into the feasibility of diverse iterations of the YOLO (You Only

Look Once) algorithm in order to detect wildlife species in the Peruvian Amazon. Our assessment covers a spectrum of YOLO versions, including YOLOv5x6, YOLOv5l6, YOLOv7-W6, YOLOv7-E6, YOLOv8l, and YOLOv8x. To empower our models, we embarked on a training journey using a dataset comprising 653 images thoughtfully collected from reputable sources in ecology and tourism marketing. This dataset encompasses six species: Ara ararauna, Ara chloropterus, Ara macao, Opisthocomus hoazin, Pteronura brasiliensis, and

Saimiri sciureus. Our efforts show the efficiency of the YOLOv5l6 model, which stands out prominently in all metrics evaluated. This model achieves a Precision rate of 86.1%; Recall of 84.7%, F1-Score measuring 85.39%, and mean Average Precision (mAP) of 88.1%. Noteworthy is the fact that this model also boasts the swiftest training time among its counterparts, with a total 30.71 minutes. These findings offer promising prospects for refining our understanding of Amazonian wildlife species and establishing proactive measures to safeguard those that face potential vulnerability or endangerment. The YOLO algorithm's capabilities underscore the confluence of technology and ecological conservation, providing optimism for the preservation of the Amazon's intricate biodiversity.

Keywords. Wildlife species, Peruvian Amazon, YOLO, object detection, transfer learning.

1 Introduction

Amazon biome is one of the main sources of biodiversity in the world's ecosystems [1]. Wildlife is an important component within its territory [2, 3]. While the species list is growing all the time, only a fraction of the Amazon's enormous biodiversity is known to science [4, 5].

According to estimates only 90-95 per cent of mammals, birds and plants are known, only 2-10 per cent of insects have been described, and only 2,500 from the approximately 6,000 – 8,000 amazon fish species have been described [6]. Eight countries share responsibility for the Amazon, one of them is Peru which is home of 11.27 per cent of the biome. Peru stands out as one of the most biodiverse countries in the world.

The extensive Amazonian forests cover 62% of the Peruvian territory and they are home to approximately 50% of the plant species registered by the Ministry of Environment-MINAM [7]. This remarkable diversity also includes numerous wildlife endemic species in the region.

For example, 115 endemic bird species have been identified (representing 6% of the world total), 109 mammal species (27.5% of the world total), 185 amphibian species (48.5% of the world total) and 59 endemic butterfly species (12.5% of the world population) [8].

The efforts to conserve and manage fauna in the Amazon Biome does not still fill gaps of knowledge about tropical fauna [3, 9]. It is needed

to strength accurate identification and monitoring of wildlife through discovering and documentation. Tradition-ally the identification is based on different biological assessments of biodiversity.

However, identification and documentation methods of wildlife in the Amazon represents a considerable challenge for professionals engaged in population biology and ecology studies, as it involves a high cognitive load and significant time consumption [10]. This difficulty is attributed to the existence of multiple types of animals that exhibit high similarity to each other, making difficult their precise classification [11].

Over the past few decades, automated species identification has brought about a revolution in conventional methodologies [12]. Recent research has demonstrated the emergent use of artificial intelligence (AI) and more specifically computer vision in the identification and monitoring biodiversity species, is the case of MobileNetV3 a deep learning model that were successfully used enabling faster and more efficient analysis at identifying mangrove species [13].

The same technology has been used to detect plant species [14]. Similar technologies have contributed to detect camels on roads [15], as well as to identify rodent species [16]. Another example is the use of AlexNet model to identify ringed seal of the Saimaa, according to results the experiment get an accuracy of 91.2% in the individual identification of species [17].

In the same line, [18] proposed a framework for animal's recognition which consisting of 2 Convolutional neural network (CNN) based models for image classification, results show values close to 90% in the identification of 3 most common animals. [19] proposed the application of a 3-branch VGG CNN in parallel with the aim of recognizing wild animal species. CNN has contributed to recognize wild boars [20] toads/frogs, lizards and snakes [21].

The transfer learning architecture called YOLOv3 has enabled the wildlife monitoring [11]. The advantages offered by CNN's applications in image recognition have been applied for different purposes [22, 23, 24], constructed the wildlife dataset of Northeast China Tiger and Leopard National Park to identify and process images captured with camera traps. Using three deep learning object detection models YOLOv5, FCOS

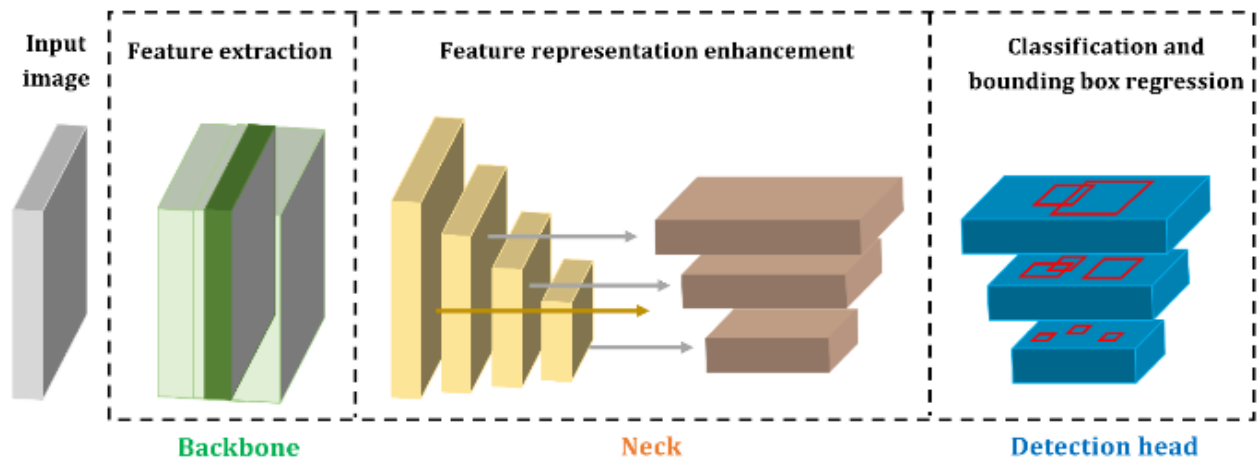


Fig. 1. General architecture of YOLO object detectors

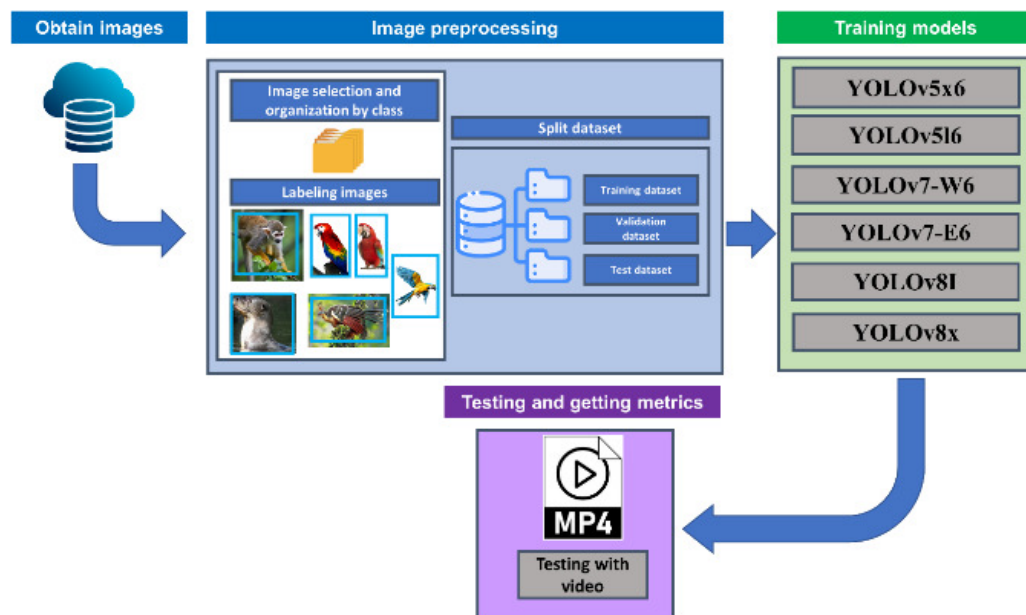


Fig. 2. Methodology used for the detection of wildlife species in the Peruvian Amazon

