REVISTA OWL (*OWL Journal*)

www.revistaowl.com.br – ISSN: 2965-2634

ENHANCING BREAST LESION DETECTION IN MAMMOGRAPHY THROUGH TEACHER–STUDENT LEARNING AND PSEUDO-LABELING

APRIMORAMENTO DA DETECÇÃO DE LESÕES MAMÁRIAS EM MAMOGRAFIA POR MEIO DE APRENDIZADO PROFESSOR-ALUNO E PSEUDO-ROTULAGEM

DOI: 10.5281/zenodo.21912263

Isaias Soares Figueiredo[1]
José Paulo Gonçalves de Oliveira[2]
Isaque Soares Figueiredo[3]
Rui Nóbrega de Pontes Filho[4]
João Victor Vieira[5]
Raquel Santos de Araújo[6]
Verlayne Kelley da Hora Rocha Araujo[7]
Rodrigo Gayer Amaro[8]

1 Mestre em Engenharia da Computação, Universidade de Pernambuco, Recife, Pernambuco, Brasil. E-mail: isf2@ecomp.poli.br.
2 Doutor em Engenharia Elétrica, Universidade Federal de Pernambuco, Recife, Pernambuco, Brasil. E-mail: jpgo@ecomp.poli.br,
3 Mestre em Engenharia Civil, Universidade Federal de Alagoas, Maceió, Alagoas, Brasil. E-mail: isaque776@hotmail.com.
4 Mestre em Engenharia de Sistemas, Universidade de Pernambuco, Recife, Pernambuco, Brasil. E-mail: rnpf@ecomp.poli.br.
5 Engenheiro Eletricista, Universidade Maurício de Nassau, Recife, Pernambuco, Brasil. E-mail: jvbav@ecomp.poli.br.
6 Biomédica, Centro Universitário do Vale do Ipojuca, Caruaru, Pernambuco, Brasil. E-mail: raquel98450@gmail.com.
7 Mestre em Ciência da Computação, Universidade Federal de Pernambuco, Recife, Pernambuco, Brasil. E-mail: verlaynne@gmail.com.
8 Mestre em Ciências Contábeis, Universidade Federal de Pernambuco, Recife, Pernambuco, Brasil. E-mail: rodrigo.amaro@ufrpe.br.

ABSTRACT

This work investigates a teacher–student learning framework with pseudo-labeling for automated breast lesion detection in mammography using YOLO (You Only Look Once) object detection models. Motivated by the scarcity of reliable bounding-box annotations and the high variability of mammography datasets, the proposed approach leverages a previously trained teacher model to generate pseudo-labels that are subsequently used to train a student model. Experiments are conducted on the public VinDr-Mammo dataset from Vietnam, in conjunction with additional mammography datasets obtained from Kaggle, which are used for pseudo-label generation. The teacher model achieves a mAP@50 of 80.55% and precision of 80.73%. The student reaches a mAP@50 of 82.50% and precision of 88.42%. Overall, pseudo-labeling improves breast lesion detection in heterogeneous mammography scenarios, including it is possible to find for the code on Github.

Keywords: Breast lesion detection; Mammography analysis; Teacher–student learning; Pseudo-labeling; YOLO models.

RESUMO

Este trabalho investiga uma estrutura de aprendizado do tipo professor-aluno (*teacher-student*) com *pseudo-labeling* (pseudo-rotulagem) para a detecção automatizada de lesões mamárias em mamografias, utilizando modelos de detecção de objetos YOLO (*You Only Look Once*). Motivada pela escassez de anotações confiáveis de caixas delimitadoras (*bounding boxes*) e pela alta variabilidade dos conjuntos de dados de mamografia, a abordagem proposta utiliza um modelo "professor" previamente treinado para gerar pseudo-rótulos, os quais são posteriormente empregados no treinamento de um modelo "aluno". Os experimentos foram realizados utilizando o conjunto de dados público VinDr-Mammo (do Vietnã) em conjunto com bases de dados de mamografia adicionais obtidas no Kaggle, utilizadas para a geração dos pseudo-rótulos. O modelo professor alcançou um mAP@50 de 80,55% e uma precisão de 80,73%, enquanto o modelo aluno atingiu um mAP@50 de 82,50% e uma precisão de 88,42%. De modo geral, a técnica de *pseudo-labeling* aprimora a detecção de lesões mamárias em cenários de mamografia heterogêneos; o código do projeto está disponível no GitHub.

Palavras-chave: Detecção de lesões mamárias; Análise de mamografias; Aprendizado professor-aluno; Pseudo-rotulagem; Modelos YOLO.

1. INTRODUCTION

The development of Computer-Aided Diagnosis (CAD) systems for mammography has advanced considerably with the emergence of deep learning, particularly through object detectors from the YOLO (You Only Look Once) family (Litjens et al., 2017; Aydemir et al., 2025). Despite their promising performance, these models still rely on large collections of manually annotated images with bounding boxes, whose acquisition is time-consuming, costly, and highly dependent on experienced radiologists (Shehzadi et al., 2026; Karimi et al., 2020).

In this context, semi-supervised learning methods have gained increasing attention by enabling the exploitation of unlabeled images during training (Shehzadi et al., 2026; Sohn et al., 2020a). Among these approaches, pseudo-labeling allows a previously trained model to automatically generate labels for new data, thereby expanding the training set without requiring additional manual annotations (Sohn et al., 2020a; Li et al., 2022a). When combined with the Teacher–Student paradigm, this strategy enables a teacher model to generate reliable pseudo-labels that are subsequently incorporated into the training of a student model, improving its generalization capability (Sohn et al., 2020a; Liu et al., 2021).

Although these strategies have demonstrated remarkable success in several computer vision tasks, their application to breast lesion detection in mammography remains limited. The high heterogeneity among mammography datasets, differences in acquisition devices, annotation protocols, low image contrast, tissue overlap, and lesion variability make the generation of reliable pseudo-labels particularly challenging, requiring confidence-based filtering mechanisms to reduce the propagation of erroneous annotations (Chen et al., 2024; Karimi et al., 2020; Figueiras et al., 2025).

Motivated by these challenges, this work proposes a Teacher–Student Learning framework based on YOLO object detectors for pseudo-label generation and utilization in automated breast lesion detection. The Teacher model is obtained through a progressive sequential fine-tuning strategy across multiple public mammography datasets, including CBIS-DDSM, INBreast, and VinDr-Mammo, enabling the model to progressively capture domain-specific characteristics from different imaging sources. The resulting teacher model is then employed to generate pseudo-labels for additional mammography images, which are incorporated into the training of the Student model. Finally, the proposed framework is evaluated using the Brazilian SPR-Mammo dataset, allowing the assessment of its generalization capability under a distinct clinical scenario (Nguyen et al., 2023; Figueiredo et al., 2026).

Experimental results demonstrate that the incorporation of pseudo-labels improves the performance of the student model compared with the teacher model, indicating that the combination of progressive sequential fine-tuning, confidence-based pseudo-labeling, and Teacher–Student Learning constitutes a promising strategy for reducing the dependence on manual annotations while improving the robustness of automated breast lesion detection systems.

