Revista Ciência Agronômica, v. 56, e202392543, 2025
Centro de Ciências Agrárias - Universidade Federal do Ceará, Fortaleza, CE
www.ccarevista.ufc.br

Scientific Article
ISSN 1806-6690

Fish oocyte morphology detection using neural networks: a comparison of YOLO architectures[1]

Yanna Leidy Ketley Fernandes Cruz[2]*, Ewaldo Eder Carvalho Santana[23], Isa Rosete Araujo Nascimento[4], Antonio Fhillipi Maciel Silva[5], Raimunda Nonata Fortes Carvalho Neta[6], José Ribamar de Souza Torres-Junior[7]

ABSTRACT - The recognition of oocytes, in their maturational stages, allow estimate the ovarian development and the type of spawning of a species. Although, distinguishing oocytes on histological images requires a visual and subjective interpretation by the specialist. With the development of deep learning techniques, automatic object detection has become an important mechanism for this task. However, studies that use deep learning techniques have not been widely explored for the analysis of fish oocyte samples so far. In this paper, we propose the use of YOLO, a family of convolutional neural networks, for oocyte morphology detection of *Centropomus undecimalis* fish. The research uses an image database with 5,680 oocytes with different maturation stadiums (PV - pre-vitellogenesis, VI - early vitellogenesis and VF - late vitellogenesis), in histological images, divided into training, testing and validation, and detection performed by YOLOv3, YOLOv4, and YOLOv5 architectures. The results obtained were promising, highlighting that the YOLOv5l model, in the detection of oocytes of the VF class, reached the best values in the metrics precision, recall, mAP@.5 and mAP@.95, with 85.4%, 95.3%, 95.7%, and 75.9%, respectively. When considering all classes, YOLOv5l was the model that obtained the best results in the analyzed metrics.

Key words: Maturational phases. Histological images. Artificial intelligence. Convolutional Neural Networks. *Centropomus undecimalis.*

DOI: 10.5935/1806-6690.20250038
Editor-in-Chief: Prof. Alexandre Holanda Sampaio - sampaioa@ufc.br
*Author for correspondence
Received for publication 07/12/2023; approved on 17/09/2024
[1]Work extracted from the thesis of the first author presented to the Graduate Program in Electrical Engineering, Federal University of Maranhão/MA
2Graduate Program in Electrical Engineering, Federal University of Maranhão (UFMA), São Luís-MA, Brazil, yanna.cruz@discente.ufma.br (ORCID ID 0000-0003-4143-8167), ewaldosantana@professor.uema.br (ORCID ID 0000-0002-8894-5353)
[3]Graduate Program in Computer and Systems Engineering, State University of Maranhão (UEMA), São Luís, Bazil, ewaldosantana@professor.uema.br (ORCID ID 0000-0002-8894-5353)
[4]Federal Institute of Education, Science and Technology of Maranhão (IFMA), São Luís-MA, Bazil, isabio@ifma.edu.br (ORCID ID 0000-0003-0170-765X)
[5]Computer Science Department, State University of Piauí (UESPI), Floriano-PI, Brazil, fhillipi.ti@gmail.com (ORCID ID 0000-0002-0410-0768)
[6]Graduate Program in Biodiversity and Biotechnology of the Bionorte Network, State University of Maranhão (UEMA), São Luís-MA, Bazil, raifortes@gmail.com (ORCID ID 0000-0002-3519-5237)
[7]Oceanography and Limnology Department, Federal University of Maranhão (UFMA), São Luís-MA, Brazil, jose.torres@ufma.br (ORCID ID 0000-0002-6356-6517)

INTRODUCTION

A Fish is a food of excellent nutritional value due to its high biological value proteins, vitamins and unsaturated fatty acids (Albuquerque; Vieira; Vieira, 2006). According to FAO data, fish consumption by the global population has been increasing about 1.5% per year, shifting from 9 kg/inhabitant in 1961 to 20.7 kg/inhabitant in 2022 (FAO, 2024). In Brazil, in 2022, there was an increase of 3% compared to the previous year, reaching 2.15 million tons, the second best performance since 2014, according to the SeafoodBrasil (2022).

Specifically, the wide capture of the species *Centropomus undecimalis* (Bloch, 1792) through artisanal, industrial and sport fishing, developed on the Amazon coast, has been causing concerns (Pereira *et al*., 2020), since most of the marine fish reserves, about 70%, are found in areas where there is intense exploration. In view of this fact, the risk of interruptions in the reproductive cycle of these fish is increased (Rosa; Lima, 2008). Such a number, in addition to affecting world marine fishing, threatens food security on a global scale.

Faced with this situation, the lack of monitoring of this fish makes it impossible to analyze information regarding biological and reproductive aspects. In this sense, studies that address the determination of the development and sexual maturation of fish are fundamental to the science of fisheries and are prerequisites for understanding the life cycle of fish (Sivakumaran *et al*., 2003). The most suitable method for determining the reproductive cycle in fish is the observation of changes in gonadal (Karlou-Riga; Panos, 1996). Especially in studies related to females, oocytes are fundamental for this analysis.

Thus, researchers require the use of histological images, that provide information on morphological changes in the germ cells of a species (Mendonça, 2004). Germ cells, better known as oocytes, are structures that have a generally spherical shape, consisting of follicular cells, yolk envelope, yolk granules, cytoplasm, and a nucleus (Rhody *et al*., 2013).

The maturational stages described by Vazzoler (1996) are fundamental to understanding the reproductive development of organisms. The first stage, called immature, is characterized by the absence of evident secondary sexual characteristics and by the initial development of the reproductive organs, with no signs of reproductive activity and with an immature gonadal structure. In the developing stage, the reproductive organs begin to show signs of maturation, with the formation of gametes, such as eggs in the ovaries and sperm in the testes, although there is still no effective reproductive activity. The mature stage is marked by the full development and functionality of the reproductive organs, with ovaries containing eggs ready for spawning and testes producing sufficient sperm for fertilization. Finally, in the post-reproductive or senescent stage, a reduction in reproductive activity is observed, with regression of the reproductive organs and a decrease in reproductive capacity. The classification of these stages is crucial for studies of biology and ecology, particularly in the management and conservation of species populations.

Object detection is an important computer vision task that, aims to identify and locate objects (Redmon *et al*., 2016). For this purpose, techniques that address convolutional neural networks have been used that aim to bring machine intelligence closer to the human level, making them capable of solving any problem in a specific subject (Fracarolli *et al*., 2021). One of these is YOLO (You Only Look Once), developed originally by Joseph Redmon in 2016, which consists of a method to identify and locate objects with a single forward pass, which speeds up predictions in real time. YOLO was originally developed for detecting objects in natural images, but the algorithm can be applied to different domains like skin (Nie *et al*., 2019) and cell cancer detection (Aly *et al*., 2021), blood cell count (Jiang *et al*., 2021), outer space object detection (Liu; Xiao; Chengchao, 2019), zebra-fish monitoring (Barreiros *et al*., 2021). Here we propose using YOLO networks for detection and identification of oocytes.

