Evaluation of the Detection Capabilities of Ovarian Structures in Ultrasound Images of Cows Using Convolutional Neural Networks

João Moura[1], Marcela Alves[1], Hugo Kuribayashi[1], André Cascalho[2], Adam Santos[1]

[1]Instituto de Geociências e Engenharias (IGE)
Universidade Federal do Sul e Sudeste do Pará (Unifesspa) - Marabá, PA - Brasil

[2]Instituto de Estudos do Trópico Úmido (IETU)
Universidade Federal do Sul e Sudeste do Pará (Unifesspa) - Xinguara, PA - Brasil

{joaov.mour, marcela.alves, hugo}@unifesspa.edu.br

{andre.cascalho, adamdreyton}@unifesspa.edu.br

Abstract. *Ultrasound plays a crucial role in veterinary medicine, being widely used for the detection of diseases and the identification of reproductive conditions in animals in a less invasive and more cost-effective manner. However, the effectiveness of this method is highly dependent on the experience and training of the professional conducting the analysis, as the relatively low quality of the generated images can hinder accurate interpretation of the results. This study evaluates the ability to detect ovarian structures in a cow ultrasound image dataset using convolutional neural networks with the YoloV8 model. The results demonstrated high accuracy in the task, achieving approximately 90% accuracy in detecting objects of interest.*

1. Introduction

In recent years, approaches based on Artificial Intelligence (AI) have gained increasing importance in the medical field, as they are able to reduce costs by performing tasks that previously required highly trained professionals to achieve comparable results. In particular, Machine Learning (ML) techniques have been used to extract information from raw data to make accurate and reliable decisions in medicine [Rahmani et al. 2021, Chafai et al. 2024].

There are examples of the application of ML-based approaches in genomic medicine [Chafai et al. 2024], in the diagnosis of cerebrovascular accidents [Daidone et al. 2024], in the detection of malignant tumors [Lubbad et al. 2024], and in the identification of cardiovascular diseases [Friedrich et al. 2021], among others. Furthermore, several systematic reviews of the literature have investigated the potential applications of ML techniques in medicine, as exemplified by [Peng et al. 2021, Pan et al. 2024, Pantanowitz et al. 2024].

In the veterinary medicine context, ML has also gained importance in recent years, being frequently used as an auxiliary tool in various fields and purposes. The application of ML has explored tasks related both to animal welfare [Fuentes and othres 2022] and to the accurate identification of diseases [Walle Girmaw 2025, Megahed et al. 2025], malformations and general problems. The adoption of such techniques has proven to be of paramount importance for the rapid and accurate identification of problems that can contribute to public health protection strategies.

Considering the animal reproduction scenario, the *Corpus Luteum* (CL) plays a crucial role in determining fertility, since it regulates the levels of progesterone necessary for gestation [Arruda et al. 2001, Koerts et al. 2024]. Given the importance of this structure in reproductive management, the strategy of identifying the presence of the CL has been widely adopted to optimize efforts to perform reproduction in these animals [Koerts et al. 2024]. Among the techniques used for CL detection, Ultrasonography (US) stands out as a well-established and recommended method due to its non-invasive nature and the use of non-ionizing radiation [Miller et al. 2012]. However, this technique has drawbacks, such as a steep learning curve and the experience required for the correct use and identification of ovarian structures in the generated images, which require a professional with such experience to perform the detection.

In this context, Computer Vision (CV) techniques represent a significant opportunity to streamline the identification of ovarian structures by substantially reducing both the workload and the demand for manual labor. By automating the detection process, these technologies can enhance efficiency and accuracy, allowing professionals to focus primarily on reviewing and validating the results, ultimately optimizing their role within the diagnostic workflow. However, given the novelty of this application area, there is a notable scarcity of studies investigating specialized approaches or models dedicated to the detection of ovarian structures. This work aims to address this gap by developing and evaluating a specialized Deep Learning (DL) model for the automatic detection of ovarian structures in bovine ultrasound images, thus advancing the integration of computer vision techniques into veterinary reproductive diagnostics.

The remainder of this paper is organized as follows: Section 2 presents the related work, Section 3 describes the materials and methods used in this study, Section 4 offers a discussion of the results, and Section 5 concludes this work with the main findings and prospects for future work.

2. Related Work

Several systematic literature reviews have investigated the application of ML-based techniques in medical image diagnosis, including US, radiography, and even photography, using CV methods [Burti et al. 2024, Albuquerque et al. 2025]. Following a comprehensive review of the literature, no studies were identified that specifically address the same objective as this work, which focuses on the evaluation of ML methods for the detection of ovarian structures in a cow US image database. However, related works were found that discuss convergent aspects in terms of the variables collected or regarding the application of detection methods. In fact, most of the research is focused on the detection of diseases in human US examinations, as demonstrated in [Dadjouy and Sajedi 2024] and [Pham and Le 2024], resulting in a scarcity of studies applied to the veterinary field.

