Domain Adaptation: A Survey

Ashly Ajith and G. Gopakumar

Abstract In computer vision, domain shifts are a typical issue. A classifier that has been trained on a source domain will not be able to perform well on a target domain. As a result, a source classifier taught to discriminate based on a particular distribution will struggle to classify new data from a different distribution. Domain adaptation is a hot area of research due to the plethora of applications available from this technique. Many developments have been made in this direction in recent decades. In light of this, we have compiled a summary of domain adaptation research, concentrating on work done in the last few years (2015–2022) for the benefit of the research community. We have categorically placed the important research works in DA under the chosen methodologies and have critically assessed the performances of these techniques. The study covers these features at length, and thorough descriptions of representative methods for each group are provided.

Keywords Domain adaptation · Computer vision · Transfer learning

1 Introduction

Machine learning is not the same as human learning. Humans can learn from a small number of labeled instances and apply what they have learned to new examples in unique situations. On the other hand, supervised machine learning approaches only work well when the test data is from the same distribution as the training data [22, 26]. They perform poorly when the testing dataset is from a non-identical distribution [13]. This happens due to the shift between the domain distributions.

Domain adaptation has found many applications in computer vision related to applying a trained network to real-world data. It can also label a synthetic dataset related to an earlier labeled dataset with less effort. Several works like [36] have utilized domain adaptation for segmentation problems in computer vision where the

A. Ajith (✉) · G. Gopakumar
Department of Computer Science and Engineering, Amrita Vishwa Vidyapeetham, Amritapuri, Kerala 690525, India
e-mail: ashlyajith619@yahoo.com

M. Tistarelli et al. (eds.), *Computer Vision and Machine Intelligence*,
Lecture Notes in Networks and Systems 586,
https://doi.org/10.1007/978-981-19-7867-8_47

testing set of data is from a distribution dissimilar to the training set. Under such cases, traditionally trained models will perform poorly. It has also found several applications under image captioning as well. Many domain adaptation techniques were devised to neutralize the performance reduction caused by domain shift. They can be broadly categorized into:

- **Supervised**: Under supervised domain adaptation, most of the samples in the target domain are labeled.
- **Semi-Supervised**: In the target domain, a few labeled samples are supplied to learn the suitable model adaption. This form of semi-supervised learning can extract invariant features from both domains. However, it also requires a small sample of target images to be labeled.
- **Unsupervised**: Unsupervised domain adaptation (UDA) reduces the shift between domains through unlabeled target datasets while seeking to maximize the classifier's performance on them. The target images are passed simultaneously with the source images. The network tries to classify the target images on the labels provided by the source domain images [3, 14].

This study aims to assess current progress in domain adaptation techniques and to offer some inferences from research directions.

2 Datasets Used for Domain Adaptation

Datasets used for domain adaptation simulate the condition where the data is from a different but related distribution. Datasets are therefore developed to generalize models across different domains. We have listed a few of the popular datasets used in domain adaptation. These datasets are used as the benchmark for calculating the domain adaptation of a technique.

2.1 Office 31

Office 31 [25] is a benchmark dataset with 4110 images with 31 categories and 3 domains. It contains Amazon, which has images extracted from Amazon, DSLR has images taken by a DSLR camera, and Webcam consists of images taken using a Web camera under various photographic settings. Office dataset was created in 2010 and has since been the benchmark dataset for DA problems. Sample images of the datasets are given in Fig. 1.

Fig. 1 Office 31 dataset samples

Fig. 2 Office-Caltech-256 dataset samples

2.2 Caltech

Caltech-256 [15] is a 257-class object recognition dataset that contains 30,607 real-world photos of various sizes. There are at least 80 photos for each class. The Caltech-101 dataset is a subset of this dataset. Caltech-256 is more intricate and demanding due to the more significant variability in size, backdrop, and other factors. We have added some samples from Office and Caltech, which share overlapping classes in Fig. 2. However, recently, more extensive datasets (Sects. 2.3 and 2.4) have been created to explore the scope of adaptation learning.

2.3 Office Home

Office Home [32] has four domains and 15,500 pictures or snapshots from 65 distinct categories. The four domains are Art, i.e., Ar, Clipart, i.e., Cl, Product, i.e., Pr,

Fig. 3 Office Home dataset samples

and Real-World, i.e., Rw. Snapshots from drawings, canvas and various creative renderings of images are included in the art domain. The Clipart domain is a collection of clipart pictures. Real-World is made up of typical photographs acquired with a camera, while Product comprises images without a background (refer Fig. 3).

