YOLO Based Object Detection in SAR Images: Embedded Implementation and Performance Analysis

Emirhan Küçük[1], Muhammed Kadir Aksoy[1], Mustafa Kabak[1], Burak Catalbas[3], Ali Alp Akyol[2], Bahadir Catalbas[1]

Abstract— **Object detection using synthetic aperture radar (SAR) imagery is a significant challenge. This paper presents a comparative analysis of YOLO26x, YOLO26x-OBB, and YOLOv8x-OBB models using the FAIR-CSAR dataset for object detection in synthetic aperture radar (SAR) imagery. While many studies evaluate SAR object detection models in offline workstation environments, their performance across different computational platforms remains largely unexplored. To address this gap, the models are deployed and evaluated on multiple platforms, including the Jetson AGX Orin, NVIDIA RTX A4000, and NVIDIA RTX PRO 6000 Blackwell Max-Q. The study also focuses on the performance of the oriented bounding box (OBB) detection, which is essential to accurately localize objects in SAR imagery. In addition to detection accuracy, the performance of the real-time inference on these platforms is analyzed. Experimental results highlight the trade-off between detection accuracy and inference latency across different computational systems, demonstrating the feasibility of deploying advanced SAR object detection models on both edge and high-performance platforms.**

***Index Terms*—Synthetic Aperture Radar (SAR), Neural Networks, Object Detection, YOLO, Oriented Bounding Box (OBB), Edge Computing, Quantization**

I. INTRODUCTION

Object detection and classification in synthetic aperture radar (SAR) images is a challenging task, there are many different applications in the literature [1], [2]. Unlike optical images, SAR images contain speckle noise and complex background patterns, making accurate object detection difficult [3]. In addition, targets may appear at different orientations, scales and numbers. The motion of the radar platform and external disturbances such as vibrations can introduce image distortions and low-resolution patches, further complicating target detection. These factors can significantly reduce the performance of traditional detection methods. To address these challenges, recent studies have employed deep learning techniques, multi-scale feature extraction, and attention mechanisms to improve feature representation [4], [5]. Oriented bounding box (OBB) approaches are also used to better represent object orientation, improving both detection accuracy and classification performance.

[1]E. Küçük, M. K. Aksoy, M. Kabak, and B. Catalbas are with the Department of Control and Automation Engineering, Yildiz Technical University, Istanbul, Turkey. {emirhankucuk47, kadiraksoy434, mkabak.work}@gmail.com, bahadir.catalbas@yildiz.edu.tr

[2]A. A. Akyol is with ASELSAN, Ankara, Turkey. alialpakyol@gmail.com

[3]B. Catalbas is with the Department of Computer Engineering, Bahcesehir University, Istanbul, Turkey. burak.catalbas@bau.edu.tr

Artificial neural networks are employed for many purposes. From medical diagnosis [6] to robot locomotion [7], voice generation [8] and object classification [9], differing tasks are accomplished thanks to the significant achievements of neural network literature. Using various network architectures and application configurations, neural networks becoming even more successful by the increasing amounts of data being allocated for their training. Given their significant increase of success in comparison to other machine learning methods, neural network-based architectures are chosen for the goals of this work.

Most SAR object detection studies are conducted on powerful computing platforms such as high-performance workstations with powerful Graphic Processing Units (GPUs). These systems provide sufficient computational resources to process large SAR images and employ complex deep learning models. However, deploying such models on resource-constrained edge devices remains challenging due to limitations in processing power, and high-memory consumption. Consequently, SAR object detection is typically performed in offline environments, rather than directly on embedded systems. Therefore, deployment and performance evaluation of SAR detection models on edge platforms remains relatively underexplored. With the development of modern AI accelerators, however, it has become increasingly feasible to run lightweight or moderately sized models on edge devices when optimized architectures and appropriate datasets are used [10].