In [Walle Girmaw 2025], the authors propose a ML-based approach for the recognition of bovine skin disease, where three recent DL models (EfficientNetB7, MobileNetV2 and DenseNet201) were evaluated. The results indicate that the color and texture of the animal skin are crucial for the DL models for feature extraction and to recognize the affected area. In addition, the authors used a SoftMax classifier due to the presence of multiple classes, and their experimental results demonstrated that EfficientNetB7 achieved an accuracy of 99.01%, outperforming other models in classifying the four investigated classes.

The authors of [Hassan et al. 2024] propose an approach employing DL models for the counting of antral follicles and detection of the CL, aiming to study reproductive dynamics. The study uses images obtained from thin sections of ovarian tissue analyzed under a microscope, which provide high cellular detail, characterizing an *in vitro* and highly controlled setting. This automated quantification and counting of the follicles and CL seeks to accelerate traditional manual, labor intensive, and prone to errors analyzes, thus contributing to research on folliculogenesis and the development of novel treatments. The study achieved promising results using the You Only Look Once (YOLO) and RetinaNet networks, with average precision on the set of tests 75% and 83%, respectively. Furthermore, the study employed Data Augmentation (DA) techniques to augment the training dataset, which consists of approximately 1000 images. However, due to the relatively small size of the dataset, the models may be susceptible to *overfitting*.

In [Ninphet et al. 2024], the authors present an approach for predicting and detecting estrus in dairy cows through image analysis, optimizing reproductive management in agricultural practices. Detecting estrus in cows is crucial for efficient reproductive management in livestock, particularly in dairy and beef production. Estrus is the period of the reproductive cycle during which the female is receptive to mating and occurs close to ovulation, that is, the optimal moment for artificial insemination or natural mating. The proposed approach utilizes Convolutional Neural Networks (CNNs), employing the YOLOv5 model. The results indicate that the proposed approach achieved high precision in real-time detection, with an average F1-Score of 0.993 and mAP50 of 0.995, demonstrating promising performance in estrus prediction and detection in dairy cows.

The identification of pregnancy in cattle using ovarian US images collected 30 days after timed artificial insemination is addressed by the authors of [Andrade et al. 2023]. The proposed approach employs CNNs and investigates the performance of different architectures to accurately classify pregnant and non-pregnant cows, while also considering the impact of image quality on model accuracy. The results demonstrated that the DenseNet121 architecture exhibited the best overall performance, achieving an accuracy of up to 78.8%, a sensitivity of 75.4%, and a specificity of 82.2% in detecting pregnancy from ovarian US images, particularly when applied to images of good and regular quality. These findings indicate the potential of CNNs to support reproductive management in bovine herds.

Several CNN architectures have been investigated for image analysis, particularly in US applications. However, challenges remain with respect to the accurate detection of ovarian structures in cow US images due to the low image quality and variability inherent in this modality. Therefore, this work seeks to address these challenges by evaluating the ability of CNN models to detect key ovarian characteristics in an annotated US image database. Thus, our approach differs from previous research in the following ways:

- Unlike earlier research that primarily addressed classification or basic detection, this study focuses exclusively on advanced detection techniques to enhance diagnostic accuracy and reduce the reliance on operator expertise.
- Developing a novel dataset of real-world veterinary US images, annotated by spe-

cialists, to validate models and support further research in this area.

3. Research Design and Methods

The methodology adopted in the present study is derived from the Cross-Industry Standard Process for Data Mining (CRISP-DM) [Hayat Suhendar and Widyani 2023] framework, which is structured into six main phases: business understanding, data understanding, data preparation, modeling, evaluation, and deployment. These stages also serve as mutual validation mechanisms, allowing iterations back to previous phases whenever necessary to ensure proper interpretation and adequate handling of the data. For this study, the methodology was adapted as follows (as represented in Figure 1):

- **Problem Understanding:** Identification of the objectives and challenges associated with the manual detection of ovarian structures;
- **Data Understanding:** Definition of the structures to be localized by the model and analysis of their complexity, aiming to ensure the solution's usability;
- **Data Preparation:** Construction of the dataset comprising images containing the Regions of Interest (ROI), along with the corresponding annotations;
- **Modeling:** Selection of the most suitable detection models and architectures, taking into account their known advantages and limitations;
- **Evaluation:** Performance analysis of the models based on metrics such as precision, mAP50, and mAP50-95, with the objective of ensuring compliance with the expected results.