2.4 MNIST and MNIST-M

MNIST [18] dataset is made up of 70,000 hand-written digits images in gray-scale. MNIST-M [10] is a dataset made from the MNIST dataset with the background containing varied patches of color from colored photos. It contains around 59,001 training and 90,001 test images. The MNIST/MNIST-M dataset samples have been given in Fig. 4.

3 Methods of Deep Domain Adaptation

We have divided distinct domain adaption techniques into categories depending on their methodology in this section.

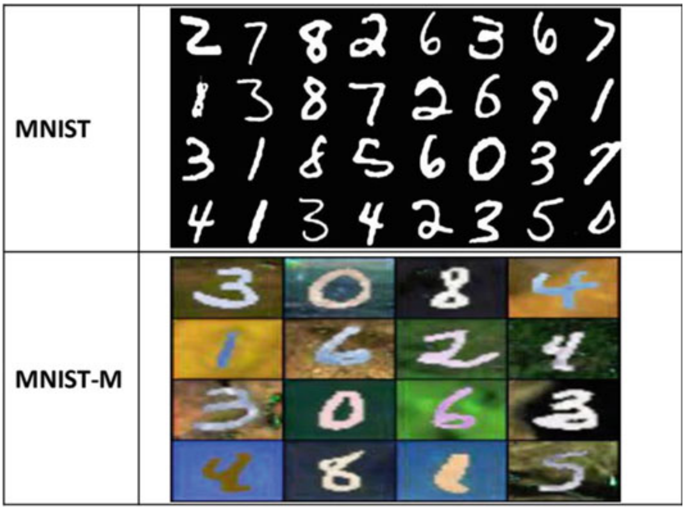

Fig. 4 MNIST/MNIST-M dataset samples

3.1 Discrepancy Based

Recent works show that deep domain adaptation networks, which perform discrepancy-based domain adaptation, produce better results than earlier multi-step works [29, 30]. Under discrepancy-based works, the commonly seen criteria to increase domain adaptation are.

Maximum Discrepancy Loss (MMD) MMD [21, 35] used a residual block attached to the end of the source network. The loss function was the sum of minimum mean discrepancy (MMD) loss [20] and entropy loss. The high representational features were extracted using the residual block, while the domain alignment was improved using MMD loss. The principle behind maximum mean discrepancy loss (MMD loss) is that distances between distributions are represented as distances between mean embeddings of features. The difference between the two alternative projections of the mean is the MMD. The maximum mean discrepancy (MMD) loss uses the kernel trick to find infinite moments across the distributions and reduce the distance between the expectation of distribution from the samples.

Correlational Alignment (CORAL) Similar to MMD, DeepCORAL [27, 28] was developed to improve domain adaptation using second-order distributions statistics. CORAL [27] explores the second-order statistics between the source and target domains using the higher representational features to align them. It was used within a DA network, reducing the domains' correlational alignment (CORAL) by minimizing the covariance between the higher representational tensors within the fully connected layers. This was done by using the CORAL loss obtained from the network layers. The domains were aligned by reducing the coral loss within the network.

Optimal Transport (OT) Optimal transport (OT) was proposed as a discrepancy technique to improve domain adaptation in joint distribution optimal transport (JDOT) [7]. Optimal transport is unique because it operates as a tool for converting one (continuous) probability distribution into another with minimal effort. The OT solution determines the most efficient technique to transform one distribution into another. The solution may be used to interpolate between them and acquire intermediate transformations seamlessly. The source data was transformed to a subspace with the shortest Wasserstein distance [2] across domains using OT. However, it scaled quadratically in computational cost with the sample size. This drawback was overcome in DeepJDOT [8] by implementing the OT stochastically in a deep adaptation network (DAN).

3.2 Adversarial Based

Adversarial network-based domain adaptation has recently attracted much attention since it provides state-of-the-art accuracy that outperforms discrepancy-based approaches. These networks follow a generator-discriminator architecture where the generator will learn to produce domain invariant and discriminative characteristics from the domains. A discriminator, which functions as a domain classifier, will predict the source of the image from the generator. This creates domain confusion within the network. Thus, working adversarially, these networks will produce domain invariant features which can be used to perform UDA.