Among deep learning–based object detection approaches, the You Only Look Once (YOLO) family has attracted significant attention due to its high detection speed and strong performance. Several studies have explored YOLO models for SAR object detection [11], [12], demonstrating a good compromise between detection accuracy and computational efficiency. In addition, benchmark datasets such as FAIR-CSAR are introduced to support object detection tasks in SAR imagery [13]. YOLO architectures were selected for this study because their single-stage nature offers a superior balance between spatial accuracy and inference speed compared to two-stage detectors like Faster R-CNN. While OBB-based variants increase computational overhead by adding a rotation parameter to the loss function, they significantly mitigate bounding box overlap in dense SAR scenes. Several studies have explored YOLO models for SAR object detection, demonstrating a good compromise between accuracy and efficiency. In this work, the YOLO26x-OBB model was

TABLE I

COMPARISON OF THE PROPOSED YOLO26X-OBB WITH STATE-OF-THE-ART MODELS ON THE FAIR-CSAR DATASET.

Model	mAP50 (%)	F1 Score
RetinaNet [13]	31.8	0.39
Faster-RCNN [13]	35.3	0.44
GWD [13]	34.5	0.40
RoI Transformer [13]	38.0	0.47
YOLO26x-OBB (Proposed)	**44.54**	**0.467**

specifically chosen as the primary architecture based on the benchmark results presented in Table I. In addition to the proposed methods on the FAIR-CSAR dataset, a method of YOLO26x-OBB is proposed in Table I. As shown in the comparative analysis, YOLO26x-OBB achieves a mAP50 of 44.54%, significantly outperforming established state-of-the-art models such as RoI Transformer (38.0%) and Faster-RCNN (35.3%) on the FAIR-CSAR dataset. Furthermore, with an F1 score of 0.467, YOLO26x-OBB maintains a competitive balance between precision and recall, proving its effectiveness for diverse SAR object categories while remaining suitable for real-time deployment.

In this work, we evaluate the performance of several YOLO-based models for SAR object detection using the FAIR-CSAR dataset. Specifically, YOLOv8x, YOLO26x-OBB, YOLO26x, YOLOv8m, and YOLOv8s models are analyzed and compared in terms of both detection accuracy and computational efficiency. To the best of our knowledge, this study is the first application of YOLO26x to detect objects in FAIR-CSAR dataset. In addition to algorithmic performance, the deployment feasibility of these models is also investigated across different computing environments.

This study investigates the feasibility of deploying advanced SAR object detection models across diverse computing environments, ranging from workstation-class NVIDIA RTX A4000 and PRO 6000 GPUs to the embedded Jetson AGX Orin platform. The main contributions of this work are three-fold: (i) the first comprehensive benchmarking of YOLO26x and OBB variants on the FAIR-CSAR dataset, (i) a cross-platform performance analysis across high-end and power-constrained devices, and (iii) a systematic investigation of resolution scaling and quantization impacts on real-time SAR detection. This evaluation provides a practical roadmap for balancing detection accuracy and inference speed in real-world processing environments.

II. DATASET AND PROPOSED METHOD

This section presents the dataset, detection models, and deployment platforms used in this study. YOLO-based object detection models are trained and evaluated on the FAIR-CSAR dataset [13], and their deployment feasibility on different hardware is also investigated.

A. FAIR-CSAR Dataset

The experiments in this study are conducted using the FAIR-CSAR dataset, a large-scale benchmark designed for synthetic aperture radar (SAR) object detection and recognition tasks. The dataset is constructed from Gaofen-3 satellite SAR imagery and contains more than 340,000 annotated object instances collected from multiple scenes across different geographic regions. The targets belong to five major categories with several fine-grained subcategories, providing a diverse set of object classes for detection and recognition.

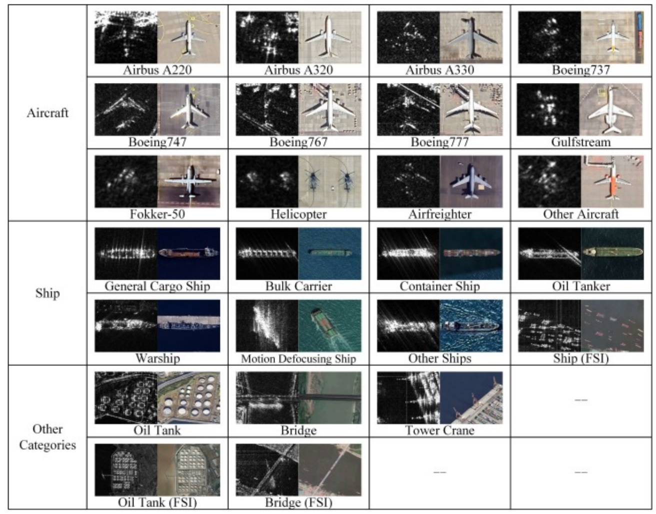

Fig. 1. Representative visualization of the five main categories and 22 detailed subcategories included in the FAIR-CSAR dataset, along with their corresponding SAR samples and paired optical reference images [13].