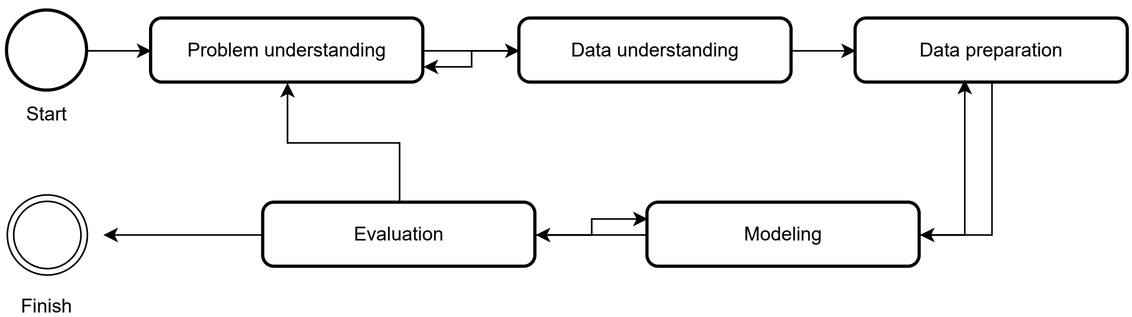


Figure 1. Representation of the methodology used in the study.

3.1. Dataset

For the execution of the experiments, this study involves the development of a dataset comprising US images of female bovines. The dataset contains a total of 212 samples, filtered to include only images featuring ROIs. All captured images are in Portable Network Graphic (PNG) format, with an average resolution of 640x480 pixels.

It is important to note that all images were manually annotated by experts, with the aid of the Computer Vision Annotation Tool (CVAT)[1] tool. Furthermore, each selected image contains one or more ROIs, classified into two categories: "Ovary" and "*Corpus Luteum*". To our knowledge, this is the only available dataset that contains annotated ultrasound images of bovine ovarian structures, specifically labeled for ovary and CL regions of interest. The complete dataset is available at [Andrade 2025].

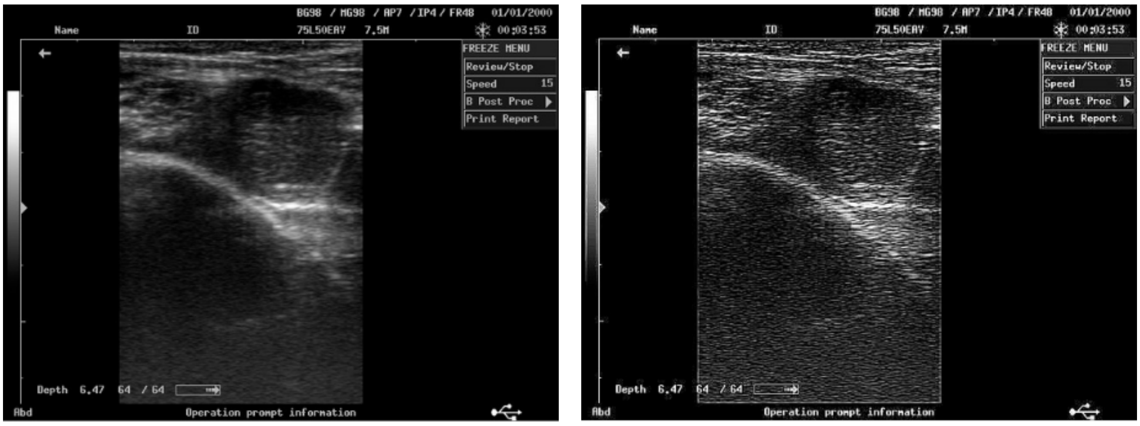
[1] Available at: https://www.cvat.ai/

The dataset was partitioned into three subsets, with 80% allocated to training (171 samples) and the remaining 20% distributed between the validation sets (21 samples) and the test sets (20 samples). Additionally, the adopted methodology aimed to preserve the class distribution across the subsets, to prevent bias in the model evaluation metrics.

3.2. Data Preparation

During the data preparation stage, the sharpening technique was applied with the objective of improving image visualization and model performance. This technique applies a filter to the image pixel grid that accentuates the edge pixels, rendering the image sharper, as illustrated in Figure 2. Specifically, Figure 2a depicts an US image prior to the sharpening procedure, while Figure 2b shows the result after the application of this procedure.

Furthermore, the color channels of the images were converted to grayscale, ensuring the standardization of the inputs provided to the model. This procedure is fundamental for reducing the variability of the input data, thereby facilitating the modeling learning process by focusing on relevant texture and shape features without interference from color information.



(a) Original US. **(b) US after sharpening transformation.**

Figure 2. US image before and after the application of the sharpening procedure.

3.3. YOLO for Ovarian Structure Detection

The YOLO model is a single stage object detector that employs a single CNN to directly identify the bounding boxes and classify them within an input image. The image is divided into a predetermined grid, with each grid cell predicting a certain number of bounding boxes along with an associated confidence score. This score reflects both the model's certainty regarding the presence of an object within the corresponding bounding box and the accuracy of the bounding box's fit relative to the ground truth. Bounding boxes with class probabilities that exceed a specified threshold are selected for subsequent object localization [Dadjouy and Sajedi 2024].

