

# Domain Adaptive Semantic Segmentation of Remote Sensing Images

Subjects: Others

Contributor: Zhihao Xi, Yu Meng, Jingbo Chen, Yupeng Deng, Diyou Liu, Yunlong Kong, Anzhi Yue

Semantic segmentation techniques for remote sensing images (RSIs) have been widely developed and applied. When a large change occurs in the target scenes, model performance drops significantly. Therefore, unsupervised domain adaptation (UDA) for semantic segmentation is proposed to alleviate the reliance on expensive per-pixel densely labeled data.

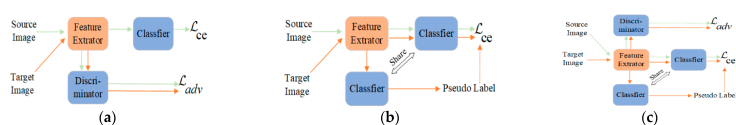
Keywords: adversarial perturbation consistency ; self-training ; semantic segmentation ; remote sensing

## 1. Introduction

Image segmentation has been widely researched as a basic remote sensing intelligent interpretation task [1][2][3][4]. In particular, semantic segmentation based on deep learning plays an important role as a pixel-level classification method in remote sensing interpretation tasks, such as building extraction [5], landcover classification [6] and change detection [7][8]. However, the prerequisite for good performance in existing fully supervised deep learning approaches is sufficiently annotated data. It is also essential that the training and test data follow the identical distributions [9]. Once applied to unseen scenarios with different data distributions, model performance can degrade significantly [10][11][12]. This means new data might be annotated and retrained for performance requirements, which requires considerable labor and time [13].

In practical applications, the domain discrepancy problem is prevalent in remote sensing images (RSIs) [14][15]. Different remote sensing platforms, payload imaging mechanisms, and photographic angles will induce variations in image spatial resolution and object features [16]. Due to the variation in seasons, geographic locations, illumination, and atmospheric radiation conditions, the same source images may also show significant feature distribution differences [17]. The data distribution shift caused by the mix of these complex factors leads the segmentation network to behave poorly in the unseen target domain.

As a transfer learning paradigm [18], unsupervised domain adaptation (UDA) can improve the domain generalization performance of the model by transferring knowledge from the source domain data with annotations to the target domain [19]. This method has been extensively researched in computer vision to address the domain discrepancy issue in natural image scenes [20]. Domain adaptive (DA) methods have also gained intensive attention in remote sensing [21]. Compared with natural images, RSIs contain more complex spatial detail information and object boundary situation, and homogeneous and heterogeneous phenomena are more common in images. Additionally, the factors that generate domain discrepancies are more complex and diverse. Thus, solving the problem of domain discrepancies in RSIs became more challenging. Currently, existing research works focus on three main approaches: UDA based on image transfer [17][22], UDA based on deep adversarial training (AT), and UDA based on self-training (ST) [23][24]. Image transfer methods achieve image-level alignment based on generative adversarial networks. AT-based methods (as shown in **Figure 1a**) reduce the feature distribution in the source and target domains by minimizing the adversarial loss to achieve feature-level alignment [25]. The ST approach (as shown in **Figure 1b**) focuses on generating high-confidence pseudolabels in the target domain and then participating in the iterative training of the model to achieve the progressive transfer process [26][27].



**Figure 1.** General paradigm description of existing DA training methods. (a) AT based DA approach. (b) Self-training (ST) based DA approach. (c) A combined ST and AT for DA methods.

One general conclusion about the DA performance of the model is:  $AT + ST > ST > AT$  [27]. However, as shown in **Figure 1c**, combining ST and AT methods typically requires strong coupling between submodules, which leads to a poorly stabilized model during training [28]. Therefore, fine-tuning the network structure and the submodules parameters is generally needed, so that model performance depends on specific scenarios and loses its scalability and flexibility. Recently, several studies have been conducted to optimize and improve the process, such as decoupling AT and ST methods functionally by constructing dual-stream networks [28], and using exponential moving average (EMA) techniques to construct teacher networks to smooth instable features in the training process [29]. However, it also complicates the network architecture, increasing the spatial computational complexity, and reducing training efficiency.