Distinguishing objects on histological images is a time-consuming and error-prone procedure, as it requires visual and subjective interpretation by the expert. In this paper, a method is proposed to detect oocytes at different stages in histological images using the YOLOv3, YOLOv4 and YOLOv5 architectures. Thus, the development of computational methods should speed up and improve the identification of oocytes, which on its turn which on it is turn is an important step on monitoring of the economically significant *Centropomus undecimalis* fish species.

MATERIAL AND METHODS

Our method has four steps: acquisition of images, database partitioning, data augmentation, and application of YOLO architectures for object detection and classification. We evaluated YOLO architectures to identify the stages of oocyte development pre-vitellogenesis, early vitellogenesis and final vitellogenesis. Figure 1 shows the steps of the proposed method for automatic oocyte detection.

Acquisition of Images

The samples of histological images of the fish gonads were obtained from the Graduate Program in Biodiversity and Biotechnology of the Legal Amazon (BIONORTE) and the Graduate Program in Computer and Systems Engineering of the State University of Maranhão (UEMA).

The samples of fish of the species *C. undecimalis* were obtained through direct purchase at a fish merchant, found in the city of Tutóia (2°45'44"' S; 42°16'28"' W), the state of Maranhão (Figure 2). All samples were bought monthly during the period from January 2019 to February 2020.

The fish were stored in Styrofoam with ice, sealed with tape and sent to the Laboratory of Fisheries and Aquatic Ecology at UEMA, where they were identified. The specimens were then counted, measured, weighed, and processed to remove gonads and other viscera. Some gonads were stored in flasks with Bouin's solution for 24 hours and then transferred to a 70% alcohol solution for further histological analysis.

Figure 1 - Steps of the method used in this work

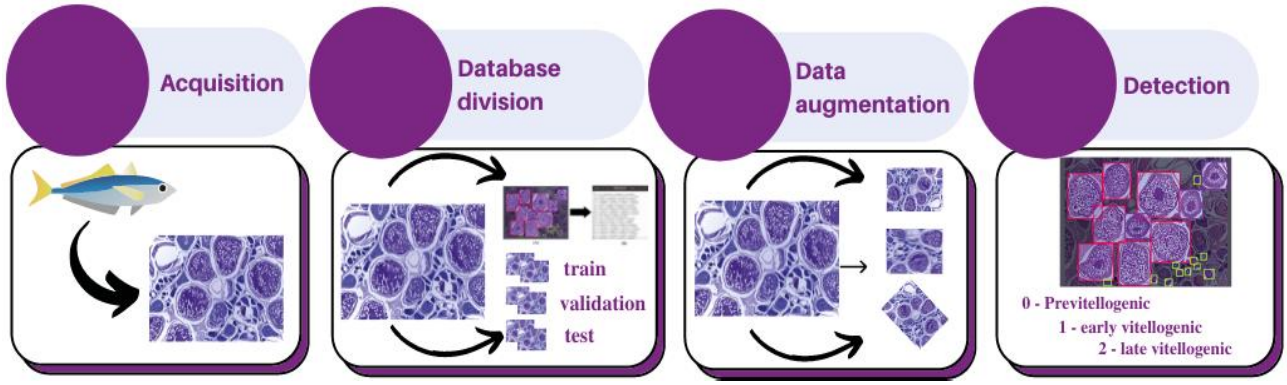
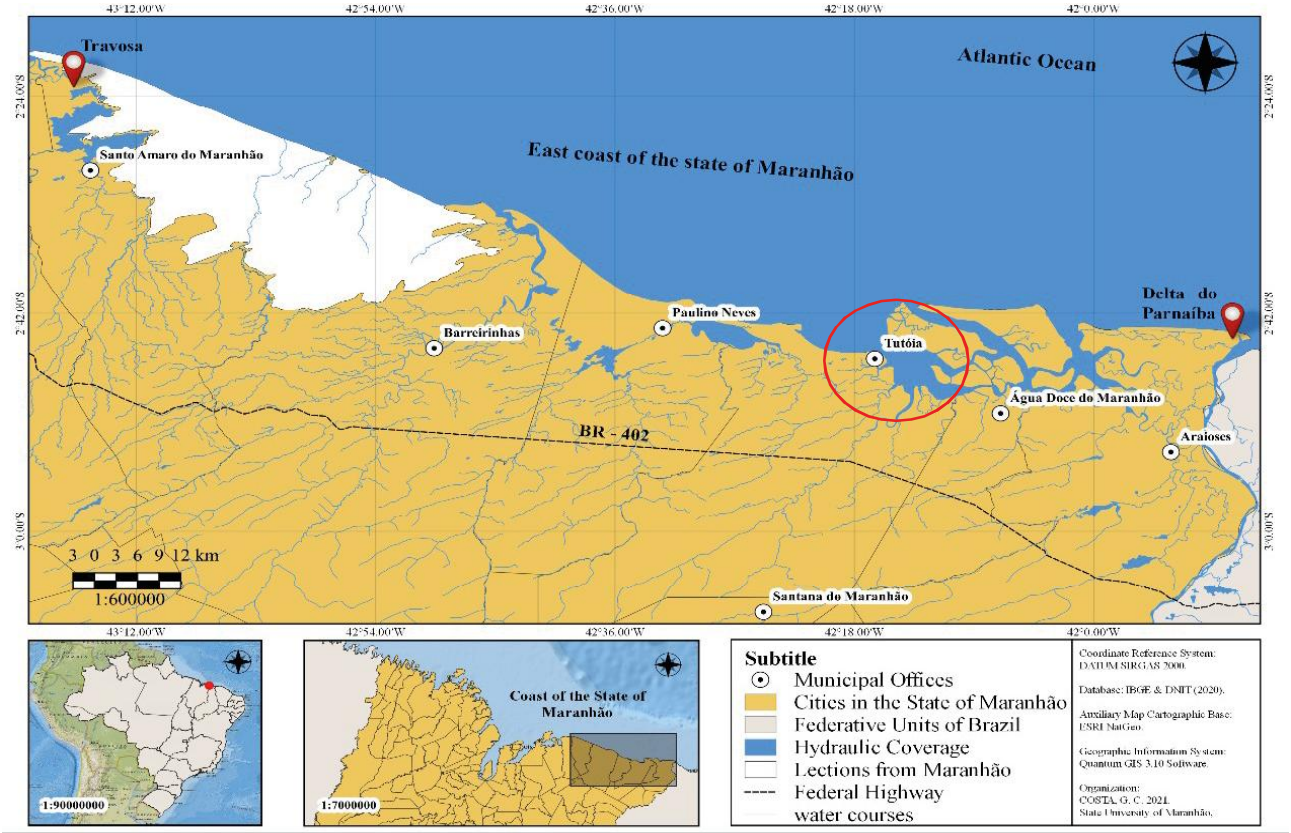


Figure 2 - Map of the study area highlighting with location points the sample capture range (from Travosa to the Parnaíba Delta) and with a red circle the municipality of Tutóia, place of purchase of *C. undecimalis* sea bass specimens on the north coast of Brazil



In this study, a total of 15 females were evaluated, with the aim of investigating the dynamics and maturation of female gametes. The choice to focus exclusively on female gametes was due to the fact that detailed analysis of the eggs can provide crucial information about the reproductive processes and fertilization potential of the population studied. In addition, the study focused on females allows a more specific approach to the management and conservation of aquatic species (Vazzoler, 1996).