Domain-adversarial neural network (DANN) [9] model was developed based on this principle of using GANs to reduce domain shift between the domains. Unsupervised domain adaptation is achieved by linking the feature extractor to a domain classifier through a gradient reversal layer that, during backpropagation-based training, doubles the gradient by a negative constant. The gradient reversal feature within the network was responsible for extracting the domain invariant features from both the domains within the networks.

Coupled generative adversarial networks (CoGANs) [19] use two GANs through which the domains are simultaneously passed, and the weights are shared between the networks. This encourages learning the joint distribution from both domains without using target domain labels.

PixelDA [4] uses adversarial architecture to reduce the shift between source and target distribution, to enable the classification of the target distribution samples. Instead of attempting to capture the domain's invariant features, this approach shifts the source domain to the target domain using GAN. Once trained, the model was able to produce and classify samples in the target domain. However, they are computationally expensive and limited by the dimensions of the input sample.

Adversarial discriminative domain adaptation (ADDA) [31] model uses a pretrained source encoder which is pretrained using labeled source domain images. The model uses the target and source images such that the discriminator cannot identify the origin of the domain from the images. The joint distribution of domains is learned from this and used for feature transformation and testing. Unsupervised classification of the target domain images is conducted based on parameters learned from these domain invariant features.

Selective adversarial networks (SANs) [6] are a type of deep adversarial adaptation proposed to perform partial transfer learning. It is used in applications where the target label is a subset of the source label category. Consequently, all the source domain labels are not present within the target domain. SAN can match source and target data distributions in a common subspace while segregating the outlier classes in the source domain. It is accomplished by increasing the alignment between similar data distributions in the latent space to the maximum degree possible.

3.3 Reconstruction Based

Deep reconstruction domain adaptation [12] is a method to improve adaptation which creates domain invariant features from each of the domains using an external reconstruction. It uses an encoder-decoder network architecture to classify and reconstruct the target distribution images. The higher representation of the target distribution is learned by the model trained on the source distribution and is used to categorize the target. After training, the reconstructed images from the original distribution show characteristics similar to the target distribution samples. Thus, it shows that the network has learned the joint distribution from both domains. This is then used to classify the target samples.

In the work [11] using an MTAE (Multitask autoencoder), the model reconstructs the images in different domains. This enables us to find the naturally occurring inter-domain variability when images are present across different domains. Such a process enables to find invariant features from both distributions. These features are then extracted from the different domains.

Domain separation networks (DSNs) [5] learn the higher-order representations and partition them into two subspaces. One of the subspaces is private to that particular domain, while the other subspace is shared across the domains. The partial representation is used to reproduce the image from both domains. The private and source representations are used in a shared decoder which learns to replicate the input domain distribution. The discriminability of the classes is increased by increasing the orthogonality constraints between the private and the shared subspaces components. The alignment of the domains is improved using the similarity loss between the shared subspace loss components.

3.4 Combination Based

Minimum discrepancy estimation [24] had used a combination of discrepancy-based losses of MMD and CORAL within the neural net to reduce the domain shift jointly.