2. RELATED WORKS

The development of deep learning models for breast lesion detection has been driven by the increasing availability of public mammography datasets, such as CBIS-DDSM, INBreast, and VinDr-Mammo. However, differences in image acquisition protocols, image quality, annotation standards, and patient populations introduce substantial domain heterogeneity, limiting the generalization capability of models trained on a single dataset. Consequently, several studies have investigated domain adaptation strategies to reduce these discrepancies and improve the robustness of object detectors (Nguyen et al., 2023; Figueiredo et al., 2026; Zhang et al., 2025).

To address this challenge, transfer learning and fine-tuning have become widely adopted in medical imaging applications, enabling previously learned representations to be reused and adapted to new domains. More recently, progressive sequential fine-tuning has been proposed as an effective strategy for gradually adapting object detectors across multiple mammography datasets, facilitating knowledge transfer while reducing domain discrepancies throughout the training process (Fatima et al., 2025; Figueiredo et al., 2026; Zhang et al., 2025).

Beyond domain adaptation, the reliability of model evaluation is strongly influenced by data partitioning and training strategies. In this context, K-fold cross-validation has been extensively employed to reduce the variance of performance estimates when dealing with limited medical datasets, whereas data augmentation techniques artificially increase training data diversity, mitigating overfitting and improving model robustness (Teodorescu and

Brașoveanu, 2025; Qi et al., 2025). In semi-supervised learning, data augmentation also plays an additional role by enforcing prediction consistency across transformed versions of the same image, providing an important criterion for evaluating prediction reliability (Sohn et al., 2020b; Lu et al., 2023).

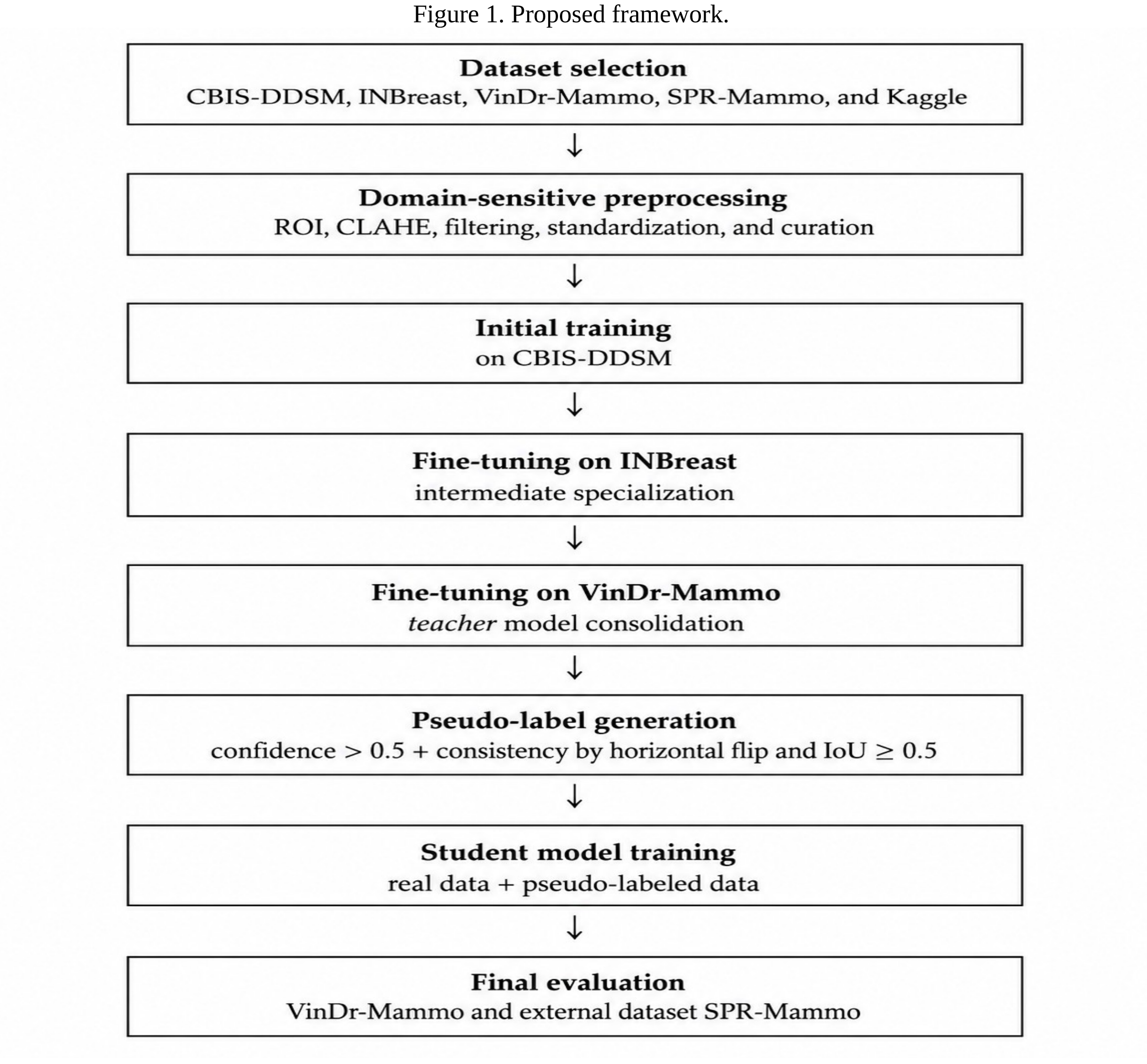
In parallel, semi-supervised learning has emerged as a promising alternative for reducing the dependence on large manually annotated datasets. Among the available approaches, pseudo-labeling enables a previously trained model to automatically generate labels for unlabeled images, expanding the training set without requiring additional expert annotation. Nevertheless, the quality of pseudo-labels depends directly on the reliability of the generated predictions, requiring selection mechanisms based on confidence thresholds, prediction consistency, and spatial criteria such as the Intersection over Union (IoU) to minimize the propagation of noisy annotations (Shehzadi et al., 2026; Silva, 2025; Sohn et al., 2020a; Liu et al., 2021; Li et al., 2022a).

Within this context, the Teacher–Student paradigm has become one of the most widely adopted frameworks for semi-supervised object detection. In this approach, a previously trained Teacher model generates pseudo-labels that are subsequently incorporated into the training of a Student model, allowing labeled and unlabeled data to be jointly exploited during the learning process. Despite the progress achieved in this field, relatively few studies have investigated the integration of progressive domain adaptation, confidence-based pseudo-label generation, and Teacher–Student learning within a unified framework for breast lesion detection in mammography (Li et al., 2022b; Chen et al., 2024; Lu et al., 2025).

3. METHODOLOGY

The proposed methodology is based on a Teacher–Student framework for automated breast lesion detection in mammography using YOLO object detectors. Unlike our previous study, whose primary contribution was the development of a robust Teacher model through progressive sequential fine-tuning, the present work investigates the use of this pretrained model to generate pseudo-labels for training Student models. The proposed framework consists of three main stages: (i) adoption of a previously trained Teacher model; (ii) pseudo-label generation from unlabeled mammograms; and (iii) training of Student models using both manually annotated and pseudo-labeled images. Figure 1 illustrates the overall workflow of the proposed methodology.