In the process of preparing slides, histological procedures were used to section the ovaries. Sections were stained with Hematoxylin and Eosin (H&E), then images were captured by a LEICA DM 500 microscope in an enlargement of x200. The additional equipment used were a LEICA EC4 digital camera directly connected to the microscope and LEICA LAS EZ image acquisition software, compatible with the equipment and a Sony monitor for viewing.

The image base developed consists of 305 histological images. This base can be made available later, allowing more research to be conducted in this field, expanding the monitoring of this species, since the fish under study presents a peculiar characteristic of sexuality, the protandrous hermaphroditism, are individuals that first mature as males and later change sex and remain mostly as females for the rest of their lives (Sousa, 2011).

Ethics Committee

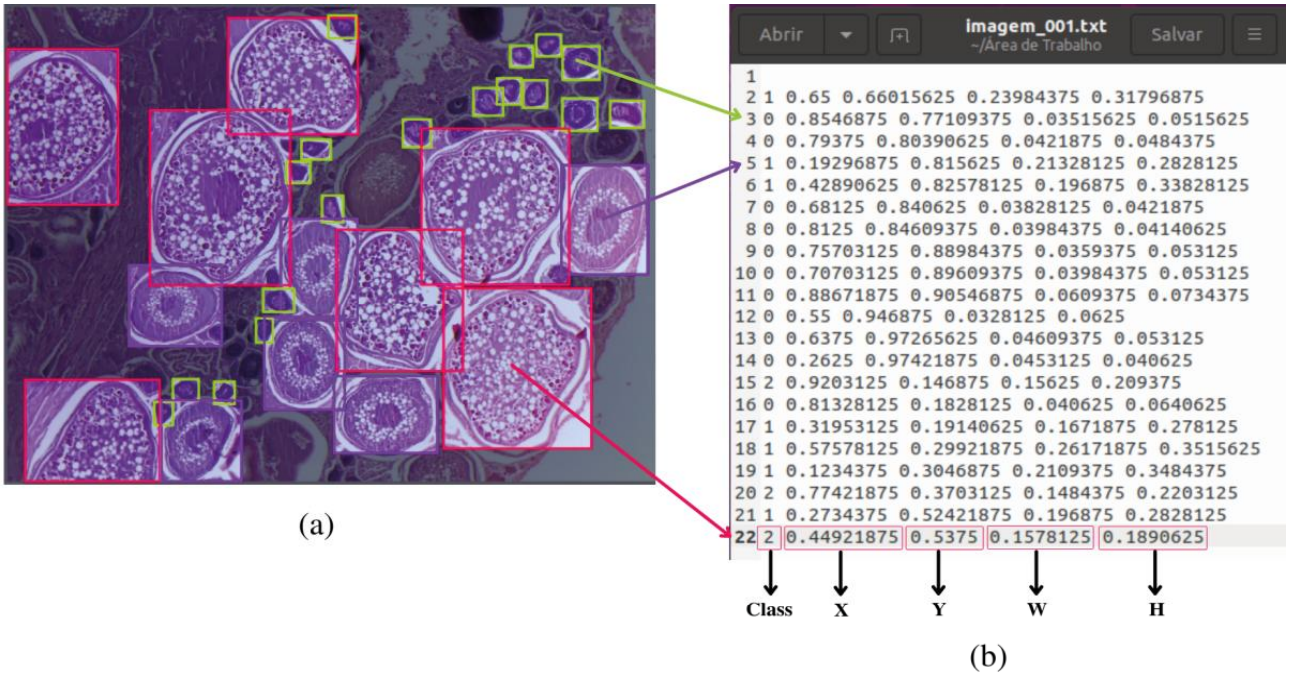
All methodological procedures for fish management were approved by the Ethics Committee of the State University of Maranhão (n°. 4.476.902/2020). The species of *C. undecimalis* were acquired monthly through commercial purchases in the city of Tutóia, Maranhão, from January 2019 to February 2020 (Nascimento *et al.*, 2022).

Division of the Image Base

The Database images were used in order to organize, label and divide the images with histological cuttings of the fish C. *undecimalis*. The software Roboflow was used to support this task. After the acquisition step, the expert, making use of the delimiter box, labeled in TXT format, 305 images, following specific patterns for YOLOv3, YOLOv4 and YOLOv5. The generated file consists of a set of coordinates that identify the objects in an image with annotations in a standardized way (Everingham *et al.*, 2015).

In the structure of the TXT file one object per line is presented, in each line information is provided such as class number, center in x, center in y, width in w and height in h of the object. The classes are numbered, starting with zero: Class 0 (PV - pre-vitellogenesis), class 1 (VI - early vitellogenesis) and class 2 (VF - late vitellogenesis). According to Redmon and Farhadi (2018), the coordinates of the object bounding box should be arranged in xywh normalized between [0,1]. Figure 3 shows the labeled oocytes in the image with their respective classes and coordinates.

Figure 3 - Process of labeling the oocytes in the image. (a) Image labeled by the specialist. (b) File in TXT format for YOLOv5 pytoch



One of the attributes present in the labels represents the stages of oocyte maturation development. The phases used in this research were chosen due to the frequency of these oocytes in the images, because according to Chlap *et al*. (2021) deep learning-based techniques require a large number of objects for a more efficient detection.

To validate the detection, model the image base was randomly subdivided into three sets, according to what is specified in the research of Nie *et al*. (2019):

Training: the training set contains 244 images (2,393 PV; 633 VI; 1,500 VF);

Validation: the validation set contains 31 images (299 PV; 51 VI; 139 VF);

Test: the test set contains 30 images (374 PV; 98 VI; 193 VF).

A set of 5,680 oocytes were labeled by an expert. The class 0 (PV) has 3,066 oocytes, the class 1 (VI) has 782 oocytes and the class 2 (VF) has 1,832 oocytes. Table 1 shows in detail the division of the data.