and Cascade R-CNN, they found an average Accuracy of 97.9%, along with an approximate mAP50 of 81.2% for all three models. Along the same lines, [25] used a transfer learning approach to detect the presence of four endangered mammals in the forests of Negros Island (*Viverra tangalunga*, *Prionailurus javanensis sumatranus*, *Rusa alfredi* and *Sus cebifron*).

The authors used the YOLOv5 model as a detection method. The trained model yielded a mean mAP50 of 91%. [26] proposed the

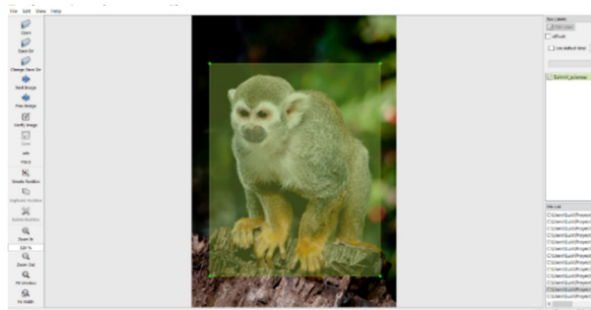
identification of snake, lizard and toad/frog species from camera trap images using CNN. The results obtained by accuracy were 60% in the validation stage. [27] proposed a system to detect animals on the road and avoid accidents.

To do this, the animals were classified into groups of capybaras and donkeys. The authors used two variants of pre-trained CNN models: YOLOv4 and YOLOv4-tiny.

The results showed an accuracy of 84.87% and 79.87% for YOLOv4 and YOLOv4-tiny, respectively.

Table 1. Number of images per wildlife species

Species	Quantity
Ara ararauna	232
Ara chloropterus	50
Ara Macao	52
Opisthocomus hoazin	100
Pteronura brasiliensis	110
Saimiri sciureus	109
Total	653

**Fig. 3.** Image labeling process in the labeling tool

As shown, there are important advances at animal detection using deep learning technologies.

However, they have not been applied to detect wildlife species in the Peruvian part of the Amazon biome yet. The main objective of this study was to evaluate the application of the YOLO algorithm in its versions YOLOv5x6, YOLOv5l6, YOLOv7-W6, YOLOv7-E6, YOLOv8l and YOLOv8x in the detection of wildlife species in the Peruvian Amazon.

To achieve this, we have evaluated the aforementioned models using the following metrics: Precision, Recall, F1-Score and mAP50, applied to six species: Ara ararauna, Ara chloropterus, Ara macao, Opisthocomus hoazin, Pteronura brasiliensis and Saimiri sciureus.

In addition, as part of our contribution to the scientific community, we provide a labelled dataset for classification and/or detection of these species. We have structured the remaining contents of the paper as follows: In Section 2, the methodology adopted to conduct the experiments is presented in detail.

Results and discussions are addressed in Section 3, while our conclusions are presented in Section 4.

2 Material and Methods

We performed our experiments using the machine learning technique called transfer learning, which consists of using previously learned knowledge trained on large volumes of public images [28, 29]. Specifically, we have used the object detection algorithm in images and video called YOLO (You Only Look Once) in its versions YOLO-v5 [30], YOLO-v7 [31] and YOLO-v8 [32].

We trained and evaluated our models on a computer with these characteristics: AMD A12-9700P RADEON R7, 10 COMPUTE CORES 4C+6G at 2.50 GHz, 12 GB RAM, Windows 10 Home 64-bit operating system and x64 processor.

The development environment used was Google Colab with GPU accelerator type A100. Figure 1 shows the general architecture of YOLO, taking as reference the study presented by [33, 34]. The main components of YOLO are listed below:

- **Backbone:** The backbone is usually a convolutional neural network that extracts useful features from the input image [33].
- **Neck:** The neural network neck is used to extract features from images at different stages of the backbone. YOLOv4 make use of Spatial Pyramid Pooling (SPP) [35] and Path Aggregation Network (PAN) [36].
- **Head:** Is the final component of the object detector; this component is responsible for making the predictions from the features provided by the spine and neck [33].

In Figure 2, we show the methodology used for the detection of wildlife species in the Peruvian Amazon using transfer learning. It is composed of four phases, they are: Obtain images, images preprocessing, training models and testing and getting metrics. Each step of the proposed methodology is detailed below:

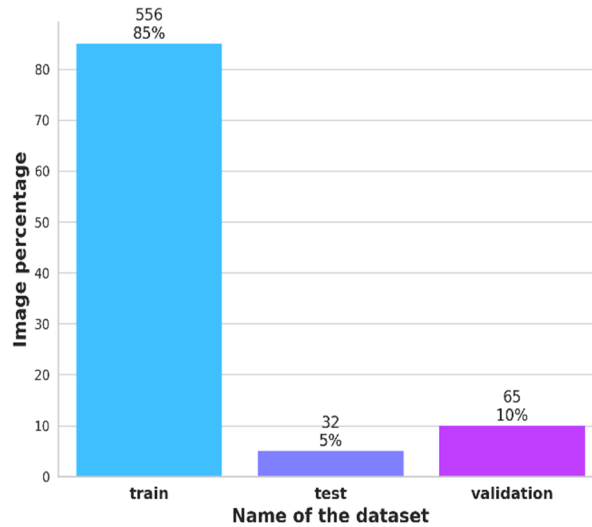


Fig. 4. Dataset division for learning transfer process

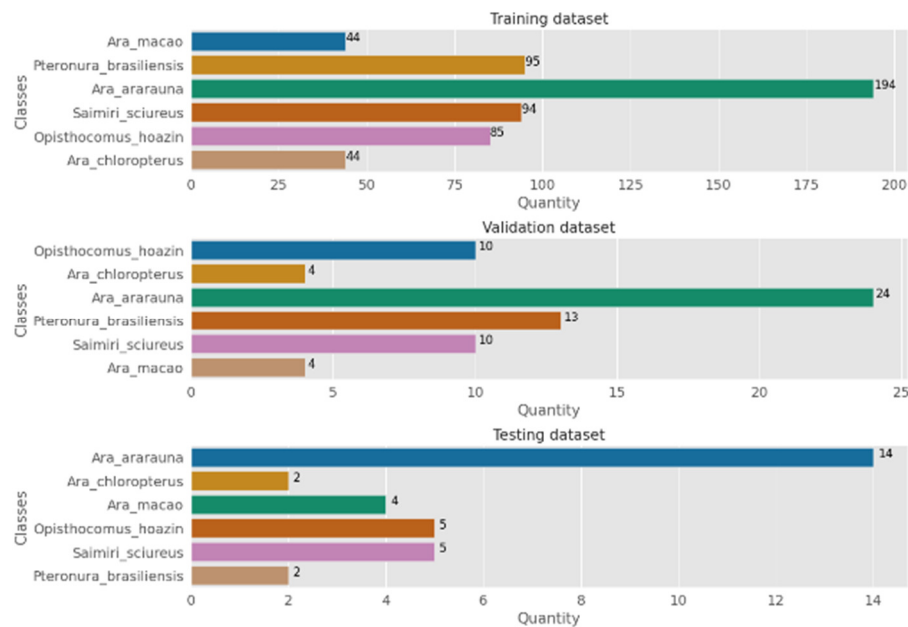


Fig. 5. Images distribution by classes in datasets training, validation and test

2.1 Image Obtaining

At this stage, we have searched for wild animals' images by their scientific name. The images were collected from websites related to ecology studies and tourism marketing; such as Rainforest

expeditions [37], Go2peru [38], Ararauna Tambopata [39].