## 2. Image-Level Alignment for UDA

Image-level alignment reduces the data distribution shift between the source and target domains through image transfer methods [30][31]. This scheme generates pseudo images that are semantically identical to the source images, but whose spectral distribution is similar to that of the target images [17]. Cycle-consistent adversarial domain adaptation (CyCADA) improves the semantic consistency of the image transfer process through cycle consistency loss [32]. To preserve the semantic invariance of RSIs after being transferred, ColorMapGAN designs a color transformation method without a convolutional structure [17]. Many UDA schemes adopt GAN-based style transfer methods [33] to align data distributions in the source and target domains. ResiDualGAN [22] introduces scale information of RSIs based on DualGAN [34]. Some work also leverages non-adversarial optimization transform methods, such as Fourier transform-based FDA [35] and Wallis filtering methods [36], to reduce image domain discrepancies.

## 3. Feature-Level Alignment by AT

Adversarial-based feature alignment methods train additional domain discriminators [19][37] to distinguish target samples from source samples and then train the feature network to fool the discriminator, thus generating a domain-invariant feature space [38]. Many works have made significant progress using AT to align the feature space distribution to reduce the domain variance in RSIs. Wu et al. [39] focused on interdomain category differences and proposed class-aware domain alignment. Deng et al. [23] designed a scale discriminator to detect scale variation in RSIs. Considering regional diversity, Chen et al. [40] focused on difficult-to-align regions through a region adaptive discriminator. Bai et al. [20] leveraged contrast learning to align high-dimensional image representations between different domains. Lu et al. [41] designed global-local adversarial learning methods to ensure local semantic consistency in different domains.

## 4. Self-Training for UDA

Self-training acts as a kind of semi-supervised learning [42], which involves high-confidence prediction as easy-to-transfer pseudolabels, and participates in the next iteration of training together with the corresponding target images, progressively realizing the knowledge transfer process [26][27]. Yao et al. [36] used the ST paradigm to improve the performance of the model for building extraction on unseen data. CBST [26] designs class-balanced selectors for pseudolabels to avoid the easy-to-predict classes becoming dominant. ProDA [43] computes representation prototypes that represent the centers of category features to correct pseudolabels. CLUDA [44] constructs contrast learning between different classes and different domains by mixing source and target domain images. Additionally, several works have attempted to combine ST and adversarial methods to improve domain generalization performance. However, these models are difficult to optimize and often require fine-tuning of the model parameters. Zhang et al. [45] established the two-stage training process of AT followed by ST. DecoupleNet [28] decouples ST and AT through two network branches to alleviate the difficulty of model training.

## 5. Consistency Regularization

Consistency regularization is generally employed to solve semi-supervised problems, where the essential idea is to preserve the output consistency of the model under different versions of input perturbations, thus improving the generalization ability of the model for test data [46][47]. FixMatch [48] establishes two network flows, which include weak perturbation augmentation and strong perturbation augmentation at the image level, using the weak perturbation to ensure the high quality of the output and using the strong perturbation to provide better training of the model. FeatMatch [49] extracts class representative prototypes for feature-level augmentation transformations. Liu et al. [50] constructed dual-teacher networks to provide more rigorous pseudolabels for unlabeled test data. UniMatch [47] provides an auxiliary feature perturbation stream using a simple dropout mechanism. Several recent regularization models have been designed

under the ST paradigm, but fail to account for domain discrepancy scenes, which has led to the fact that pure consistency regularization has not behaved remarkably well in cross-domain scenes.