Data augmentation

Data augmentation is a set of techniques used to create new training examples from existing data (Hussain *et al.*, 2018). These techniques are intended to auxiliary in the generalization of the machine learning model. For this research, the image base built contains 305 samples, which is considered small for training a deep learning model. Making use of the Roboflow software (Bochkovskiy; Wang; Liao, 2020), we employed the following data augmentation techniques for robust prediction, Flip, 90° Rotation, Crop, and Random Rotation.

Flipping, or inversion, consists of creating mirrored versions of the images, helping the model to recognize objects in different orientations and perspectives, thus increasing generalization (Shorten; Koehring, 2019). 90° rotation applies fixed rotations to provide the model with a diverse view of the images, promoting invariance to specific rotations and facilitating the identification of objects at any angle (Tamminen *et al*., 2018). Cropping involves randomly cutting out parts of the images, forcing the model to learn relevant features from different regions, which improves its ability to deal with partially visible images and vary in scale (Zhao *et al.*, 2020). Finally, random rotation applies rotations at varying angles, increasing data variability and allowing the model to become more robust in recognizing objects regardless of their orientation (Hong; Bai; Cao, 2016).

Machine learning sample splitting typically follows the practice of separating data into three distinct sets: training, validation, and testing. The common setup for these splits is to use approximately 80% of the data for training, 10% for validation, and 10% for testing (Kohavi; John, 1997). Therefore, after using these techniques, a set of 759 images were generated, which were divided into 708 images for training, 31 images for validation and 30 images for testing.

Evaluation Metrics

Various evaluation metrics are used to find out the performance of the classifier. The evaluation metrics used in the study are explained below.

A confusion matrix was used to express classification accuracy numerically. The confusion matrix is one of the most commonly used techniques in machine learning, and it includes information about the actual and predicted classes obtained by a classification system. The confusion matrix has two dimensions: actual and predicted classes. While each row represents an actual class example, each column represents the state of a predicted class. In the confusion matrix, TP is the number of true positive, TN is the number of true negative, FP is the number of false positive, and FN is the number of false negative (Taner; Öztekin; Duran, 2021).

Precision

Precision shows how much of the data predicted as positive are predicted correctly. In other words, high precision means fewer false positives.

Table 1 - Database division

Dataset	Number images	Number of oocytes per image			Total oocytes
		PV	VI	VF	
Train	244	2.393	633	1.500	4.526
Validation	31	299	51	139	489
Test	30	374	98	193	665
Total	305	3.066	782	1.832	5.680

PV - pre-vitellogenesis, VI - early vitellogenesis, VF - late vitellogenesis

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

Recall

The recall is the metric of determining the completeness of the classifier. Higher recall indicates lower false negatives, while lower recall indicates higher false negatives. Precision often decreases with an improvement in recall.

$$\text{Re}\,call = \frac{TP}{TP + FN} \quad (2)$$

Mean Average Precision (mAP)

The Mean Average Precision (mAP) metric is widely used to evaluate the performance of image detection systems. To calculate the mAP (Equation 3), the Precision and Recall for each class are initially obtained, where the Average Precision (AP) is calculated as the area under the precision-recall curve. The AP metric evaluates the accuracy of the model at different recall points, capturing the model's ability to correctly identify relevant objects while minimizing false positives. The mAP is then calculated as the average of the APs of all classes, providing an overview of the system's ability to deal with multiple object categories (Everingham *et al*., 2015).

$$mAP = \frac{1}{N}\sum_{i=1}^{N} Ap_i \quad (3)$$

Architectures adopted

YOLO is a method for object detection that is based on the architecture of Convolutional Neural Networks. It divides the image into a grid of cells and predicts bounding boxes, confidence intervals for these boxes and the class probability for each cell, allowing a unique assessment of the image (Redmon *et al*., 2016).

The detection architectures were used with the aim of classify and localize the oocytes in images according to their phases. YOLO versions 3, 4 and 5 were selected for the experiments due to the fact that these architectures have the ability to identify variations of objects of the same class in different dimensions (Golatkar; Anand; Sethi, 2018; Redmon *et al*., 2016; Redmon; Farhadi, 2018).

In the YOLOv3 version, the models YOLOv3-tiny and YOLOv3-ssp were applied (Redmon; Farhadi, 2018). In the YOLOv4 version, the models YOLOv4-tiny and YOLOv4-csp were applied (Bochkovskiy; Wang; Liao, 2020). In the YOLOv5 version, the models YOLOv5s, YOLOv5m, YOLOv5n, YOLOv5l and YOLOv5x were applied (Ultralytics, 2022). For the choice of these models the high-performance characteristics in real-time applications were considered, which according to Silva, Narciso and Gonçalves (2019) stand out by the use of independent logic classifiers for multiple classification with class overlap. Luo *et al.* (2019) points out that these versions present higher performance when dealing with object of small sizes.

A common point of all YOLO object detection architectures is that the resources of the input image are compressed through a resource extractor called the backbone and then forwarded to an object detector (Detection Neck and Detection Head). Neck works as a resource aggregator that has the task of mixing and matching the resources formed in the backbone and preparing them for the detection step to be carried out by the Detection Head, responsible for carrying out the detection, including the location and classification of objects (Redmon *et al*., 2016).

Table 2 - Composition of the YOLOv3, YOLOv4, and YOLOv5 architectures

Model	Backbone	Neck	Head
	YOLOv3		
YOLOv3-tiny	Tiny	FPN	YOLO layer
YOLOv3-ssp	Darknet53	FPN	
	YOLOv4		
YOLOv4-tiny	Tiny	FPN	YOLO layer
YOLOv4-csp	CSPDarket53	PANet	
	YOLOv5		
YOLOv5l	CSPDarket53 with an SPPF Layer	PANet	YOLO layer
YOLOv5m			
YOLOv5n			
YOLOv5s			
YOLOv5x			

To extract essential information from the input image, the YOLOv3-tiny and YOLOv3-ssp models use the Tiny and Darknet53 backbone respectively. In the YOLOv4-tiny and YOLOv4-csp models, the Tiny and CSPDarket53 backbones are used, respectively. The YOLOv5l, YOLOv5m, YOLOv5n, YOLOv5s and YOLOv5x models use a CSPDarket53 backbone plus a Spatial Pyramid Pooling SPPF layer.

To increase the receptive field and separate the most important features, the YOLOv3-tiny, YOLOv3-ssp and YOLOv4-tiny models use the Neck Filter Pyramid Network (FPN). While the YOLOv4-csp, YOLOv5l, YOLOv5m, YOLOv5n, YOLOv5s and YOLOv5x models use the Neck Path Aggregation Network (PANet).

To perform the final prediction, the YOLOv3-tiny, YOLOv3-ssp, YOLOv4-tiny, YOLOv4-csp, YOLOv5l, YOLOv5m, YOLOv5n, YOLOv5s and YOLOv5x models use a YOLO layer to generate a vector containing the bounding box coordinates: width, height, class label and class probability. Table 2 summarizes the composition of the YOLOv3, YOLOv4, and YOLOv5 architectures.