Then, in order to download the images with high resolution, we have used the Fatkun Batch Download Image extension in its version 5.7.7 by the Google Chrome [40], as it was used in previous

Table 2. Confusion matrix for three classes

True	Prediction			
	A	B	C	FN
A	n_{11}	n_{12}	n_{13}	$n_{12} + n_{13}$
B	n_{21}	n_{22}	n_{23}	$n_{21} + n_{23}$
C	n_{31}	n_{32}	n_{33}	$n_{31} + n_{32}$
FP	$n_{21} + n_{31}$	$n_{12} + n_{32}$	$n_{13} + n_{23}$	

Table 3. Configuration and training times for models

Model	Epochs	Batch	Input (resolution)	Training time in minutes
YOLOv5x6	70	16	640x640	57.94
YOLOv5l6	70	16	640x640	30.71
YOLOv7-W6	70	16	640x640	57.48
YOLOv7-E6	70	16	640x640	60.36
YOLOv8l	70	16	640x640	37.14
YOLOv8x	70	16	640x640	62.20

studies [41]. Finally, we have selected and filter manually only the images in jpg format.

2.2 Image Processing

This stage encompassed the process of image curation, organization and labeling. The curation of images was made according to the species to be identified, since the search yielded images related to the keyword. In Table 1, It is shown the total summary of images by species in the first dataset after selection.

We have performed the images labeling manually. For this purpose, we have used the labellmg tool [42]. In Figure 3, we show an example of this labeling task in the specie *Saimiri sciureus*. As a result of this process, a textual file is generated and it fulfills the mission to designate each image.

Internally this file encompasses both the class to which the image belongs and the coordinates delimiting the bounding boxes containing the image. Finally, we have divided the dataset as follow:

We divided the dataset into 85% (556 images) for training, 10% (65 images) for validation and 5% (32 images) for testing. Figure 4 shows this division graphically.

In Figure 5 we show a summary of the distribution of classes in the datasets used to train,

validate and test the species detection models based on the YOLO architecture.

2.3 Training Model

During this phase we conducted out our experiments with the object detection algorithm YOLO in its versions: YOLOv5x6, YOLOv5l6, YOLOv7-W6, YOLOv7-E6, YOLOv8l and YOLOv8x. Within the file called 'custom_data.yaml' we have defined the configuration of the path to the training, validation and test images.

In the same file, we have additionally configured the classes as follow: Scientific names: [Ara_ararauna, Ara_chloropterus, Ara_macao, Opisthocomus_hoazin, Pteronura_brasiliensis, Saimiri_sciureus].

2.4 Testing and Getting Metrics

We performed our tests with a video adapted from the public videos: "Vive como sueñas | Reserva Nacional Tambopata" from the Ministry of Environment [43, 44] and the video "Manu & Tambopata" from the Antara-Peru travel agency [45]. Figure 6 and Figure 7, respectively, show a screenshot of these videos.



Fig. 6. Screenshot of video: "Vive como sueñas Reserva Nacional Tambopata"



Fig. 7. Screenshot of video: "Manu and Tambopata"

We have used the metrics: Precision, Recall, F1-Score and Mean Average Precision (mAP), and additionally the confusion matrix in order to evaluate the performance of the models. Subsequently we proceed to detail each of these metrics: Confusion matrix: It is an $n \times n$, dimension table where n represents the number of classes or objects to be detected.

This metric allows to evaluate the performance of a classification algorithm by counting the hits and mistakes in each one of the model classes. A confusion matrix for three classes is observed in Table 2, which can be extrapolated to object detection and classification problems with n classes.

where:

FN represents false negatives.

FP represents false positives.

n_{11} represents the true positives (TP) for class A.

n_{22} represents the true positives (TP) for class B.

n_{33} represents the true positives (TP) for class C.

The true negatives (TN) for class:

$$A = n_{22} + n_{33} + n_{32} + n_{23}.$$

The true negatives (TN) for class:

$$B = n_{11} + n_{33} + n_{31} + n_{13}.$$

The true negatives (TN) for class:

$$C = n_{11} + n_{22} + n_{21} + n_{12}.$$

Precision: This metric, also known as positive predictive value (PPV) indicates the proportion of cases correctly identified as belonging to a specific class (e.g., class C) among all cases where the classifier claims to belong to that class.

In other words, accuracy answers the question: Considering that the classifier predicts that a sample belongs to class C, what is the probability that the sample actually belongs to class C? [46, 47]. Equation 1 illustrates the calculation of this metric:

$$\text{Precision} = \frac{TP}{TP+FP}, \quad (1)$$

where:

TP=True Positive.

FP=False Positive.

Recall: This metric, is also referred to as Sensitivity or True Positive Rate (TPR) measures the ratio of positive correctly identified positive cases (for our case study it represents the species to be identified) by the algorithm [48] Equation 2 shows the formula for calculating this metric:

$$\text{Recall} = \frac{TP}{TP+FN}. \quad (2)$$

F1-Score: It is defined as a harmonic mean of precision and recall. The F1 score reaches its best value at 1 and its worst value at 0. Equation 3 shows the formula for calculating this metric:

$$F1 = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}. \quad (3)$$

Average Precision (AP): This metric represents the relationship between precision and recall at different confidence thresholds, in addition to quantifying the ability of the detection model to discriminate between positive and negative classes. It is calculated from the Precision-Recall curve (PR Curve). Its value varies between 0 and 1, where an AP of 1 indicates perfect detection and AP of 0 indicates random detection [49, 51].

