

DOMAIN ADAPTIVE OBJECT DETECTION FOR SPACE APPLICATIONS WITH REAL-TIME CONSTRAINTS

Samet Hicsonmez, Abd El Rahman Shabayek, Arunkumar Rathinam, and Djamila Aouada

*University of Luxembourg, Luxembourg, Luxembourg,
(samet.hicsonmez,abdelrahman.shabayek,arunkumar.rathinam,djamila.aouada)@uni.lu*

ABSTRACT

Object detection is essential in space applications targeting Space Domain Awareness and also applications involving relative navigation scenarios. Current deep learning models for Object Detection in space applications are often trained on synthetic data from simulators, however, the model performance drops significantly on real-world data due to the domain gap. However, domain adaptive object detection is an overlooked problem in the community. In this work, we first show the importance of domain adaptation and then explore Supervised Domain Adaptation (SDA) to reduce this gap using minimal labeled real data. We build on a recent semi-supervised adaptation method and tailor it for object detection. Our approach combines domain-invariant feature learning with a CNN-based domain discriminator and invariant risk minimization using a domain-independent regression head. To meet real-time deployment needs, we test our method on a lightweight Single Shot Multibox Detector (SSD) with MobileNet backbone and on the more advanced Fully Convolutional One-Stage object detector (FCOS) with ResNet-50 backbone. We evaluated on two space datasets, SPEED+ and SPARK. The results show up to 20-point improvements in average precision (AP) with just 250 labeled real images.

Key words: Domain adaptation; Object Detection.

1. INTRODUCTION

Vision-based object detection plays a central role in a wide range of space applications (1), including Space Domain Awareness missions and also missions that involve relative navigation or autonomous rendezvous such as in-orbit servicing and active debris removal. Reliable detection of spacecraft or orbital objects is critical for enabling safe and autonomous operations in space. Modern object detection models, particularly those based on deep learning, require large amounts of annotated data to achieve high accuracy. However, collecting labeled data in the space domain is extremely difficult due to constraints such as limited accessibility, high operational costs, and

the complexity of manual annotation under space conditions.

To mitigate the challenge of data scarcity, synthetic datasets (2; 3) generated through simulation environments (4) have become the primary training source. Synthetic datasets are scalable and easy to generate and label. However, models trained solely on synthetic data often perform poorly when deployed in real-world scenarios (5). This performance drop is mainly attributed to the domain gap between the simulated training data (source) and the real orbital imagery (target), which includes differences in lighting, background, object textures, and noise characteristics.

Domain Adaptation (DA) (6) has emerged as a promising solution to bridge this gap. Unsupervised Domain Adaptation (UDA) methods (7) aim to align source and target domains without requiring labeled target images. Although they offer improvements over naive transfer, their ability to fully close the domain gap remains limited (8). In contrast, semi-supervised domain adaptation methods (9; 10), which rely on a small set of labeled real images, have shown greater potential. Even limited real supervision, on the order of a few hundred annotated samples, can significantly enhance detection performance when properly integrated into the training process.

In semi-supervised domain adaptation (SDA), it is typically assumed that the target domain provides a small set of labeled samples together with a large collection of unlabeled data. While this assumption is reasonable in many application areas, it rarely holds in the space domain. Real-world space datasets usually contain only a few hundred images, making large-scale unlabeled target data unavailable and thus challenging the standard SDA setting.

Although object detection is a critical component in many space applications, the use of domain adaptation techniques for this task has received limited attention, until recently (11; 12). This may be due to the common assumption that the problem is relatively simple, as scenarios often involve a single target object against a relatively static or uniform background. However, both recent studies and our findings demonstrate that this assumption is considerably more challenging than commonly assumed.

In this paper, we build upon a recent semi-supervised domain adaptation approach and adapt it to become a Supervised Domain Adaptive (SDA) object detection. We employ a simple CNN-based domain discriminator to encourage feature alignment between domains and introduce a domain-independent regression head to enforce consistency in object localization. This setup allows the model to generalize better across domains, even with limited real data.

To evaluate the practicality of our approach for real-world deployment, we experiment with both lightweight and modern object detectors. Specifically, we use SSD (13) with a MobileNetv2 (14) backbone for real-time performance and FCOS (15) with ResNet-50 (16) for architectural generality. Experiments are conducted on two publicly available space datasets, SPEED+ (2) and SPARK (3). Our results show that the proposed approach significantly closes the domain gap and improves performance.