Environments for the Experiments

The models were trained for 300 epochs at the most, as per the YOLO documentation and as observed in correlated work (Abas; Abdulazeez; Zeebaree, 2022; Ultralytics, 2023; Wang *et al.*, 2020). Additionally, training was interrupted at various times as no improvement was observed after the maximum number epoch run. The following hyperparameters were used in these models: optimizer = SGD, lr0 = 0.01, lrf = 0.2, momentum = 0.937, weight_decay = 0.0005, warmup_epochs = 3.0, warmup_momentum = 0.8, warmup_bias_lr = 0.1, box = 0.05, cls = 0.5, cls_pw = 1.0, obj = 1.0, obj_pw = 1.0, iou_t = 0.2, anchor_t = 4.0, fl_gamma = 0.0, hsv_h = 0.015, hsv_s = 0.7, hsv_v = 0.4, degrees = 0.0, translate = 0.1, scale = 0.5, shear = 0.0, perspective = 0.0, flipud = 0.0, fliplr = 0.5, mosaic = 1.0, mixup = 0.0.

The experiments were conducted using the free Google Collaboratory platform (Colab), which is a cloud storage service for laptops focused on creating and running code in Python, directly in a browser (Bisong *et al.*, 2019) running on a Graphics Processing Unit (GPU). The hardware configurations include a NVIDIA Tesla P100 tensor-core graphics processing unit (GPU), 12GB of RAM and 68GB of disk. Colab provides an environment configured with Python 3, and the manual installation of the Keras, OpenCV, TensorFlow, Matplotlib, NumPy and SciPy libraries.

RESULTS AND DISCUSSIONS

Model of training and testing

To locate the objects correctly, during the training phase, the weights of the connections were adjusted according to the patterns present in the images. A comparison of weight values, training time, mean mAP@.5 and inference time are presented in Table 3. As a result, a model was generated containing trainable and untrainable parameters, expressed in millions (M). After training, the investigated models showed substantial variations in weight generation. Regarding the number of weights were generated for the models YOLOv5x 86 (M) and YOLOv5n 7 (M). In terms of training time, the models reached 5h 35min (longer time) for the YOLOv5x model and 0h 37min (shortest time) for the YOLOv5n model.

Table 3 - Comparison of the Different Detection Models

Model	Model Weight (M)	Training Time (h min)	mAP@.5 (%)	Inference Time (ms)
YOLOv3				
YOLOv3-tiny	61	0 h 46 min	72.6	2.1
YOLOv3-ssp	62	4 h 19 min	71.8	2.2
YOLOv4				
YOLOv4-tiny	68	4 h 25 min	72.3	1.2
YOLOv4-csp	72	3 h 16 min	74.8	1.6
YOLOv5				
YOLOv5l	46	3 h 32 min	75.0	0.9
YOLOv5m	20	2 h 16 min	71.4	0.8
YOLOv5n	17	0 h 37 min	68.9	0.9
YOLOv5s	7	0 h 54 min	74.1	0.9
YOLOv5x	86	5 h 35 min	71.6	0.7

Measure mAP@.5 allows validation of the classification with the assessment of the fitness of the bounding box to the detected object (Kubera *et al*., 2022). The mean accuracy for 50% confidence (mAP@.5) in the detection of oocytes, of each model, was calculated in relation to all stages of development (PV, VI and VF). The values obtained for mAP@.5 showed little variation, with 75% to 71.4%, however, for the YOLOv5n model there was an average accuracy of 68.9%, considered the lowest among all models.

In the test stage, the specialist should perform the detection of oocytes in histological images quickly and accurately, for which a model capable of automating the visual process of identifying oocytes on a slide is applied. This process is calculated considering the inference time (expressed in milliseconds, ms) in the test images. Thus, in this research, results were obtained that varied in inference time between 2.1 ms for the YOLOv3-tiny model and 0.7 ms for the YOLOv5x model.

Comparison of the Different Detection Models

To verify the effectiveness of the architectures proposed in this work, the models YOLOv3, YOLOv4, and YOLOv5 were comparatively evaluated, considering the metrics precision (P), recall (R) and measure AP (mAP) for each class. All models were trained and tested using the same dataset and environment in the experiments. In Table 4, a comparison of the performance of the models is presented.

Nagahama and Yamashita (2008) point out that early in oocyte development, the initial vitellogenesis phase is essential for the final quality of the oocyte, as it is in this phase that there is a large accumulation of substances that will be responsible for part of the yolk production in the subsequent phases of oocyte development.

The models, for the most part, presented values greater than 90% in the recall metric in the detection of oocytes in phase VI. YOLOv5m reached 97.5% recall, while YOLOv3-ssp reached the lowest value, only 89.1%. The importance of precise identification of this phase is highlighted, as in this phase, the synthesis and accumulation of the substance responsible for the survival of embryos occurs Nagahama and Yamashita (2008).

The YOLO architecture uses two important elements to detect an object: location (bounding box) and classification. Thus, when applying the mAP metric, a mechanism for evaluating an object detected from an established threshold is provided. In the case of oocyte detection, thresholds between 50% and 95% are considered to assess the true location and classification of objects. Considering the mAP@.5 metric, the YOLOv5l model presented the highest detection rate with 95.7% for the VF phase and 78% and 90.2% for the PV and VI phases, respectively. When increasing the threshold to 95%, it was possible to observe that there was a decrease in the results, however, the YOLOv5l model remained with the highest rate, reaching 75.9% of detection.

It is possible to observe in Table 4, highlighted in gray, that the class VF, in general, was the class in which the models reached the highest rates to Recall, mAP@.5 and mAP@.95. This occurred because the classifiers handle better with oocytes that present, on average, diameters of 388.3 μm (Mendonça, 2004). This confirms that objects with larger dimensions show more satisfactory results in the detection task (Kubera *et al*., 2022; Wang *et al*., 2020).

