Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 11, No. 1, November : 2023

AN ANALYTICAL RESEARCH ON UNSUPERVISED DEEP DOMAIN ADAPTATION

Md. Talib Ahmad Assistant Professor, Katihar Engineering College, Katihar, Bihar, India
Email id- talib@keck.ac.in
Mr. Pranav Kumar Assistant Professor, Government Engineering College, Arwal, Bihar, India
Email id- pranavkmr79@gmail.com

Abstract-
For many tasks, deep learning has yielded state-of-the-art outcomes. Although these methods for supervised learning have shown good results, they rely on the assumption that training and testing data come from the same distribution, which isn't always the case. In addition to this problem, single-source unsupervised domain adaptation may deal with scenarios in which the objective of achieving good test-time performance on the target domain is achieved by training a network on labelled data from a source domain and unlabeled data from a related but distinct target domain. In order to lessen dependency on potentially expensive target data labels, a number of single-source and generally homogenous unsupervised deep domain adaptation algorithms have been developed. These approaches combine domain adaptation with the potent, hierarchical representations from deep learning. By looking at different techniques, the distinctive and similar components, outcomes, and theoretical insights, this survey will compare various strategies. We then examine potential application areas and avenues for future study.
Keywords— Supervised Learning, Domain Adaptation Algorithms, Expensive Target Data Labels, Approaches Combine Domain Adaptation.

INTRODUCTION

Perhaps the most common kind of machine learning is supervised learning, which has shown great promise in a wide range of application domains. The training and testing data are taken from the same distribution, which is a typical assumption made by many supervised learning techniques. A classifier trained on the source domain is likely to perform worse when tested on the target domain when this restriction is broken because of the differences between the two domains. The objective of single-source domain adaptation is to learn a concept from labeled data in a source domain that excels on a distinct but related target domain. Unsupervised domain adaptation focuses on the scenario in which training is conducted using labeled source data and solely unlabeled target data. Because of its ability to adapt labeled data for use in a new application, domain adaptation can reduce the need for costly labeled data in the target domain. As an example, consider the problem of semantically segmenting images. Each real image in the Cityscapes dataset required approximately 1.5 hours to annotate for semantic segmentation. In this case, human annotation time could be spared by training an image semantic segmentation model on synthetic street view images (the source domain) since these can be cheaply generated, then adapting and testing for real street view images (the target domain, here the Cityscapes dataset). An undeniable trend in machine learning is the increased usage of deep neural networks. Deep networks have produced many state-of-the-art results for a variety of machine learning tasks such as image classification, speech recognition, machine translation, and image generation. When trained on large amounts of data, these many-layer neural networks can learn powerful, hierarchical representations and can be highly scalable. At the same time, these networks can also experience performance drops due to domain shifts. Thus, much research has gone into adapting such networks from large labeled datasets to domains where little (or possibly no) labeled training data are available. These single-source and typically homogeneous unsupervised deep domain adaptation approaches, which combine the benefit of deep learning with the very practical use of domain adaptation to remove the reliance on potentially costly target data labels, will be the focus of this survey. A number of surveys have been created on the topic of domain adaptation and more generally transfer learning, of which domain adaptation can be viewed as a special case. Previous domain adaptation surveys lack depth of coverage and comparison of unsupervised deep domain adaptation

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 11, No. 1, November : 2023
approaches. In some cases, prior surveys do not discuss domain mapping, normalization statistic-based, or ensemble-based methods. In other cases, they do not survey deep learning approaches. Still others are application-centric, focusing on a single use case such as machine translation. One earlier survey focuses on the multi-source scenario, while we focus on the more prevalent single-source scenario. Transfer learning is a broader topic to cover, thus surveys provide minimal coverage and comparison of the deep learning methods that have been designed for unsupervised domain adaptation, or they focus on tasks such as activity recognition or reinforcement learning. The goal of this survey is to discuss, highlight unique components, and compare approaches to single-source homogeneous unsupervised deep domain adaptation. We first provide background on where domain adaptation fits into the more general problem of transfer learning. We follow this with an overview of generative adversarial networks (GANs) to provide background for the increasingly widespread use of adversarial techniques in domain adaptation. Next, we investigate the various domain adaptation methods, the components of those methods, and the results. Then, we overview domain adaptation theory and discuss what we can learn from the theoretical results. Finally, we look at application areas and identify future research directions for domain adaptation.