In Section 2, we review existing domain adaptation methods in both real-world and space-related contexts. Section 3 provides a detailed description of our proposed approach. Experimental results are presented in Section 4, followed by concluding remarks in Section 5.

2. RELATED WORK

In this section, we first explain the main approaches for domain adaptation in natural images and then describe the methods targeted for space applications in the context of *object detection*.

2.1. Domain Adaptation for Natural Images

Unsupervised approaches (17; 18; 19; 20; 21; 22; 23) currently dominate research in Domain Adaptive Object Detection (DAOD). Only a few methods explore semi-supervised DA (9; 24), and there is virtually no work focused on SDA training. However, unsupervised or semi-supervised methods can often be adapted to supervised settings, potentially yielding better performance in the target domain. Therefore, the following review focuses on UDA methods.

Following the taxonomy presented in a recent survey on UDA for object detection (25), the dominant DAOD approaches can be broadly grouped into three categories: 1) adversarial feature learning (18; 26; 27; 28; 29; 30; 31), 2) image-to-image translation (32; 33; 34; 35; 36; 37; 38), and 3) mean teacher training (19; 22; 39; 40; 17).

Adversarial feature learning aims to match the features coming from the source and target domains with the help of a discriminator. In this setting, usually a single object detector model (e.g. Faster-RCNN (41), SSD (13) or FCOS (15)) is trained using images from both domains,

and depending on the method, one or many domain discriminators are employed with GRLs.

The main advantage of this approach is that it improves the baseline source-only training with a very small number of parameters, i.e. shallow domain classifiers. However, since the single detection network is trained using images from both the source and target domains, the network cannot capture specific features of the target, which limits overall performance.

Image to image translation aims to close the gap between the source and target domains by using an unpaired image-to-image translation (im2im) model such as CycleGAN (42), UNIT (43), MUNIT (44) or CUT (45). Generative Adversarial Networks (GAN) (46)-based im2im models learn a mapping from a source domain X to a target domain Y . Paired im2im models have ground truth pairs in both domains such as an image and its edge map, and unpaired im2im models have no ground truth pairs and employ unsupervised training. After translation, an intermediate image representation that resembles the target domain more than the initial source domain is obtained, and since the method works at the pixel level, it is possible to use the source annotations on the intermediate dataset.

The image translation part of this architecture is trained offline, which adds little or no overhead to the main object detection method training, and the target translated images show better resemblance than using special image augmentations, for example, contrast adjustment for day to night domain adaptation. In addition, im2im is widely used in the training of other DAOD approaches.

Mean teacher (MT) based training is proposed to overcome the limitations of training a single method (i.e. biasing towards the source dataset) using both source and target datasets. In this setting, a student model is trained using labeled source data with supervised training, the teacher model is trained unsupervised using only the target dataset, and the student model updates the weights of the teacher model by exponential moving average. Finally, inconsistency between the detections of the student and teacher models is penalized to support the learning. The weights of the teacher model are used for inference on the target dataset.

MT based methods show superior performance compared to the above two approaches. Most, if not all, of the MT methods employ an offline trained im2im method to augment the source and target datasets. Moreover, the MT's teacher network is trained using only target images, making the training more stable. Adversarial learning is also included in some recent MT approaches (39; 17).

2.2. Domain Adaptation for Space Images

Domain adaptation (DA) for space imagery has gained attention because of the difficulty of collecting real space

data. However, existing DA research has focused primarily on pose estimation tasks (5; 47; 48; 49; 50), while the object detection task remains largely overlooked. Recently, fANOVA (11) explored the impact of data augmentations in the UDA setting. Their framework identifies optimal combinations of image augmentations to significantly enhance source-only training performance. Another recent work (12) incorporates DA using image-to-image translation. The authors train a translation model to generate target-style images, which are then used to improve the generalization of the model across domains.

In contrast to these works, we propose a DA framework in a more realistic SDA setting, where a limited number of labeled target-domain images are available. Our approach explicitly leverages this limited supervision to improve object detection performance under domain shift.