Figure 3 illustrates the architecture of *YOLOv8*, which serves as the primary tool employed in this study. The model processes an input image by applying pooling layers to resize the image, followed by convolutional layers that extract features used as weights for the final decision-making. Subsequently, the image undergoes additional pooling layers aimed at restoring the original dimensions and performing the annotation of the bounding boxes. At this stage, the image contains numerous annotated bounding boxes, each associated with class confidence scores ranging from 0 to 100%. The model then applies a threshold to discard annotations with low confidence levels, thereby producing the final image with class detections and their corresponding confidence scores.

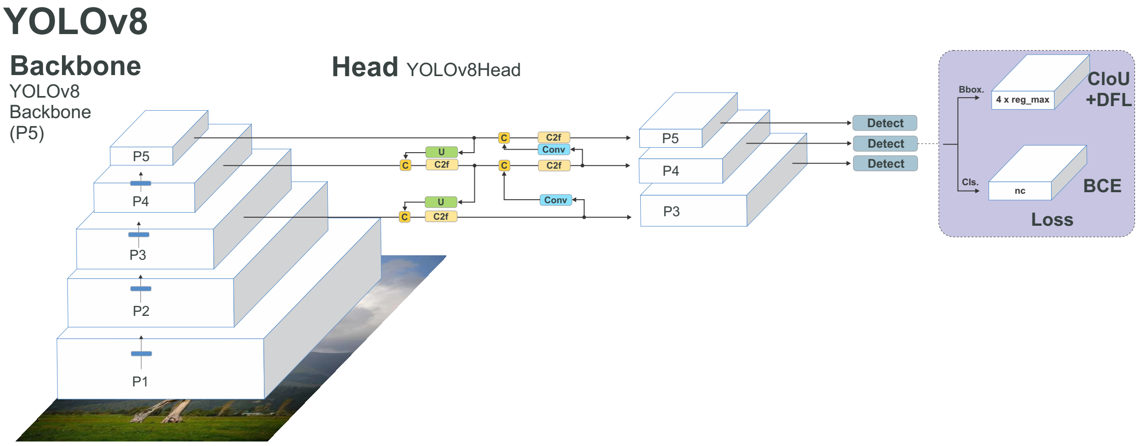


Figure 3. Representation of the YOLOv8 Architecture.

Furthermore, Figure 4 presents a representation in which an image initially contains multiple annotations in the limiting box (Figure 4a). The YOLO model performs the filtering process, resulting in the final image shown in Figure 4b, which displays the detected classes along with their respective confidence intervals.

(a) Multiple annotated bounding boxes. **(b) Final version after YOLO filtering.**

Figure 4. Bounding box annotation and final filtered detections by YOLO model.

3.4. Hyperparameter Settings

The methodological design adopted in this study involves the evaluation of hyperparameters of the YOLOv8 model considering three distinct sets of hyperparameter variations. This step was incorporated to provide insight into the impact of different configurations on model performance, allowing the identification of a combination that optimizes precision and robustness in the detection of ovarian structures, as detailed in Table 1.

Thus, the training phase involved the training of three sets of hyperparameters, designated as Models **A**, **B**, and **C**. Models **A** and **B** employ the Learning Rate Decay (LRD) technique, characterized by a higher learning rate at the beginning of training that gradually decreases over epochs, while Model **C** uses fixed initial and final training rates. Furthermore, the Stochastic Gradient Descent (SGD) optimizer was selected for training Model **A**, while for training Models **B** and **C**, the optimizer was set to the Adaptive Moment Estimation with Weight Decay algorithm (AdamW).

All models were implemented using Python 3.10.12, in accordance with the Ultralytics platform[2]. The data preparation stage was carried out with the help of the scikit-learn library, which provided essential tools for preprocessing and partitioning data sets. For all models, the number of epochs was set at 200. However, all models were configured with Early Stopping to prevent overfitting and ensure training efficiency. Furthermore, the weight decay rate was standardized to 0.0005 in all three models, ensuring consistency in regularization during training.

The batch size parameter defines the number of samples processed simultaneously during training, directly influencing how the model learns data patterns. Higher values of this parameter tend to reduce variability among batches, which may compromise the model's generalization capability, especially when dealing with smaller or more complex datasets. In the trained models, three different batch size values were tested. Model **C** utilized the highest value, which was reduced in Model **B**. For Model **A**, a value of *-1* was adopted, indicating that the batch size was automatically adjusted based on the available processing capacity, to optimize the use of computational resources.

Table 1. Summary of hyperparameters used in the models.