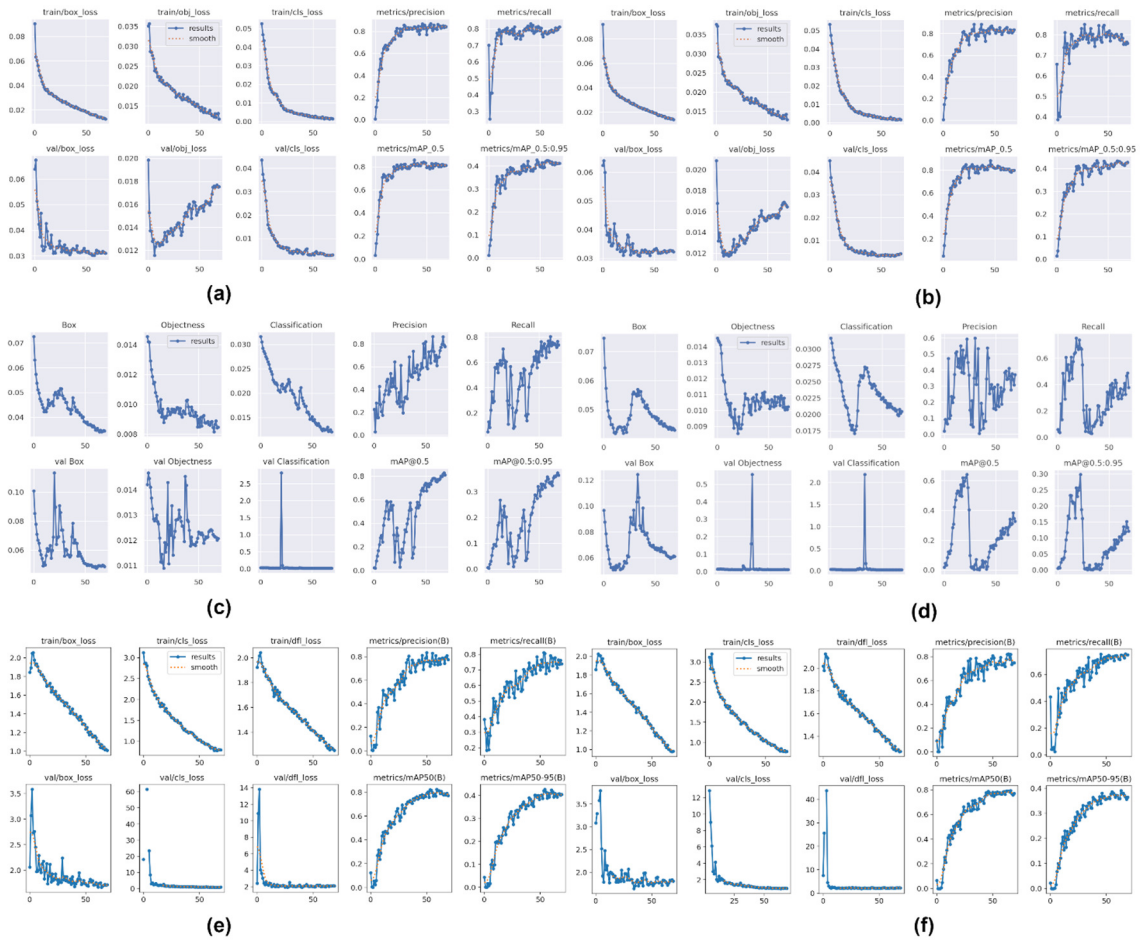


Fig. 8. Evaluation and performing metrics of the models during training. (a) YOLOv5x6 model, (b) YOLOv5l6 model, (c) YOLOv7-W6 model, (d) YOLOv7-E6 model, (e) YOLOv8l model and (f) YOLOv8x model

The mathematical operation for this calculation is shown in Equation 4:

$$AP_i = \int_0^1 P_i(R_i) dR_i, \tag{4}$$

where P and R refer to the precision and recall of the detection model that we detail in Equations (1) and (2), respectively. Mean average precision (mAP): We have calculated this metric by averaging the Average Precision (AP) values for all classes present in the dataset.

Its value fluctuates between 0 and 1, where 1 indicates perfect performance, i.e., all detections are correct and there are no false positives or false

negatives. In Equation 5 we show the formula for the calculation of this metric:

$$mAP = \frac{1}{n} \sum_{i=1}^n AP_i. \tag{5}$$

3 Results and Discussions

Table 3, shows the configurations we have made for each YOLO version we used within our experimental framework, along with their corresponding training duration measured in minutes. Our selection process adhered to the guidelines outlined in the official documentation.

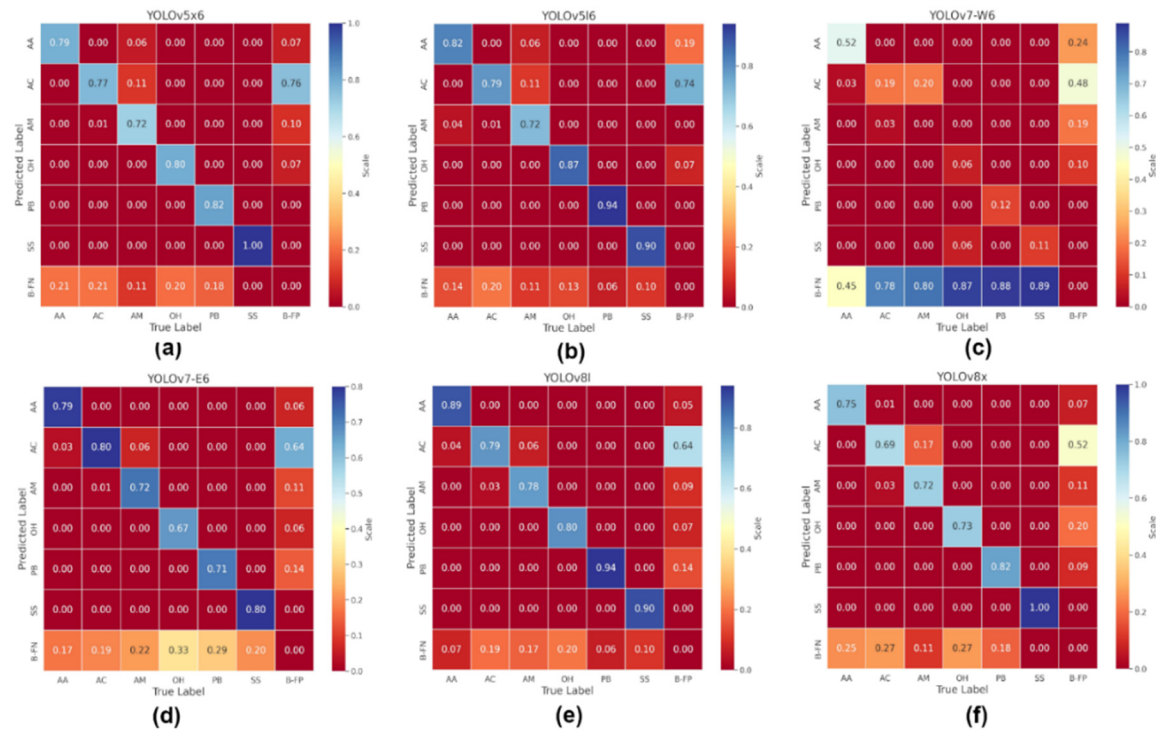


Fig. 9. Standard confusion matrices. (a) YOLOv5x6 model, (b) YOLOv5l6 model, (c) YOLOv7-W6 model, (d) YOLOv7-E6 model, (e) YOLOv8l model and (f) YOLOv8x model

According to this, we have chosen the two best pre-trained models per version [52]–[55]. We have made a particular exception for YOLOv7 because the YOLOv7-D6 and YOLOv7-E6E versions, which have a slightly better AP value, required a high computational cost for training, so we opted to use the YOLOv7-W6 and YOLOv7-E6 versions.

Figure 8 presents the metrics for evaluating and monitoring of the selected model's performance during training. The metrics are focused on the prediction accuracy of the object bounding boxes (box_loss) coordinates, error in the prediction of the classes of the detected objects (cls_loss), precision and recall.