All targets are annotated using oriented bounding boxes (OBBs), which accurately represent object locations and orientations in SAR imagery. In addition, the dataset includes objects with varying scales, orientations, and complex background conditions such as clutter, speckle noise, and imaging artifacts. The large scale and diversity of the dataset make it suitable not only for evaluating detection accuracy but also for assessing the computational feasibility of deploying deep learning models on embedded and resource-constrained platforms.

B. YOLO-Based Detection Models

In this work, several YOLO-based detection models are evaluated: YOLOv8x, YOLOv8x-OBB, YOLO26x-OBB, YOLO26x, YOLOv8m, and YOLOv8s. The YOLO architecture performs object detection in a single stage, where object localization and classification are predicted simultaneously. This design enables efficient inference and makes YOLO suitable for real-time applications.

The YOLOv8x model represents a high-capacity version of the YOLOv8 architecture. In addition, oriented bounding box (OBB) variants are used to better represent the orientation of targets in SAR images. The YOLOv8x-OBB and YOLO26x-OBB models extend the standard detection framework by predicting rotated bounding boxes, allowing improved localization of objects with arbitrary orientations.

In addition to these four architectures, lighter YOLO variants are also evaluated to analyze their suitability for deployment on embedded platforms. The YOLO26x model represents a standard bounding-box detector, while YOLOv8m and YOLOv8s correspond to medium and lightweight versions of the YOLOv8 architecture. These models provide reduced computational complexity and faster inference speeds, making them more suitable for edge and embedded AI applications.

C. Hardware Implementations

To evaluate the deployment feasibility of SAR object detection models across different computing environments, trained YOLO-based models are tested on three hardware platforms with varying computational capabilities. First, experiments are conducted on an NVIDIA RTX A4000 GPU and NVIDIA RTX PRO 6000 Blackwell Max-Q, which serve as high-performance reference platforms. The models are then deployed on the NVIDIA Jetson AGX Orin to assess their performance on a powerful embedded AI system, to experiment on an edge device to obtain respective benchmarks.

III. APPLICATION AND PERFORMANCE ANALYSIS

This section analyzes the application and performance of the proposed detection framework. Experiments are conducted hierarchically across different computing platforms, from high-performance GPU environments to an embedded edge system.

A. OBB and HBB

Firstly, the difference between horizontal bounding boxes (HBB) and oriented bounding boxes (OBB) is illustrated in Fig. 2. In SAR imagery, targets such as ships and bridges frequently appear with arbitrary orientations. While HBB annotations enclose objects using axis-aligned rectangles, OBB annotations align with the actual orientation of the targets and therefore provide more precise localization for elongated or rotated structures.

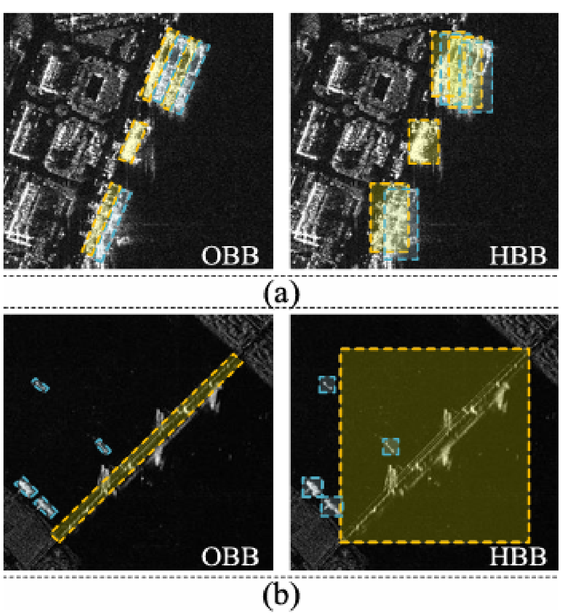

Fig. 2. Comparison between horizontal bounding boxes (HBB) and oriented bounding boxes (OBB) in SAR imagery, shown separately for two images at (a) and (b) [13].