Hyperparameter	Model A	Model B	Model C
Initial Learning Rate	0.01	0.01	0.01
Final Learning Rate	0.0001	0.001	0.01
Optimizer	SGD	AdamW	AdamW
Momentum	0.937	0.937	0.937
Epochs	200	200	200
Batch Size	-1	8	16
Weight Decay	0.0005	0.0005	0.0005

3.5. Evaluation Metrics

We evaluated the performance of the detection of ovarian structures with Precision, Recall and F1-Score (Equations 1a, 1b and 1c). In the context of this work, correctly identified ovarian structures in ultrasound images were considered true positives (TP), non-ovarian structures mistakenly identified as ovarian structures were false positives (FP), and ovarian structures that were not detected by the model were false negatives (FN).

$$\text{Precision} = \frac{TP}{TP + FP}, \quad (1a) \qquad \text{Recall} = \frac{TP}{Tp + FN}, \quad (1b) \qquad \text{F1} = \frac{Precision \cdot Recall}{Precision + Recall}, \quad (1c)$$

where TP refers to true positives, and FP and FN indicate false positive and negative cases.

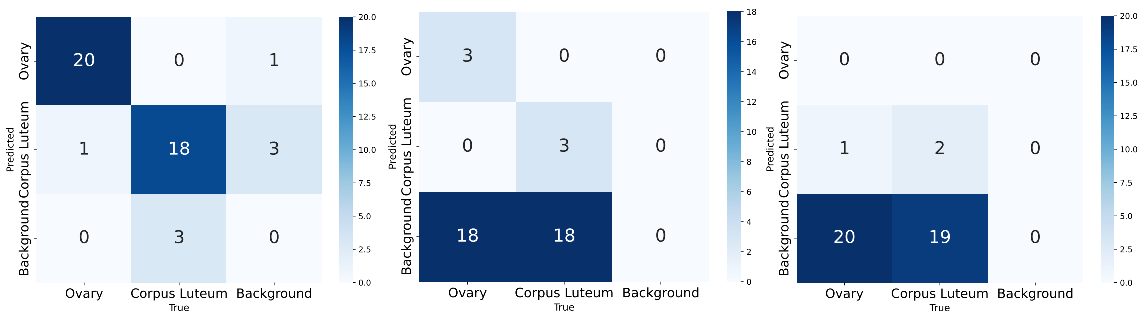
[2]Available at: https://docs.ultralytics.com/

For the YOLOv8 models, we employed the mean Average Precision (mAP) metric evaluated at an Intersection over Union (IoU) threshold of 50% (mAP50), as well as the average mAP across multiple IoU thresholds ranging from 50% to 95% in increments of 5% (mAP50-95). The IoU measures the degree of overlap between the predicted bounding box and the manually annotated bounding box, while the mAP represents the area beneath the precision-recall curve corresponding to a specific IoU threshold.

4. Results and Discussion

Figure 5 presents the confusion matrix, which illustrates the number of samples identified within the classes Ovary, CL, and background (other regions of the ovarian structure). It is important to note that the test dataset contains images exclusively featuring the Ovary and CL structures. Therefore, instances where these structures are classified as background constitute misclassifications, as the model failed to correctly identify the presence of ovarian structures in those regions. In this context, any marking as background represents an error of omission, indicating that the model did not detect an ovarian or CL region where it was actually present.

Furthermore, the models produced results consistent with expectations. However, given the complexity of the USs images, some misclassifications were expected. Notably, Model A (Fig. 5a) demonstrated significant advantages, achieving an accuracy of 82.61% across all three classes and 97.44% when considering only the Ovary and CL classes. In contrast, Models B and C exhibited unsatisfactory performance, with overall accuracies of 14.29% and 4.76%, respectively. In particular, Models B (Fig. 5b) and C (Fig. 5c), all instances of the background class were misclassified. This indicates a substantial overlap or lack of discriminative features between the background and the other two classes within the model's representation. In Model C, an instance of CL was misclassified as Ovary, suggesting some degree of confusion between these biologically related classes.



(a) Model A. (b) Model B. (c) Model C.

Figure 5. Comparison of confusion matrices for the evaluated models.

Furthermore, Table 2 summarizes the evaluation metrics for the evaluated models. Model A demonstrated better overall performance, achieving promising results for both classes, along with the highest mAP50 and mAP50-95 for the Ovary class. Model A exhibited a precision of 0.952 and an F1-Score of 0.952 for Ovary, and a precision of 0.857 with an F1-Score of 0.837 for CL. In contrast, Model B showed perfect recall for both classes, however, at the expense of substantially lower precision and F1-Scores, reflecting a higher rate of false positives. Moreover, Model C exhibited the poorest performance, particularly for the Ovary class, while CL class achieved marginally better but still limited

Table 2. Summary of performance metrics for the assessed models.