It is noted that the YOLOv5 (Figure 8a and 8b) and YOLOv8 (Figure 8e and 8f) models exhibit remarkable stability and out-standing performance in the task of accurate bounding box localization. These models demonstrate a consistent tendency to reduce the loss associated with accuracy, which holds important relevance in image object detection applications.

Specifically, in our domain study related to wildlife species identification, the results obtained by YOLOv5 and YOLOv8 show a superior ability to accurately localize the bounding boxes of the interest objects.

Moreover, during the validation stage, it is observed that these models maintain their stability, which confirms their robustness and their potential for practical applications in the field of computer vision.

With regard to the "cls_loss" metric, which reflects the discrepancy between model's classification predictions and the actual labels of the object classes, it is graphically observed in Figure 8a and 8b that the YOLOv5 model manages to efficiently reduce this value down to epoch 50.

This suggests that as the model is trained its performance is better in the classification accuracy of wildlife species in the Peruvian Amazon.

Figure 9 shows the normalized confusion matrices during the training phase for the models utilized in our experiments. AA corresponds Ara

Table 4. Results obtained per evaluation metric

Model	Class	Precision	Recall	F1-Score	mAP50
YOLOv5x6	All	0.822	0.787	0.804	0.839
	Ara_ararauna	0.842	0.786	0.813	0.857
	Ara_chloropterus	0.743	0.760	0.751	0.790
	Ara_macao	0.848	0.778	0.811	0.852
	Opisthocomus_hoazin	0.819	0.733	0.774	0.747
	Pteronura_brasiliensis	0.812	0.763	0.787	0.812
	Saimiri_sciureus	0.870	0.900	0.885	0.978
YOLOv5l6	All	0.861	0.847	0.854	0.881
	Ara_ararauna	0.842	0.857	0.849	0.867
	Ara_chloropterus	0.727	0.745	0.736	0.816
	Ara_macao	0.781	0.722	0.750	0.759
	Opisthocomus_hoazin	0.864	0.849	0.856	0.871
	Pteronura_brasiliensis	0.951	0.941	0.946	0.980
	Saimiri_sciureus	1.000	0.967	0.983	0.995
YOLOv7-W6	All	0.371	0.378	0.374	0.327
	Ara_ararauna	0.326	0.571	0.415	0.514
	Ara_chloropterus	0.373	0.480	0.420	0.349
	Ara_macao	0.080	0.218	0.117	0.065
	Opisthocomus_hoazin	0.240	0.267	0.253	0.225
	Pteronura_brasiliensis	0.493	0.229	0.313	0.327
	Saimiri_sciureus	0.714	0.500	0.588	0.483
YOLOv7-E6	All	0.776	0.739	0.757	0.808
	Ara_ararauna	0.851	0.821	0.836	0.864
	Ara_chloropterus	0.700	0.779	0.737	0.765
	Ara_macao	0.753	0.722	0.737	0.792
	Opisthocomus_hoazin	0.832	0.662	0.737	0.791
	Pteronura_brasiliensis	0.613	0.647	0.630	0.681
	Saimiri_sciureus	0.909	0.800	0.851	0.952
YOLOv8l	All	0.743	0.806	0.773	0.817
	Ara_ararauna	0.828	0.893	0.859	0.910
	Ara_chloropterus	0.604	0.720	0.657	0.645
	Ara_macao	0.736	0.722	0.729	0.732
	Opisthocomus_hoazin	0.737	0.733	0.735	0.823
	Pteronura_brasiliensis	0.654	0.891	0.754	0.863
	Saimiri_sciureus	0.897	0.877	0.887	0.930
YOLOv8x	All	0.819	0.745	0.780	0.790
	Ara_ararauna	0.843	0.714	0.773	0.836
	Ara_chloropterus	0.711	0.624	0.665	0.698
	Ara_macao	0.820	0.758	0.788	0.795
	Opisthocomus_hoazin	0.784	0.726	0.754	0.694
	Pteronura_brasiliensis	0.856	0.765	0.808	0.820
	Saimiri_sciureus	0.898	0.885	0.891	0.898

ararauna species; AC, Ara cholopterus; AM, Ara macau; OH, Opisthocomus hoazin; PB, Pteronura brasiliensis and SS; Saimiri sciureus.

In this graphic it is also observed that YOLOv8l y YOLOv5l6 models attain the highest values along the main diagonal. This diagonal represents

instances in which the predicted labels by the aforementioned models align with the actual labels and suggests that in our dataset of Peruvian Amazon wildlife, the YOLOv8l and YOLOv5l6 models effectively identify the six species studied. Table 4 presents the results of the obtained metrics

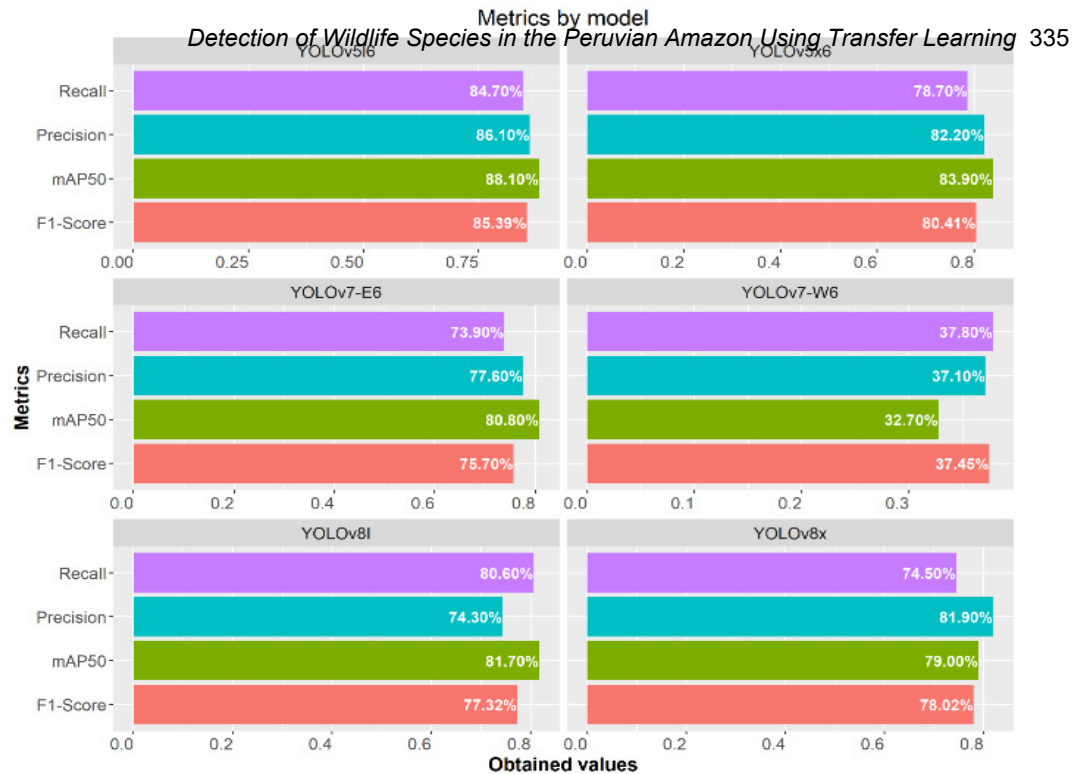


Fig. 10. Overview of the metrics for the models analyzed

for each model in our experiments in detection of the six Peruvian Amazon wildlife species. The best values per species and metric are highlighted.

