Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Vehicle Detection in SAR Imagery

Semester Project

Jérôme Bonvin

Monday 29th July, 2024

Advisors: Dr. Elías Méndez Dominguez, Prof. Dr. Daniel Razansky

Department of Electrical Engineering, ETH Zürich

List of Abbreviations

CNN	Convolutional Neural Network
IC	Image Contrast
OBB	Oriented Bounding Box
RADAR	Radio Detection and Ranging
SAR	Synthetic Aperture RADAR
YOLO	You Only Look Once

Contents

Vehicle Detection in SAR Imagery

1 Introduction

Synthetic Aperture Radar (SAR) offers a powerful tool for Earth observation, capturing high-resolution images independent of daylight or weather conditions. This capability makes SAR ideal for large-scale monitoring applications, such as disaster response [1], infrastructure management [2, 3], and environmental monitoring. SAR has already proven valuable in real-world scenarios, such as maritime surveillance for illegal activities [3] and aircraft detection for traffic management and military applications [4]. By enabling real-time car detection in SAR imagery, this system has the potential to significantly contribute to traffic management efforts, enhance surveillance capabilities for search, and facilitate rapid response deployments in critical rescue missions. For instance, during natural disasters where traditional infrastructure might be damaged, SAR imagery can provide crucial information about road accessibility and stranded vehicles, enabling faster intervention and improved coordination of rescue efforts.

Manually analyzing vast SAR datasets for objects like cars is a tedious and time-consuming task, rendering it impractical for large-scale applications. To address this challenge, machine learning object detection algorithms can be employed. Traditional methods have achieved success with objects exhibiting distinct characteristics in SAR images. For example, ships are readily identifiable due to their size and contrasting background (water) [2, 3]. Similarly, airplanes can be distinguished by their unique shapes [4]. Car detection in SAR images, however, presents a greater challenge compared to these other objects. Cars often have a simple rectangular shape and appear in diverse environments, leading to significant variability in their appearance within SAR data. Additionally, their relatively small size compared to boats and airplanes makes them more difficult to distinguish from background clutter and other objects. Furthermore, compared to ship and aircraft detection in SAR imagery, car detection has received considerably less research attention [5], highlighting the need for further exploration in this area.

Deep learning has emerged as a powerful approach to address these complexities. In the context of car detection, deep learning models can be trained on vast collections of SAR images containing labeled cars. By processing these images through the network's layers, the model learns to extract features that are characteristic of cars. This ability to learn intricate patterns allows deep learning models to overcome the challenges associated with car detection in SAR data, where traditional methods struggle due to the variability and small size of the target objects. By leveraging the capabilities of deep learning models, we can automate car detection in SAR images, enabling faster and more efficient analysis of large datasets.

This report explores the development and evaluation of a deep learning-based car detection system for SAR imagery. We will detail the chosen model, labeling process, and algorithm performances. Additionally, the report will discuss the system's strengths, limitations, and potential future improvements.

2 Methods

2.1 Object Detection Model

The object detection model employed in this study is the YOLOv8 (You Only Look Once) algorithm, developed by Ultralytics [6]. This convolutional neural network (CNN) is known for its efficient balance between speed and accuracy. Given the extensive existing literature on YOLOv8, we will not delve into its detailed architecture [7].

It is important to acknowledge the inherent advantages of CNNs for object detection tasks. Notably, CNNs exhibit spatial invariance, meaning they can detect objects regardless of their location within the image. Additionally, CNNs share weights between filters, leading to faster training and detection.

In the context of car detection in SAR imagery, YOLOv8 utilizes Oriented Bounding Boxes (OBBs) for localization. OBBs can capture the orientation of a car in addition to its size, providing a more accurate representation of non-axis-aligned objects. This is particularly relevant for SAR images, where the cars' signatures can be oriented in all possible directions.

2.2 Datasets

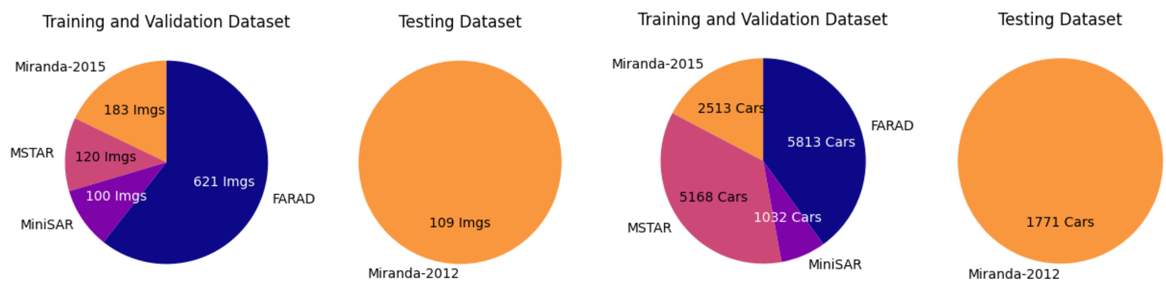
Two primary groups of datasets were used for training and validation of the car detection models. The first dataset consisted of SAR images acquired during flight campaigns in 2015 using the MIRANDA Sensor. These images were manually labeled using the *Image Labeler* [8] app in MATLAB [9]. To enhance the model's ability to detect various cars, we additionally incorporated data from publicly available datasets like MSTAR, MiniSAR, and FARAD [5].