Model	Class	Precision	Recall	F1-Score	mAP50	mAP50-95
A	Ovary	**0.952**	**0.952**	**0.952**	**0.995**	**0.786**
	Corpus Luteum	**0.857**	**0.818**	**0.837**	**0.991**	**0.571**
B	Ovary	0.143	1.000	0.250	0.577	0.219
	Corpus Luteum	0.143	1.000	0.250	0.505	0.314
C	Ovary	0.000	0.000	0.000	0.0895	0.564
	Corpus Luteum	0.091	0.667	0.154	0.0306	0.284

results. These results suggest that Model A provides the most reliable detection capabilities among the evaluated models, effectively balancing sensitivity and specificity, while Models B and C require further optimization to reduce misclassifications.

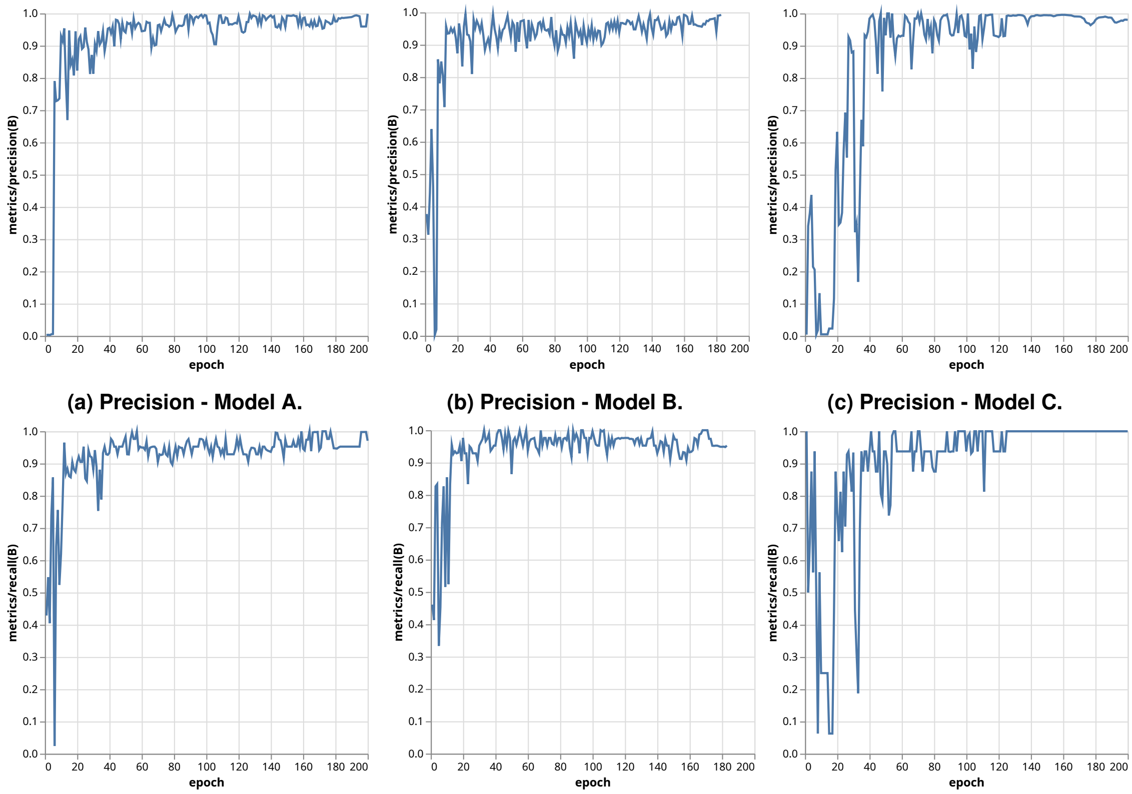
Figure 6 presents the results related to the precision and recall metrics throughout the model training. In all cases, all models achieved values that exceeded 90% for precision or recall, particularly after 40 epochs, indicating a high level of reliability in the correct identification of classes among the detections performed. It should be noted that, although early stopping was employed in all configurations, only Model B experienced early termination of training, approximately at epoch 180. In veterinary medical imaging, high recall is crucial to minimize false negatives, thereby ensuring that critical ovarian structures are accurately detected and preventing potential misdiagnoses that could adversely affect animal treatment outcomes.

Figure 7 presents the results of mAP50 and mAP50–95 throughout the training epochs. Specifically, it is observed, particularly in Figures 7c and 7f, that Model C exhibits pronounced instabilities until approximately epoch 80, indicating initial learning difficulties and limited generalization capability for this configuration. In contrast, consistent with the patterns observed in the precision and recall results, Model A demonstrates the best training performance among the configurations evaluated.