Figure 1. Proposed framework.



Source: Prepared by the author.

3.1 Teacher Model

The Teacher model employed in this study was previously developed in our earlier work (Figueiredo et al., 2026). In that study, a progressive sequential fine-tuning strategy was proposed using the public mammography datasets CBIS-DDSM, INBreast, and VinDr-Mammo, enabling the detector to progressively adapt to different mammography domains while improving its robustness and generalization capability.

Since the development of the Teacher model is not the objective of the present work, the pretrained detector was adopted without further modifications and used exclusively to generate pseudo-labels. Details regarding image preprocessing, training procedures, hyperparameter configuration, and the sequential fine-tuning strategy can be found in our previous work (Figueiredo et al., 2026).

3.2 Pseudo-Label Generation

Pseudo-labels were generated using public mammography datasets available on the Kaggle platform. As these datasets do not provide manual annotations compatible with supervised object detection, the pretrained Teacher model was employed to automatically detect breast lesions and generate the corresponding bounding boxes.

To improve the reliability of the generated annotations, each mammogram was processed in its original form and after applying a horizontal flip transformation. Predictions obtained from the transformed image were projected back to the original image coordinates and compared with the detections produced from the original image.

Pseudo-labels were retained only when satisfying two complementary criteria: (i) a prediction confidence score greater than 0.5, and (ii) an Intersection over Union (IoU) greater than or equal to 0.5 between the detections obtained from the original and transformed images. By combining confidence estimation with spatial consistency, this filtering strategy reduces the incorporation of noisy annotations into the pseudo-labeled dataset.

The accepted detections were converted into the YOLO annotation format and stored together with their corresponding mammograms, producing an expanded pseudo-labeled dataset that was subsequently used during the Student training stage.

3.3 Student Model Training

The proposed framework was initially applied to the VinDr-Mammo dataset. Following the same experimental protocol adopted for the Teacher model, the dataset was partitioned using five-fold cross-validation. During each iteration, the training subset was expanded through the incorporation of the pseudo-labeled mammograms generated in the previous stage, whereas the validation subset remained composed exclusively of manually annotated images (ground truth).

The Student model was trained using the same detector architecture, training configuration, and hyperparameter settings adopted for the Teacher model, differing only in the use of the expanded training set containing the pseudo-labeled samples. This strategy allows the performance improvements to be attributed exclusively to the proposed Teacher–Student learning framework rather than to modifications in the detector architecture or optimization procedure.

The trained Student model was evaluated using the standard object detection metrics of Precision, Recall, F1-score, mAP@0.5, and mAP@0.5:0.95.

3.4 Framework Validation on an Independent Dataset

To further investigate the generalization capability of the proposed methodology, the complete Teacher–Student framework was also applied to the independent SPR-Mammo dataset. Following the same experimental protocol adopted for the VinDr-Mammo experiments, a pretrained Teacher model was used to generate pseudo-labels for the SPR-Mammo images.

The generated pseudo-labels were filtered using the same confidence (confidence > 0.5) and spatial consistency (IoU ≥ 0.5) criteria before being incorporated into the training subset. Similarly, the validation subset remained composed exclusively of manually annotated images, ensuring an unbiased performance evaluation.

Subsequently, a new Student model was trained using the same detector architecture, training protocol, and hyperparameter configuration employed in the VinDr-Mammo experiments. By maintaining an identical experimental procedure across both datasets, it was possible to assess the robustness and generalization capability of the proposed Teacher–Student framework under different mammography domains. Finally, the results obtained on the SPR-Mammo dataset were compared with those achieved on VinDr-Mammo to evaluate the effectiveness of the proposed pseudo-labeling strategy.

4. RESULTS AND DISCUSSION

This section presents the experimental results obtained with the proposed Teacher–Student framework. Initially, the performance of the previously developed Teacher model is presented as the baseline for comparison. Subsequently, the results of the pseudo-label generation stage and the training of the Student model using the VinDr-Mammo dataset are analyzed. Finally, the complete framework is validated on the independent SPR-Mammo dataset to evaluate its generalization capability to a different mammography domain.

4.1 Teacher Model Performance

The Teacher model employed in this study was previously developed through a progressive sequential fine-tuning strategy using the public CBIS-DDSM, INBreast, and VinDr-Mammo datasets, as described in our previous work. To provide context for the experiments conducted in this second stage, Table 1 summarizes the datasets used during the construction of the Teacher model.

Table 1 – Datasets used during the construction of the Teacher model.

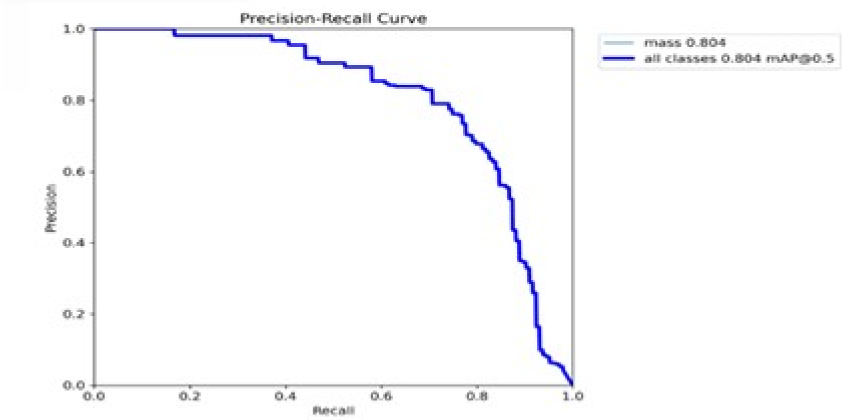
Dataset	Initial Images	Removed Images	Final Images	Strategy
CBIS-DDSM	1,592	0	1,592	Initial Training
INBreast	300	176	124	Fine-tuning
VinDr-Mammo	1,107	393	714	Fine-tuning

Source: Prepared by the author.

During the initial stage, the CBIS-DDSM dataset was divided into training, validation, and testing subsets following an 80/10/10 split, and the model was trained for 50 epochs. Subsequently, fine-tuning was performed using the INBreast dataset, which contained 124 images after the filtering process, adopting a 5-fold cross-validation strategy. Finally, the model was further refined using the VinDr-Mammo dataset, consisting of 714 images, of which 571 were used for training and 143 for validation, also following the 5-fold cross-validation protocol.