Table 1 Table showing the summary view of all works

Method	Title	Year	Dataset	Acc%
Discrepancy-based	Learning transferable features with deep adaptation networks [20]	2015	Office 31	72.9
	Unsupervised domain adaptation with residual transfer networks [21]	2016	Office 31	73.7
	Mind the class weight bias: weighted maximum mean discrepancy for unsupervised domain adaptation [35]	2017	Office 31	72.1
	Return of frustatingly easy domain adaptation [27]	2016	Office 31	69.4
	Deep coral: Correlation alignment for deep domain adaptation [28]	2016	Office 31	72.1
	Joint distribution optimal transportation for domain adaptation [7]	2017	Caltech-office	80.04
	Deepjdot: Deep joint distribution optimal transport for unsupervised domain adaptation [8]	2018	MNIST MNIST-M	92.4
Adversarial-based	Unsupervised domain adaptation by backpropagation [9]	2015	MNIST MNIST-M	89.01
	Coupled generative adversarial networks [19]	2016	MNIST USPS	91.2
	Adversarial Discriminative Domain Adaptation [31]	2017	MNIST USPS	89.4
	Partial transfer learning with selective adversarial networks [6]	2017	Office 31	87.27
	Unsupervised pixel-level domain adaptation with generative adversarial networks [4]	2017	MNIST MNIST-M	98.2
Reconstruction-based	Deep reconstruction-classification networks for unsupervised domain adaptation [12]	2016	MNIST USPS	91.8
	Domain generalization for object recognition with multi-task autoencoders [11]	2015	Office Caltech	86.29
	Domain separation networks [5]	2016	MNIST MNIST-M	83.2
Combination-based	On minimum discrepancy estimation for deep domain adaptation [24]	2020	Office 31	74.6
	Dynamic Weighted Learning for Unsupervised Domain Adaptation [34]	2021	Office 31	87.1
	The domain shift problem of medical image segmentation and vendor-adaptation by unet-gan [36]	2019	Philips Siemens	83.5
	Cross-domain contrastive learning for unsupervised domain adaptation [33]	2022	Office 31	90.6
	Category contrast for unsupervised domain adaptation in visual tasks [17]	2022	Office 31	87.6
	Online unsupervised domain adaptation via reducing inter- and intradomain discrepancies adaptation [37]	2022	Office 31	85.3
Transformation-based	Direct domain adaptation through reciprocal linear transformations [1]	2021	MNIST MNIST-M	70
	Learning transferable parameters for unsupervised domain adaptation [16]	2022	Office 31	90.9

The source and target samples are sent simultaneously into the network. The encoded outputs from the network are taken from the higher representational layer and passed to the discrepancy-based adaptation losses. The losses from the MMD and CORAL loss are combined and backpropagated within the network for improved alignment within the domains.

Weighted DWL [34] utilizes a combination of criteria inspired by earlier works of adversarial and discrepancy-based models to build a DAN for improving the alignment. The network model focuses on two fronts: improving domain alignment and improving class discriminability within the domains. The class imbalance problem is also considered here by reweighing the samples before passing them into the network for training. This reduces the model bias due to the class imbalance of the training samples. The MMD is used to improve the alignment, and the linear discriminate analysis (LDA) loss improves the class discriminability.

Application of domain adaptation in medical research has been implemented by combining U-Net and adversarial-based GAN networks [36] to perform domain adaptation and image segmentation. The GAN is used for domain adaptation of the images (MRI scans) from different vendors, and U-Net is used to image segmentation.

Another application of the UDA combining technique is followed in OUDA [37], wherein the target domain is taken from the online streaming source. The methodology followed has two parts: The initial method involves reducing the discrepancy-based difference between the domains. A trained subspace is obtained from the initial phase, which captures domain invariant features through feature-level, sample-level, and domain-level adaptation. In the second part, the online classification uses the lower-dimensional alignment of the incoming target samples to the trained subspace. This is then used to reduce the intradomain distance between the online samples (target domain), which is then classified.

Unsupervised methods have also been developed and used with the earlier mentioned losses to overcome the target domain's label-free and source-free classifications. The work in [33] used a self-supervised method of implementation where the target was initially clustered based on KNN [23] and given pseudo-labels based on the source domain. These were then taken to perform contrastive learning by reducing the distance between the same classes from the target and increasing the distance between different classes. This was performed by decreasing intra-class and improving the inter-class distance, respectively.

A similar method that uses contrastive learning is followed in [17]. This approach creates a dictionary-like structure consisting of samples from the labeled source and unlabeled target domains. The samples from the unlabeled target domain are given pseudo-labels based on the source domain categories. The category contrast (CaCo) approach utilizes contrastive learning on the dictionary to reduce the distance between the same classes. The dictionary created for this purpose will also focus on class balance and class discriminability between the categories to minimize bias. Thus, the technique also accounts for class imbalance while training.

3.5 Transformation Based

Other than discrepancy-based loss functions, reconstruction, and GAN-based end-to-end architectures discussed earlier, another significant contribution to domain adaptation is the transformation applied as preprocessing on the input domain samples. A seminal work in this domain is the direct domain adaptation (DDA [1]). DDA using reciprocal linear transformation is a recent method based on prepossessing the input data to reduce the domain shift before the samples are passed through the network. The technique matches the signal-to-noise ratio between domains. The samples from the target domain are convoluted with the source domain and vice versa to reduce the shift. Thus, the domain shift is reduced outside the network before the training and testing.