Furthermore, as evidenced by Figure 7, Models B and C maintained an average performance below 60%, both during the training and testing phases. These findings indicate that among the models evaluated, only Model A exhibits robust and stable performance at different levels of detection precision requirements, making it the most suitable for practical applications demanding reliable localization and identification of ovarian structures. Moreover, Figure 4b shows sample results from Model A with annotations. The confidence scores above each bounding box were confirmed by experts, validating the performance of the model in the identification of structures.

5. Conclusions and Implications

This work developed a deep learning-based approach for the automatic detection of ovarian structures in bovine US images. The methodology adopted was grounded in the CRISP-DM framework, enabling a structured approach that encompassed all stages from problem definition to model evaluation, incorporating adaptations tailored to the specific context of the animal reproduction domain. The dataset utilized, manually annotated by



(a) Precision - Model A. (b) Precision - Model B. (c) Precision - Model C.

(d) Recall - Model A. (e) Recall - Model B. (f) Recall - Model C.

Figure 6. Analysis comparing Precision and Recall over epochs.

experts, constituted a novel and valuable contribution to the field, given its specificity and technical curation.

The study implemented the YOLOv8 architecture, which was shown to be well suited for the detection task in medical imaging. The evaluation of the models, based on metrics such as precision, mAP50, and mAP50-95, demonstrated that the models achieved satisfactory results, localizing the structures of interest within the images. These findings confirm the feasibility of employing computer vision techniques to assist in the automated detection of ovarian structures in US images, representing a significant advance in the integration of artificial intelligence into veterinary practice and livestock health monitoring.

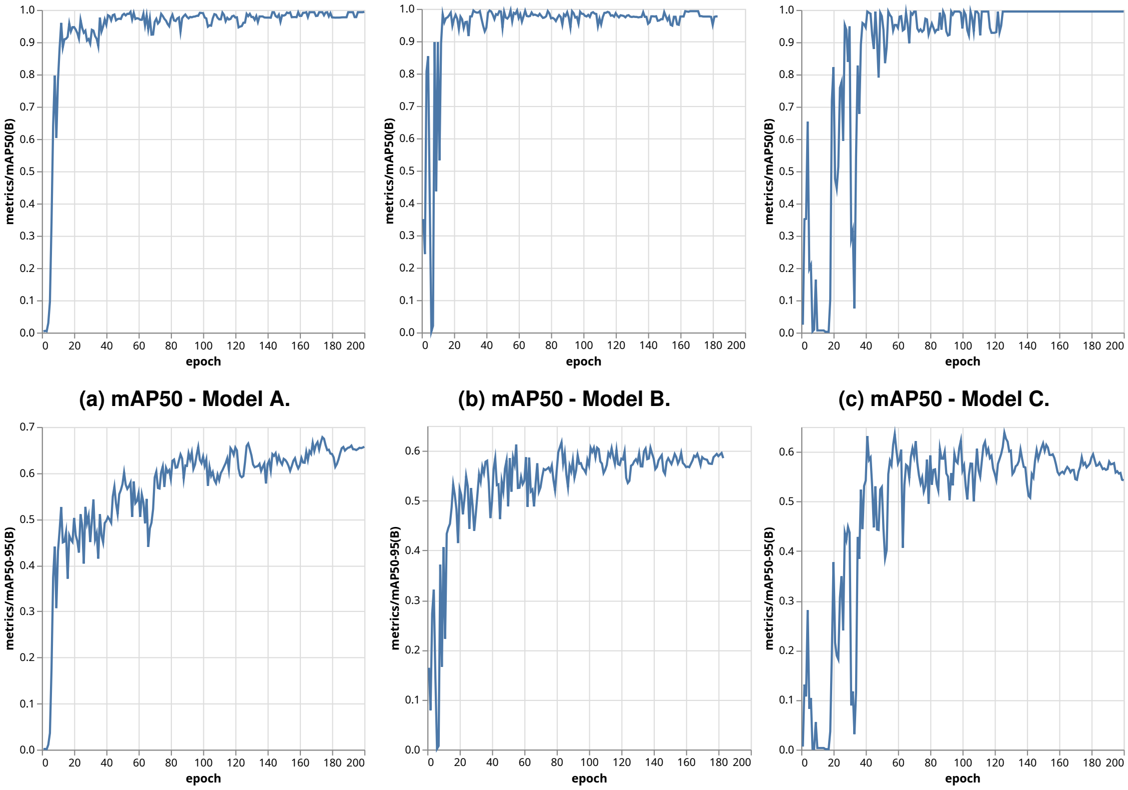
Acknowledgment

The authors acknowledge the partial funding of this research by the Amazon Foundation for Support to Studies and Research (FAPESPA).

References

Albuquerque, C., Henriques, R., and Castelli, M. (2025). Deep learning-based object detection algorithms in medical imaging: Systematic review. *Heliyon*, 11(1).