---

## References

1. Zhu, Q.; Sun, X.; Zhong, Y.; Zhang, L. High-Resolution Remote Sensing Image Scene Understanding: A Review. In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium, Yokohama, Japan, 28 July–2 August 2019; pp. 3061–3064.
2. Kotaridis, I.; Lazaridou, M. Remote Sensing Image Segmentation Advances: A Meta-Analysis. *ISPRS J. Photogramm. Remote Sens.* 2021, 173, 309–322.
3. Zhao, C.; Qin, B.; Feng, S.; Zhu, W.; Sun, W.; Li, W.; Jia, X. Hyperspectral Image Classification with Multi-Attention Transformer and Adaptive Superpixel Segmentation-Based Active Learning. *IEEE Trans. Image Process.* 2023, 32, 3606–3621.
4. Zhao, C.; Zhu, W.; Feng, S. Superpixel Guided Deformable Convolution Network for Hyperspectral Image Classification. *IEEE Trans. Image Process.* 2022, 31, 3838–3851.
5. Yang, X.; Li, S.; Chen, Z.; Chanussot, J.; Jia, X.; Zhang, B.; Li, B.; Chen, P. An Attention-Fused Network for Semantic Segmentation of Very-High-Resolution Remote Sensing Imagery. *ISPRS J. Photogramm. Remote Sens.* 2021, 177, 238–262.
6. Marcos, D.; Volpi, M.; Kellenberger, B.; Tuia, D. Land Cover Mapping at Very High Resolution with Rotation Equivariant CNNs: Towards Small yet Accurate Models. *ISPRS J. Photogramm. Remote Sens.* 2018, 145, 96–107.
7. Soto Vega, P.J.; da Costa, G.A.O.P.; Feitosa, R.Q.; Ortega Adarme, M.X.; de Almeida, C.A.; Heipke, C.; Rottensteiner, F. An Unsupervised Domain Adaptation Approach for Change Detection and Its Application to Deforestation Mapping in Tropical Biomes. *ISPRS J. Photogramm. Remote Sens.* 2021, 181, 113–128.
8. Deng, Y.; Chen, J.; Yi, S.; Yue, A.; Meng, Y.; Chen, J.; Zhang, Y. Feature-Guided Multitask Change Detection Network. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 2022, 15, 9667–9679.
9. Wilson, G.; Cook, D.J. A Survey of Unsupervised Deep Domain Adaptation. *arXiv* 2018, arXiv:1812.02849.
10. Zhao, C.; Qin, B.; Feng, S.; Zhu, W.; Zhang, L.; Ren, J. An Unsupervised Domain Adaptation Method Towards Multi-Level Features and Decision Boundaries for Cross-Scene Hyperspectral Image Classification. *IEEE Trans. Geosci. Remote Sens.* 2022, 60, 1–16.
11. Jiang, Z.; Li, Y.; Yang, C.; Gao, P.; Wang, Y.; Tai, Y.; Wang, C. Prototypical Contrast Adaptation for Domain Adaptive Semantic Segmentation. *arXiv* 2022, arXiv:2207.06654.
12. Yan, L.; Fan, B.; Xiang, S.; Pan, C. Adversarial Domain Adaptation with a Domain Similarity Discriminator for Semantic Segmentation of Urban Areas. In Proceedings of the 25th IEEE International Conference on Image Processing (ICIP), Athens, Greece, 7–10 October 2018; pp. 1583–1587.
13. Liu, W.; Su, F. Unsupervised Adversarial Domain Adaptation Network for Semantic Segmentation. *IEEE Geosci. Remote Sens. Lett.* 2020, 17, 1978–1982.
14. Tong, X.Y.; Xia, G.S.; Zhu, X.X. Enabling Country-Scale Land Cover Mapping with Meter-Resolution Satellite Imagery. *ISPRS J. Photogramm. Remote Sens.* 2023, 196, 178–196.
15. Wang, D.; Zhang, J.; Du, B.; Tao, D.; Zhang, L. Scaling-up Remote Sensing Segmentation Dataset with Segment Anything Model. *arXiv* 2023, arXiv:2305.02034.
16. Zhang, L.; Xia, G.S.; Wu, T.; Lin, L.; Tai, X.C. Deep Learning for Remote Sensing Image Understanding. *J. Sens.* 2016, 2016, 7954154.
17. Tasar, O.; Happy, S.L.; Tarabalka, Y.; Alliez, P. ColorMapGAN: Unsupervised Domain Adaptation for Semantic Segmentation Using Color Mapping Generative Adversarial Networks. *IEEE Trans. Geosci. Remote Sens.* 2020, 58, 7178–7193.
18. Jiang, J.; Shu, Y.; Wang, J.; Long, M. Transferability in Deep Learning: A Survey. *arXiv* 2022, arXiv:2201.05867.
19. Zhao, S.; Yue, X.; Zhang, S.; Li, B.; Zhao, H.; Wu, B.; Krishna, R.; Gonzalez, J.E.; Sangiovanni-Vincentelli, A.L.; Seshia, S.A.; et al. A Review of Single-Source Deep Unsupervised Visual Domain Adaptation. *IEEE Trans. Neural Networks Learn. Syst.* 2022, 33, 473–493.
20. Bai, L.; Du, S.; Zhang, X.; Wang, H.; Liu, B.; Ouyang, S. Domain Adaptation for Remote Sensing Image Semantic Segmentation: An Integrated Approach of Contrastive Learning and Adversarial Learning. *IEEE Trans. Geosci. Remote Sens.*