AN OVERVIEW OF DOMAIN ADAPTATION

(i) Transfer Learning- The focus of this survey is domain adaptation. Because domain adaptation can be viewed as a special case of transfer learning, we first review transfer learning to highlight the role of domain adaptation within this topic. Transfer learning is defined as the learning scenario where a model is trained on a source domain or task and evaluated on a different but related target domain or task, where either the tasks or domains (or both) different. For instance, we may wish to learn a model on a handwritten digit dataset (e.g., MNIST) with the goal of using it to recognize house numbers (e.g., SVHN). Or, we may wish to learn a model on a synthetic, heap-to-generate traffic sign dataset with the goal of using it to classify real traffic signs (e.g., GTSRB). In these examples, the source dataset used to train the model is related but different from the target dataset used to test the model both are digits and signs respectively, but each dataset looks significantly different. When the source and target differ but are related, then transfer learning can be applied to obtain higher accuracy on the target data.

(ii) Generative Adversarial Networks- Many deep domain adaptation methods that we will discuss in the next section incorporate adversarial training. We use the term adversarial training broadly to refer to any method that utilizes an adversary or an adversarial process during training. Before other adversarial methods were developed, the term was narrowly applied to training designed to improve the robustness of a model by
utilizing adversarial examples, e.g., image inputs with small worst-case perturbations that lead to misclassification. Subsequently, other techniques have arisen that also utilize an adversary during training, including generative-adversarial training of generative adversarial networks (GANs) and domain-adversarial training of domain adversarial neural networks (DANN), both of which have been used for domain adaptation. To provide background for the domain adaptation methods utilizing these techniques, we will first discuss GANs and later when discussing DANN note the differences.

Figure 1- Realistic but entirely synthetic images of human faces generated by a GAN trained on the CelebA-HQ dataset

METHODOLOGY

A rising number of novel unsupervised domain adaptation techniques have been put out in recent years, with a focus on neural network-based methods. Different research strands have evolved. These include focusing on making the model target discriminative by shifting the decision border into areas of lower data density, mapping between domains, separating normalization statistics, and aligning the distributions of the source and target domains. Furthermore, other people have looked at combining these strategies. We will outline each of these classifications along with the most recent techniques that fit inside them. As is often investigated, homogeneous domain adaptation that is, adaptation with a single source domain and a single destination domain will be the major topic of this survey. Multi-source domain adaptation is an additional scenario in which there are several source domains but only one destination domain. Since Sun et al.'s survey on multi-source domain adaptation, several more techniques have been created specifically for this situation. Multi-target domain adaptation is another option, however it is much less frequently explored. Similar to this, even if certain heterogeneous approaches have been established, we concentrate on homogeneous domain adaptation because of its predominance.

