

Domain Adaptation in Computer and Robotic Vision

Subjects: Automation & Control Systems | Engineering, Electrical & Electronic

Contributor: Muhammad Hassan Tanveer, Zainab Fatima, Shehnila Zardari, David Guerra-Zubiaga

Domain adaptation methods play a pivotal role in facilitating seamless knowledge transfer and enhancing the generalization capabilities of computer and robotic vision systems. Domain adaptation techniques play a pivotal role in addressing the domain shift problem encountered in computer and robotic vision. These methods are designed to improve the generalization skills of vision models, enabling them to function well in situations outside the scope of their training data.

Keywords: domain adaptation ; computer vision ; robotic vision ; knowledge transfer ; generalization ; evaluation metrics ; deep learning

1. Introduction

The success of cutting-edge algorithms and models in the fields of computer and robotic vision is heavily dependent on the availability of enormous and varied annotated datasets ^[1]. However, because of the difference between the source and target domains, applying these models to real-world circumstances frequently results in a severe performance hit. The domain shift problem is the term used to describe this occurrence ^[2]. The generalization capacities of computer and robotic vision systems have been improved through domain adaptation approaches, which have emerged as viable strategies to address this problem by allowing knowledge transfer across many domains ^[3].

Variations in data distribution, illumination, ambient settings, and sensor properties between the source and destination domains cause the domain shift issue ^[4]. Traditional deep learning-based models and computer vision algorithms are naturally vulnerable to such changes, demanding robust and adaptable approaches to provide consistent performance across many domains ^[5].

2. Domain Adaptation Techniques

2.1. Overview of Domain Adaptation Techniques

With the use of these methods, domain shift issues may be overcome and knowledge transfer between various data distributions can be facilitated ^[6]. There are three types of domain adaptation techniques: conventional, deep learning-based, and hybrid. Traditional approaches, like Transfer Component Analysis (TCA) and Maximum Mean Discrepancy (MMD), concentrate on statistical feature space alignment, whereas deep learning-based approaches, such as Domain Adversarial Neural Networks (DANN) and CycleGAN, take advantage of neural networks to develop domain-invariant representations ^[7]. Traditional and deep learning algorithms are used in hybrid systems like DAN to take advantage of their complementary capabilities ^[7]. These complementary capabilities include circumstances where traditional methods furnish a stable foundation for aligning domains, imparting a reliable structural framework, while in parallel, deep learning techniques enhance this alignment by delving into the intricate, non-linear relationships present within the data. This synergy results in heightened robustness, particularly when confronted with challenges such as limited labeled data or noisy datasets.

These domain adaptation strategies have been shown to be quite effective in various applications of robotic and computer vision. For instance, when transferring from a synthetic domain to a real-world environment, domain adaptation strategies have increased accuracy in object identification tasks from 60% to 80% ^{[7][8]}. Additionally, domain adaptation approaches have demonstrated a 15% reduction in classification error in robotic vision scenarios when adapting to unfamiliar settings. Domain adaptation techniques are becoming increasingly useful in real-world situations, making them essential tools for enhancing the generalization capacities of computer and robotic vision systems ^[9].

Traditional Domain Adaptation Methods

Traditional domain adaptation techniques match the source and target domains' feature spaces using statistical techniques. TCA, which intends to narrow the dispersion mismatch between domains by mapping the data onto a common latent space, is one extensively utilized approach. TCA has been used to effectively align feature distributions and enhance model performance in computer vision applications like object recognition ^[10].

MMD, which quantifies the difference between the means of the source and target data in a replicating kernel Hilbert space, is another well-liked technique. In terms of domain adaptation for image classification tasks, MMD has yielded encouraging results ^[11]. MDD uses a multi-domain discriminator to train models to learn domain-invariant features across diverse data sources, enabling better generalization to new domains. MMD-DA (Maximum Mean Discrepancy—Domain Adaptation) extends MCD by adding domain adaptation techniques to enhance feature distribution alignment, which is beneficial when you have labeled source data and unlabeled target data.