Data	Source	Band	Pixel Spacing
MIRANDA-2012	Frauenhofer FHR	Ka	0.17m
MIRANDA-2015	Frauenhofer FHR	Ka	0.1m-0.8m
FARAD	Sandia National Lab.	X & Ka	0.1m
MiniSAR	Sandia National Lab.	Ku	0.1m
MSTAR	U.S. Air Force	X	0.3m

Table 1. Properties of various datasets

For the final model evaluation, a separate dataset was employed. This dataset comprised SAR images captured during a 2012 MIRANDA flight campaign. Utilizing a distinct dataset for evaluation helps ensure a more objective assessment of the model's performance on unseen data.

To enrich the training dataset and improve model robustness to variations in car appearance, we employed various data augmentation techniques. These techniques included 90-degree rotations, translations, horizontal flipping, and mirroring. This process effectively augmented the training data, allowing the model to learn from a wider range of car orientations and positions within the SAR images. The final dataset comprised 6,437 images for training and 614 images for validation. This translates to a training-validation split of approximately 90.5% and 9.5%, respectively.



(a) Number of Images per Source

(b) Number of Cars per Source

Figure 1. Overview of the Dataset Provenance in Pie-Charts before Data Augmentation

2.3 Algorithm Performances

Label Quality Assessment

Recognizing the critical role of training data quality, we implemented a two-pronged approach to verify car labels in the SAR dataset. First, we leveraged corresponding high-resolution optical data (when available) for visual confirmation, as human identification of cars is easier in optical imagery. However, limitations in the existing optical data (Table 1) necessitated the introduction of the MIRANDA-2022 dataset. This new dataset contains high-resolution optical imagery as well as SAR data. Secondly, we verified car dimensions by comparing their labeled lengths and widths to expected ranges based on a reference source [10] (4.49 m length, 1.84 m width and 1.58m height). Since our labels represent the projected car area, we adjusted these expected dimensions using the sensor's depression angle (23°) to account for potential height information included in the labels. This comprehensive verification process aimed to ensure the accuracy of car labels within the SAR data. Therefore, we can estimate an expected average length of 4.76 m and width of 2.84 m.

Model Detection Analysis

To understand the model's detection patterns and improve its performance, we conducted several analyses. We investigated various factors that might influence the model's ability to detect cars in SAR imagery. We examined the relationship between a car's rotation relative to the flight path (heading) and its detection outcome (correct, missed, false positive). The orientation of a car can drastically change the back-scattering of the signal and consequently change its appearance.

In addition to car rotation, we analyzed the distribution of pixel intensities for cars and their backgrounds, categorized by detection state (correct, missed, false positive). This is evaluating how much the car stands out compared to its background. The background was defined as a 5-pixel wide frame surrounding the car's bounding box.

We further investigated the relationship between image contrast (IC) and detection state. Image contrast was calculated for each car using the following formula [11]:

$$IC(p) = \frac{\sqrt{E\{[I(p) - E\{I(p)\}]^2\}}}{E\{I(p)\}} \tag{1}$$

where I(p) represents the pixel intensities of the car's bounding box and E denotes the expected value (mean) of those intensities.

3 Results

This chapter presents the performance of the best-performing object detection model trained for car detection in SAR imagery. The primary evaluation metric was the correct detection rate. A detection was considered correct if the Intersection Over Union (IOU) between the labeled OBB and the predicted OBB exceeded 0.3, and the confidence score of the predicted OBB was greater than 10%. As shown in Figure 2, the model achieved a detection accuracy of 60.4% for cars in the testing dataset. It also achieved a precision of 81.6% and a recall of 69.9%.

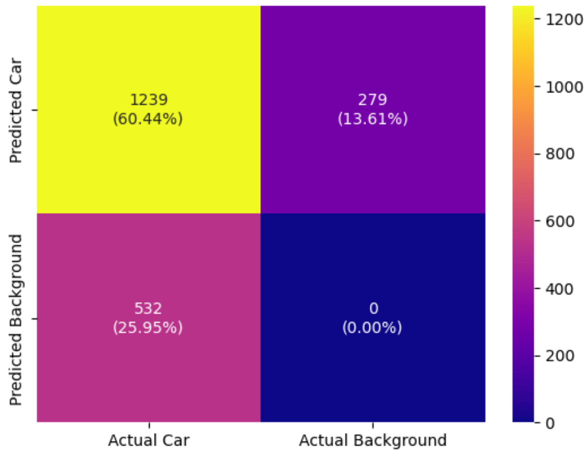


Figure 2. Confusion Matrix of the Testing Data

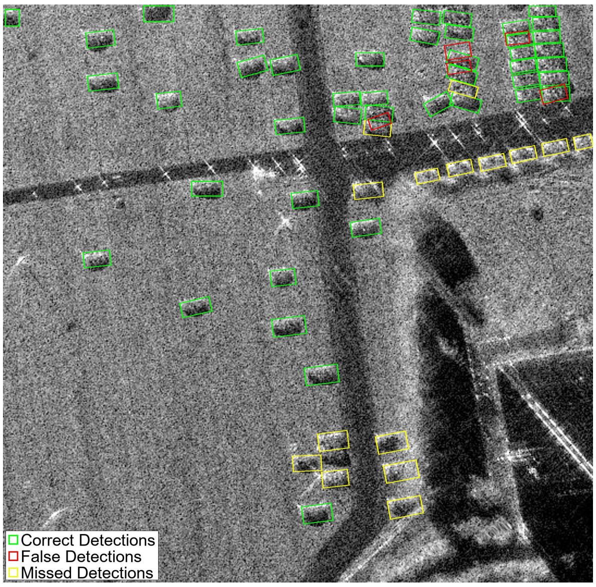
To gain a deeper understanding of the model's performance in various scenarios, we will present a series of case studies. These case studies will analyze the model's detection capabilities under different circumstances, such as parking lot densities, and highlight potential limitations.