21. Tuia, D.; Persello, C.; Bruzzone, L. Recent Advances in Domain Adaptation for the Classification of Remote Sensing Data. *IEEE Geosci. Remote Sens. Mag.* 2021, 4, 41–57.
22. Zhao, Y.; Guo, P.; Sun, Z.; Chen, X.; Gao, H. ResiDualGAN: Resize-Residual DualGAN for Cross-Domain Remote Sensing Images Semantic Segmentation. *Remote Sens.* 2023, 15, 1428.
23. Deng, X.; Zhu, Y.; Tian, Y.; Newsam, S. Scale Aware Adaptation for Land-Cover Classification in Remote Sensing Imagery. In *Proceedings of the 2021 IEEE Winter Conference on Applications of Computer Vision (WACV)*, Waikoloa, HI, USA, 3–8 January 2021; pp. 2159–2168.
24. Zhao, Q.; Lyu, S.; Liu, B.; Chen, L.; Zhao, H. Self-Training Guided Disentangled Adaptation for Cross-Domain Remote Sensing Image Semantic Segmentation. *arXiv* 2023, arXiv:2301.05526.
25. Tsai, Y.H.; Hung, W.C.; Schuster, S.; Sohn, K.; Yang, M.H.; Chandraker, M. Learning to Adapt Structured Output Space for Semantic Segmentation. *arXiv* 2018, arXiv:1802.10349.
26. Zou, Y.; Yu, Z.; Vijaya Kumar, B.V.K.; Wang, J. Unsupervised Domain Adaptation for Semantic Segmentation via Class-Balanced Self-Training. In *Proceedings of the 15th European Conference, Munich, Germany, 8–14 September 2018*; pp. 297–313.
27. Mei, K.; Zhu, C.; Zou, J.; Zhang, S. Instance Adaptive Self-Training for Unsupervised Domain Adaptation. In *Proceedings of the 16th European Conference, Glasgow, UK, 23–28 August 2020*; pp. 415–430.
28. Lai, X.; Tian, Z.; Xu, X.; Chen, Y.; Liu, S.; Zhao, H.; Wang, L.; Jia, J. DecoupleNet: Decoupled Network for Domain Adaptive Semantic Segmentation. In *Proceedings of the 17th European Conference, Tel Aviv, Israel, 23–27 October 2022*; pp. 369–387.
29. Chen, R.; Rong, Y.; Guo, S.; Han, J.; Sun, F.; Xu, T.; Huang, W. Smoothing Matters: Momentum Transformer for Domain Adaptive Semantic Segmentation. *arXiv* 2022, arXiv:2203.07988.
30. Xu, Q.; Ma, Y.; Wu, J.; Long, C.; Huang, X. CDAda: A Curriculum Domain Adaptation for Nighttime Semantic Segmentation. In *Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW)*, Montreal, BC, Canada, 11–17 October 2021; pp. 2962–2971.
31. Tasar, O.; Happy, S.L.; Tarabalka, Y.; Alliez, P. SEMI2I: Semantically Consistent Image-to-Image Translation for Domain Adaptation of Remote Sensing Data. In *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium, Waikoloa, HI, USA, 26 September–2 October 2020*; pp. 1837–1840.
32. Hoffman, J.; Tzeng, E.; Park, T.; Phillip, J.Z.; Kate, I.; Alexei, S.; Darrell, T.; Chang, W.G.W.L.; Wang, H.P.; Peng, W.H.; et al. CyCADA: Cycle-Consistent Adversarial Domain Adaptation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, CA, USA, 15–20 June 2019; pp. 7346–7354.
33. Tasar, O.; Giros, A.; Tarabalka, Y.; Alliez, P.; Clerc, S. DAUGNet: Unsupervised, Multisource, Multitarget, and Life-Long Domain Adaptation for Semantic Segmentation of Satellite Images. *IEEE Trans. Geosci. Remote Sens.* 2021, 59, 1067–1081.
34. Yi, Z.; Zhang, H.; Tan, P.; Gong, M. DualGAN: Unsupervised Dual Learning for Image-to-Image Translation. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, Venice, Italy, 22–29 October 2017; pp. 2868–2876.
35. Yang, Y.; Soatto, S. FDA: Fourier Domain Adaptation for Semantic Segmentation. In *Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, WA, USA, 13–19 June 2020; pp. 4084–4094.
36. Peng, D.; Guan, H.; Zang, Y.; Bruzzone, L. Full-Level Domain Adaptation for Building Extraction in Very-High-Resolution Optical Remote-Sensing Images. *IEEE Trans. Geosci. Remote Sens.* 2022, 60, 1–17.
37. Ganin, Y.; Ustinova, E.; Ajakan, H.; Germain, P.; Larochelle, H.; Laviolette, F.; Marchand, M.; Lempitsky, V. Domain-Adversarial Training of Neural Networks. *Adv. Comput. Vis. Pattern Recognit.* 2017, 17, 189–209.
38. Wang, H.; Shen, T.; Zhang, W.; Duan, L.Y.; Mei, T. Classes Matter: A Fine-Grained Adversarial Approach to Cross-Domain Semantic Segmentation. In *Proceedings of the 16th European Conference, Glasgow, UK, 23–28 August 2020*; pp. 1–17.
39. Xu, Q.; Yuan, X.; Ouyang, C. Class-Aware Domain Adaptation for Semantic Segmentation of Remote Sensing Images. *IEEE Trans. Geosci. Remote Sens.* 2020, 60, 1–18.
40. Chen, J.; Zhu, J.; Guo, Y.; Sun, G.; Zhang, Y.; Deng, M. Unsupervised Domain Adaptation for Semantic Segmentation of High-Resolution Remote Sensing Imagery Driven by Category-Certainty Attention. *IEEE Trans. Geosci. Remote Sens.* 2022, 60, 1–15.