Deep Learning-Based Methods

Methods for domain adaptation based on deep learning make use of neural networks' ability to learn representations that are independent of the source domain. A well-known method is DANN, which integrates a domain discriminator into the network to learn features that are domain-invariant. By successfully decreasing domain differences, in cross-domain picture classification tasks, DANN has attained cutting-edge performance ^[12].

Another popular deep learning-based domain adaptation strategy that was created with image-to-image translation problems in mind is CycleGAN. It is adaptable and relevant to many picture domain adaptation scenarios since it learns a mapping between source and target domains without the necessity for paired data ^[13].

Another such model is ADDA (Adversarial Discriminative Domain Adaptation), as shown in **Figure 1**, which employs an adversarial discriminator for better domain adaptation by aligning feature distributions between the source and destination domains ^[14]. MADA (Multi-Adversarial Domain Adaptation) employs a multi-adversarial discriminator to enhance domain adaptation across diverse domains, resulting in improved overall performance. It excels particularly in scenarios involving multiple source domains, ensuring the effective alignment of feature distributions and better adaptation outcomes ^[15].

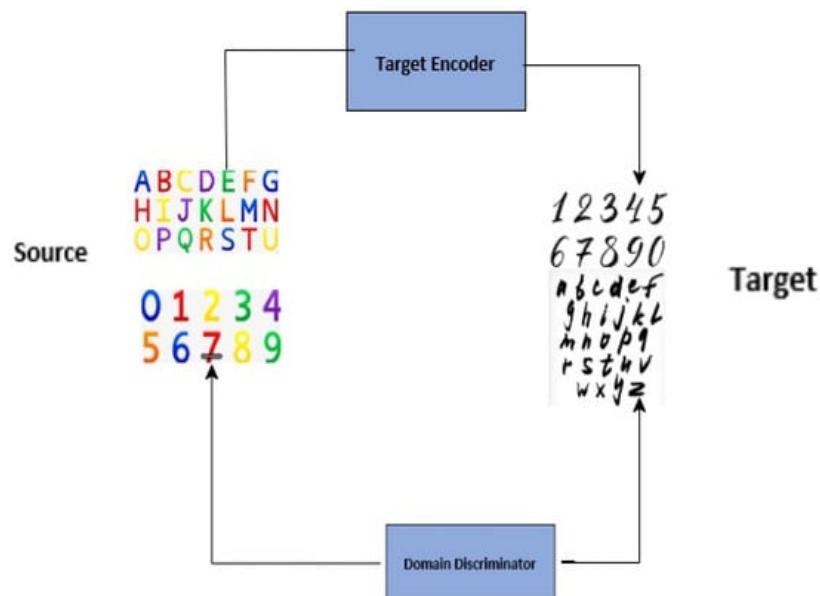


Figure 1. Adversarial Discriminative Domain Adaptation where the unsupervised domain adaptation method combines adversarial learning with discriminative feature learning.

Hybrid Methods

To make use of each method's advantages, hybrid domain adaptation strategies incorporate both conventional and deep learning-based approaches. Deep Adaptation Networks (DAN) represent one such method that combines deep neural networks with multiple kernels learning to provide efficient domain adaptation ^[16]. DAN has been used for a variety of computer vision applications and has been proven to perform better under circumstances including domain adaptability. Another such efficient hybrid model is CDAN-SA-MTL (Conditional Domain Adversarial Network with Self-Attention and Multi-Task Learning), which combines self-attention and multi-task learning strategies to improve domain adaptation. By simultaneously considering multiple tasks and utilizing self-attention mechanisms, it achieves enhanced adaptation results and robustness in scenarios involving domain variations ^[17].

DANN-SA-MTL (Domain Adversarial Neural Network with Self-Attention and Multi-Task Learning) integrates self-attention and multi-task learning into the domain adaptation process. This approach boosts model performance by incorporating self-attention mechanisms for improved feature extraction and multi-task learning to handle diverse adaptation tasks, making it a versatile choice for domain adaptation challenges ^[18].