It is important to stand out that the YOLOv5l6 model has demonstrated an outstanding performance by obtaining the highest values of Precision, Recall, F1-Score and mAP50 for three of the six species analyzed: *Opisthocomus hoazin*, *Pteronura brasiliensis* and *Saimiri sciureus*.

The reference model demonstrated remarkable performance across evaluated metrics. It achieved a Precision rate of 86.4%, Recall of 84.9%, and an F1-Score of 85.6%, and a mAP50 of 87.1% for the first specie.

For the second one, the model exhibited exceptional Precision rate of 95.1%, paired with a Recall of 94.1%, an F1-Score of 94.6%, and an mAP50 of 98.0%.

The third specie yielded unprecedented results, attaining a perfect Precision of 100%, while sustaining a Recall of 96.7%, contributing to an impressive F1-Score of 98.3%. Remarkably, the mAP50 reached an astounding 99.5%. It's worth

noting that this model consistently outperforms others, particularly evident in its remarkable mAP50 of 81.6% for the *Ara chloropterus* species.

Summarizing, our experiments have evidenced that the YOLOv5l6 model proves highly effective in specific species detection and it is notably excelling in the cases of *Opisthocomus hoazin*, *Pteronura brasiliensis* and *Saimiri sciureus*, such as in *Ara chloropterus* specie registering the highest mAP50 value.

On the other hand, the YOLOv5x6 model has also yielded remarkable results, particularly for the *Ara macao* species, in addition to obtaining higher values in the precision and F1 score metrics for *Ara chloropterus*. In Figure 10, we present a visual summary of the evaluated metrics for the six analyzed models. Our experiments highlight that the YOLOv5l6 model achieves the highest values in all of the evaluated metrics with a Precision rate of 86.1%, a recall of 84.7%, an F1-Score of 85.39% and an mAP50 of 88.1%.

Attained values are closely trailed by those reported by the YOLOv5x6 model. Notably, the overall averages of the metrics used in this study



Fig. 11. Detection outcomes with confidence values within bounding boxes for models. (a) YOLOv5x6, (b) YOLOv5l6, (c) YOLOv7-W6, (d) YOLOv7-E6, (e) YOLOv8l and (f) YOLOv8x

firmly indicate the YOLOv5l model as the optimal choice for our specific case.

These findings resonate with the conclusions drawn by [24], who explored animal detection and classification through camera trap images. Their evaluation of YOLOv5, FCOS, and Cascade_R-CNN_HRNet32 models yielded an impressive average Precision of 97.9%, along with an approximate mAP50 of 81.2% across all models.

Likewise, [25] introduced a framework for detecting four endangered mammal species *Viverra zangalunga*, *Prionailurus javanensis sumatranus*, *Rusa alfredi*, and *Sus cebifrons* in the forests of Negros Island using the YOLOv5 model.

Their effort resulted in an average mAP50 of 91% and a commendable Precision of 91%. We posit that the disparities in the metric values observed are attributed to the size discrepancy between the species examined in our research and those in the referenced studies.

Specifically, the species within our study context are comparably smaller than those encompassed in the aforementioned research. Our findings hold high significance for applications in wildlife monitoring and conservation.

They underscore the effectiveness of the YOLOv5l6 model in detecting wildlife species within the expanse of the Peruvian Amazon. Furthermore, these outcomes establish robust basis for forthcoming research endeavors and

initiatives aimed at safeguarding the biodiversity of Amazonian regions.

Figure 11 provides an enlightening overview of numerical results of confidence in bounding boxes in the detection of wildlife species in the Peruvian Amazon. These results are screenshot captures obtained from the execution of our models on a video file in mp4 format (minute 1:00), derived from the videos "Vive como sueñas | Reserva Nacional Tambopata" [56] and "Manu & Tambopata" [57].

It is noteworthy that the YOLOv7-E6 model is notable for its ability to identify the highest number of species in the video under analysis. However, a more detailed evaluation of the confidence associated with the object bounding boxes yields more accurate data.

We have observed that the YOLOv5l6 (b) model achieves a confidence of 83% for the *Opisthocomus hoazin* species, which represents the highest figure within this specific category. Concerning the detection of *Ara araucana* and *Saimiri sciureus* species, confidence levels of 90% and 75% are attained respectively.

It is worth noting that these values are remarkably close to those obtained by the YOLOv5x6 version. At the captured moment depicted in the figure, it is evident that both the YOLOv5l6 and YOLOv5x6 models failed to identify the *Ara macao* species.

However, it is important to note that the YOLOv7-E6, YOLOv8l and YOLOv8x models successfully detected this species at the same minute of capture. These findings provide a visual and tangible perspective on how the models perform in wildlife detection in the Peruvian Amazon.

These observations are crucial for understanding the applicability of the models in real world situations and highlight the significance of prediction confidence, as well as the selection of appropriate models based on detection objectives.

4 Conclusions

In this study we comprehensively assess the effectiveness of the YOLO algorithm in its versions YOLOv5x6, YOLOv5l6, YOLOv7-W6, YOLOv7-E6, YOLOv8l, and YOLOv8x.

Our evaluation centers on their suitability for detecting six wildlife species within the Peruvian Amazon, namely: *Ara ararauna*, *Ara chloropterus*, *Ara macao*, *Opisthocomus hoazin*, *Pteronura brasiliensis*, and *Saimiri sciureus*.

To build a robust foundation for our analysis, we curated a specialized dataset by sourcing images from ecological and tourism outlets, including Rainforest Expeditions, Go2Peru, and Ararauna Tambopata. The meticulous curation process, executed under the guidance of a wildlife specialist.

Our findings prominently showcase the prowess of the YOLOv5l6 model, which exhibits exceptional performance across all evaluated metrics. Notably, it achieves an impressive accuracy of 86.1%, a recall rate of 84.7%, an F1-Score of 85.39%, and a mean Average Precision (mAP) of 88.1%.

Remarkably, this model also boasts the shortest training duration at a mere 30.71 minutes among all models scrutinized. Furthermore, our experimental outcomes reveal a striking similarity between the achievements of the YOLOv5l6 model and those of the YOLOv5x6 model.

This convergence of results underscores the consistency and reliability of our evaluation

methodology. These outcomes stand as promising and auspicious strides forward in fortifying initiatives aimed at identifying Amazonian wildlife species and vigilantly monitoring those that could potentially slip into states of vulnerability or endangerment.

By showcasing the potential of advanced algorithms, we not only demonstrate the power of technology but also emphasize the significance of collective efforts in safeguarding the rich biodiversity of the Amazon rainforest.

Data Availability Statement: We make our dataset and source code available to the scientific community at the following e-mail address¹ Additionally, you can watch the video used for the tests. There are also 6 videos that were generated for each model in the validation stage, these videos show the bounding boxes and their respective confidence value².

References

1. **Andrade-Silva, J., Baccaro, F. B., Prado, L. P., Guénard, B., Warren, D. L., Kass, J. M., Economo, E. P., Silva, R. R. (2022).** A large-scale assessment of ant diversity across the Brazilian Amazon Basin: Integrating geographic, ecological and morphological drivers of sampling bias. *Ecography*, Vol. 2022, No. 9, pp. e06295. DOI:10.1111/ecog.06295.
2. **Peres, E. A., Pinto-da-Rocha, R., Lohmann, L. G., Michelangeli, F. A., Miyaki, C. Y., Carnaval, A. C. (2020).** Patterns of species and lineage diversity in the Atlantic Rainforest of Brazil. *Neotropical diversification: patterns and processes*, pp. 415–447. DOI: 10.1007/978-3-030-31167-4_16.
3. **Geldmann, J., Manica, A., Burgess, N. D., Coad, L., Balmford, A. (2019).** A global-level assessment of the effectiveness of protected areas at resisting anthropogenic pressures. *Proceedings of the National Academy of Sciences*, Vol. 116, No. 46, pp. 23209–23215. DOI: 10.1073/pnas.1908221116.