3. METHOD

The proposed approach is built on top of a recent semi-supervised domain adaptation framework titled LIRR (9) (Learning Invariant Representations and Risks). LIRR proposes to jointly learn invariant representations and risks at the same time to better mitigate the accuracy discrepancy across domains. LIRR presents a generic framework which could be used in both classification and regression tasks. In our work, we utilize it for a regression task, i.e. object detection, in a fully-supervised setting considering all the target data is labeled.

Invariant representation ensures that the features of the images of the source and target domain extracted using a feature extractor backbone $g(\cdot)$ are not separable and is formulated as follows:

$$\mathcal{L}_{\text{rep}}(g, \mathcal{C}) = \mathbb{E}_{X \sim \mathcal{D}_S(X)} [\log(\mathcal{C}(g(X)))] + \mathbb{E}_{X \sim \mathcal{D}_T(X)} [\log(1 - \mathcal{C}(g(X)))] \quad (1)$$

where \mathcal{C} is the domain classifier.

Invariant risk aims to minimize domain-invariant task losses, i.e. cross-entropy losses of domain-invariant predictor f_i and domain-dependent predictor f_d . Considering that we have a domain-invariant loss \mathcal{L}_i formulated as follows:

$$\min_{g, f_i} \mathcal{L}_i = \mathbb{E}_{(x, y) \sim \mathcal{D}_S, \mathcal{D}_T} [L(y, f_i(g(x)))], \quad (2)$$

where (x, y) pairs denote input data and its ground truth label, and L denotes task-dependent loss. In our case, L is the sum of the classification (cross entropy) and localization (smooth L_1) losses. The domain-dependent loss \mathcal{L}_d is formulated as follows:

$$\min_{g, f_d} \mathcal{L}_d = \mathbb{E}_{d \sim \mathcal{D}} \mathbb{E}_{(x, y) \sim \mathcal{D}_d} [L(y, f_d(g(x), d))], \quad (3)$$

Table 1: Details of SPARK and SPEED+ datasets used in the experiments.

Dataset	Source		Target	
	Train	Val	Train	Val
SPARK (3)	22500	7500	250/500	1600
SPEED+ (2) Sunlamp	47966	11994	250/500	2200
SPEED+ (2) Lightbox	47966	11994	250/500	6200

where the additional d signifies the domain from which the data is sampled.

Thus, the invariant risk loss is formulated as follows:

$$\min_{g, f_i} \max_{f_d} \mathcal{L}_{\text{risk}} = \mathcal{L}_i + \lambda_{\text{risk}}(\mathcal{L}_i - \mathcal{L}_d). \quad (4)$$

The final loss function used to train the methods could be described as:

$$\min_{g, f_i} \max_{\mathcal{C}, f_d} \mathcal{L}_{\text{LIRR}}(g, f_i, f_d, \mathcal{C}) = \mathcal{L}_{\text{risk}}(g, f_i, f_d) + \lambda_{\text{rep}} \mathcal{L}_{\text{rep}}(g, \mathcal{C}). \quad (5)$$

where λ_{rep} is set to 0.1, and λ_{risk} is set to 1.0 in our experiments.

Note that all the equations are directly taken from LIRR (9) and presented here for completeness.

4. EXPERIMENTS

4.1. Datasets

SPEED+ (2) is a well-known dataset in the spacecraft pose estimation domain that features synthetic images generated from a simulator along with two distinct hardware-in-the-loop (HIL) image domains, *sunlamp* and *lightbox*. The *lightbox* domain contains 6,740 images of the mockup of the Tango spacecraft, illuminated using albedo light boxes to simulate diffuse lighting conditions typical of Earth orbit. We split this dataset into train and test sets. The test set contains 6200 images, and the remaining images were sampled into two train subsets with 250 and 500 images from the remaining images. The *sunlamp* domain includes 2,791 images of the same mockup, captured under a metal halide arc lamp to replicate direct sunlight. These domain-specific lighting conditions make SPEED+ a valuable benchmark for studying domain adaptation in space-based vision tasks. Similarly, we split the images in the sunlamp domain into a test set with 2200 images and two training sets with 250 and 500 images. Note that in both domains, the larger train set completely overlaps with the smaller training split. Additionally, this dataset contains close to 60,000 synthetic images.