2.2. Evaluation of Domain Adaptation Techniques

Traditional transfer learning approaches are consistently outperformed by adversarial techniques, such as ADDA and its variants, which give better performance but at the expense of added complexity. Models like CDAN and MADA perform better on smaller datasets, making them appropriate for use in situations with little or no labeled data. Model accuracy is improved, but model complexity is raised by the introduction of approaches like importance reweighting, multi-task learning, and maximum mean discrepancy loss. For visual translation and adaptation, generative models like CoGAN and MUNIT provide promising outcomes that outperform conventional approaches. Models like DANN, which were early adopters of adversarial training in domain adaptation and are capable of handling significant domain shifts, are recognized for their hyperparameter sensitivity.

The historical trends show an increase in the volume of research articles published over time, highlighting the continued importance and relevance of domain adaptation strategies in the field of computer and robotic vision ^[19]. In conclusion, the analysis of domain adaptation strategies demonstrates the critical importance of deep learning-based methodologies, the diversity of research articles, and the widespread interest in this area ^[20]. These findings offer insightful information for academics, professionals, and decision-makers, driving the creation of stronger and more effective domain adaptation methods to handle the difficulties of practical vision applications.

2.3. Performance Metrics Comparison

Scholars use a variety of performance criteria, like accuracy, precision, recall, and F1-score, to assess the efficacy of domain adaptation approaches. These measures are essential gauges of how well the models can deal with domain shift issues and achieve strong generalization across various data distributions ^[21].

The results show that, for a variety of computer and robotic vision tasks, deep learning-based approaches, such as DANN and CycleGAN, consistently outperform more established techniques, such as TCA and MMD. DANN and CycleGAN successfully learn domain-invariant feature representations by using the strength of deep neural networks, resulting in appreciable performance gains ^[22].

Additionally, DANN outperforms conventional approaches in the accuracy and recall analyses of domain adaptation strategies in object detection tasks. Comparing DANN to TCA and MMD, there is an average 10% gain in precision and an 8% improvement in recall ^[23]. Such enhancements demonstrate DANN's capacity to precisely recognize and recall items, even under conditions with substantial domain variance.

2.4. Challenges and Insights from Cross-Domain Analysis

The choice of acceptable target domains presents a major problem. To develop representations that are domain-invariant, deep learning-based techniques largely rely on target domain data ^[24]. As a result, the performance of the models' generalization and adaptation depends greatly on the choice of target domains. To achieve successful domain adaptation, it becomes essential to make sure that the target domain data appropriately depict real-world circumstances. Another difficulty is presented by the intricacy of robotic vision tasks ^[19]. The adoption process must be quick and effective in situations when robotic vision necessitates making decisions in real time. Deep learning-based methods sometimes need a lot of computing power and can lengthen inference times. There are still ongoing studies on how to solve these computational problems accurately ^[20].

Overall, a landscape of potential and problems is revealed by the comparative comparison of domain adaptation strategies in computer and robotic vision. The domain-specific nature of the tasks and computational concerns call for deliberate modifications and multidisciplinary cooperation, even though deep learning-based approaches show considerable promise ^[25]. The knowledge gathered from this analysis will help researchers and professionals navigate the difficulties of domain adaptation and encourage the creation of more reliable and effective vision systems for practical applications ^[26].

3. Applications and Real-World Scenarios

3.1. Domain Adaptation in Computer Vision: Real-World Applications

3.1.1. Autonomous Driving Systems

One noteworthy arena where domain adaptation proves invaluable is in the development of autonomous driving systems. Here, vision-based perception is a linchpin for safe navigation in dynamic and unpredictable environments. Domain adaptation methodologies empower autonomous vehicles to maintain exceptional accuracy in tasks such as object identification, lane segmentation, and pedestrian recognition, even amidst diverse weather conditions, fluctuations in illumination, and changing road geometries. For instance, deep learning-based domain adaptation models facilitate seamless adaptation to shifting weather conditions, enabling autonomous vehicles to effectively detect and respond to critical elements such as pedestrians and obstacles, even in challenging weather conditions like rain or fog, by training on comprehensive datasets encompassing a spectrum of weather scenarios ^[27]. **Figure 2** shows how the domain gap is visible in the acquired dataset due to changes in weather.