First Case Example

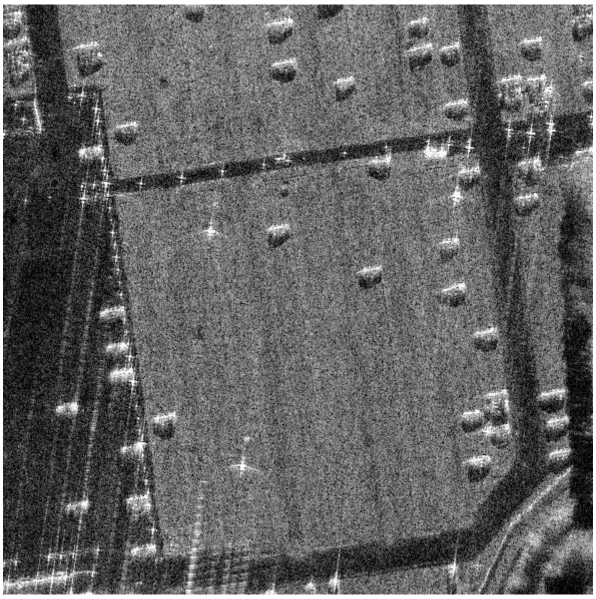
The first case study examines car detection in parking lots with sparse car arrangements. As shown in Figures 3(b) and 3(d), the model successfully detected a significant number of cars. However, most false detections occurred near the image edges. These edge detections can often be disregarded during post-processing steps.

A closer inspection of the bottom portion of Figure 3(d) reveals three instances where the model correctly identified true car detections that were unlabeled.

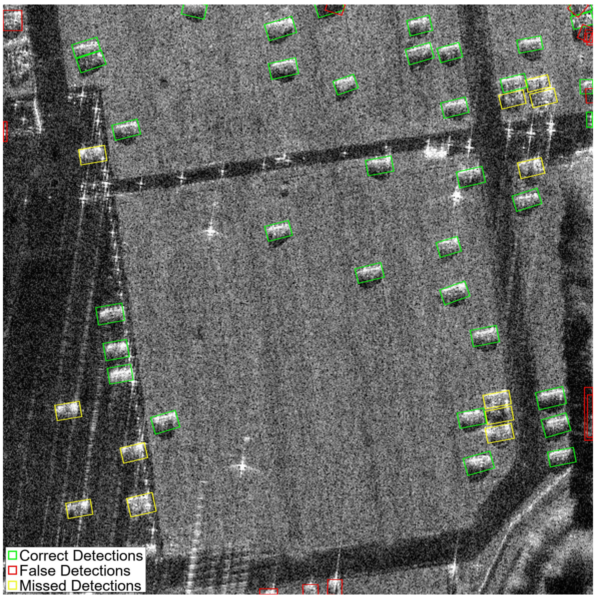
(a) Original Image #1



(b) Image #1 with detections



(c) Original Image #2



(d) Image #2 with detections

Figure 3. Example Case 1 of Car Detection

Second Case Example

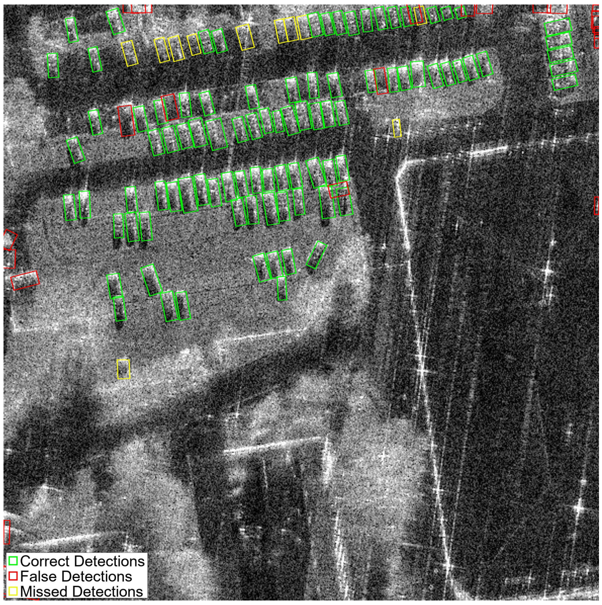
The second case study explores car detection in parking lot scenarios with a high density of parked cars. Similar to the first case study (sparsely parked cars), the model achieved a high success rate in identifying a significant number of vehicles (Figures 4(b) and 4(d)).

However, the nature of missed detections is significantly lower in this scenario. In addition to the edge-based false positives observed in sparse parking lots, missed detections in high-density scenarios tend to occur between correctly identified cars.