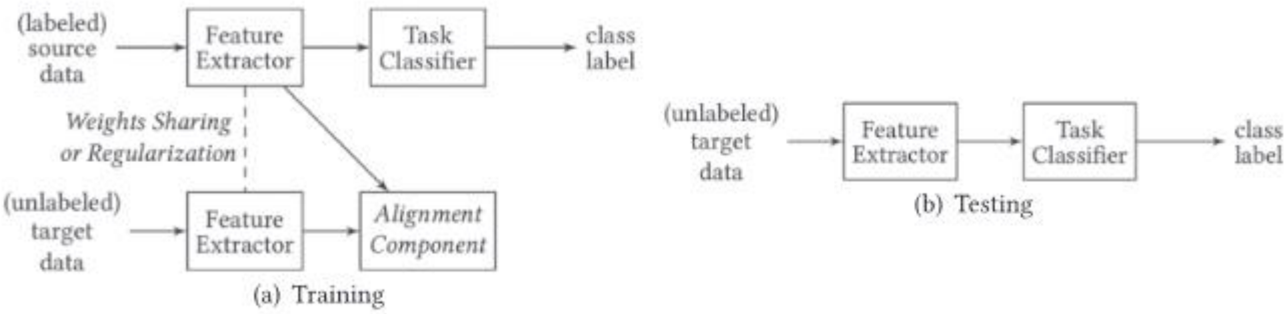

Figure 2- General network setup for domain adaptation methods learning domain-invariant features

(i) Domain-Invariant Feature Learning- Most recent domain adaptation methods align source and target domains by creating a domain- invariant feature representation, typically in the form of a feature extractor neural network. A feature representation is domain-invariant if the features follow the same distribution regardless of whether the input data are from the source or target domain. If a classifier can be trained to perform well on the source data using domain-invariant features, then the classifier may generalize well to the target domain since the features of the target data match those on which the classifier was trained. The general training and testing setup of these methods is illustrated in Figure 2. Methods differ in how they align the domains. Some minimize divergence, some perform reconstruction, and some employ adversarial training.

(ii) Domain Mapping- Mapping between domains is an option to building a domain-invariant feature representation. As will be covered at the conclusion of this section, the mapping is usually generated adversarially and at the pixel level (pixel-level adversarial domain adaptation). However, this is not always the case. A conditional GAN can be utilised to achieve this mapping. The generator converts a source input image to an image that closely mimics the target distribution in order to carry out adaptation at the pixel level. As seen in Figure 3, the GAN may, for instance, transform a picture of a synthetic car travelling into one that appears realistic. Then, using the known source labels, a classifier may be jointly trained with the GAN or trained on the source data mapped to the target domain. We will first go over the operation of a conditional GAN and then how domain adaptation can be achieved with it.

Figure 3- Synthetic vehicle driving image (left) adapted to look realistic (right)

(iii) Normalization Statistics- Most neural networks utilise normalisation layers, such batch norm. These offer advantages such as decreasing initialization sensitivity, enabling deeper networks to converge, flattening the optimisation landscape and making the gradients more Lipschitz, and enabling greater learning rates and therefore quicker training. Every batch norm layer sets a zero mean and unit variance for its input. Running averages of the batch norm parameters might be utilised during testing. Other alternatives have been created, such as group norm, which eliminates the dependency on batch size, and instance norm, which permits usage in recurrent neural networks. Nevertheless, do-main adaptation was not considered in the development of any of these normalisation strategies.

(iv) Ensemble Methods- When a base model, such a decision tree or neural network, is used, an ensemble of several models can frequently perform better than a single model by collecting votes for certain tasks, like classification, or by averaging the outputs of the models (like regression). This is due to the fact that various models are more likely to have distinct errors made by each model. Ensembles are widespread for some use situations, such as contests, but uncommon for comparing models since this performance advantage is correlated with an increase in computing cost due to the huge number of models to assess for each ensemble prediction. Several ensemble-based techniques have been developed for domain adaptation despite the associated costs; these techniques either use the ensemble predictions to direct learning or the ensemble to calculate the prediction confidence for target data pseudo-labeling.