	OD	Method	Ims	AP	AP_{50}	AP_{75}	Ims	AP	AP_{50}	AP_{75}
Sunlamp	SSD	Source Only	-	44.4	85.2	42.1	-	44.4	85.2	42.1
		Oracle	250	53.0	98.0	50.8	500	63.4	98.9	76.5
		SDA	250	79.4	99.0	97.6	500	82.0	99.0	97.7
	FCOS	Source Only	-	52.1	95.1	52.6	-	52.1	95.1	52.6
		Oracle	250	72.8	98.9	90.8	500	77.1	99.0	94.8
		SDA	250	52.1	98.9	81.7	500	66.3	98.9	81.7
Lightbox	SSD	Source Only	-	26.2	81.1	35.8	-	26.2	81.1	35.8
		Oracle	250	46.8	95.8	37.3	500	60.5	98.6	69.6
		SDA	250	83.4	99.0	96.3	500	85.7	99.0	97.6
	FCOS	Source Only	-	11.2	36.6	4.2	-	11.2	36.6	4.2
		Oracle	250	56.8	94.8	62.0	500	65.8	97.2	76.0
		SDA	250	54.6	94.6	60.7	500	61.4	96.3	70.1

Table 2: Results on the SPEED+ dataset for the sunlamp and lightbox subsets.

SPARK (3) dataset is collected in the SnT Zero-G Lab (4) at the Interdisciplinary Centre for Security, Reliability, and Trust (SnT), University of Luxembourg. The source domain comprises 100 trajectory groups, each containing 300 labeled synthetic images. The target domain includes four real trajectory sequences, RT001, RT002, RT003, and RT004, consisting of 681, 424, 678, and 340 time-sequential images, respectively. We created non-overlapping train and test splits for SPARK as well.

The statistics of the dataset in terms of the number of images in each split are presented in Table 1.

4.2. Setup

In order to evaluate the performance of the proposed method, we devise a three-stage evaluation pipeline, see Table 3 and 2. First, we quantify the domain gap by training a model using only the source (synthetic) training dataset and evaluate it on the target (lab) test set. This experiment is labeled as *source only*. Second, we train a model using only the target training set to measure performance without domain adaptation, and it is labeled as *Oracle*. Finally, we train our proposed model using both the source and target training sets to see the effect of domain adaptation. Note that all models are evaluated on the newly created target test set mentioned in the above section.

To evaluate our approach under limited supervision, we select subsets of 250 and 500 labeled real images from both datasets to train the Oracle and SDA models. We conduct experiments using two object detection frameworks: SSD (13) with a MobileNetV2 (14) backbone, chosen for its suitability in resource-constrained deployment scenarios, and FCOS (15) with a ResNet-50 (16) backbone, used for general purpose performance evaluation. We adapt our approach to the official implementation of each method.

The images are resized to 300×300 pixels for SSD train-

ing, and for the FCOS, the images are resized to have a shorter side of 800 pixels. During inference, the images are resized similarly to the training, and we do not employ any data augmentation.

4.3. Metrics

We evaluated object detection performance using standard COCO-style (51) metrics: Average Precision (AP), AP_{50} , and AP_{75} . The AP score measures the mean average precision over multiple IoU thresholds (ranging from 0.50 to 0.95 in 0.05 increments), providing a comprehensive assessment of overall detection quality. AP_{50} reports average precision at a fixed IoU threshold of 0.50 and reflects the model’s ability to locate objects coarsely. In contrast, AP_{75} evaluates the performance at a stricter IoU threshold of 0.75, highlighting the precise localization capability. These metrics allow us to assess both robustness and localization accuracy under domain-shift conditions.

4.4. Quantitative Results

We present SPEED+ (2) and SPARK (3) quantitative results in Table 2 and Table 3, respectively. In each table, we show the metrics for the 250 and 500 images, and for SSD (13) and FCOS (15) object detectors.

First, the *source only* baseline shows the severity of the domain gap in both datasets, as the performance using only the labeled source dataset achieves dramatically low performance on the target data.