(a) Original Image #3

(b) Image #3 with detections

(c) Original Image #4

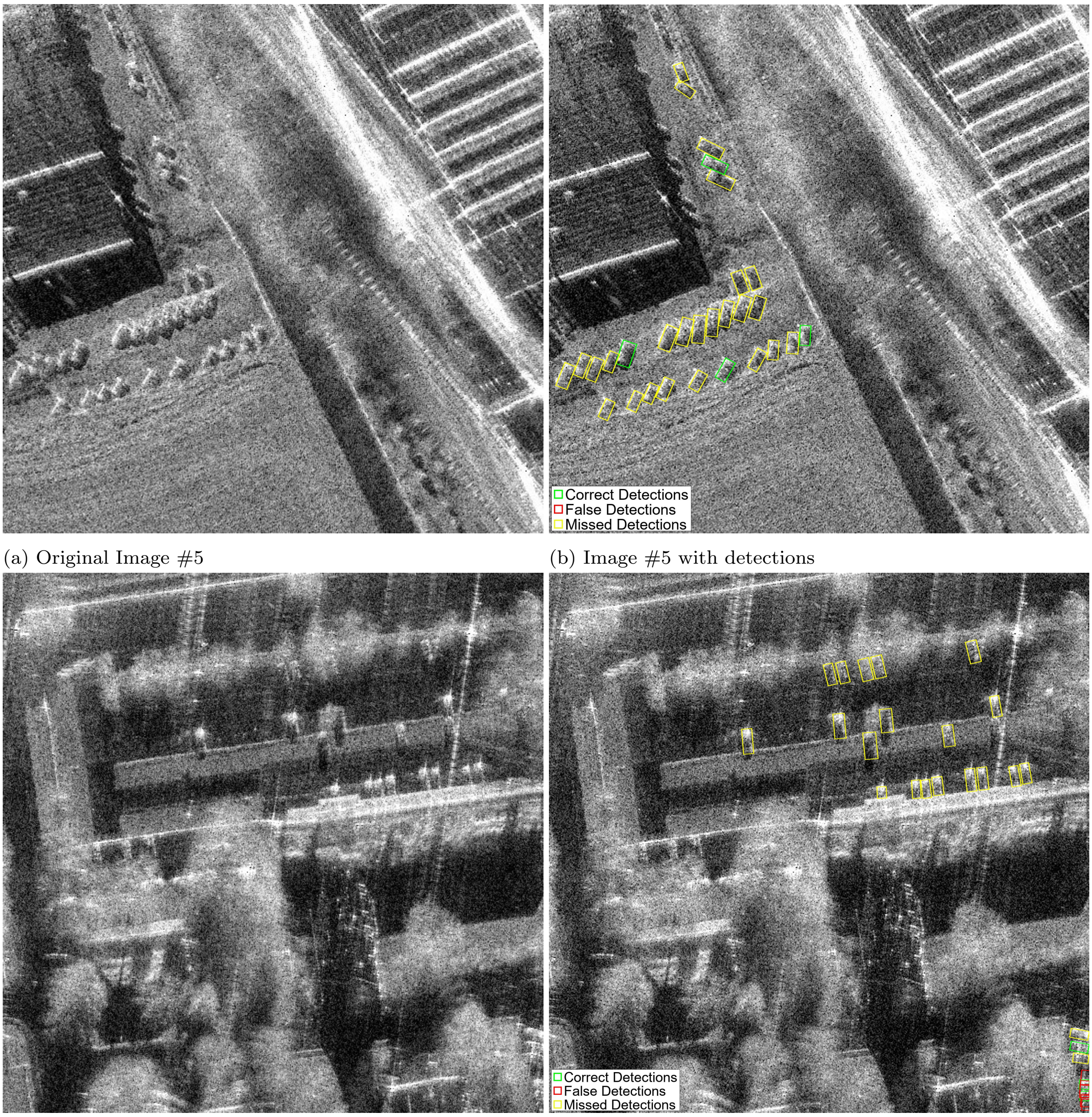


(d) Image #4 with detections

Figure 4. Example Case 2 of Car Detection

Third Case Example

Finally, we will present a challenging example where the model performed poorly. As illustrated in Figure 5, the model struggles to detect a significant number of cars present in the images.



(a) Original Image #5

(b) Image #5 with detections

(c) Original Image #6

(d) Image #6 with detections

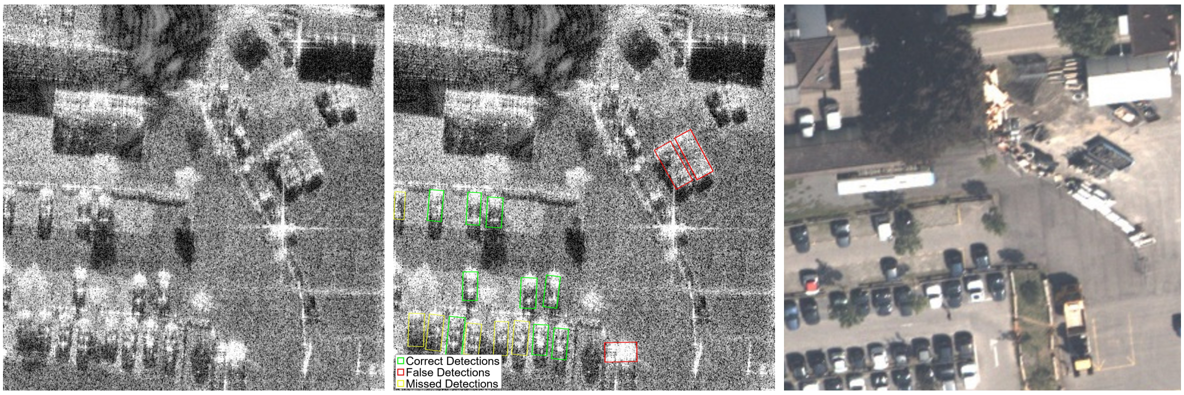
Figure 5. Poor Performance of Model Detection

3.1 Label Quality Assessment

In deep learning, the quality of training data significantly impacts model performance. Therefore, a thorough review of the SAR image labels was conducted using two approaches: comparison with high-resolution optical data and analysis of label dimensions.