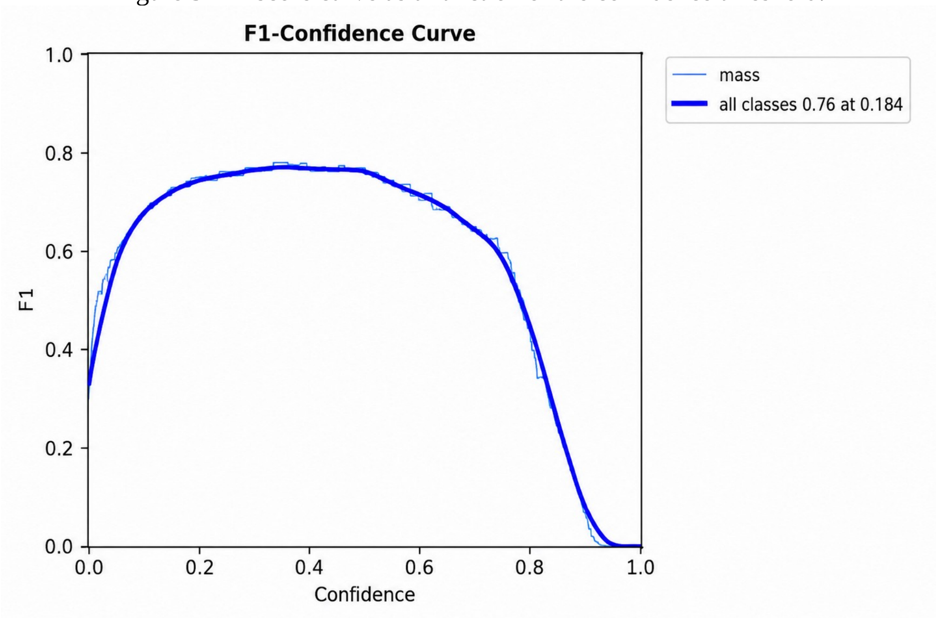
After the sequential adaptation process, the Teacher model achieved a Precision of 80.73%, Recall of 70.60%, mAP@0.5 of 80.55%, and mAP@0.5:0.95 of 51.80%, establishing the baseline for the experiments presented in this work.

Figure 2-Precision–Recall curve of the Teacher model.



Source: Prepared by the author.

Figure 3- F1-score curve as a function of the confidence threshold.



Source: Prepared by the author.

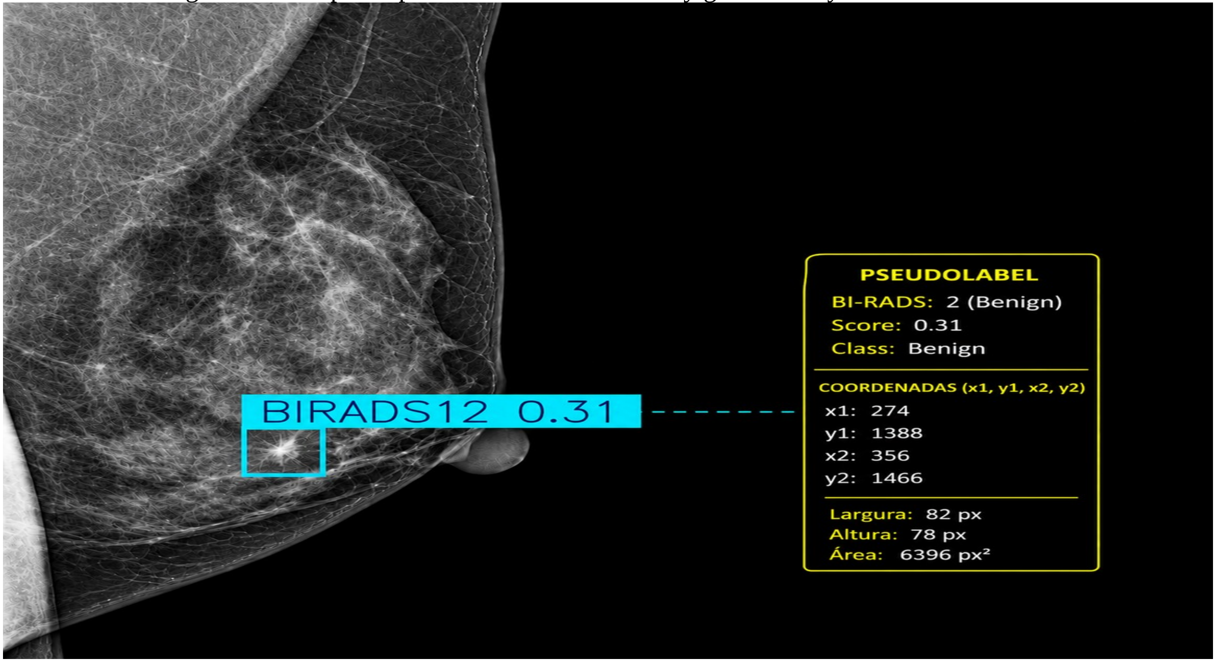
4.2 Pseudo-Label Generation

After obtaining the Teacher model, pseudo-labels were generated using public mammography datasets available on the Kaggle platform. Since these datasets do not provide manual annotations compatible with supervised object detection, the Teacher model was employed to automatically detect breast lesions and generate the corresponding bounding boxes.

Each mammogram was processed in its original form and after applying a horizontal flip transformation. The detections obtained from both inferences were then compared using two selection criteria: confidence greater than 0.5 and Intersection over Union (IoU) greater than or equal to 0.5. Only detections satisfying both conditions were retained as valid pseudo-labels. This procedure resulted in 7128 pseudo-labeled images, which were incorporated into the expanded training set used for the Student model.

Insert Figure 4

Figure 4- Example of pseudo-labels automatically generated by the Teacher model.



Source: Prepared by the author.

4.3 Student Model Performance on VinDr-Mammo

The **Student** model was trained using the **VinDr-Mammo** dataset while maintaining the same experimental protocol adopted for the Teacher model. The dataset was partitioned using **5-fold cross-validation**, with pseudo-labeled images being incorporated exclusively into the training subset. The validation subset remained composed only of manually annotated images (ground truth), ensuring an unbiased evaluation of the model's performance.

Table 2 summarizes the quantitative results obtained during the last ten training epochs of the Student model.

Table 2 – Performance metrics obtained during the last ten training epochs of the Student model.