On the SPEED+ dataset (Table 2), applying our SDA strategy to the SSD detector yields a substantial gain, exceeding 20 points in AP, over the *Oracle* baseline. This highlights the critical role of domain adaptation, particularly for shallow detectors. In contrast, deeper architectures such as FCOS exhibit degraded performance when

	OD	Method	Ims	AP	AP_{50}	AP_{75}	Ims	AP	AP_{50}	AP_{75}
Spark	SSD	Source Only	-	27.9	88.1	7.1	-	27.9	88.1	7.1
		Oracle	250	80.9	98.9	98.9	500	86.2	98.9	98.9
		SDA	250	86.6	98.9	98.9	500	87.2	98.9	98.9
	FCOS	Source Only	-	27.2	85.6	6.5	-	27.2	85.6	6.5
		Oracle	250	95.9	99.0	99.0	500	97.9	99.0	99.0
		SDA	250	96.6	99.0	99.0	500	98.0	99.0	99.0

Table 3: Results on the SPARK dataset.

trained with source-domain data, with their *Oracle* results surpassing SDA in both the Sunlamp and Lightbox splits.

A further noteworthy finding is that SSD with SDA outperforms the *Oracle* performance of FCOS on both splits by more than 5 AP points. This underscores the importance of domain adaptation when using lightweight detectors, likely because SDA mitigates overfitting in models with fewer parameters.

On the SPARK dataset (Table 3), we observe performance gains for both detectors under low-data regimes. However, as the number of training images increases, the gap between the *Oracle* and SDA models diminishes, and their performances converge to a comparable level.

4.5. Qualitative Results

We present visual results of *ORACLE* and our method in Figure 1 for all datasets. Our supervised domain adaptive object detection method significantly improves the *ORACLE* performance in all datasets by getting affected from the background as in the first two rows and external light sources as in the fourth row.

5. CONCLUSION

We presented a supervised domain adaptation approach for spacecraft object detection using limited real-world annotations. By combining domain-invariant feature learning and invariant risk minimization, our method improves performance under domain shift with minimal supervision. Experiments on two space datasets and two detector architectures show up to 20-point gains in average precision with only 200 labeled real images. These findings demonstrate the practical value of the method for real-world space applications with constrained annotation resources.

ACKNOWLEDGMENTS

This research has been conducted in the context of the DIOSSA project, supported by the European

Space Agency (ESA) under contract no. [4000 1 4 4941241 NL/KK/adu]. The simulations were performed on the Luxembourg National Supercomputer MeluXina. The authors gratefully acknowledge the Lux-Provide teams for their expert support and the partners of the consortium for their contributions.