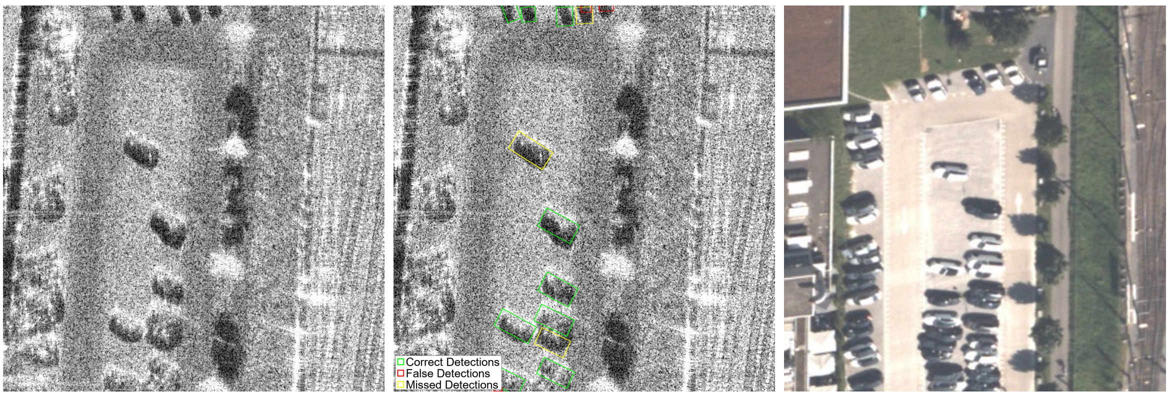
Comparison with Optical Data

A subset of the testing data was compared with corresponding high-resolution optical images (Figure 6 and Figure 7). Twenty image pairs were analyzed. From the 102 cars visually identifiable in the optical data, 67 were correctly labeled in the SAR data. One car was a false positive (incorrectly labeled), and 35 cars were missed due to factors like noise or human error in labeling. This resulted in a labeling accuracy of 65%. This comparison highlights the inherent limitations of SAR data for car detection compared to the richer visual detail provided by optical imagery. For instance, if we look closer in Figure 7, the left part of Figure 7(a) has some indistinguishable cars even though we can clearly see them in Figure 7(c). The optical images represented in Figure 6(c) and Figure 7(c) are left in their original state and are not distorted to fit exactly like their corresponding SAR images.



(a) Original SAR Image (b) Detected SAR Image (c) Optical Image

Figure 6. Example#1 Comparison between an Optical Image and a SAR Image



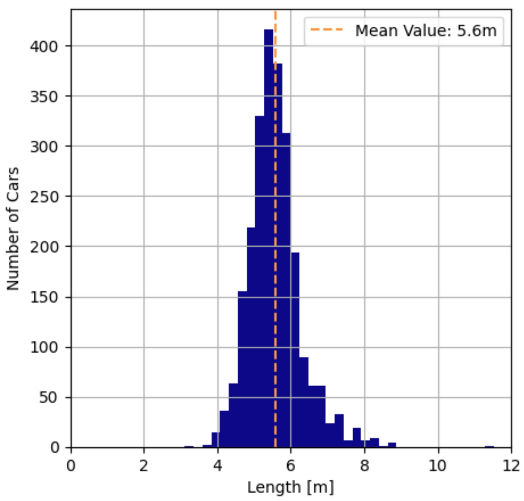
(a) Original SAR Image (b) Detected SAR Image (c) Optical Image

Figure 7. Example#2 Comparison between an Optical Image and a SAR Image

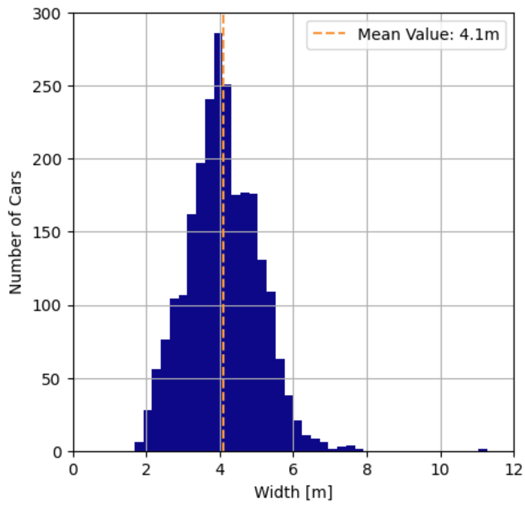
Label Dimension Analysis

To further assess the consistency of our labeled bounding boxes, we analyzed their dimensions (length and width). The histograms of the car's length and width are presented in Figure 8.

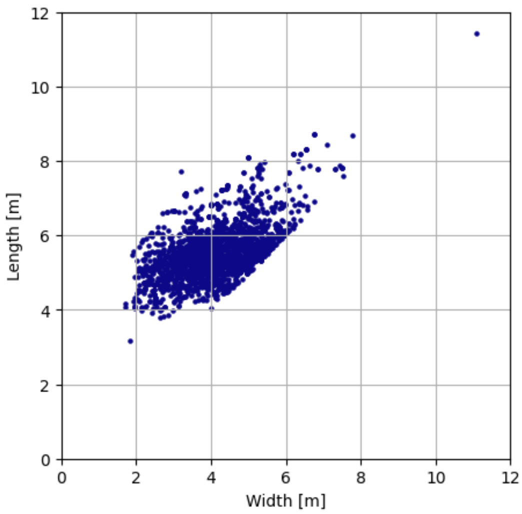
The mean values of length and width are 5.6m and 4.1m respectively. While these values are larger than the predicted values in Subsection 2.3, they still fall short of the actual average length and width observed in our labeled data. This discrepancy will be further explored in the discussion chapter.