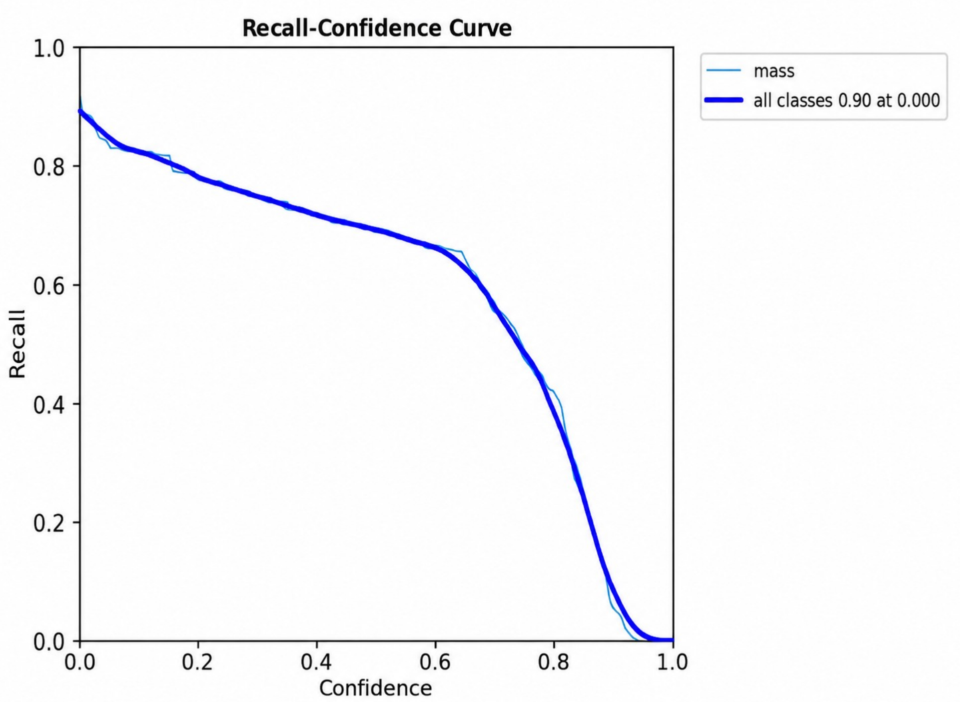
Epoch	Precision	Recall	mAP@50	mAP@50-95
61	0.80738	0.76056	0.81590	0.49577
62	0.80506	0.76056	0.82263	0.48554
63	0.93239	0.67989	0.83606	0.50163
64	0.80621	0.76056	0.83398	0.49309
65	0.78502	0.77149	0.82601	0.49811
66	0.83206	0.73944	0.82340	0.49067
67	0.90755	0.66901	0.82146	0.48641
68	0.88384	0.70423	0.82966	0.50385
69	0.89854	0.69718	0.83727	0.50293
70	0.88420	0.69718	0.82500	0.49722

Source: Prepared by the author.

The best experimental configuration was obtained at **epoch 70**, achieving **Precision of 88.42%**, **Recall of 69.71%**, **mAP@0.5 of 82.50%**, and **mAP@0.5:0.95 of 49.72%**.

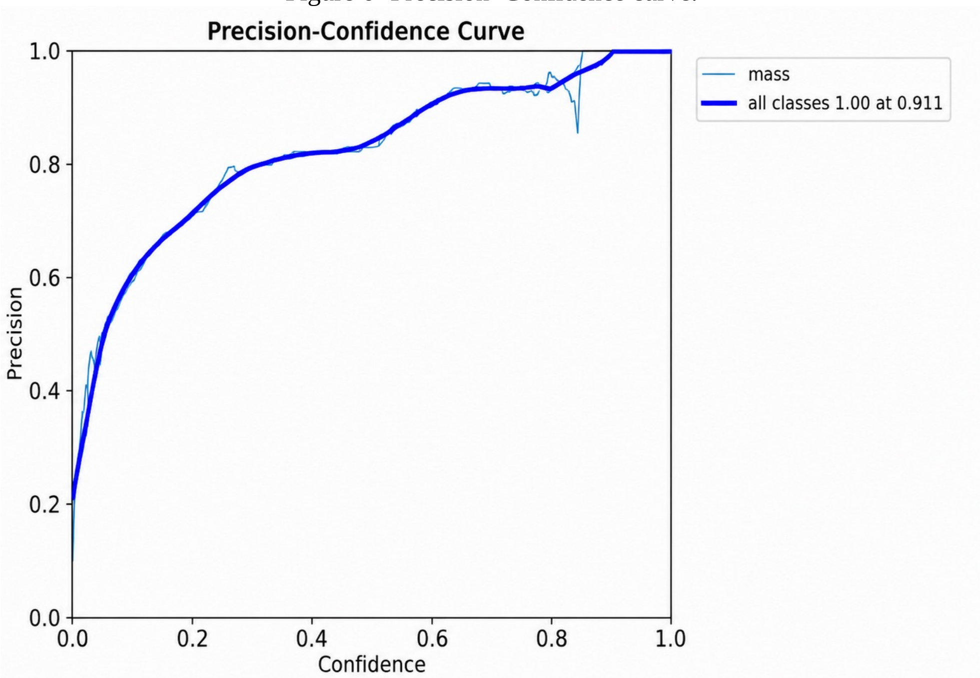
Compared with the Teacher model, the Student model achieved an increase of approximately **7.7 percentage points in Precision** and nearly **2 percentage points in mAP@0.5**. The results indicate that incorporating pseudo-labels reduced the number of false-positive detections while preserving the overall lesion detection capability.

Figure 5- Recall–Confidence curve.



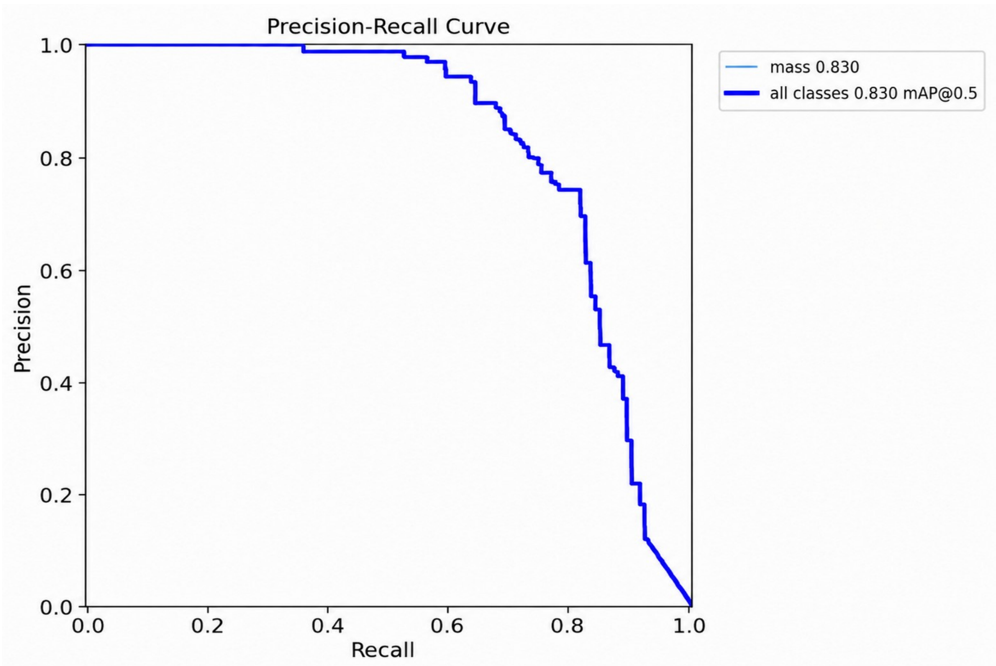
Source: Prepared by the author.

Figure 6- Precision–Confidence curve.