Andrade, A. (2025). Ultrasonographic Images of Female Bovines. Available at: `https://doi.org/10.7910/DVN/YC9KAX`.



(a) mAP50 - Model A. **(b) mAP50 - Model B.** **(c) mAP50 - Model C.**

(d) mAP50–95 - Model A. **(e) mAP50–95 - Model B.** **(f) mAP50–95 - Model C.**

Figure 7. Analysis comparing mAP50 and mAP50–95 over epochs.

Andrade, A. C. et al. (2023). Identifying pregnancy in cows using ovarian ultrasound images and convolutional neural networks - a proof-of-concept study. *Computers and Electronics in Agriculture*, 206:107674.

Arruda, R. et al. (2001). Existem relações entre tamanho e morfoecogenicidade do corpo lúteo detectados pelo ultra-som e os teores de progesterona plasmática em receptoras de embriões eqüinos? *Brazilian Journal of Veterinary Research and Animal Science*, 38(5):233–239.

Burti, S. et al. (2024). Artificial intelligence in veterinary diagnostic imaging: Perspectives and limitations. *Research in Veterinary Science*, 175:105317.

Chafai, N. et al. (2024). Emerging applications of machine learning in genomic medicine and healthcare. *Critical Reviews in Clinical Laboratory Sciences*, 61(2):140–163. PMID: 37815417.

Dadjouy, S. and Sajedi, H. (2024). Gallbladder cancer detection in ultrasound images based on yolo and faster r-cnn. In *2024 10th International Conference on Artificial Intelligence and Robotics (QICAR)*, pages 227–231.

Daidone, M., Ferrantelli, S., and Tuttolomondo, A. (2024). Machine learning applications in stroke medicine: advancements, challenges, and future prospectives. *Neural Regeneration Research*, 19(4).

Friedrich, S. et al. (2021). Applications of artificial intelligence/machine learning approaches in cardiovascular medicine: a systematic review with recommendations. *Eu-*

ropean Heart Journal - Digital Health, 2(3):424–436.

Fuentes, S. and othres (2022). Animal biometric assessment using non-invasive computer vision and machine learning are good predictors of dairy cows age and welfare: The future of automated veterinary support systems. *Journal of Agriculture and Food Research*, 10:100388.

Hassan, M. H., Reiter, E., and Razzaq, M. (2024). Automatic ovarian follicle detection using object detection models. *Scientific Reports*, 14(1):31856.

Hayat Suhendar, M. T. and Widyani, Y. (2023). Machine learning application development guidelines using crisp-dm and scrum concept. In *2023 IEEE International Conference on Data and Software Engineering (ICoDSE)*, pages 168–173.

Koerts, J. J. et al. (2024). Impact of corpus luteum number on maternal pregnancy and birth outcomes: the rotterdam periconception cohort. *Fertility and Sterility*.

Lubbad, M. et al. (2024). Machine learning applications in detection and diagnosis of urology cancers: a systematic literature review. *Neural Computing and Applications*, 36(12):6355–6379.

Megahed, A. A. et al. (2025). Using supervised machine learning algorithms to predict bovine leukemia virus seropositivity in dairy cattle in florida: A 10-year retrospective study. *Preventive Veterinary Medicine*, 235:106387.

Miller, D. L. et al. (2012). Overview of therapeutic ultrasound applications and safety considerations. *Journal of Ultrasound in Medicine*, 31(4):623–634.

Ninphet, W., Amdee, N., and Sangsongfa, A. (2024). Adaptive deep learning for image-based estrus prediction and detection in dairy cows. *PriMera Scientific Engineering*, 5(2):12–37.

Pan, Y. et al. (2024). Applications of hyperspectral imaging technology combined with machine learning in quality control of traditional chinese medicine from the perspective of artificial intelligence: A review. *Critical Reviews in Analytical Chemistry*, 54(8):2850–2864.

Pantanowitz, L. et al. (2024). Non-generative artificial intelligence (ai) in medicine: Advancements and applications in supervised and unsupervised machine learning. *Modern Pathology*, page 100680.

Peng, J., Jury, E. C., Dönnes, P., and Ciurtin, C. (2021). Machine learning techniques for personalised medicine approaches in immune-mediated chronic inflammatory diseases: Applications and challenges. *Frontiers in Pharmacology*, 12.

Pham, T. and Le, V. (2024). Ovarian Tumors Detection and Classification from Ultrasound Images based on YOLOv8. *Journal of Advances in Information Technology*, 15(2):264–275.

Rahmani, A. M. et al. (2021). Machine learning (ml) in medicine: Review, applications, and challenges. *Mathematics*, 9(22).

Walle Girmaw, D. (2025). Livestock animal skin disease detection and classification using deep learning approaches. *Biomedical Signal Processing and Control*, 102:107334.