A different form of transformation technique is explored under TransPar [16]. This work follows a method of identifying parameters in a network that learns the domain invariant features while training. Based on the lottery ticket hypothesis, the approach finds a network's transferable and untransferable parameters. The ratio of transferable parameters is inversely related to the domain shift distance between the distributions. The backpropagated weights from the loss identify the transferable parameters. The model then updates both parameters separately, focusing on the parameters that can generalize across the domains better. It is focused on reducing the domain-specific information learned by the network and can be integrated into current UDA-based models.

A summary of all works and their results has been given in Table 1. We have categorized them based on their approaches to aligning the domains. The performance comparison of these models is also compared to each other in Table 1. Their performance shows that adversarial methods produce the best results from all the models. Combination-based models also have results close to adversarial methods. These combinational models are often a combination of techniques involving adversarial and discrepancy-based methods. Reconstruction networks based on encoder-decoder architecture are also producing good results on diverse datasets.

4 Conclusions

This work has surveyed several papers on different approaches to domain adaptation techniques. In most real-world cases, it is likely that the target domain is not labeled and shares only a few classes similar to the source domain. This promises extensive research in unsupervised domain adaptation in the coming years.

5 Future Works

A promising direction in domain adaptation is adversarial networks and combinational models of different domain adaptation techniques. These types of approaches produce better results than most techniques in DA. Due to the more significant number of applications available for domain adaptation, much work in this direction is expected in the future.

References

1. Alkhalifah, T., Ovcharenko, O.: Direct Domain Adaptation Through Reciprocal Linear Transformations (2021)
2. Arjovsky, M., Chintala, S., Bottou, L.: Wasserstein generative adversarial networks. In International Conference on Machine Learning, pp. 214–223. PMLR (2017)
3. Ashokkumar, P., Don, S.: High dimensional data visualization: a survey. J. Adv. Res. Dyn. Control Syst. **9**(12), 851–866 (2017)
4. Bousmalis, K., Silberman, N., Dohan, D., Erhan, D., Krishnan, D.: Unsupervised pixel-level domain adaptation with generative adversarial networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3722–3731 (2017)
5. Bousmalis, K., Trigeorgis, G., Silberman, N., Krishnan, D., Erhan, D.: Domain separation networks. Adv. Neural Inform. Process. Syst. **29** (2016)
6. Cao, Z., Long, M., Wang, J., Jordan, M.I.: Partial transfer learning with selective adversarial networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2724–2732 (2018)
7. Courty, N., Flamary, R., Habrard, A., Rakotomamonjy, A.: Joint Distribution Optimal Transportation for Domain Adaptation (2017). arXiv preprint arXiv:1705.08848
8. Damodaran, B. B., Kellenberger, B., Flamary, R., Tuia, D., Courty, N.: Deepjdot: deep joint distribution optimal transport for unsupervised domain adaptation. In: Proceedings of the European Conference on Computer Vision (ECCV), pp. 447–463 (2018)
9. Ganin, Y., Lempitsky, V.: Unsupervised domain adaptation by backpropagation. In: International Conference on Machine Learning, pp. 1180–1189. PMLR (2015)
10. Ganin, Y., Ustinova, E., Ajakan, H., Germain, P., Larochelle, H., Laviolette, F., Marchand, M., Lempitsky, V.: Domain-adversarial training of neural networks. J. Mach. Learn. Res. **17**(1), 2096–2030 (2016)
11. Ghifary, M., Kleijn, W.B., Zhang, M., Balduzzi, D.: Domain generalization for object recognition with multi-task autoencoders. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 2551–2559 (2015)
12. Ghifary, M., Kleijn, W.B., Zhang, M., Balduzzi, D., Li, W.: Deep reconstruction-classification networks for unsupervised domain adaptation. In: European Conference on Computer Vision, pp. 597–613. Springer (2016)
13. Gopika, P., Sowmya, V., Gopalakrishnan, E.A., Soman, K.P.: Transferable approach for cardiac disease classification using deep learning. In: Deep Learning Techniques for Biomedical and Health Informatics, pp. 285–303. Elsevier (2020)
14. Gressel, G., Hrudya, P., Surendran, K., Thara, S., Aravind, A., Prabaharan, P.: Ensemble learning approach for author profiling. In: Notebook for PAN at CLEF, pp. 401–412 (2014)
15. Griffin, G., Holub, A., Perona, P.: Caltech-256 Object Category Dataset (2007)
16. Han, Z., Sun, H., Yin, Y.: Learning transferable parameters for unsupervised domain adaptation. IEEE Trans, Image Proces (2022)