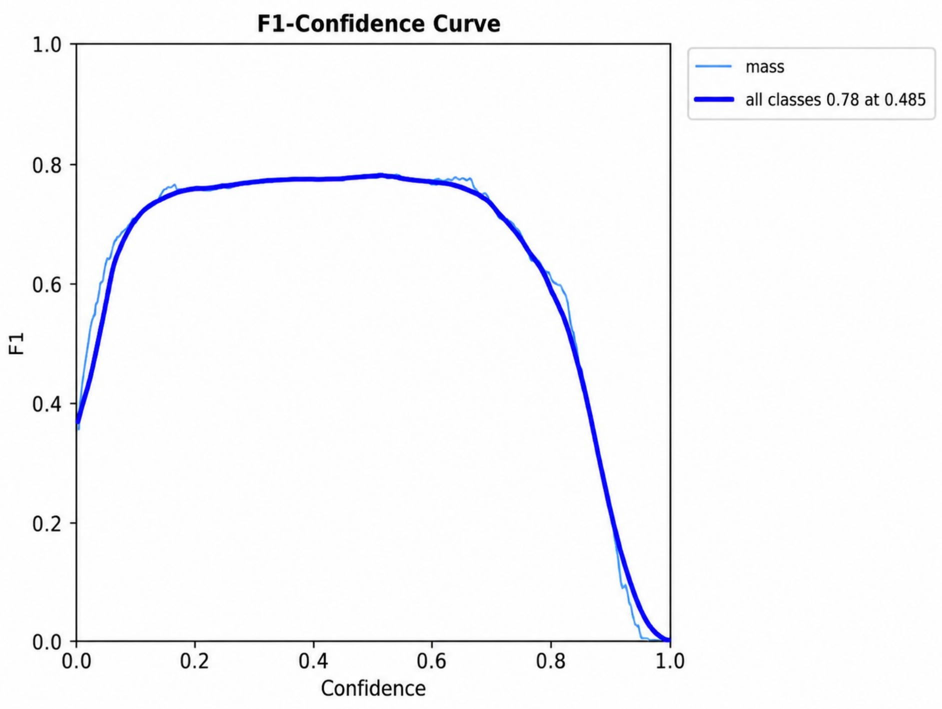
Source: Prepared by the author.

Figure 7 Precision–Recal curve.



Source: Prepared by the author.

Figure 8- F1-score curve as a function of the confidence threshold.



Source: Prepared by the author.

4.4 Framework Validation on the SPR-Mammo Dataset

To further investigate the generalization capability of the proposed approach, the complete Teacher–Student framework was also applied to the independent SPR-Mammo dataset, consisting of 24334 mammographic images. Initially, a pretrained Teacher model was employed to generate pseudo-labels for the entire dataset using the same filtering strategy based on confidence greater than 0.5 and IoU greater than or equal to 0.5.

This procedure produced 2231 pseudo-labels, which were subsequently incorporated into the training of a new Student model while maintaining exactly the same experimental protocol adopted for the VinDr-Mammo experiments.

Table 3 summarizes the quantitative results obtained during the final training epochs.

REVISTA OWL (*OWL Journal*)
www.revistaowl.com.br – ISSN: 2965-2634

Table 3 – Performance metrics obtained by the Student model on the SPR-Mammo dataset.

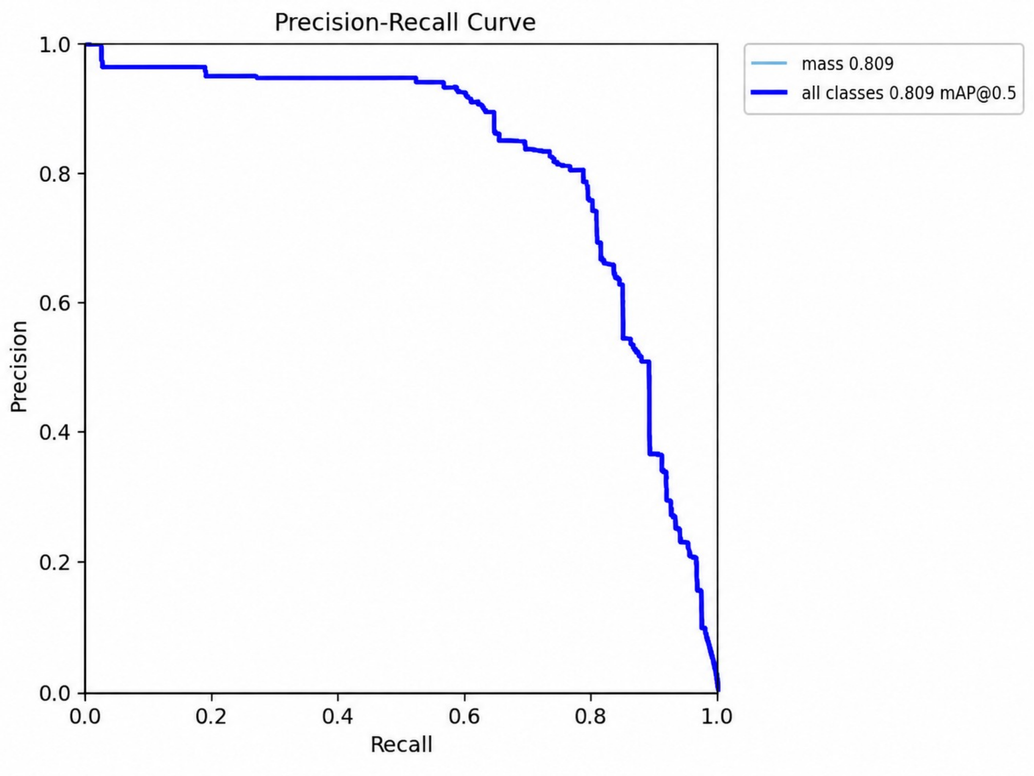
Epoch	**Precision**	**Recall**	**mAP@50**	**mAP@50-95**
61	0.79791	0.74825	0.81561	0.45279
62	0.79098	0.74126	0.78364	0.46042
63	0.76687	0.75910	0.77460	0.46226
64	0.77201	0.75777	0.79069	0.45940
65	0.82128	0.76923	0.81467	0.46487
66	0.79615	0.75524	0.80901	0.46914
67	0.79279	0.74915	0.79892	0.46489
68	0.79978	0.75420	0.79163	0.46432
69	0.83939	0.73096	0.80273	0.46270
70	0.83126	0.73427	0.80674	0.46736

Source: Prepared by the author.

The best experimental configuration achieved Precision of 83.13%, Recall of 73.43%, mAP@0.5 of 80.67%, and mAP@0.5:0.95 of 46.74%.