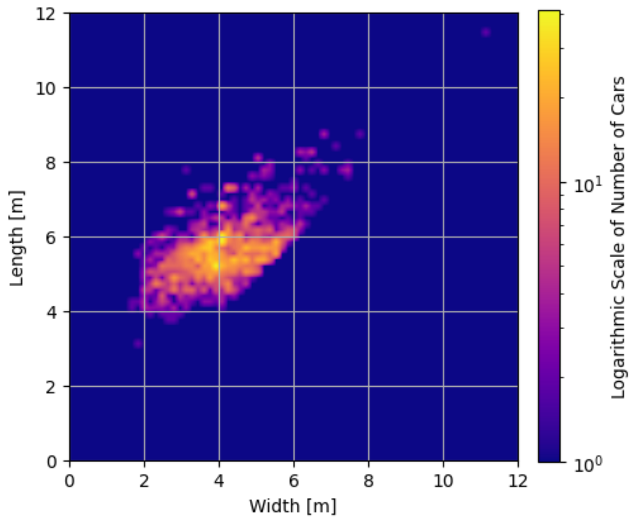
(a) Length Distribution of Labeled Data



(b) Width Distribution of Labeled Data



(c) Width-Length Scatter Plot of Labeled Data



(d) Width-Length Heatmap of Labeled Data

Figure 8. Histograms and scatter plots of size distribution of the labeled vehicles

3.2 Model Detection Analysis

To understand the model's detection behavior and improve its performance, we conducted several analyses. We investigated various factors that might influence the car detection capability in SAR imagery. One key factor we explored was the influence of car orientation on detection probability. The analysis of car orientation is represented by a polar distribution plot (Figure 9). The angle on the polar axis represents the car's heading direction (where it is pointing) or 180° opposite (since differentiating the hood from the trunk in SAR data can be difficult). As evident in the figure, a higher concentration of both labeled and detected cars falls around the -12° and 81° mark.

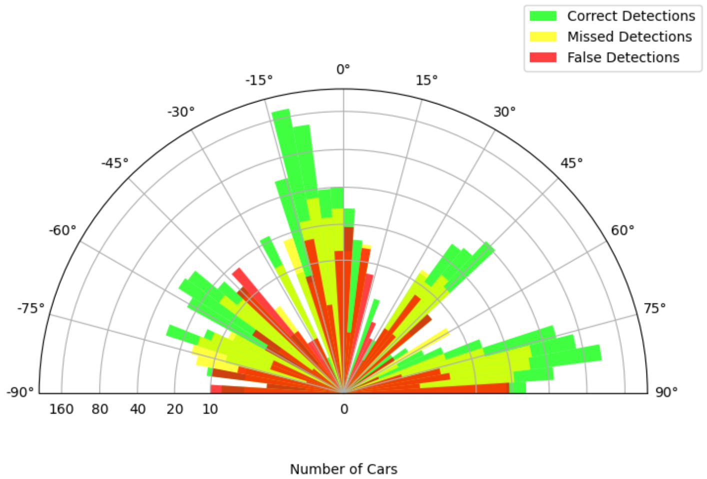
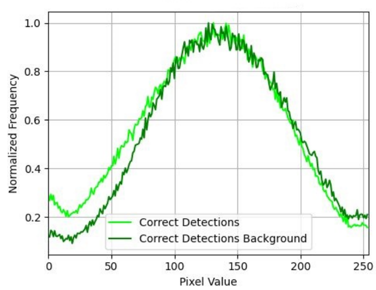
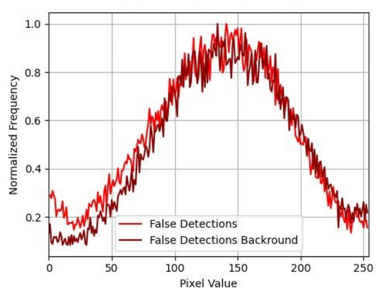


Figure 9. Rotation of car in function of its detection value

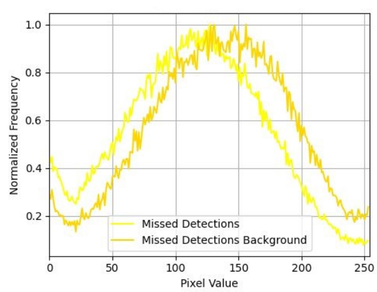
Another analysis revealed distinct patterns in the pixel intensity distributions (Figures 10). For correctly and falsely detected cars (respectively Figure 10(a) and Figure 10(b)), the car pixel intensity distribution closely resembled the background distribution. Conversely, missed detections (Figure 10(c)) exhibited a clear shift in intensity compared to their backgrounds. The missed detections appeared to have a lower average intensity (darker) than their surroundings.