¹github.com/cheshire21/wild_species_detection_cnn

² www.youtube.com/@luisalbertoholgadoapaza919/videos

4. **de-Souza, V. D. C., de-Oliveira, R. E., Sais, A. C. (2022).** Agro e biodiversidade na agricultura familiar: potencial de diversificação e conservação em paisagens desmatadas na Amazônia. *Desenvolvimento e Meio Ambiente*, Vol. 60. DOI: 10.5380/dma.v60i0.73625.
5. **Fidelis, E. G., Querino, R. B., Adaime, R. (2023).** The Amazon and its biodiversity: a source of unexplored potential natural enemies for biological control (Predators and parasitoids). *Neotropical Entomology*, Vol. 52, No. 2, pp. 152–171. DOI: 10.1007/s13744-022-01024-y.
6. **Charity, S., Dudley, N., Oliveira, D., Stolton, S. (2016).** *Living Amazon, Report 2016. A regional approach to conservation*, Brazilia and Quito, 2016.
7. **Torres-Montenegro, L. A., Ríos-Paredes, M. A., Pitman, N. C., Vriesendorp, C. F., Hensold, N., Mesones-Acuy, Í., Trujillo-Calderón, W. (2019).** Sesenta y cuatro nuevos registros para la flora del Perú a través de inventarios biológicos rápidos en la Amazonía peruana. *Revista Peruana de Biología*, Vol. 26, No. 3, pp. 379–392, DOI: 10.15381/RPB.V26I3.16780.
8. **SERNANP (2016).** 'PERÚ: PAÍS MEGADIVERSO', 2022. <https://old.sernanp.gob.pe/sernanp/archivos/imagenes/vida/Peru-Pais-Megadiverso.pdf>.
9. **Dudley, N., Phillips, A., Amend, T., Brown, J., Stolton, S. (2016).** Evidence for biodiversity conservation in protected landscapes. *Land*, Vol. 5, No. 4, p. 38. DOI: 10.3390/land5040038.
10. **Pollicelli, M. D., Delrieux, C. A., Coscarella, M. A. (2020).** Avances en foto-identificación automatizada de fauna silvestre. *Repositorio Institucional CONICET Digital*, No. 5, DOI: 10.33414/AJEA.5.719.2020.
11. **Manasa, K., Paschyanti, D. V., Vanama, G., Vikas, S. S., Kommineni, M., Roshini, A. (2021).** Wildlife surveillance using deep learning with YOLOv3 model. 2021 6th International Conference on Communication and Electronics Systems ICCES, pp. 1798–1804. DOI: 10.1109/ICCES51350.2021.9489121.
12. **de-Souza, V. L., Costa, F. B., Martins, T. F., de-Oliveira, P. R., Lima, J., Guimarães, D. P., Alencar-dos-Santos, E., Oliveira-de-Moura, M. N., Pereira-Sato, T., Pais-Borsoi, A. B., Bitencourth, K., Ribamar-Lima-de-Souza, J., Salles-Gazeta, G., Guilherme, E., Glauco-de-Araújo-Santos, F. (2023).** Detection of Rickettsia tamurae-like and other spotted fever group rickettsiae in ticks (Acari: Ixodidae) associated with wild birds in the Western Amazon, Brazil. *Ticks and Tick-borne Diseases*, Vol. 14, No. 4, p. 102182. DOI: 10.1016/j.ttbdis.2023.102182.
13. **Viodor, A. C. C., Aliac, C. J. G., Santos-Feliscuzo, L. T. (2023).** Identifying mangrove species using deep learning model and recording for diversity analysis: A mobile approach. 2023 IEEE Open Conference of Electrical, Electronic and Information Sciences (eStream), pp. 1–6. DOI: 10.1109/eStream59056.2023.10134992.
14. **Carranza-Rojas, J., Mata-Montero, E. (2017).** Identificación automática de especies de plantas de Costa Rica usando visión por computadora. *Memorias de congresos TEC. III Jornadas Costarricenses de Investigación en Computación e Informática*. DOI: 10.18845/MCT.V0I0.4527.
15. **Alnujaidi, K., AlHabib, G. (2023).** Computer vision for a camel-vehicle collision mitigation system. Vol. 12, No. 1, pp. 141–149. DOI: 10.5121/ijci.2023.120111.
16. **Seijas, C., Montilla, G., Frassato, L. (2019).** Identificación de especies de roedores usando aprendizaje profundo. *Computación y Sistemas*, Vol. 23, No. 1, pp. 257–266. DOI: 10.13053/CYS-23-1-2906.
17. **Nepovinnykh, E., Eerola, T., Kälviäinen, H., Radchenko, G. (2018).** Identification of Saimaa ringed seal individuals using transfer learning. *Advanced Concepts for Intelligent Vision Systems: 19th International Conference, LNCS*, Vol. 11182, pp. 211–222. DOI: 10.1007/978-3-030-01449-0_18.
18. **Nguyen, H., Maclagan, S. J., Nguyen, T. D., Nguyen, T., Flemons, P., Andrews, K., Ritchie, E. G., Phung, D. (2017).** Animal recognition and identification with deep convolutional neural networks for automated