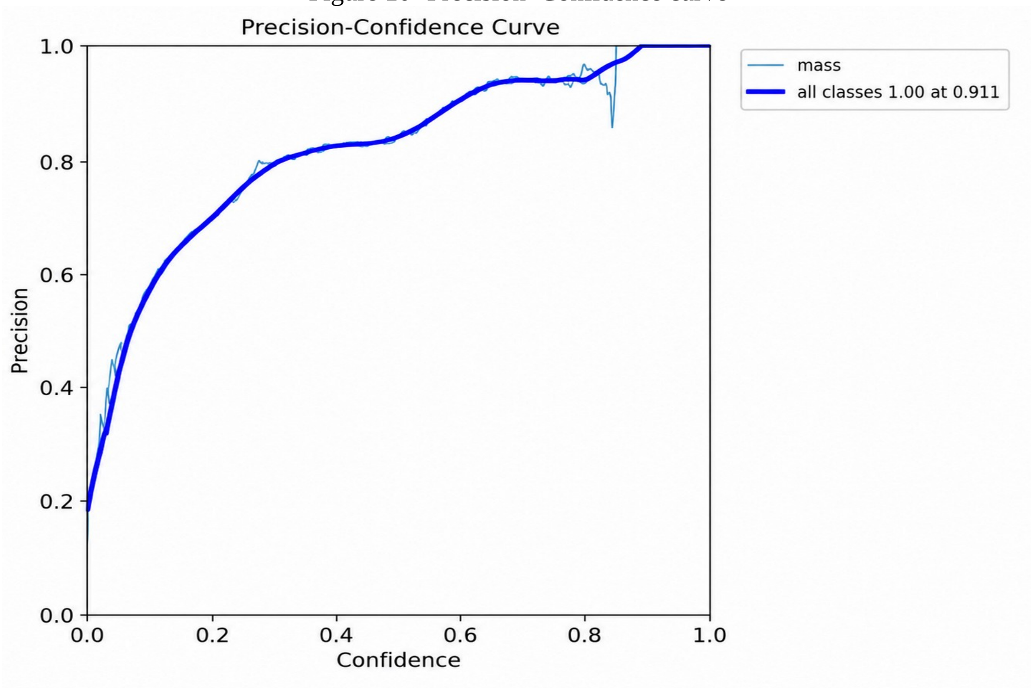
Although a slight reduction in Precision was observed compared with the VinDr-Mammo experiments, Recall increased, indicating that the proposed framework maintained a high lesion detection capability even when applied to a different clinical domain.

Figure 9- Precision–Recall curve on the SPR-Mammo dataset.



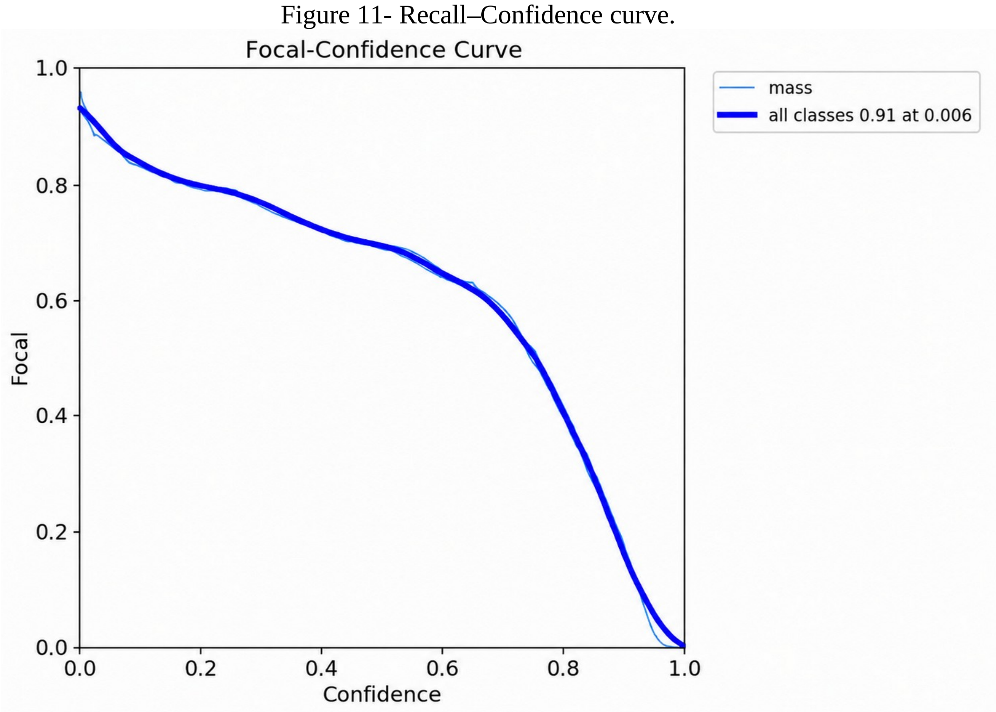
Source: Prepared by the author.

Figure 10- Precision–Confidence curve



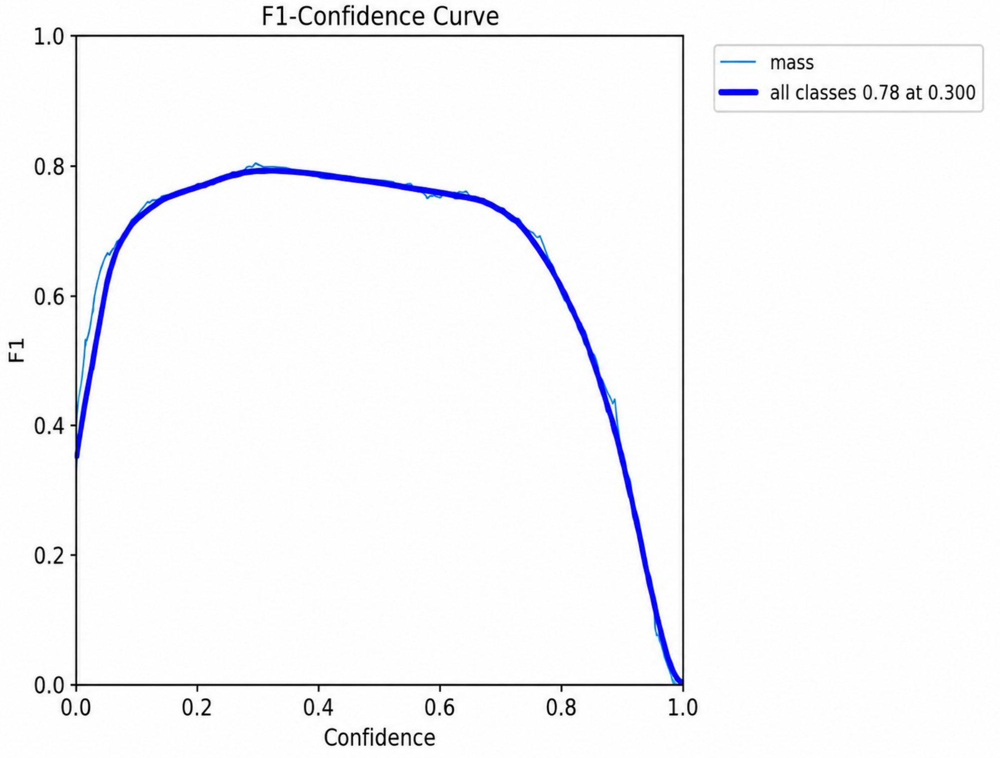
Source: Prepared by the author.

Figure 11- Recall–Confidence curve.



Source: Prepared by the author.

Figure 12- F1-score curve as a function of the confidence threshold.



Source: Prepared by the author.

4.5 Comparative Discussion

Table 4 summarizes the overall performance achieved throughout the proposed framework, allowing a direct comparison between the **Teacher** model and the **Student** models trained on the **VinDr-Mammo** and **SPR-Mammo** datasets.