Although the FAIR-CSAR dataset provides OBB annotations and several recent YOLO-based detectors support oriented bounding box detection, this study also considers standard HBB-based YOLO models. This choice allows the evaluation of lightweight detection models that are more suitable for deployment on resource-constrained embedded platforms.

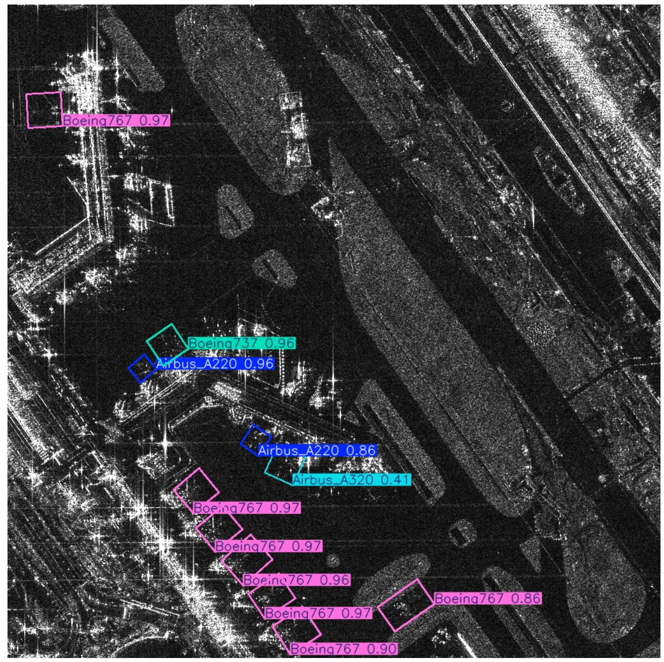

Fig. 3. Sample object detection results on SAR imagery using the YOLO26x-OBB model on the NVIDIA RTX A4000 platform. Oriented bounding boxes (OBB) demonstrate high localization accuracy for various aircraft classes.

B. Evaluation Framework

In this study, multiple YOLO-based object detection models are evaluated using a consistent experimental framework across different computing platforms. To maintain software consistency and ensure the reproducibility of the experimental results, the Ultralytics framework (version 8.4.19) was employed throughout the implementation, training, and inference phases of all five YOLO variants. The evaluation results presented in Table II, Table III, and Table IV are based on the Average Precision (AP) metric computed under the mAP50 and mAP50–95 and F1 score.

Here, mAP50 metric represents the mean Average Precision at an IoU threshold of 0.50. The mAP50-95 metric measures detection accuracy by averaging precision values across multiple intersection-over-union (IoU) thresholds from 0.50 to 0.95, providing a stricter and more comprehensive assessment of localization performance for each class than mAP50. The F1 Score corresponds to the harmonic mean of precision and recall, reflecting the balance between correct detections and missed targets.

C. High-Performance Workstation Evaluation

Table II and IV report both detection accuracy and computational efficiency metrics for the evaluated YOLO models tested in RTX A4000 and Jetson AGX Orin devices, respectively. Latency (ms) indicates the average inference time required to process a single image, whereas frames per second (FPS) measures real-time processing speed. Together, these metrics provide a unified comparison of model accuracy and runtime performance.

TABLE II
DETECTION PERFORMANCE AND INFERENCE SPEED OF YOLO MODELS ON THE RTX A4000 PLATFORM FOR SAR DATASET

YOLO Model	mAP50	mAP50-95	Latency (ms)	FPS	F1 Score
YOLO26x-OBB	0.4454	0.3102	43.65	22.91	0.4672
YOLOv8x-OBB	0.4412	0.2987	49.00	20.41	0.4585
YOLO26x	0.3963	0.2330	43.80	22.83	0.4172
YOLOv8m	0.3308	0.1858	26.08	38.35	0.3550
YOLOv8s	0.2918	0.1529	19.08	52.42	0.3538

The initial experiments were conducted on a workstation equipped with an NVIDIA RTX A4000 GPU to establish a high-performance reference platform for evaluating the proposed detection framework.