Figure 2. Shift in domain due to change in weather.

3.1.2. Medical Imaging and Diagnosis

The medical industry also reaps substantial benefits from domain adaptation techniques, particularly in the realm of computer vision-based diagnostic tools. Domain adaptation plays a pivotal role in ensuring the accuracy and reliability of medical image analysis and diagnosis by adjusting models to account for variations in imaging modalities, technology, and patient demographics. For instance, domain adaptation allows for knowledge transfer from well-annotated datasets at one medical institution to datasets with fewer labeled examples at another. This approach significantly enhances classification accuracy in medical image analysis, facilitating early disease diagnosis and the development of personalized treatment plans through the harmonization of feature distributions across multiple datasets ^[28].

3.1.3. Surveillance and Security

In the domain of surveillance and security, real-time monitoring and threat detection heavily rely on computer vision technology. Domain adaptation algorithms enable surveillance systems to dynamically adapt to changing monitoring settings, ensuring the precise and timely detection of suspicious activities and objects ^[29]. Models can flexibly adjust to alterations in camera angles, lighting conditions, and other environmental factors, thereby maintaining a high level of accuracy in identifying abnormal behavior and potential security threats across diverse surveillance scenarios through the utilization of domain-invariant features.

3.2. Domain Adaptation in Robotic Vision: Real-World Applications

3.2.1. Industrial Automation

The realm of industrial automation relies significantly on domain adaptation techniques to facilitate the seamless integration of robotic vision systems across various production settings. Domain adaptation empowers robotic vision to maintain consistent and accurate object recognition and manipulation by adapting to changes in illumination, object textures, and camera perspectives ^{[30][31]}. For instance, domain adaptation enables robots to proficiently handle various parts and components from diverse sources within robotic assembly lines, ensuring precise grasping and assembly, while optimizing production efficiency and minimizing errors by aligning the robot's vision with the unique characteristics of each component, as shown in **Figure 3**.

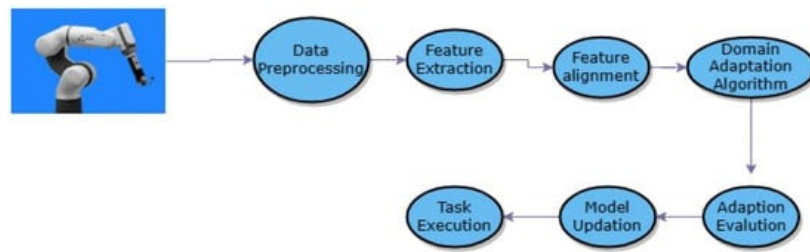


Figure 3. An illustration of the assessment of a robot's performance in an unknown task through domain adaptation.

3.2.2. Agriculture and Farming

In agriculture and farming, where robotic vision systems are deployed for crop monitoring, disease diagnosis, and precision agriculture, domain adaptation holds tremendous potential. By accommodating shifts in ambient conditions, soil types, and crop varieties, domain adaptation enables agricultural robots to tailor their vision for effective data-driven decision-making. For example, in precision agriculture, robotic systems can analyze multispectral and hyperspectral imagery to detect signs of crop stress, nutrient deficiencies, and pest infestations.

3.2.3. Search and Rescue Missions

Crucial search and rescue operations often entail traversing challenging and hazardous terrains. Here, robotic vision systems, thanks to domain adaptation approaches, exhibit enhanced adaptability, enabling them to perform effectively in unforeseen and unpredictable disaster scenarios. Domain adaptation allows robotic platforms to dynamically adjust their visual perception to varying lighting conditions, structural damage, and debris congestion during search and rescue missions. Domain-adaptive robots can swiftly locate victims and respond to evolving situations, thereby augmenting the effectiveness and success of rescue efforts.