17. Huang, J., Guan, D., Xiao, A., Lu, S., Shao, L.: Category contrast for unsupervised domain adaptation in visual tasks. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 1203–1214 (2022)
18. LeCun, Y., Bottou, L., Bengio, Y., Haffner, P.: Gradient-based learning applied to document recognition. Proc. IEEE **86**(11), 2278–2324 (1998)
19. Liu, M.-Y., Tuzel, O.: Coupled generative adversarial networks. Adv. Neural Inform. Proces. Syst. **29** (2016)
20. Long, M., Cao, Y., Wang, J., Jordan, M.: Learning transferable features with deep adaptation networks. In: International Conference on Machine Learning, pp. 97–105. PMLR (2015)
21. Long, M., Zhu, H., Wang, J., Jordan, M.I.: Unsupervised Domain Adaptation with Residual Transfer Networks (2016). arXiv preprint arXiv:1602.04433
22. Murugaraj, B., Amudha, J.: Performance assessment framework for computational models of visual attention. In: The International Symposium on Intelligent Systems Technologies and Applications, pp. 345–355. Springer (2017)
23. Peterson, L.E.: K-nearest neighbor. Scholarpedia **4**(2), 1883 (2009)
24. Rahman, M.M., Fookes, C., Baktashmotlagh, M., Sridharan, S.: On minimum discrepancy estimation for deep domain adaptation. In: Domain Adaptation for Visual Understanding, pp. 81–94. Springer (2020)
25. Saenko, K., Kulis, B., Fritz, M., Darrell, T.: Adapting visual category models to new domains. In: European Conference on Computer Vision, pp. 213–226. Springer (2010)
26. Sai, B.N.K., Sasikala, T.: Object detection and count of objects in image using tensor flow object detection api. In: 2019 International Conference on Smart Systems and Inventive Technology (ICSSIT), pp. 542–546. IEEE (2019)
27. Sun, B., Feng, J., Saenko, K.: Return of frustratingly easy domain adaptation. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 30 (2016)
28. Sun, B., Saenko, K.: Deep coral: Correlation alignment for deep domain adaptation. In: European Conference on Computer Vision, pp. 443–450. Springer (2016)
29. Tamuly, S., Jyotsna, C., Amudha, J.: Deep learning model for image classification. In: International Conference On Computational Vision and Bio Inspired Computing, pp. 312–320. Springer (2019)
30. Thampi, S.M., Piramuthu, S., Li, K.-C., Berretti, S., Wozniak, M., Singh, D.: Machine Learning and Metaheuristics Algorithms, and Applications: Second Symposium, SoMMA 2020, Chennai, India, 14–17 Oct 2020, Revised Selected Papers, vol. 1366. Springer Nature (2021)
31. Tzeng, E., Hoffman, J., Saenko, K., Darrell, T.: Adversarial discriminative domain adaptation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 7167–7176 (2017)
32. Venkateswara, H., Eusebio, J., Chakraborty, S., Panchanathan, S.: Deep hashing network for unsupervised domain adaptation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 5018–5027 (2017)
33. Wang, R., Wu, Z., Weng, Z., Chen, J., Qi, G.-J., Jiang, Y.-G.: Cross-domain contrastive learning for unsupervised domain adaptation. IEEE Trans, Multimed (2022)
34. Xiao, N., Zhang, L.: Dynamic weighted learning for unsupervised domain adaptation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 15242–15251 (2021)
35. Yan, H., Ding, Y., Li, P., Wang, Q., Xu, Y., Zuo, W.: Mind the class weight bias: Weighted maximum mean discrepancy for unsupervised domain adaptation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2272–2281 (2017)
36. Yan, W., Wang, Y., Gu, S., Huang, L., Yan, F., Xia, L., Tao, Q.: The domain shift problem of medical image segmentation and vendor-adaptation by unet-gan. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 623–631. Springer (2019)
37. Ye, Y., Pan, T., Meng, Q., Li, J., Tao Shen, H.: Online unsupervised domain adaptation via reducing inter-and intra-domain discrepancies. IEEE Trans. Neural Netw. Learn. Syst. (2022)