41. Lu, X.; Zhong, Y.; Zheng, Z.; Wang, J. Cross-Domain Road Detection Based on Global-Local Adversarial Learning Framework from Very High Resolution Satellite Imagery. *ISPRS J. Photogramm. Remote Sens.* 2021, 180, 296–312.
42. Chapelle, O.; Schölkopf, B.; Zien, E.A. Semi-Supervised Learning. *IEEE Trans. Neural Netw.* 2009, 20, 2015975.
43. Zhang, P.; Zhang, B.; Zhang, T.; Chen, D.; Wang, Y.; Wen, F. Prototypical Pseudo Label Denoising and Target Structure Learning for Domain Adaptive Semantic Segmentation. *arXiv* 2021, arXiv:2101.10979.
44. Vayyat, M.; Kasi, J.; Bhattacharya, A.; Ahmed, S.; Tallamraju, R. CLUDA: Contrastive Learning in Unsupervised Domain Adaptation for Semantic Segmentation. *arXiv* 2022, arXiv:2208.14227.
45. Zhang, L.; Lan, M.; Zhang, J.; Tao, D. Stageswise Unsupervised Domain Adaptation With Adversarial Self-Training for Road Segmentation of Remote-Sensing Images. *IEEE Trans. Geosci. Remote Sens.* 2021, 60, 1–13.
46. Chen, X.; Yuan, Y.; Zeng, G.; Wang, J. Semi-Supervised Semantic Segmentation with Cross Pseudo Supervision. In *Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Nashville, TN, USA, 20–25 June 2021; pp. 2613–2622.
47. Yang, L.; Qi, L.; Feng, L.; Zhang, W.; Shi, Y. Revisiting Weak-to-Strong Consistency in Semi-Supervised Semantic Segmentation. *arXiv* 2022, arXiv:2208.09910.
48. Sohn, K.; Berthelot, D.; Zizhao, C.L.; Nicholas, Z.; Cubuk, E.D.; Kurakin, A.; Zhang, H.; Raffel, C. FixMatch: Simplifying Semi-Supervised Learning with Consistency and Confidence. *arXiv* 2020, arXiv:2001.07685.
49. Kuo, C.W.; Ma, C.Y.; Huang, J.B.; Kira, Z. FeatMatch: Feature-Based Augmentation for Semi-Supervised Learning. In *Proceedings of the 16th European Conference, Glasgow, UK, 23–28 August 2020*; pp. 479–495.
50. Liu, Y.; Tian, Y.; Chen, Y.; Liu, F.; Belagiannis, V.; Carneiro, G. Perturbed and Strict Mean Teachers for Semi-Supervised Semantic Segmentation. In *Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, New Orleans, LA, USA, 18–24 June 2022; pp. 4248–4257.

---

Retrieved from <https://encyclopedia.pub/entry/history/show/123929>