Table 4- Efficiency comparison

Model	Precision	Recall	mAP@0.5	mAP@0.5:0.95
Teacher (VinDr-Mammo)	80.73%	70.60%	80.55%	**51.80%**

Student (VinDr-Mammo)	**88.42%**	69.71%	**82.50%**	49.72%
Student (SPR-Mammo)	83.13%	**73.43%**	80.67%	46.74%

Source: Prepared by the author.

Overall, the experimental results demonstrate that the proposed Teacher–Student framework effectively improves automated breast lesion detection through the incorporation of confidence-filtered pseudo-labels. Training the Student model using both manually annotated and pseudo-labeled data significantly increased Precision while maintaining competitive Recall and mAP values. Furthermore, applying the same framework to the independent SPR-Mammo dataset demonstrated that the proposed methodology preserves its performance under a different mammography domain, highlighting the robustness and generalization capability of the proposed pseudo-labeling strategy.

5. CONCLUSION

This study presented a **Teacher–Student** framework based on confidence-filtered pseudo-labeling for automated breast lesion detection in mammography. A previously developed **Teacher** model was employed to generate pseudo-labels from public mammography datasets, allowing the expansion of the training set without requiring additional manual annotations. The generated pseudo-labels were subsequently incorporated into the training of a **Student** model, following the same experimental protocol adopted during the construction of the Teacher model.

The experimental results demonstrated that the proposed strategy improved the overall detection performance. When evaluated on the **VinDr-Mammo** dataset, the Student model achieved higher **Precision** and **mAP@0.5** than the Teacher model, indicating that the incorporation of pseudo-labeled images contributed to reducing false-positive detections while preserving competitive detection capability. Furthermore, the application of the complete Teacher–Student framework to the independent **SPR-Mammo** dataset demonstrated that the proposed methodology maintains consistent performance in a different mammography domain, highlighting its robustness and generalization capability.

Although the obtained results are encouraging, some limitations should be considered. The quality of the generated pseudo-labels is directly dependent on the performance of the Teacher model, meaning that inaccurate detections may propagate errors during Student training. In addition, the pseudo-label filtering strategy adopted in this work relies exclusively on confidence and spatial consistency criteria, which may still retain incorrect annotations in challenging clinical scenarios.

Future work will investigate more sophisticated pseudo-label refinement strategies, including uncertainty estimation, model ensembles, and adaptive confidence thresholds. Additional experiments involving larger and more diverse mammography datasets, as well as external clinical validation, are also necessary to further assess the applicability of the proposed framework in real-world computer-aided diagnosis systems.

Overall, the proposed Teacher–Student framework demonstrates that confidence-based pseudo-labeling represents an effective strategy for improving breast lesion detection while reducing the dependence on manually annotated datasets. The proposed methodology provides a practical alternative for leveraging large collections of unlabeled mammograms and represents a promising direction toward scalable and robust computer-aided diagnosis systems for breast cancer screening.

6. REFERENCES

CHEN, Yuanhong et al. BRAIxDet: Learning to detect malignant breast lesion with incomplete annotations. **Medical image analysis**, v. 96, p. 103192, 2024.

FATIMA, Maria; ZIA, Razia; USMANI, Irfan Ahmed. A hybrid framework integrating deep learning and XAI for breast cancer detection with stratified cross-validation. **Systems and Soft Computing**, p. 200412, 2025.

FIGUEIRAS, Hugo et al. Self-supervised learning for breast cancer detection: A review. **Computers in Biology and Medicine**, v. 198, p. 111245, 2025.

FIGUEIREDO, I. S., OLIVEIRA, J. P. G. de, GALINDO, J. M. da S., & VIEIRA, J. V. (2026). A progressive fine-tuning strategy for robust breast lesion detection across multiple mammography datasets. **Caderno Pedagógico**, 23(1), e22794.

KARACA AYDEMIR, Büşra Kübra et al. Detecting and classifying breast masses via YOLO-based deep learning. **Neural Computing and Applications**, v. 37, n. 17, p. 11555-11582, 2025.

KARIMI, Davood et al. Deep learning with noisy labels: Exploring techniques and remedies in medical image analysis. **Medical image analysis**, v. 65, p. 101759, 2020.

LI, Gang et al. Pseco: Pseudo labeling and consistency training for semi-supervised object detection. In: **European Conference on Computer Vision**. Cham: Springer Nature Switzerland, 2022. p. 457-472.

LI, Hengduo et al. Rethinking pseudo labels for semi-supervised object detection. In: **Proceedings of the AAAI conference on artificial intelligence**. 2022. p. 1314-1322.

LITJENS, Geert et al. A survey on deep learning in medical image analysis. **Medical image analysis**, v. 42, p. 60-88, 2017.

LIU, Yen-Cheng et al. Unbiased teacher for semi-supervised object detection. **arXiv preprint arXiv:2102.09480**, 2021.

LU, Liyun et al. Uncertainty-aware pseudo-label and consistency for semi-supervised medical image segmentation. **Biomedical Signal Processing and Control**, v. 79, p. 104203, 2023.

LU, Yunyao et al. Enhancing dual network based semi-supervised medical image segmentation with uncertainty-guided pseudo-labeling. **Knowledge-Based Systems**, p. 114454, 2025.

NGUYEN, Hieu T. et al. VinDr-Mammo: A large-scale benchmark dataset for computer-aided diagnosis in full-field digital mammography. **Scientific Data**, v. 10, n. 1, p. 277, 2023.

QI, Xuyin et al. Mediaug: Exploring visual augmentation in medical imaging. In: **Annual Conference on Medical Image Understanding and Analysis**. Cham: Springer Nature Switzerland, 2025. p. 218-232.

SHEHZADI, Tahira et al. Semi-supervised object detection: A survey on progress from cnn to transformer. **Sensors (Basel, Switzerland)**, v. 26, n. 1, p. 310, 2026.

SILVA, Luiz Miguel Santos. Uma estratégia robusta de reavaliação de pseudo-rótulos para o algoritmo Self-Training. 2025.

SOHN, Kihyuk et al. Fixmatch: Simplifying semi-supervised learning with consistency and confidence. **Advances in neural information processing systems**, v. 33, p. 596-608, 2020.

SOHN, Kihyuk et al. A simple semi-supervised learning framework for object detection. **arXiv preprint arXiv:2005.04757**, 2020.

TEODORESCU, Vlad; OBREJA BRAȘOVEANU, Laura. Assessing the validity of k-fold cross-validation for model selection: Evidence from bankruptcy prediction using random forest and XGBoost. **Computation**, v. 13, n. 5, p. 127, 2025.

ZHANG, Ziqi et al. Robust Source-Free Domain Adaptation for Medical Image Segmentation based on Curriculum Learning. **arXiv preprint arXiv:2510.08393**, 2025.

Apêndice

Github: https://github.com/isf2enge/Mammo-dataset-complete-articles

Recebido em: 23/06/2026
Aprovado em: 18/07/2026
Publicado em: 13/08/2026