(v) Target Discriminative Methods- The cluster assumption, which states that data points are dispersed in distinct clusters and that samples within each cluster share a common label, is one premise that has contributed to the success of semi-supervised learning methods. Should this be the case, then decision borders ought to be located in areas with low population density (that is, they shouldn't cross over areas with a lot of data points). Different domain adaptation techniques have been investigated to shift decision boundaries into lower density density areas. Usually, they have had adversarial training. Variational adversarial training (VAT), created by Miyato et al, and conditional entropy loss are used by Shu et al. in virtual adversarial domain adaptation (VADA) and Kumar et al. in co-regularized alignment (Co-DA). They are used in tandem because, in the absence of the entropy loss, VAT may lead to overfitting to the unlabeled data points, and in the presence of the entropy loss, VAT may prevent the network from becoming locally Lipschitz, which would prevent the decision boundary from shifting away from the data points.

THEORY OF UNSUPERVISED DEEP DOMAIN ADAPTATION

Having surveyed domain adaptation methods, we now address the question of when adaptation may be beneficial. Ben-David et al. develop a theory answering this in terms of an ideal predictor on both domains, Zhao et al. further this theory by removing the dependence on a joint ideal predictor while focusing on domain-invariant feature learning methods, and Le et al. develop theory looking beyond domain-invariant methods. These theoretical results can help answer two questions: (1) when will a classifier (or other predictor) trained on the source data per- form well on the target data and (2) given

a small number of labeled target examples, how can they best be used during training to minimize target test error? Answering the first question, labeled source data and unlabeled target data are both required (unsupervised). Answering the second question, additionally some labeled target data are required (semi-supervised). We will first review the theoretical bounds followed by a discussion of what insights these bounds provide into answering the above two questions. Ben-David et al. also address the case of multiple source domains, as do Mansour et al. In this article, we have focused on the cases containing only one source and one target (as is common in the methods we survey).
(i) Unsupervised- shared Hypothesis Space. Ben-David et al. propose setting a bound on the target error based on the source error and the divergence between the source and target domains. The empirical source error is easy to obtain by first training and then testing a classifier. However, the divergence between the domains cannot be directly obtained with standard methods like Kullback-Leibler divergence due to only having a finite number of samples from the domains and not assuming any particular distribution. Thus, an alternative is to measure it using a classifier induced divergence called H ΔH -divergence. Estimates of this divergence with finite samples converges to the real H ΔH -divergence. This divergence can be estimated by measuring the error when getting a classifier to discriminate between the unlabeled source and target exam-
ples; though, it is often intractable to find the theoretically required divergence upper bound.
(ii) Semi-Supervised- In the semi-supervised case, a linear combination of the source and target errors is computed,called the α-error. A bound can be calculated on the true α-error based on the empirical α-error. Finding the minimum α-error depends on the empirical α-error, the divergence between source and target, and the number of labeled source and target examples. Experimentation can be used to empirically determine the values of α that will perform well. Ben-David et al. also demonstrate the process on sentiment classification, illustrating that the optimum uses non-trivial values.

CONCLUSION

Deep neural networks are widely used for supervised learning; nevertheless, their training requires huge labelled datasets. Deep networks may be adapted to potentially smaller datasets with no goal labels by using unsupervised domain adaptation. To achieve this, a number of approach categories have been created, including ensemble-based, normalisation statistics-based, domain mapping, and domain-invariant feature learning techniques. As we have shown, there are certain distinctive and common aspects among these different strategies. Theoretical findings also provide some light on actual discoveries. Some of the approaches seem highly promising, but more work is needed for unique method combinations, enhanced bi-directional adaptation, direct comparisons, and usage with new datasets and applications.

REFERENCES

[1] Hana Ajakan, Pascal Germain, Hugo Larochelle, François Laviolette, and Mario Marchand. 2014. Domain-adversarial neural networks. arXiv preprint arXiv:1412.4446 (2014).
[2] Kei Akuzawa, Yusuke Iwasawa, and Yutaka Matsuo. 2019. Adversarial invariant feature learning with accuracy constraint for domain generalization. arXiv preprint arXiv:1904.12543 (2019).
[3] Youssef Alami Mejjati, Christian Richardt, James Tompkin, Darren Cosker, and Kwang In Kim. 2018. Unsupervised attention-guided image-to-image translation. In Advances in Neural Information Processing Systems 31, S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett (Eds.). Curran Associates, Inc., 3693–3703.
[4] Asha Anoosheh, Eirikur Agustsson, Radu Timofte, and Luc Van Gool. 2018. ComboGAN: Unrestrained scalability for image domain translation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops.
[5] Martin Arjovsky, Soumith Chintala, and Léon Bottou. 2017. Wasserstein generative adversarial networks. In Proceedings of the 34th International Conference on Machine Learning (Proceedings of