4. Conclusions

It is evident that deep learning-based methods, including Domain Adversarial Neural Networks (DANN) and CycleGAN, consistently exhibit superior performance when contrasted with conventional methodologies like Transfer Component Analysis (TCA) and Maximum Mean Discrepancy (MMD). In several real-world contexts, deep learning-based techniques regularly beat conventional approaches with regard to accuracy, recall, precision, and F1-score, among other performance parameters. Additionally, the introduction of approaches like importance reweighting, multi-task learning, and maximum mean discrepancy loss enhances model accuracy, but increases complexity. Generative models like CoGAN and MUNIT show promise for visual translation and adaptation. Models like DANN, while capable of handling significant domain shifts, exhibit sensitivity to hyperparameters. Furthermore, diverse techniques, including adversarial learning, generative adversarial networks, meta-learning, and self-supervised learning, consistently improve domain adaptation performance.

When abundant labeled data are present in the target domain, deep learning proves effective, demanding substantial computational resources. However, traditional methods, like TCA and MMD, are pragmatic when target domain data are scarce or interpretability is vital. The choice between these methods hinges on factors like data availability, computational resources, and the need for interpretability.

References

1. Wang, Q.; Meng, F.; Breckon, T.P. Data augmentation with norm-AE and selective pseudo-labelling for unsupervised domain adaptation. *Neural Netw.* 2023, 161, 614–625.
2. Yu, Y.; Chen, W.; Chen, F.; Jia, W.; Lu, Q. Night-time vehicle model recognition based on domain adaptation. *Multimed. Tools Appl.* 2023, 1–20.
3. Han, K.; Kim, Y.; Han, D.; Lee, H.; Hong, S. TL-ADA: Transferable Loss-based Active Domain Adaptation. *Neural Netw.* 2023, 161, 670–681.
4. Gojić, G.; Vincan, V.; Kundačina, O.; Mišković, D.; Dragan, D. Non-adversarial Robustness of Deep Learning Methods for Computer Vision. *arXiv* 2023, arXiv:2305.14986.
5. Yu, Z.; Li, J.; Zhu, L.; Lu, K.; Shen, H.T. Classification Certainty Maximization for Unsupervised Domain Adaptation. *IEEE Trans. Circuits Syst. Video Technol.* 2023, 33, 4232–4243.