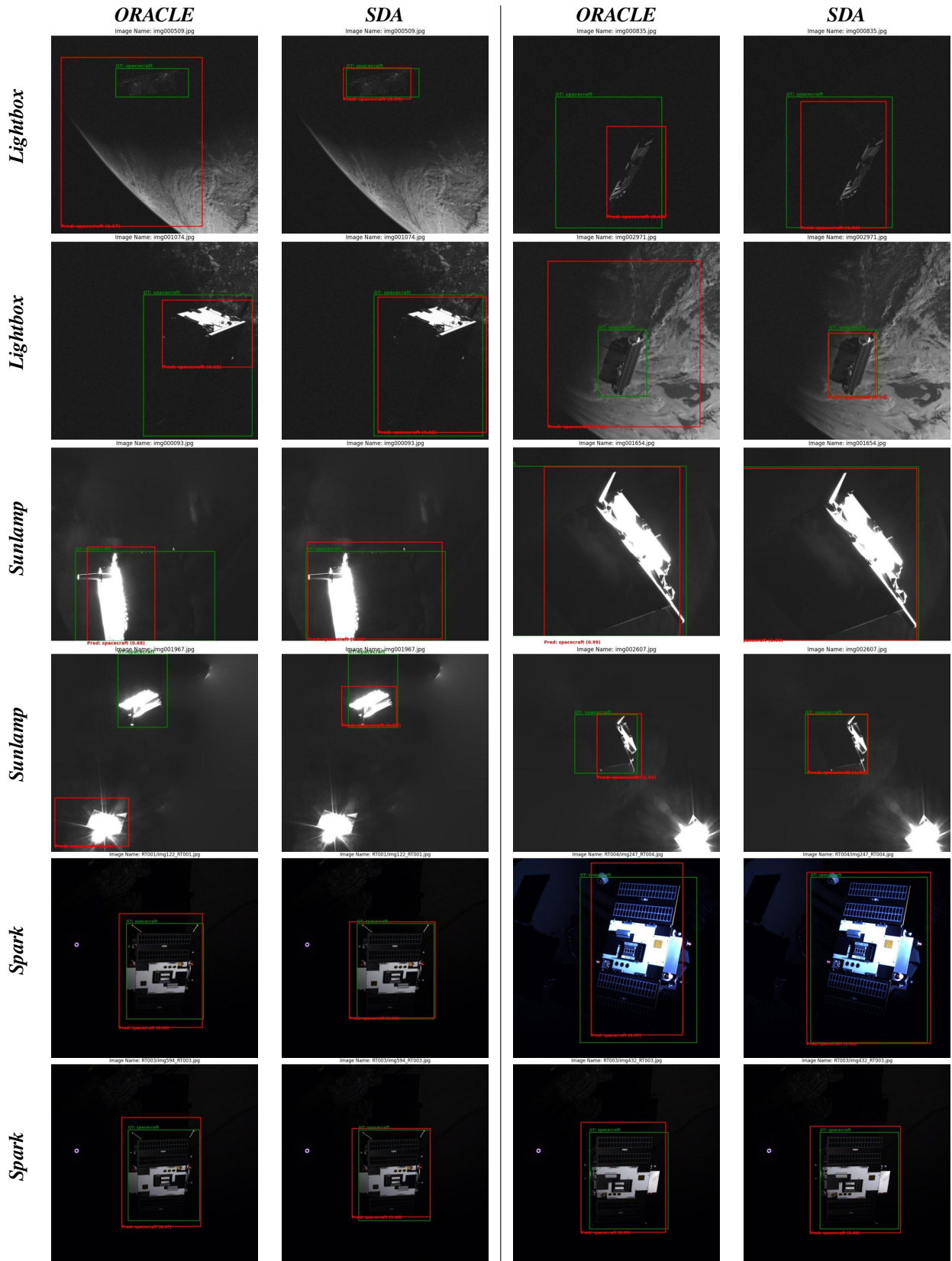


Figure 1: Visual comparison of *ORACLE* and our method on the three datasets used in experiments. Green boxes denote the ground truth and red boxes shows the predictions. Our supervised domain adaptation significantly improves the prediction performance across various cases such as earth background, another light source and close up imagery.