(a) Correct Detections



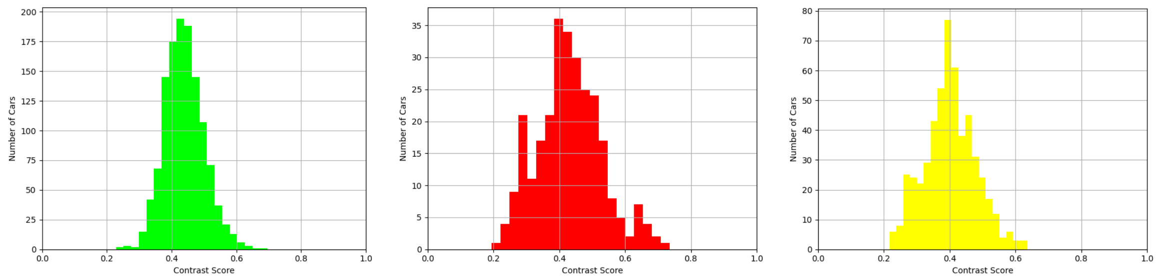
(b) False Detections



(c) Missed Detections

Figure 10. Histogram of Pixel Intensities (0 represents black and 255 white)

Finally, the image contrast was calculated for every car according to Equation 1 and grouped by detection state. The analysis results are presented in Figure 11. As observed, there's a slight difference in image contrast between correctly detections, false detections and missed detections. The correct detections histogram in Figure 11(a) is skewed, the left side of the mean value is more compressed than its right part. On the contrary, the missed detections in Figure 11(c) have a more symmetrical distribution around its mean value.



(a) Correct Detections (b) False Detections (c) Missed Detections

Figure 11. Image Contrast Distribution

4 Discussion

This section discusses the limitations identified in the model's performance and explores potential explanations for these limitations.

Impact of Scene Context

Our analysis revealed a significant influence of scene context on car detection performance. Sparsely populated scenes, like those containing isolated cars in parking lots, exhibited a higher rate of missed detections compared to densely populated areas. We hypothesize that the lack of surrounding cars in sparse scenes hinders the model's ability to detect individual vehicles. In SAR imagery, cars often lack distinct features, and their rectangular shape might be confused with background clutter. However, the presence of surrounding cars provides a structured pattern that the model can leverage for detection. This is analogous to human perception, where recognizing a single object in isolation can be more challenging than identifying the same object within a familiar context with other reference points. This explains why high-density parking lots, offering a richer contextual pattern, yielded a higher correct detection rate compared to sparse parking lot scenes.

Car-to-Car Interference

Another limitation observed is the model's difficulty in handling situations with close proximity between cars with a certain rotation, as shown in Figures 5(a) and 7(a). In such cases, the overlapping features from adjacent cars become indistinguishable, leading to both missed detections and potential false positives. While the analysis in Figure 9 suggests a preference for cars oriented around -12° and 81° relative to the flight path, densely packed scenes often contain cars at various orientations. This highlights the SAR imagery susceptibility to car-to-car occlusion, particularly for non-perpendicular orientations.

High Reflectance from External Influences

Our analysis of pixel intensity distribution (Figure 10(c)) suggests a potential link between the shift of the pixel intensity of the cars compared to its background and missed detections. Missed detections often exhibit darker pixel intensity compared to their surroundings. This could be attributed to the proximity of highly reflective objects like buildings, which can cast shadows or alter the backscattered signal received by the SAR sensor. This hypothesis is further supported by the presence of missed detections near bright buildings in Figure 5(d). Beyond buildings, other natural objects such as trees can introduce external variations that deviate the car's radar signature from the patterns learned by the model.

Label Dimension Discrepancy

The analysis revealed a discrepancy between the expected average car dimensions (length: 4.76 m, width: 2.84 m) and the labeled data (length: 5.6 m, width: 4.1 m). Two main factors contribute to this difference. First, the bounding boxes used for labeling encompass the entire projected area of the car on the ground. This area might include some surrounding background, leading to larger label dimensions compared to the actual car size. Second, the conversion from pixel units to meters relies on the average pixel-to-meter ratio (0.1-0.8 m/pixel). Even slight variations in this conversion (2-3 pixels) can significantly affect the perceived dimensions of the labeled cars. It is important to note that despite these discrepancies, the labeling process maintained consistency throughout the dataset. This consistency ensures a reliable baseline for evaluating the model's performance.

Overall Detection Performance

While a thorough evaluation of human performance using optical data falls outside the scope of this study, we can estimate the model's effectiveness by combining two key metrics. The first metric, detailed in Section 3, is the Detection Rate. This metric represents the percentage of cars correctly identified by the model (60.4%). The second metric, Labeling Accuracy, discussed in Subsection 2.3, indicates the percentage of correctly labeled cars within the dataset (65%). By multiplying these two percentages (60.4% x 65%), we arrive at a combined detection percentage of approximately 39.3%. This value provides a preliminary assessment of how the model's performance might compare to human identification in ideal conditions, assuming perfect labeling. However, it is important to acknowledge that this is a simplified estimation. Further studies could explore more robust methods for comparing model performance to human capabilities.

5 Conclusion & Outlook

This study successfully developed and evaluated a YOLOv8 model for car detection in SAR imagery. By leveraging deep learning techniques, we achieved a preliminary model capable of identifying cars in this challenging data domain. Our analysis explored the strengths and limitations of the model's performance. We identified several factors influencing detection accuracy, including scene context, car-to-car interference, color shifts due to external factors, and discrepancies between labeled and actual car dimensions. Understanding these limitations is crucial for guiding future improvements. This research sets the stage for further development of robust and accurate car detection models in SAR imagery. Here are some promising avenues for future work:

- Incorporating Scene Context: Our findings highlight the impact of scene context on detection performance. Future research could explore multi-stage detection architectures that incorporate scene information during the training process. This could enable the model to leverage contextual cues and improve its ability to identify cars in diverse environments.