- wildlife monitoring. 2017 IEEE international conference on data science and advanced Analytics, Vol. 2018, pp. 40–49. DOI: 10.1109/DSAA.2017.31.
19. **Favorskaya, M., Pakhirka, A. (2019).** Animal species recognition in the wildlife based on muzzle and shape features using joint CNN. *Procedia Computer Science*, Vol. 159, pp. 933–942, DOI: 10.1016/J.PROCS.2019.09.260.
 20. **Silva, L. C., Pádua, M. B., Ogusuku, L. M., Keese-Albertini, M., Pimentel, R., Backes, A. R. (2021).** Wild boar recognition using convolutional neural networks. *Concurrency and Computation: Practice and Experience*, Vol. 33, No. 22. DOI: 10.1002/CPE.6010.
 21. **Binta-Islam, S., Valles, D., Hibbitts, T. J., Ryberg, W. A., Walkup, D. K., Forstner, M. R. (2023).** Animal species recognition with deep convolutional neural networks from ecological camera trap images. *Animals*, Vol. 13, No. 9, p. 1526. DOI: 10.3390/ANI13091526.
 22. **Gomez Villa, A., Salazar, A., Vargas, F. (2017).** Towards automatic wild animal monitoring: Identification of animal species in camera-trap images using very deep convolutional neural networks. *Ecological Informatics*, Vol. 41, pp. 24–32. DOI: 10.1016/j.ecoinf.2017.07.004.
 23. **Tan, M., Chao, W., Cheng, J. K., Zhou, M., Ma, Y., Jiang, X., Ge, J., Yu, L., Feng, L. (2022).** Animal detection and classification from camera trap images using different mainstream object detection architectures. *Animals*, Vol. 12, No. 15, p. 1976. DOI: 10.3390/ani12151976.
 24. **Castañeda, J. A. J., De-Castro, A. L., Sy, M. A. G., Aïdahoul, N., Tan, M. J. T., Karim, H. A. (2022).** Development of a detection system for endangered mammals in Negros Island, Philippines using YOLOv5n. *International Conference on Computational Science and Technology Singapore: Springer Nature Singapore*, Vol. 983, pp. 435–447. DOI: 10.31219/osf.io/62vje.
 25. **Islam, S. B., Valles, D. (2020).** Identification of wild species in Texas from camera-trap images using deep neural network for conservation monitoring. 2020 10th Annual Computing and Communication Workshop and Conference IEEE, pp. 537–542. DOI: 10.1109/CCWC47524.2020.9031190.
 26. **Sato, D., Zanella, A. J., Costa, E. J. X. (2021).** Computational classification of animals for a highway detection system. *Brazilian Journal of Veterinary Research and Animal Science*, Vol. 58, pp. e174951–e174951. DOI: 10.11606/ISSN.16784456.BJVRAS.2021.174951.
 27. **Alnujaidi, K., AlHabib, G. (2023).** Computer vision for a camel-vehicle collision mitigation system. Vol. 12, No. 1, pp. 141–149. DOI: 10.5121/ijci.2023.120111.
 28. **Abel, L. A. J., Oconer, T. C. N., Cruz, J. C. D. (2022).** Realtime object detection of pantry objects using YOLOv5 transfer learning in varying lighting and orientation. 2022 2nd International Conference on Innovative Research in Applied Science, Engineering and Technology IEEE, pp. 1–7. DOI: 10.1109/IRASET52964.2022.9738370.
 29. Jocher, G. et al. (2022). Ultralytics/yolov5: v7.0 - YOLOv5 SOTA Realtime Instance Segmentation', DOI: 10.5281/ZENODO.7347926.
 30. **Wang, C. Y., Bochkovskiy, A., Liao, H. Y. M. (2023).** YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 7464–7475.
 31. Dorfer, T. A. (2023). Enhanced object detection: How to effectively implement YOLOv8. <https://towardsdatascience.com/enhanced-object-detection-how-to-effectively-implement-yolov8-afd1bf6132ae>.
 32. **Terven, J., Cordova-Esparza, D. (2023).** A comprehensive review of YOLO: From YOLOv1 to YOLOv8 and beyond. *arXiv preprint arXiv:2304.00501*.
 33. **Rodriguez-Conde, I., Campos, C., Fdez-Riverola, F. (2022).** Optimized convolutional neural network architectures for efficient on-device vision-based object detection. *Neural Computing and Applications*, Vol. 34, No. 13, pp. 10469–10501. DOI: 10.1007/S00521-021-06830-W.

34. **He, K., Zhang, X., Ren, S., Sun, J. (2015).** Spatial pyramid pooling in deep convolutional networks for visual recognition. *IEEE transactions on pattern analysis and machine intelligence*, Vol. 37, No. 9, pp. 1904–1916. DOI: 10.1109/TPAMI.2015.2389824.
35. **Liu, S., Qi, L., Qin, H., Shi, J., Jia, J. (2018).** Path aggregation network for instance segmentation. *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 8759–8768. DOI: 10.1109/CVPR.2018.00913.
36. **Rainforest Expeditions (2022).** Guacamayo Azul y Amarillo. <https://rainforestexpeditions.com/es/animales/guacamayo-azul-y-amarillo/>.
37. **Cueto-Salazar, M. R. (2018).** Ararauna Tours and Expeditions. <http://araraunatambopata.com/nosotros/>.
38. **Chrome-stats (2021).** 'Fatkun Batch Descargar imagen. <https://chrome-stats.com/d/nnjahlikiabnchcpehckpckdeckfgnohf?hl=es>.
39. **Seetala, K., Birdsong, W., Reddy, Y. B. (2019).** Image classification using tensorflow. *16th International Conference on Information Technology-New Generations ITNG 2019*, Springer International Publishing, Vol. 800, pp. 485–488. DOI: 10.1007/978-3-030-14070-0_67.
40. **MINAM-Perú (2020).** Vive como sueñas. Reserva Nacional Tambopata, YouTube, https://www.youtube.com/watch?v=cDbkiWhs7SU&t=613s&ab_channel=MinisteriodelAmbiente-Perú.
41. **MINAM-Perú (2023).** Ministerio del ambiente. Plataforma del Estado Peruano, 2023. <https://www.gob.pe/minam>.
42. **Antara-Perú (2023).** Manu & Tambopata. YouTube, [https://www.youtube.com/watch?v=hxoDoBH_7mA&t=3s&ab_channel=Antara Perú](https://www.youtube.com/watch?v=hxoDoBH_7mA&t=3s&ab_channel=AntaraPerú).
43. **Beleites, C., Salzer, R., Sergio, V. (2013).** Validation of soft classification models using partial class memberships: An extended concept of sensitivity & co. applied to grading of astrocytoma tissues. *Chemometrics and Intelligent Laboratory Systems*, Vol. 122, pp. 12–22. DOI: 10.1016/j.chemolab.2012.12.003.
44. **Palma-Ttito, L., Holgado-Apaza, L., Ccapa-Luque, R., Canaza-Canqui, E., Cornejo-Aparicio, V. (2021).** Detection of oligonucleotide microarray mutations by multiclass support vector machine. *RISTI Revista Iberica de Sistemas e Tecnologias de Informacao*, Vol. 2021, No. E39. pp. 643–657.
45. **Corrêa-Krüger, J. G., de-Souza-Britto-Jr, A., Barddal, J. P. (2023).** An explainable machine learning approach for student dropout prediction. *Expert Systems with Applications*, Vol. 233, p. 120933. DOI: 10.1016/J.ESWA.2023.120933.
46. **Kumar, A., Kalia, A., Kalia, A. (2022).** ETL-YOLO v4: A face mask detection algorithm in era of COVID-19 pandemic. *Optik*, Vol. 259, p. 169051. DOI: 10.1016/J.IJLEO.2022.169051.
47. **Suo, R., Gao, F., Zhou, Z., Fu, L., Song, Z., Dhupia, J., Cui, Y. (2021).** Improved multi-classes kiwifruit detection in orchard to avoid collisions during robotic picking. *Computers and Electronics in Agriculture*, Vol. 182, p. 106052. DOI: 10.1016/J.COMPAG.2021.106052.
48. **Fu, L., Yang, Z., Wu, F., Zou, X., Lin, J., Cao, Y., Duan, J. (2022).** YOLO-banana: A lightweight neural network for rapid detection of banana bunches and stalks in the natural environment. *Agronomy*, Vol. 12, No. 2. DOI: 10.3390/AGRONOMY12020391.
49. **Jocher, G. (2020).** YOLOv5 by ultralytics. DOI: 10.5281/zenodo.3908559.
50. **Li, C., Li, L., Jiang, H., Weng, K., Geng, Y., Li, L., Wei, X. (2022).** YOLOv6: A single-stage object detection framework for industrial applications. *arXiv preprint arXiv:2209.02976*.
51. **Wang, C. Y., Bochkovskiy, A., Liao, H. Y. M. (2023).** YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 7464–7475.
52. **Jocher, G., Chaurasia, A., Qiu, J. (2023).** YOLOv8 by Ultralytics. Ultralytics. <https://github.com/ultralytics/ultralytics>.

Article received on 11/10/2023; accepted on 11/03/2024.

**Corresponding author is Luis Alberto Holgado-Apaza.*