REFERENCES

- [1] H. A. Dung, B. Chen, and T.-J. Chin, "A spacecraft dataset for detection, segmentation and parts recognition," in *CVPR*, pp. 2012–2019, 2021.
- [2] T. H. Park, M. Märtens, G. Lecuyer, D. Izzo, and S. D'Amico, "Speed+: Next-generation dataset for spacecraft pose estimation across domain gap," in *2022 IEEE Aerospace Conference (AERO)*, pp. 1–15, 2022.
- [3] "Spark 2024: Datasets for spacecraft semantic segmentation and spacecraft trajectory estimation."
- [4] M. Olivares-Mendez, M. R. Makhdoomi, B. C. Yalçın, Z. Bokal, V. Muralidharan, M. O. Del Castillo, V. Gaudilliere, L. Pauly, O. Borgue, M. Alandihallaj, *et al.*, "Zero-g lab: A multi-purpose facility for emulating space operations," *Journal of Space Safety Engineering*, vol. 10, no. 4, pp. 509–521, 2023.
- [5] S. Hashimoto, D. Hirano, H. Higuchi, K. Fujimoto, M. Ito, S. Iizuka, Y. Sugimoto, R. Proffitt, Y. Kobayashi, and N. Ishihama, "Domain adaptation for 6-dof pose estimation: Deep learning with uncertainty," in *2022 IEEE Aerospace Conference (AERO)*, pp. 1–12, 2022.
- [6] S. Ben-David, J. Blitzer, K. Crammer, and F. Pereira, "Analysis of representations for domain adaptation," *NeurIPS*, vol. 19, 2006.
- [7] G. Wilson and D. J. Cook, "A survey of unsupervised deep domain adaptation," *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 11, no. 5, pp. 1–46, 2020.
- [8] J. Kay, T. Haucke, S. Stathatos, S. Deng, E. Young, P. Perona, S. Beery, and G. Van Horn, "Align and distill: Unifying and improving domain adaptive object detection," *Transactions on Machine Learning Research*, 2025.
- [9] B. Li, Y. Wang, S. Zhang, D. Li, K. Keutzer, T. Darrell, and H. Zhao, "Learning invariant representations and risks for semi-supervised domain adaptation," in *CVPR*, 2021.
- [10] H. Zhou, F. Jiang, and H. Lu, "Ssda-yolo: Semi-supervised domain adaptive yolo for cross-domain object detection," *Computer Vision and Image Understanding*, vol. 229, p. 103649, 2023.
- [11] M. Ulmer, L. Klüpfel, M. Durner, and R. Triebel, "How important are data augmentations to close the domain gap for object detection in orbit?," in *2025 IEEE Aerospace Conference*, pp. 1–12, IEEE, 2025.
- [12] H. Zhang, Y. Zhang, Q. Feng, and K. Zhang, "A semi-supervised object detection method for close range detection of spacecraft and space debris," *International Journal of Aeronautical and Space Sciences*, vol. 26, no. 2, pp. 773–784, 2025.
- [13] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, "Ssd: Single shot multibox detector," in *ECCV*, pp. 21–37, 2016.
- [14] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "Mobilenetv2: Inverted residuals and linear bottlenecks," in *CVPR*, pp. 4510–4520, 2018.
- [15] Z. Tian, C. Shen, H. Chen, and T. He, "Fcos: Fully convolutional one-stage object detection," in *ICCV*, 2019.
- [16] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *CVPR*, pp. 770–778, 2016.
- [17] M. Kennerley, J.-G. Wang, B. Veeravalli, and R. T. Tan, "Cat: Exploiting inter-class dynamics for domain adaptive object detection," in *CVPR*, 2024.
- [18] Y. Chen, W. Li, C. Sakaridis, D. Dai, and L. Van Gool, "Domain adaptive faster r-cnn for object detection in the wild," in *CVPR*, pp. 3339–3348, 2018.
- [19] Q. Cai, Y. Pan, C.-W. Ngo, X. Tian, L. Duan, and T. Yao, "Exploring object relation in mean teacher for cross-domain detection," in *CVPR*, pp. 11457–11466, 2019.
- [20] M. Chen, W. Chen, S. Yang, J. Song, X. Wang, L. Zhang, Y. Yan, D. Qi, Y. Zhuang, D. Xie, *et al.*, "Learning domain adaptive object detection with probabilistic teacher," pp. 3040–3055, 2022.
- [21] S. Cao, D. Joshi, L.-Y. Gui, and Y.-X. Wang, "Contrastive mean teacher for domain adaptive object detectors," in *CVPR*, 2023.
- [22] J. Deng, W. Li, Y. Chen, and L. Duan, "Unbiased mean teacher for cross-domain object detection," in *CVPR*, pp. 4091–4101, 2021.
- [23] O. Krishna, H. Ohashi, and S. Sinha, "Mila: Memory-based instance-level adaptation for cross-domain object detection," in *BMVC*, 2023.
- [24] Z. Yan, Y. Wu, G. Li, Y. Qin, X. Han, and S. Cui, "Multi-level consistency learning for semi-supervised domain adaptation," in *IJCAI*, 2022.
- [25] P. Oza, V. A. Sindagi, V. V. Sharmini, and V. M. Patel, "Unsupervised domain adaptation of object detectors: A survey," *IEEE TPAMI*, 2023.
- [26] K. Saito, Y. Ushiku, T. Harada, and K. Saenko, "Strong-weak distribution alignment for adaptive object detection," in *CVPR*, pp. 6956–6965, 2019.
- [27] V. A. Sindagi, P. Oza, R. Yasarla, and V. M. Patel, "Prior-based domain adaptive object detection for hazy and rainy conditions," in *ECCV*, pp. 763–780, Springer, 2020.
- [28] C.-C. Hsu, Y.-H. Tsai, Y.-Y. Lin, and M.-H. Yang, "Every pixel matters: Center-aware feature alignment for domain adaptive object detector," in *ECCV*, pp. 733–748, Springer, 2020.
- [29] Y. Zheng, D. Huang, S. Liu, and Y. Wang, "Cross-domain object detection through coarse-to-fine feature adaptation," in *CVPR*, 2020.
- [30] V. Vs, V. Gupta, P. Oza, V. A. Sindagi, and V. M. Patel, "Mega-cda: Memory guided attention for category-aware unsupervised domain adaptive object detection," in *CVPR*, pp. 4516–4526, 2021.
- [31] L. Zhao and L. Wang, "Task-specific inconsistency alignment for domain adaptive object detection," in *CVPR*, pp. 14217–14226, 2022.
- [32] H.-K. Hsu, C.-H. Yao, Y.-H. Tsai, W.-C. Hung, H.-Y. Tseng, M. Singh, and M.-H. Yang, "Progressive domain adaptation for object detection," pp. 749–757, 2020.
- [33] D. Zhang, J. Li, L. Xiong, L. Lin, M. Ye, and S. Yang, "Cycle-consistent domain adaptive faster rcnn," *IEEE Access*, vol. 7, pp. 123903–123911, 2019.