- Model Selection and Optimization: Investigating a wider range of deep learning models could identify architectures better suited for car detection in SAR imagery. Evaluating and comparing the performance of different models would be beneficial for selecting the most effective approach for this specific task.

- Labeling Refinement: Refining the labeling process to achieve higher dimensional accuracy could significantly enhance the model's performance. Techniques like leveraging high-resolution optical data (when available) for more precise labeling could be valuable in this regard.

By addressing these limitations and exploring the proposed future directions, researchers can develop increasingly robust and accurate car detection models that can unlock the full potential of SAR imagery for various applications, such as traffic management, disaster response, and infrastructure monitoring.

Bibliography

[1] *Integrating SAR Data for Improved Resilience and Response to Weather-Related Disasters — NASA Applied Sciences.* URL: http://appliedsciences.nasa.gov/what-we-do/projects/integrating-sar-data-improved-resilience-and-response-weather-related-disasters (visited on 07/04/2024).

[2] *Infrastructure Maintenance Management Using SAR Satellite: Monitoring Displacement By Combining Bird's Eye and Worm's Eye View.* NEC. URL: https://www.nec.com/en/global/insights/article/2021030004/index.html (visited on 07/04/2024).

[3] David Hemphill Manager Senior Product. *Automating Vessel Detection in SAR Satellite Imagery.* Capella Space. Aug. 17, 2022. URL: https://www.capellaspace.com/automating-vessel-detection-in-sar-satellite-imagery/ (visited on 07/04/2024).

[4] Jin Xing et al. "A Comparison of Deep Neural Network Architectures in Aircraft Detection from SAR Imagery". In: *Synthetic Aperture Radar (SAR) Data Applications.* Ed. by Maciej Rysz et al. Cham: Springer International Publishing, 2022, pp. 91–111. ISBN: 978-3-031-21225-3. DOI: 10.1007/978-3-031-21225-3_5. URL: https://doi.org/10.1007/978-3-031-21225-3_5 (visited on 07/04/2024).

[5] Xin Lin et al. "SIVED: A SAR Image Dataset for Vehicle Detection Based on Rotatable Bounding Box". In: *Remote Sensing* 15.11 (Jan. 2023). Number: 11 Publisher: Multidisciplinary Digital Publishing Institute, p. 2825. ISSN: 2072-4292. DOI: 10.3390/rs15112825. URL: https://www.mdpi.com/2072-4292/15/11/2825 (visited on 06/17/2024).

[6] *Home - Ultralytics YOLO Docs.* URL: https://docs.ultralytics.com/ (visited on 06/17/2024).

[7] *YOLOv8 Architecture: A Deep Dive into its Architecture - YOLOv8.* Section: Blog. Jan. 15, 2024. URL: https://yolov8.org/yolov8-architecture/ (visited on 07/08/2024).

[8] *Get Started with the Image Labeler - MATLAB & Simulink - MathWorks Switzerland.* URL: https://ch.mathworks.com/help/vision/ug/get-started-with-the-image-labeler.html (visited on 06/25/2024).

[9] *MATLAB.* URL: https://ch.mathworks.com/products/matlab.html (visited on 06/25/2024).

[10] *Les Suisses achètent des voitures toujours plus grandes. Nos chiffres inédits.* rts.ch. Last Modified: 2024-04-17T18:50:02Z Section: Suisse. Apr. 17, 2024. URL: https://www.rts.ch/info/suisse/2024/article/les-suisses-achetent-des-voitures-toujours-plus-grandes-nos-chiffres-inedits-28471442.html (visited on 06/25/2024).

[11] Emiliano Casalini et al. *Studies in the domain of Ground Moving Target Indication (GMTI). Annual report to armasuisse.* ARAMIS-Nr. R-3210/040-25 (Nr. 8003517876). Remote Sensing Laboratories (RSL), University of Zurich, 2018.

ETH

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Declaration of originality

The signed declaration of originality is a component of every written paper or thesis authored during the course of studies. In consultation with the supervisor, one of the following three options must be selected:

- [ ] I confirm that I authored the work in question independently and in my own words, i.e. that no one helped me to author it. Suggestions from the supervisor regarding language and content are excepted. I used no generative artificial intelligence technologies[1].

- [ ] I confirm that I authored the work in question independently and in my own words, i.e. that no one helped me to author it. Suggestions from the supervisor regarding language and content are excepted. I used and cited generative artificial intelligence technologies[2].

- [x] I confirm that I authored the work in question independently and in my own words, i.e. that no one helped me to author it. Suggestions from the supervisor regarding language and content are excepted. I used generative artificial intelligence technologies[3]. In consultation with the supervisor, I did not cite them.

Title of paper or thesis:

Vehicle Detection in SAR Imagery

Authored by:
If the work was compiled in a group, the names of all authors are required.

Last name(s):	First name(s):
Bonvin	Jérôme

With my signature I confirm the following:
- I have adhered to the rules set out in the Citation Guide.
- I have documented all methods, data and processes truthfully and fully.
- I have mentioned all persons who were significant facilitators of the work.

I am aware that the work may be screened electronically for originality.

Place, date	Signature(s)
Zürich, 27.07.2024	

If the work was compiled in a group, the names of all authors are required. Through their signatures they vouch jointly for the entire content of the written work.

[1] E.g. ChatGPT, DALL E 2, Google Bard
[2] E.g. ChatGPT, DALL E 2, Google Bard
[3] E.g. ChatGPT, DALL E 2, Google Bard