As summarized in Table II, OBB-based architectures achieve the highest detection accuracy, with mAP50 values of 0.4454 and 0.4412, respectively. This gain stems from the ability of OBB-based detectors to capture object orientation, which is important for SAR imagery where targets often appear with arbitrary rotations. However, this improvement comes at the cost of increased computational complexity and higher inference latency. In contrast, as stated in Table II, lightweight YOLO variants such as YOLOv8s and YOLOv8m achieve significantly higher inference speeds, reaching up to 52.42 FPS, while exhibiting a moderate decrease in detection accuracy. The YOLO26x model demonstrates a balanced trade-off between these two objectives, maintaining competitive accuracy with improved efficiency.

TABLE III
PER-CLASS DETECTION PERFORMANCE (MAP50 AND MAP50-95) FOR YOLO26X-OBB ON THE FAIR-CSAR DATASET (RTX A4000).

Class	mAP50	mAP50-95
Airbus A220	0.445	0.381
Airbus A320	0.175	0.129
Airbus A330	0.500	0.369
Boeing737	0.587	0.418
Boeing747	0.605	0.494
Boeing767	0.414	0.309
Boeing777	0.131	0.103
Gulfstream	0.153	0.102
Fokker-50	0.868	0.703
Helicopter	0.807	0.498
Airfreighter	0.649	0.473
Other Aircraft	0.0737	0.0527
Bulk Carrier	0.678	0.569
Container Ship	0.336	0.233
General Cargo Ship	0.517	0.323
Oil Tanker	0.271	0.233
Warship	0.230	0.155
Motion Defocusing Ship	0.393	0.204
Other Ship	0.394	0.222
Tank	0.500	0.297
Bridge	0.570	0.247
Tower Crane	0.505	0.285

Overall, these results indicate that while OBB-based models provide a significant improvement in detection accuracy, high-capacity architectures maximize overall performance on workstation-level hardware. However, practical deployment scenarios require careful consideration of the trade-off between accuracy and inference speed. This highlights the importance of selecting model architectures based not only on performance metrics but also on application-specific computational constraints.

D. Embedded Deployment and Edge Analysis

In order to evaluate the feasibility of executing SAR object detection models on an embedded computing platform, the trained YOLO-based detectors were also deployed on the NVIDIA Jetson AGX Orin system. The same experimental setup used in the workstation experiments was preserved to enable a fair comparison with the RTX A4000 platform.

TABLE IV
DETECTION PERFORMANCE AND INFERENCE SPEED OF YOLO MODELS ON THE NVIDIA JETSON AGX ORIN PLATFORM (FP32) FOR SAR DATASET

YOLO Model	mAP50	mAP50-95	Latency (ms)	FPS	F1 Score
YOLO26x-OBB	0.4454	0.3102	88.831	11.26	0.4665
YOLOv8x-OBB	0.4412	0.2985	104.412	9.57	0.4586
YOLO26x	0.3962	0.2332	86.52	11.56	0.4173
YOLOv8m	0.3308	0.1856	43.01	23.25	0.3545
YOLOv8s	0.2919	0.1530	27.208	36.72	0.3559

As summarized in Table IV, the overall performance trends remain consistent across platforms, with OBB-based models maintaining their accuracy advantage. This confirms that orientation-aware detection is inherently beneficial for SAR imagery, independent of the underlying hardware.

However, the impact of hardware constraints becomes more pronounced on the embedded platform. Although OBB-based models achieve the highest accuracy (e.g., mAP50 of 0.4454), their increased computational complexity leads to higher latency and reduced throughput compared to the workstation setup. This highlights a key trade-off between detection performance and real-time capability.

In contrast, lightweight HBB-based models demonstrate a significant improvement in inference speed, making them more suitable for time-critical applications despite their lower accuracy. These results indicate that model selection on embedded systems should be driven not only by detection performance but also by application-specific latency requirements.