Table 4 - Comparison of performance of models based on measurements of precision (P), recall (R) and AP average (mAP) for each class

Model	P (%)			R (%)			mAP@.5 (%)			mAP@.95 (%)		
	PV	VF	VI	PV	VF	VI	PV	VF	VI	PV	VF	VI
YOLOv3												
YOLOv3-tiny	67.0	68.1	71.1	73.5	96.7	63.3	72.6	88.2	70.0	36.5	52.6	35.3
YOLOv3-ssp	85.0	78.8	77.2	74.7	89.1	77.6	80.5	90.8	82.8	48.4	59.3	57.2
YOLOv4												
YOLOv4-tiny	69.9	70.7	74.9	51.0	96.9	62.2	56.6	89.7	70.6	16.3	55.8	40.1
YOLOv4-csp	79.7	76.0	79.5	55.1	92.2	75.5	65.2	90.7	81.9	24.9	62.3	56.4
YOLOv5												
YOLOv5l	83.0	85.4	79.3	64.2	95.3	85.7	78.0	95.7	90.2	61.7	75.9	67.9
YOLOv5m	82.9	70.3	89.4	79.3	97.4	77.6	85.9	95.4	86.2	59.2	73.6	64.1
YOLOv5n	79.0	56.2	50.7	41.8	97.1	68.6	61.4	88.3	61.7	27.9	58.5	31.8
YOLOv5s	68.6	58.5	52.3	48.7	95.3	70.6	57.5	84.6	65.1	24.7	57.5	36.7
YOLOv5x	78.5	79.6	82.2	66.8	94.4	84.7	75.4	95.4	86.2	36.0	73.4	58.7

PV - pre-vitellogenesis, VI - early vitellogenesis, VF - late vitellogenesis

It is also observed that the oocytes in the PV phase, despite having a rounded shape and uniform staining, were not the phase with the highest detection rates. However, in phase VI, the external structure of the oocyte loses its spherical shape, presenting heterogeneous color and its internal structure contains tiny and whitish nucleoli. These characteristics may be an indication that even class VI, containing small sized oocytes, allows the models to achieve more expressive results than with class PV.

The authors used slides processed in different ways, at different times and under different experimental conditions. Thus, we sought to ensure that the collected samples were as heterogeneous as possible, since, due to the pioneering of this work, there are no image databases that can significantly increase the distribution of this sample. Thus, the heterogeneity of the sample aims to provide greater generalization and robustness to the results obtained, allowing the developed model to deal more effectively with unforeseen variability and contribute significantly to the advancement of the area.

The stages of development used play a fundamental role in an investigation of the reproductive cycle of the fish because, from the frequency of these in an image, it is possible to distinguish the stages of maturation of a fish and consequently identify its reproductive period. Araújo and Chellappa (2002) state that knowledge of these periods is indispensable for the rational management of stocks, to be an element capable of establishing a minimum size for the fish that be capture.

Thus, this research is the first step towards creating an automated system that can be used by domain experts in the laboratory, through a web system or mobile application, to identify fish germ cells. However, the various detection methods have achieved different results and deciding which method is more reliable is not an easy task. Therefore, Table 5 shows the models that obtained the best results for each metrics used in the evaluations corresponding to the respective classes. Finally, we highlight that the model YOLOv5l achieved the best performance in this context.

Precision vs Recall Curve

The precision vs recall curve shows the relationship between precision and the recall for different thresholds generated from the test data. A

Table 5 - Models with better performance

Classes	P	R	mAP@.5	mAP@.95
PV	YOLOv3-ssp	YOLOv5m	YOLOv5m	YOLOv5l
VF	YOLOv5l	YOLOv5m	YOLOv5l	YOLOv5l
VI	YOLOv5m	YOLOv5l	YOLOv5l	YOLOv5l

PV - pre-vitellogenesis, VI - early vitellogenesis, VF - late vitellogenesis

Figure 4 - The Precision vs Recall curve generated by the models

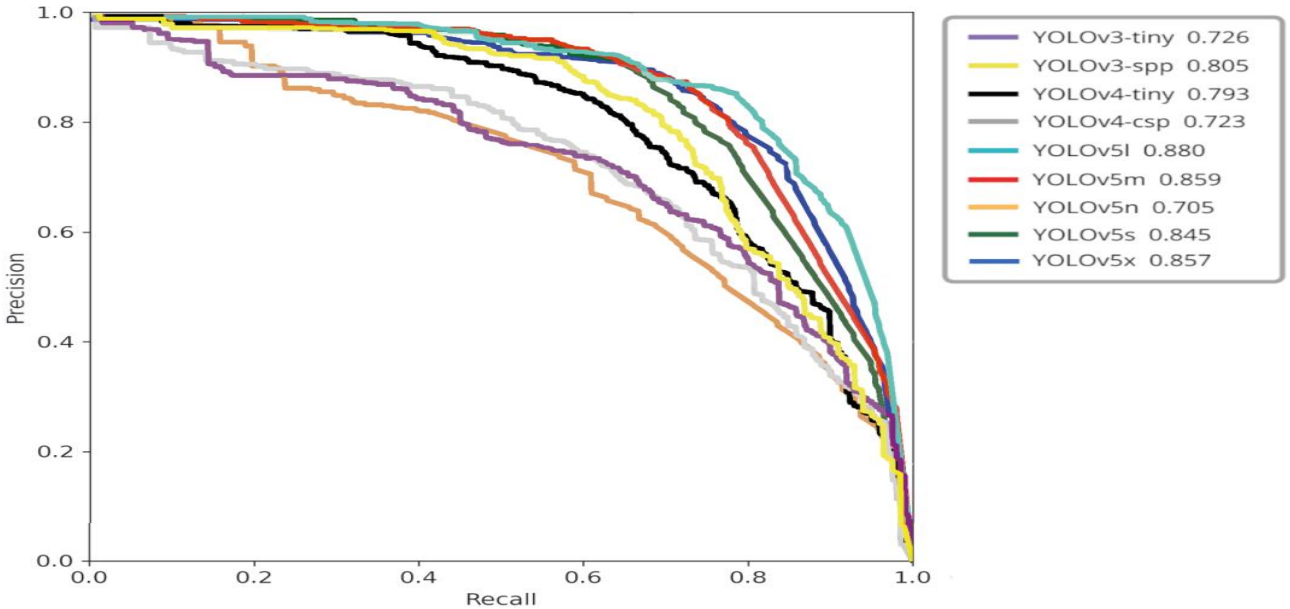
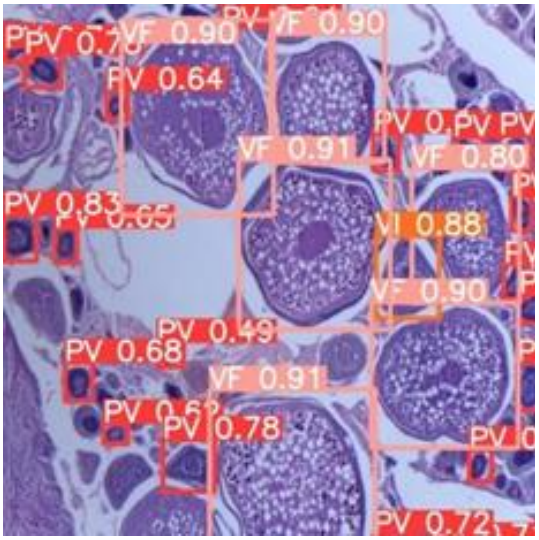
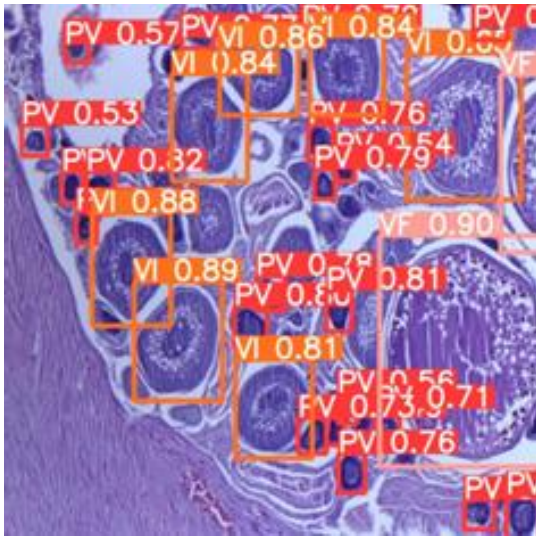