Machine Learning Research), Doina Precup and Yee Whye Teh (Eds.), Vol. 70. PMLR, 214–223. http://proceedings.mlr.press/v70/arjovsky17a.html.
[6] Sanjeev Arora, Rong Ge, Yingyu Liang, Tengyu Ma, and Yi Zhang. 2017. Generalization and equilibrium in generative adversarial nets (GANs). In Proceedings of the 34th International Conference on Machine Learning (Proceedings of Machine Learning Research), Doina Precup and Yee Whye Teh (Eds.), Vol. 70. PMLR, 224–232.
[7] Sanjeev Arora, Andrej Risteski, and Yi Zhang. 2018. Do GANs learn the distribution? Some theory and empirics. In Proceedings of the International Conference on Learning Representations. https://openreview.net/forum?id=BJehNfW0-.
[8] Amir Atapour-Abarghouei and Toby P. Breckon. 2018. Real-time monocular depth estimation using synthetic data with domain adaptation via image style transfer. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).
[9] Ben Athiwaratkun, Marc Finzi, Pavel Izmailov, and Andrew Gordon Wilson. 2019. There are many consistent explanations of unlabeled data: Why you should average. In Proceedings of the International Conference on Learning Representations. https://openreview.net/forum?id=rkgKBhA5Y7.
[10] Jimmy Lei Ba, Jamie Ryan Kiros, and Geoffrey E. Hinton. 2016. Layer normalization. arXiv preprint arXiv:1607.06450 (2016).
[11] Slawomir Bak, Peter Carr, and Jean-Francois Lalonde. 2018. Domain adaptation through synthesis for unsupervised person re-identification. In Proceedings of the European Conference on Computer Vision (ECCV).
[12] Oscar Beijbom. 2012. Domain adaptations for computer vision applications. arXiv preprint arXiv:1211.4860 (2012).
[13] Shai Ben-David, John Blitzer, Koby Crammer, Alex Kulesza, Fernando Pereira, and Jennifer Wortman Vaughan. 2010. A theory of learning from different domains. Machine Learning 79, 1 (01 May 2010), 151–175. DOI:https://doi.org/10.1007/s10994-009-5152-4
[14] Sagie Benaim and Lior Wolf. 2017. One-sided unsupervised domain mapping. In Advances in Neural Information Processing Systems 30, I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett (Eds.). Curran Associates, Inc., 752–762. http://papers.nips.cc/paper/6677-one-sided-unsupervised-domain-mapping.pdf.
[15] David Berthelot, Tom Schumm, and Luke Metz. 2017. Began: Boundary equilibrium generative adversarial networks. arXiv preprint arXiv:1703.10717 (2017).
[16] Mikołaj Bińkowski, Dougal J. Sutherland, Michael Arbel, and Arthur Gretton. 2017. Demystifying MMD GANs. In Proceedings of the International Conference on Learning Representations. https://openreview.net/forum?id=r1lUOzWCW.
[17] Gilles Blanchard, Gyemin Lee, and Clayton Scott. 2011. Generalizing from several related classification tasks to a new unlabeled sample. In Advances in Neural Information Processing Systems 24, J. Shawe-Taylor, R. S. Zemel, P. L. Bartlett, F. Pereira, and K. Q. Weinberger (Eds.). Curran Associates, Inc., 2178–2186. http://papers.nips.cc/paper/4312-generalizing-from-several-related-classification-tasks-to-a-new-unlabeled-sample.pdf.
[18] John Blitzer, Mark Dredze, and Fernando Pereira. 2007. Biographies, Bollywood, boom-boxes and blenders: Domain adaptation for sentiment classification. In Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics. 440–447.
[19] Ali Borji. 2018. Pros and cons of GAN evaluation measures. arXiv preprint arXiv:1802.03446 (2018).
[20] Konstantinos Bousmalis, Alex Irpan, Paul Wohlhart, Yunfei Bai, Matthew Kelcey, Mrinal Kalakrishnan, Laura Downs, Julian Ibarz, Peter Pastor, Kurt Konolige, Sergey Levine, and Vincent Vanhoucke. 2018. Using simulation and domain adaptation to improve efficiency of deep robotic grasping. In Proceedings of the 2018 IEEE International Conference on Robotics and Automation (ICRA). 4243–4250. DOI:https://doi.org/10.1109/ICRA.2018.8460875