Overall, the Jetson AGX Orin platform provides sufficient performance for near real-time SAR object detection, while emphasizing the importance of balancing accuracy and efficiency in practical deployment scenarios. These findings further indicate that the performance gap between workstation and embedded platforms is primarily governed by

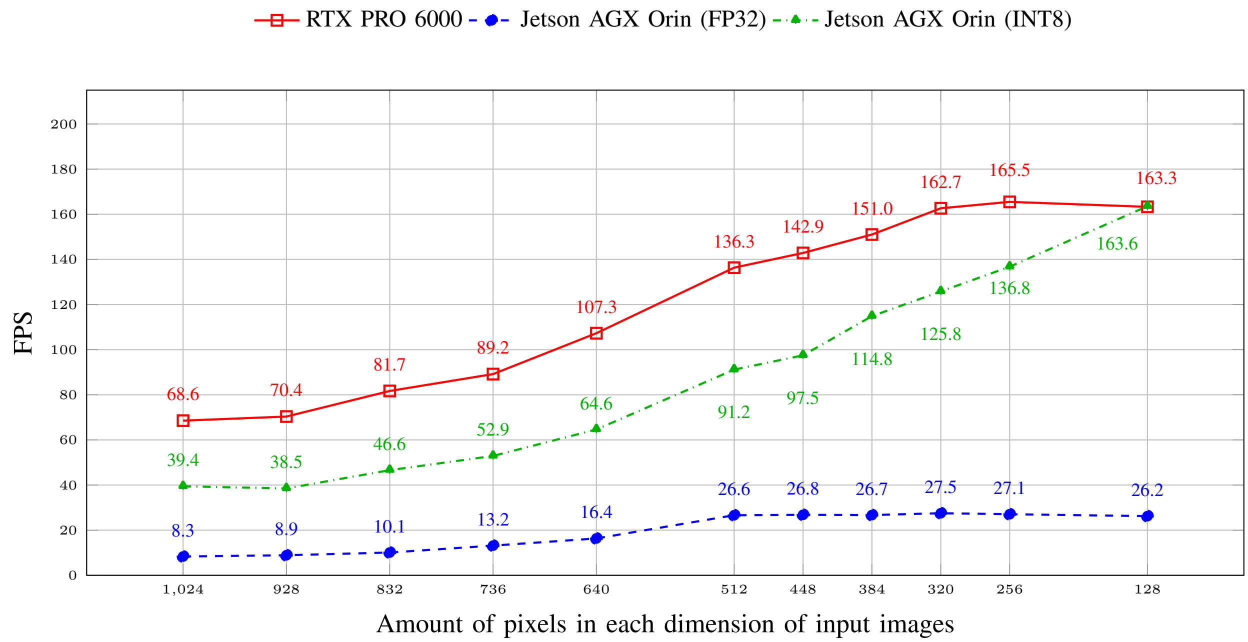

Fig. 4. Inference Speed (FPS) results compared to the decreasing image resolution.