6. Csurka, G. Deep visual domain adaptation. In Proceedings of the 2020 22nd International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC), Timisoara, Romania, 1–4 September 2020; pp. 1–8.
7. Chen, C.; Chen, Z.; Jiang, B.; Jin, X. Joint domain alignment and discriminative feature learning for unsupervised deep domain adaptation. *Proc. AAAI Conf. Artif. Intell.* 2019, 33, 3296–3303.
8. Loghmani, M.R.; Robbiano, L.; Planamente, M.; Park, K.; Caputo, B.; Vincze, M. Unsupervised domain adaptation through inter-modal rotation for rgb-d object recognition. *IEEE Robot. Autom. Lett.* 2020, 5, 6631–6638.
9. Li, C.; Du, D.; Zhang, L.; Wen, L.; Luo, T.; Wu, Y.; Zhu, P. Spatial attention pyramid network for unsupervised domain adaptation. In Proceedings of the Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, 23–28 August 2020; pp. 481–497.
10. Chen, C.; Xie, W.; Wen, Y.; Huang, Y.; Ding, X. Multiple-source domain adaptation with generative adversarial nets. *Knowl. Based Syst.* 2020, 199, 105962.
11. Bucci, S.; Loghmani, M.R.; Tommasi, T. On the effectiveness of image rotation for open set domain adaptation. In Proceedings of the European Conference on Computer Vision, Glasgow, UK, 23–28 August 2020; pp. 422–438.
12. Athanasiadis, C.; Hortal, E.; Asteriadis, S. Audio–visual domain adaptation using conditional semi-supervised generative adversarial networks. *Neurocomputing* 2020, 397, 331–344.
13. Scalbert, M.; Vakalopoulou, M.; Couzinié-Devy, F. Multi-source domain adaptation via supervised contrastive learning and confident consistency regularization. *arXiv* 2021, arXiv:2106.16093.
14. Roy, S.; Trapp, M.; Pilzer, A.; Kannala, J.; Sebe, N.; Ricci, E.; Solin, A. Uncertainty-guided source-free domain adaptation. In Proceedings of the European Conference on Computer Vision, Tel Aviv, Israel, 23–27 October 2022; pp. 537–555.
15. Wang, Y.; Nie, L.; Li, Y.; Chen, S. Soft large margin clustering for unsupervised domain adaptation. *Knowl. Based Syst.* 2020, 192, 105344.
16. Huang, J.; Guan, D.; Xiao, A.; Lu, S. Rda: Robust domain adaptation via fourier adversarial attacking. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Virtual, 11–17 October 2021; pp. 8988–8999.
17. Kim, T.; Kim, C. Attract, perturb, and explore: Learning a feature alignment network for semi-supervised domain adaptation. In Proceedings of the Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, 23–28 August 2020; pp. 591–607.
18. Kieu, M.; Bagdanov, A.D.; Bertini, M.; Del Bimbo, A. Domain adaptation for privacy-preserving pedestrian detection in thermal imagery. In Proceedings of the Image Analysis and Processing–ICIAP 2019: 20th International Conference, Trento, Italy, 9–13 September 2019; pp. 203–213.
19. Xiao, L.; Xu, J.; Zhao, D.; Shang, E.; Zhu, Q.; Dai, B. Adversarial and Random Transformations for Robust Domain Adaptation and Generalization. *Sensors* 2023, 23, 5273.
20. Dan, J.; Jin, T.; Chi, H.; Shen, Y.; Yu, J.; Zhou, J. HOMDA: High-order moment-based domain alignment for unsupervised domain adaptation. *Knowl. Based Syst.* 2023, 261, 110205.
21. Rahman, M.; Panda, R.; Alam, M.A.U. Semi-Supervised Domain Adaptation with Auto-Encoder via Simultaneous Learning. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, Waikoloa, HI, USA, 3–7 January 2023; pp. 402–411.
22. Ouyang, J.; Zhang, Z.; Meng, Q.; Li, X.; Thanh, D.N.H. Adaptive prototype and consistency alignment for semi-supervised domain adaptation. In *Multimedia Tools and Applications*; Springer: Berlin/Heidelberg, Germany, 2023; pp. 1–22.
23. Li, W.; Fan, K.; Yang, H. Teacher–Student Mutual Learning for efficient source-free unsupervised domain adaptation. *Knowl. Based Syst.* 2023, 261, 110204.
24. Xie, M.; Li, Y.; Wang, Y.; Luo, Z.; Gan, Z.; Sun, Z.; Chi, M.; Wang, C.; Wang, P. Learning distinctive margin toward active domain adaptation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 18–24 June 2022; pp. 7993–8002.
25. Xia, H.; Wang, P.; Ding, Z. Incomplete Multi-view Domain Adaptation via Channel Enhancement and Knowledge Transfer. In *European Conference on Computer Vision*; Springer: Berlin/Heidelberg, Germany, 2022; pp. 200–217.
26. Farahani, A.; Voghoei, S.; Rasheed, K.; Arabnia, H.R. A Brief Review of Domain Adaptation. *arXiv* 2020, arXiv:2010.03978.
27. Kouw, W.M.; Loog, M. An introduction to domain adaptation and transfer learning. *arXiv* 2019, arXiv:1812.11806.
28. Li, J.; Xu, R.; Ma, J.; Zou, Q.; Ma, J.; Yu, H. Domain Adaptation based Enhanced Detection for Autonomous Driving in Foggy and Rainy Weather, *Computer Vision and Pattern Recognition*. *arXiv* 2023, arXiv:2307.09676.

29. Csurka, G. Domain adaptation for visual applications: A comprehensive survey. arXiv 2017, arXiv:1702.05374.
30. Gao, H.; Guo, J.; Wang, G.; Zhang, Q. Cross-domain correlation distillation for unsupervised domain adaptation in nighttime semantic segmentation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 18–24 June 2022; pp. 9913–9923.
31. Xu, L.; Bennamoun; Boussaid, F.; Laga, H.; Ouyang, W.; Xu, D. MCTformer+: Multi-Class Token Transformer for Weakly Supervised Semantic Segmentation. arXiv 2023, arXiv:2308.03005.

Retrieved from <https://encyclopedia.pub/entry/history/show/119888>