Figure 5 - Result of detection and location of oocytes using YOLOv5l model for each oocyte phase

Model	Image 1	Image 2
YOLOv5l		

high area under the curve represents high precision and high recall, i.e., high precision is related to a low false positive rate and high recall is related to a low false negative rate (Ashraf *et al*., 2022). Figure 4 illustrates the precision vs recall curves generated for each YOLO model. The area under the curve was estimated with mAP@.5. It also can be seen that the models YOLOv5l obtained the highest area under the curve (0.880).

Result of the Detection Method

We present sample test set detections in Figure 5 presents two examples of the test dataset along with the detection process results for the YOLOv5l model. The red bounding box represents the PV phase, the orange bounding box represents the VI phase and the light pink bounding box represents the VF phase. For each of these boxes, the percentage of Intersect over Union (IoU) is generated, that is, how much this prediction of the oocyte location coincides with its true location.

When comparing images, I and II, it is possible to observe that both present equivalent results. In the images, the VF phase obtained an IoU percentage greater than 90%, while the PV phase ranged from 81 to 53%. In relation to phase VI, only one oocyte was located in image I with a percentage of 88%, in image II it presented a greater number of unidentified oocytes varying their percentage between 89 and 81%.

Despite the images showing oocytes in other phases, the model could adequately locate the objects for which it was trained with high precision. The overlap is also a factor that makes object location difficult, however, the model satisfactorily detected the present oocytes in the images.

CONCLUSIONS

This paper described a method for detecting of *C. undecimalis* oocytes from histological images using the YOLO technique in versions 3, 4, and 5. For this, a base of histological images was built containing 3 phases of oocyte development. The domain expert labeled the set of images according to the respective oocyte phase. From the results obtained it was possible to conclude that the YOLOv5l model reached the best performances in oocyte detection, when evaluated using precision, recall and mAP@.5 and mAP@.95 metrics. Considering these metrics, it was possible to observe that the YOLOv5m model achieved 89.4% accuracy for class VI. The YOLOv5m model also achieved 79.3% of recall rate for the PV class and 97.4% for the VF class. The YOLOv5m model achieved 85.9% mAP@.5 for the PV class. The YOLOv5l model achieved 85.4% accuracy for the VF class, 95.7% mean mAP@.5 and 75.9% mean mAP@.95 for this same class. the YOLOv5l model achieved 85.7%, 90.2%, and 67.9% of recall, mAP@.5, and mAP@.95 respectively for the VI class and 61.7% mean mAP@.95 for the PV class. We observed that the number of oocytes is not a relevant factor for the increase in the values of the established metrics, but the size in which the oocytes are found in the images. This fact evidences that the use of an architecture that better adapts to objects of small dimensions are necessary. The results of the research are significant and confirm the effectiveness of the proposed approach in identifying oocytes. For future work, the domain expert will validate the results of the proposed approach. The identification of the phases of

the oocytes will allow estimating the fecundity rate of this species, making it possible to control at a national level the degradation of fish stocks, in places where there is intense exploitation.

ACKNOWLEDGEMENTS

The authors would like to thank the National Council for Scientific and Technological Development (CNPq), the Coordination for the Improvement of Higher Education Personnel (CAPES) and the Foundation to Support Research and Scientific and Technological Development of Maranhão (FAPEMA) for the financial support to perform this study.

REFERENCES

ABAS, S. M.; ABDULAZEEZ, A. M.; ZEEBAREE, D. Q. A YOLO and convolutional neural network for the detection and classification of leukocytes in leukemia. **Indonesian Journal of Electrical Engineering and Computer Science**, v. 25, n. 1, p. 200-213, 2022.

ALBUQUERQUE, W. F.; VIEIRA, R. H. S.; VIEIRA, G. H. Isolamento de *Staphylococcus* aureus do gelo, água, bancadas e vendedores de pescado da feira do Mucuripe, Fortaleza, Ceará. **Revista Ciência Agronômica**, v. 37, n. 3, p. 299-303, 2006.

ALY, G. H. *et al*. YOLO based breast masses detection and classification in full-field digital mammograms. **Computer Methods and Programs in Biomedicine**, v. 200, p. 105823, 2021.

ARAÚJO, A. S.; CHELLAPPA, S. Estratégia reprodutiva do peixe voador, *Hirundichthys affinis* Günther (Osteichthyes, Exocoetidae). **Revista Brasileira de Zoologia**, v. 19, p. 691-702, 2002.

ASHRAF, A. H. *et al.* Weapons detection for security and video surveillance using cnn and YOLO-v5s. **CMC-Computers, Materials & Continua**, v. 70, p. 2761-2775, 2022.

BARREIROS, M. *et al*. Zebrafish tracking using YOLOv2 and Kalman filter. **Scientific Reports**, v. 11, n. 1, p. 3219, 2021.

BISONG, E. *et al*. Google Colaboratory. *In*: BISONG, E. **Building machine learning and deep learning models on Google Cloud Platform**: a comprehensive guide for beginners. New York: Apress Berkeley, 2019. p. 59-64.

BOCHKOVSKIY, A.; WANG, C.; LIAO, H. M. Yolov4: Optimal speed and accuracy of object detection. **arXiv preprint arXiv:2004.10934**, 2020.

CHLAP, P. *et al*. A review of medical image data augmentation techniques for deep learning applications. **Journal of Medical Imaging and Radiation Oncology**, v. 65, n. 5, p. 545-563, 2021.

EVERINGHAM, M. *et al*. The pascal visual object classes challenge: a retrospective. **International Journal of Computer Vision**, v. 111, p. 98-136, 2015.

FAO. The State of World Fisheries and Aquaculture 2024. **Blue Transformation in action**. Rome, 2024. DOI: https://doi.org/10.4060/cd0683en.

FRACAROLLI, J. A. *et al*. Computer vision applied to food and agricultural products. **Revista Ciência Agronômica**, v. 51, 2021.