[21] Konstantinos Bousmalis, Nathan Silberman, David Dohan, Dumitru Erhan, and Dilip Krishnan. 2017. Unsupervised pixel-level domain adaptation with generative adversarial networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).
[22] Konstantinos Bousmalis, George Trigeorgis, Nathan Silberman, Dilip Krishnan, and Dumitru Erhan. 2016. Domain separation networks. In Advances in Neural Information Processing Systems 29, D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon, and R. Garnett (Eds.). Curran Associates, Inc., 343–351. http://papers.nips.cc/paper/6254-domain-separation-networks.pdf.
[23] Denny Britz, Quoc Le, and Reid Pryzant. 2017. Effective domain mixing for neural machine translation. In Proceedings of the 2nd Conference on Machine Translation. 118–126.
[24] Lars Bungum and Björn Gambäck. 2011. A survey of domain adaptation in machine translation: Towards a refinement of domain space. In Proceedings of the India-Norway Workshop on Web Concepts and Technologies, Vol. 112.
[25] Jinming Cao, Oren Katzir, Peng Jiang, Dani Lischinski, Danny Cohen-Or, Changhe Tu, and Yangyan Li. 2018. Dida: Disentangled synthesis for domain adaptation. arXiv preprint arXiv:1805.08019 (2018).
[26] Zhangjie Cao, Mingsheng Long, Jianmin Wang, and Michael I. Jordan. 2018. Partial transfer learning with selective adversarial networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).
[27] Fabio Maria Carlucci, Lorenzo Porzi, Barbara Caputo, Elisa Ricci, and Samuel Rota Bulò. 2017. AutoDIAL: Automatic domain alignment layers. In Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV). IEEE, 5077–5085.
[28] Fabio M. Carlucci, Paolo Russo, Tatiana Tommasi, and Barbara Caputo. 2018. Agnostic domain generalization. arXiv preprint arXiv:1808.01102 (2018).
[29] Rich Caruana. 1997. Multitask learning. Machine Learning 28, 1 (01 Jul 1997), 41–75. DOI:https://doi.org/10.1023/A:1007379606734
[30] Olivier Chapelle and Alexander Zien. 2005. Semi-supervised classification by low density separation. In AISTATS, Vol. 2005. Citeseer, 57–64.
[31] Boxing Chen, Colin Cherry, George Foster, and Samuel Larkin. 2017. Cost weighting for neural machine translation domain adaptation. In Proceedings of the 1st Workshop on Neural Machine Translation. 40–46.
[32] Cheng Chen, Qi Dou, Hao Chen, and Pheng-Ann Heng. 2018. Semantic-aware generative adversarial nets for unsupervised domain adaptation in chest X-ray segmentation. In Machine Learning in Medical Imaging, Yinghuan Shi, Heung-Il Suk, and Mingxia Liu (Eds.). Springer International Publishing, Cham, 143–151.