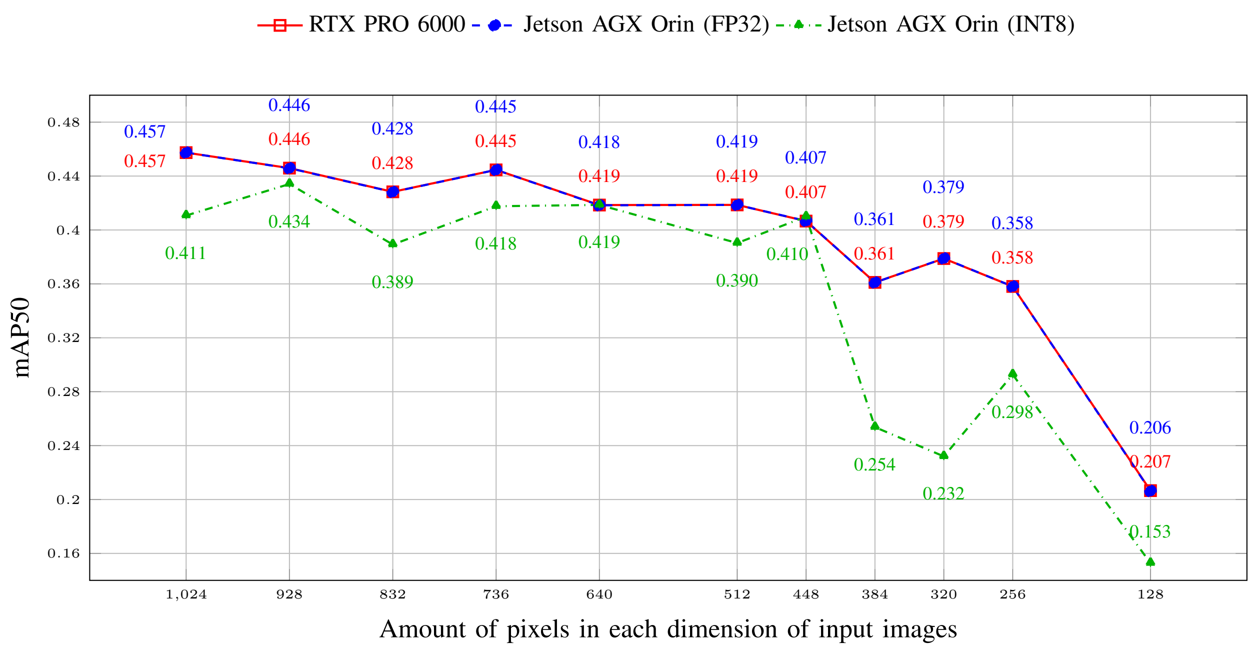

Fig. 5. Detection Accuracy (mAP50) results compared to the decreasing image resolution.

hardware constraints rather than model limitations. While OBB-based detectors consistently preserve their accuracy advantage, their increased computational complexity becomes a critical bottleneck in resource-constrained environments. This trade-off also explains the design choice in Section IV, where the YOLO26x-OBB model is selected for resolution optimization and performance evaluation. Given that OBB-based models provide superior representation of object orientation in SAR imagery, they offer a more reliable basis for analyzing the impact of input resolution across different hardware platforms. Consequently, beyond model design, deployment efficiency and hardware-awareness emerge as key factors in practical SAR detection systems, particularly when real-time operation is required.

IV. Resolution Optimization

In this section, the YOLO26x-OBB model is evaluated under varying computational configurations to analyze the impact of input resolution on performance. The model, trained on high-performance NVIDIA RTX A4000 and

PRO 6000 platforms, is deployed on the Jetson AGX Orin using both 32-bit floating-point (FP32) and 8-bit integer (INT8) precision. To ensure rigorous consistency, all experiments utilized the Ultralytics framework (v8.4.41), isolating hardware-specific capabilities as the primary variable.

The results in Fig. 4 and Fig. 5 reveal a critical non-linear trade-off: while INT8 quantization on the Jetson platform yields over a 300% increase in FPS as resolution scales down, detection accuracy mAP50 remains remarkably stable until a 448-pixel threshold. When downgrading to INT8, information loss can be inherently high and stochastic, leading to distinct fluctuations in mAP values regardless of the specific pixel resolution. From a practical standpoint, a "computational sweet spot" is identified between 512 and 640 pixels. In this range, the system maintains high precision (above 0.40 mAP50) while achieving near real-time edge performance, which is vital for time-critical autonomous sensing tasks. These findings provide a managerial roadmap: the INT8-optimized 512-pixel configuration offers the optimal balance for rapid, large-scale object detection operations on resource-constrained hardware.

V. CONCLUSION AND FUTURE WORK

To sum up, this paper focused on an object detection application using SAR imagery. The paper consisted of three main parts: selecting the target dataset and YOLO network architectures to be used, using those architectures for object detection to assess their success metrics, and performing a resolution-based performance analysis across different hardware configurations to evaluate their combined impact on detection speed and accuracy. Various YOLO versions achieved significant precision on the SAR image dataset using both OBB and HBB configurations. In comparative results, YOLO26x-OBB emerged as the top-performing model. Furthermore, inference tests conducted on the NVIDIA Jetson AGX Orin demonstrated sufficient processing speeds for edge deployment, comparable to those obtained on high-end desktop hardware. In addition to this, a resolution optimization was conducted by testing various input dimensions and quantization settings to determine the optimal computational edge. The results indicate a general trend: lower resolution leads to higher inference speeds but lower detection accuracy; however, this relationship is not strictly linear, as some resolution scales maintain high precision, while others trigger a sharp performance decline. As a result, this work deliberately addressed theoretical approach, training, and hardware implementation phases of the object detection problem for SAR imagery in the aforementioned specifications. In the future, exploring different datasets and other inference hardware could be a new research direction. After completing these tasks, it could be worthwhile to conduct a live implementation as future work.