GOLATKAR, A.; ANAND, D.; SETHI, A. Classification of breast cancer histology using deep learning. *In*: CAMPILHO, A.; KARRAY, F.; TER HAAR ROMENY, B. (ed.). **Image Analysis and Recognition**: 15 th International Conference, ICIAR 2018, Póvoa de Varzim, Portugal, June 27–29, 2018, Proceedings 15. Springer International Publishing, 2018. v. 10882, p. 837-844.

HONG, S.; BAI, S.; CAO, Y. RotNet: Rotation-invariant deep feature learning. *In*: **Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)**, 2016.

HUSSAIN, Z. *et al*. Differential data augmentation techniques for medical imaging classification tasks. *In*: AMIA annual symposium proceedings. **American Medical Informatics Association**, v. 2017, p. 979, 2018.

JIANG, Z. *et al*. Improved detection performance in blood cell count by an attention-guided deep learning method. **OSA Continuum**, v. 4, n. 2, p. 323-333, 2021.

KARLOU-RIGA, C.; PANOS, S. E. Ovarian atretic rates and sexual maturity of European horse mackerel, Trachurus trachurus (L.), in the Saronikos Gulf (Greece). **Fishery Bulletin**, v. 94, p. 66-76, 1996.

KOHAVI, R.; JOHN, G. H. *Wrappers for feature subset selection*. **Artificial Intelligence**, v. 97, n. 2, p. 273-324, 1997.

KUBERA, E. *et al*. Detection and recognition of pollen grains in multilabel microscopic images. **Sensors**, v. 22, n. 7, p. 2690, 2022.

LIU, W.; XIAO, H.; CHENGCHAO, B. Spatial multi-object recognition based on deep learning. *In*: **2019 IEEE International Conference on Unmanned Systems (ICUS)**. IEEE, 2019. p. 736-741.

LUO, H. *et al*. Contextual-YOLOV3: Implement better small object detection based deep learning. *In*: **2019 International Conference on Machine Learning, Big Data and Business Intelligence (MLBDBI)**. IEEE, 2019. p. 134-141.

MENDONÇA, M. C. F. B. de. **Autoecologia do camorim, *Centropomus undecimalis* (Bloch, 1792), (Perciformes: Centropomidae) em ambiente hipersalino em Galinhos, RN, Brasil**. 2004. Tese (Doutorado em Ciências Biológicas) – Universidade Federal de São Carlos, São Carlos, 2004.

NAGAHAMA, Y.; YAMASHITA, M. Regulation of oocyte maturation in fish. **Development, Growth & Differentiation**, v. 50, p. S195-S219, 2008.

NASCIMENTO, I. R. M. A. *et al.* Quantitative indicators of the reproductive biology of adult specimens of *Centropomus undecimalis* (Teleost: Centropomidae) obtained from commercial fishermen in the Parnaíba Delta Environmental Protection Area, north coast of Brazil. **Boletim do Instituto de Pesca**, v. 48, 2022.

NIE, Y. *et al*. Automatic detection of melanoma with yolo deep convolutional neural networks. *In*: **2019 E-Health and Bioengineering Conference (EHB)**. IEEE, 2019. p. 1-4.

PEREIRA, M. E. G. de S. *et al*. Contribution of fisheries actors to the bioecology of the Commom Snook *Centropomus undecimalis* (CENTROPOMIDAE - PERCIFORMES) captured on the Brazilian Amazon coast. **Research, Society and Development**, v. 9, n. 10, e4119108691, 2020.

REDMON, J. *et al*. You only look once: Unified, real-time object detection. *In*: **Proceedings of the IEEE conference on computer vision and pattern recognition**, 2016. p. 779-788.

REDMON, J.; FARHADI, A. Yolov3: an incremental improvement. Computer vision and pattern recognition. **arXiv preprint arXiv:1804.02767**, 2018.

RHODY, N. R. *et al*. Assessing reproductive condition in captive and wild common snook stocks: a comparison between the wet mount technique and histological preparations. **Transactions of the American Fisheries Society**, v. 142, n. 4, p. 979-988, 2013.

ROSA, R. S.; LIMA, F. C. Os peixes brasileiros ameaçados de extinção. *In*: MACHADO, A.; DRUMMOND, G. M.; PAGLIA, A. P. **Livro vermelho da fauna brasileira ameaçada de extinção**. Brasília, DF: Ministério do Meio Ambiente, 2008. v. 2, p. 9-19.

SEAFOOD BRASIL. **Veja qual é o consumo per capita de pescado no Brasil**. 2022. Seafood Brasil. Available at: https://www.seafoodbrasil.com.br/veja-qual-e-o-consumo-per-capita-de-pescado-no-brasil. Accessed on: Dec. 15, 2024.

SHORTEN, C.; KOEHRING, A. A survey of image data augmentation for deep learning. **Journal of Machine Learning Research**, v. 20, p. 1-23, 2019.

SILVA, J. G.; NARCISO, M. G.; GONÇALVES, C. Computer vision and image processing for detecting and quantifying whiteflies: a systematic review. *In*: **Embrapa Arroz e Feijao-Artigo em anais de congresso (ALICE)**, p. 382-391, 2019.

SIVAKUMARAN, K. P. *et al*. Maturation and reproductive biology of female wild carp, Cyprinus carpio, in Victoria, Australia. **Environmental Biology of Fishes**, v. 68, n. 3, p. 321-332, 2003.

SOUSA, M. L. N. M. **Caracterização morfológica de gônadas de machos do robalo, *Centropomus undecimalis* (BLOCK, 1970) oriundos de cultivo e de ambiente natural**. 2011. Dissertação (Mestrado) Faculdade de Ciências Veterinárias, Universidade Federal do Ceará, Fortaleza, 2011.

TAMMINEN, M. *et al*. Data augmentation for deep learning in medical imaging: a systematic review. **Frontiers in Robotics and AI**, v. 5, p. 58, 2018.

TANER, A.; ÖZTEKIN, Y. B.; DURAN, H. Performance analysis of deep learning CNN models for variety classification in hazelnut. **Sustainability**, v. 13, n. 12, p. 6527, 2021.

ULTRALYTICS YOLO. **v5**. Los Angeles: Ultralytics, 2015. Available at: https://www.ultralytics.com/pt/yolo. Accessed on: Dec. 17, 2023.

VAZZOLER, A. E. A. **Biologia da reprodução de peixes teleósteos**: teoria e prática. 2. ed. São Paulo: EDUEM, 1996.

WANG, C. *et al*. CSPNet: a new backbone that can enhance learning capability of CNN. *In*: **Proceedings of the IEEE/ CVF conference on computer vision and pattern recognition workshops**, 2020. p. 390-391.

ZHAO, Z. *et al*. Data augmentation techniques for improving deep learning model robustness. **Journal of Artificial Intelligence Research**, v. 69, p. 1107-1138, 2020.