- [34] A. L. Rodriguez and K. Mikolajczyk, "Domain adaptation for object detection via style consistency," in *BMVC*, 2019.
- [35] C. Chen, Z. Zheng, X. Ding, Y. Huang, and Q. Dou, "Harmonizing transferability and discriminability for adapting object detectors," in *CVPR*, pp. 8869–8878, 2020.
- [36] Z. Shen, M. Huang, J. Shi, Z. Liu, H. Maheshwari, Y. Zheng, X. Xue, M. Savvides, and T. S. Huang, "Ctdt: A large-scale cross-domain benchmark for instance-level image-to-image translation and domain adaptive object detection," *IJCV*, vol. 129, no. 3, p. 761–780, 2021.
- [37] C.-T. Lin, J.-L. Kew, C. S. Chan, S.-H. Lai, and C. Zach, "Cycle-object consistency for image-to-image domain adaptation," *Pattern Recognition*, vol. 138, p. 109416, 2023.
- [38] F. Yu, D. Wang, Y. Chen, N. Karianakis, T. Shen, P. Yu, D. Lymberopoulos, S. Lu, W. Shi, and X. Chen, "Sc-uda: Style and content gaps aware unsupervised domain adaptation for object detection," 2022.
- [39] Y.-J. Li, X. Dai, C.-Y. Ma, Y.-C. Liu, K. Chen, B. Wu, Z. He, K. Kitani, and P. Vajda, "Cross-domain adaptive teacher for object detection," in *CVPR*, pp. 7581–7590, 2022.
- [40] M. Kennerley, J.-G. Wang, B. Veeravalli, and R. T. Tan, "2pcnet: Two-phase consistency training for day-to-night unsupervised domain adaptive object detection," in *CVPR*, pp. 11484–11493, 2023.
- [41] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," *NeurIPS*, vol. 28, 2015.
- [42] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," *ICCV*, 2017.
- [43] M.-Y. Liu, T. Breuel, and J. Kautz, "Unsupervised image-to-image translation networks," in *NeurIPS*, 2017.
- [44] X. Huang, M.-Y. Liu, S. Belongie, and J. Kautz, "Multimodal unsupervised image-to-image translation," in *ECCV*, 2018.
- [45] T. Park, A. A. Efros, R. Zhang, and J.-Y. Zhu, "Contrastive learning for unpaired image-to-image translation," in *ECCV*, pp. 319–345, Springer, 2020.
- [46] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *NeurIPS*, 2014.
- [47] T. H. Park and S. D'Amico, "Robust multi-task learning and online refinement for spacecraft pose estimation across domain gap," *Advances in Space Research*, 2023.
- [48] K. Liu and Y. Yu, "Revisiting the domain gap issue in non-cooperative spacecraft pose tracking," in *CVPR*, pp. 6864–6873, 2024.
- [49] J. I. B. Pérez-Villar, Á. García-Martín, J. Bescós, and M. Escudero-Viñolo, "Spacecraft pose estimation: Robust 2d and 3d-structural losses and unsupervised domain adaptation by inter-model consensus," *IEEE Transactions on Aerospace and Electronic Systems*, 2023.
- [50] Z. Wang, M. Chen, Y. Guo, Z. Li, and Q. Yu, "Bridging the domain gap in satellite pose estimation: A self-training approach based on geometrical constraints," *IEEE Transactions on Aerospace and Electronic Systems*, 2023.
- [51] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick, "Microsoft coco: Common objects in context," in *ECCV*, 2014.