ACKNOWLEDGMENT

This research was supported by NVIDIA with RTX PRO 6000 Blackwell Max-Q and NVIDIA Jetson AGX Orin Developer Kit 64 GB grants. This work was supported in part by the AMD University Program.

REFERENCES

[1] Y. Zhou, H. Wang, F. Xu, and Y.-Q. Jin, "Polarimetric sar image classification using deep convolutional neural networks," *IEEE Geoscience and Remote Sensing Letters*, vol. 13, no. 12, pp. 1935–1939, 2016.

[2] Z. Huang, X. Yao, Y. Liu, C. O. Dumitru, M. Datcu, and J. Han, "Physically explainable cnn for sar image classification," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 190, pp. 25–37, 2022.

[3] A. W. Santoso, D. Pebrianti, L. Bayuaji, and J. M. Zain, "Performance of various speckle reduction filters on synthetic aperture radar image," in *2015 4th International Conference on Software Engineering and Computer Systems (ICSECS)*. IEEE, 2015, pp. 11–14.

[4] L. Bu, J. Zhang, Z. Zhang, Y. Yang, and M. Deng, "Deep Learning for Integrated Speckle Reduction and Super-Resolution in Multi-Temporal SAR," *Remote Sensing*, vol. 16, no. 1, p. 18, Dec. 2023. [Online]. Available: https://www.mdpi.com/2072-4292/16/1/18

[5] D. Niu, X. Zhou, and F. Zheng, "Improved YOLOv8-OBB Model for Ship Wakes Detection in SAR Imagery," in *2025 4th International Symposium on Robotics, Artificial Intelligence and Information Engineering (RAIIE)*, June 2025, pp. 140–147. [Online]. Available: https://ieeexplore.ieee.org/document/11140083/

[6] C. Chen, N. A. M. Isa, and X. Liu, "A review of convolutional neural network based methods for medical image classification," *Computers in Biology and Medicine*, vol. 185, 2025.

[7] B. Çatalbaş, B. Çatalbaş, and Ö. Morgül, "System identification and model reference adaptive control of bipedal locomotion with neural networks," *Robotics and Autonomous Systems*, vol. 198, p. 105331, 2026.

[8] R. Bogdan, A. Tatu, M. M. Crisan-Vida, M. Popa, and L. Stoicu-Tivadar, "A practical experience on the amazon alexa integration in smart offices," *Sensors*, vol. 21, no. 3, p. 734, 2021.

[9] B. Çatalbaş, B. Çatalbaş, and Ö. Morgül, "A new initialization method for artificial neural networks: Laplacian," in *26th Signal Processing and Communications Applications Conference (SIU)*, 2018, pp. 1–4.

[10] N. P. Arathy Menon, R. S. Iyer, S. Sharma, T. Suresh, S. Simi, G. Gopalakrishnan, P. N. Pournami, and P. B. Jayaraj, "Hailo-8L-Accelerated Embedded Framework for Bladder Carcinoma Detection Using Distilled YOLO Models," *IEEE Embedded Systems Letters*, pp. 1–1, 2026. [Online]. Available: https://ieeexplore.ieee.org/document/11418653/

[11] Y. Ji, D. Zhang, W. Chen, and L. Yang, "Research on SAR Image Target Recognition Using YOLO," in *2025 International Applied Computational Electromagnetics Society Symposium (ACES-China)*, Aug. 2025, pp. 1–3. [Online]. Available: https://ieeexplore.ieee.org/document/11332845

[12] Q. Wang, Y. Feng, and S. Mei, "SAR Image Small Target Detection Algorithm Based on Improved YOLOv8," in *2024 5th China International SAR Symposium (CISS)*, Nov. 2024, pp. 1–6. [Online]. Available: https://ieeexplore.ieee.org/document/11241172/

[13] Y. Wu, Y. Suo, Q. Meng, W. Dai, T. Miao, W. Zhao, Z. Yan, W. Diao, G. Xie, Q. Ke, Y. Zhao, K. Fu, and X. Sun, "FAIR-CSAR: A Benchmark Dataset for Fine-Grained Object Detection and Recognition Based on Single-Look Complex SAR Images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 63, pp. 1–22, 2025. [Online]. Available: https://ieeexplore.ieee.org/document/10806766/