Transfer learning and its application research

Yancong Zhou\textsuperscript{1*}, Xuemei Zhang\textsuperscript{2}, Yan Wang \textsuperscript{1}, Bo Zhang \textsuperscript{1}

\textsuperscript{1}School of information engineering, Tianjin University of Commerce, Tianjin 300134, China

\textsuperscript{2}School of science, Tianjin University of Commerce, Tianjin 300134, China

\*Corresponding author’s e-mail: zycong@tjcu.edu.cn

Abstract. Transfer learning relaxes the assumption’s limitation requirements of the independent and identical distribution of training data and test data in machine learning. It aims to help the target domain complete the learning task by learning one or more source domains similar to the target domain, solve the problem of scarcity of annotation data and enhance the model’s robustness and generalization performance. The article is a survey on the progress of transfer learning. According to "how to transfer", transfer learning is divided into four categories: instance-based transfer learning, feature-based transfer learning, model-based transfer learning, and relation-based transfer learning. The paper introduces the basic assumptions, main research questions, common methods, related research of various transfer learning algorithms, and the application of the transfer learning. Finally, we try to point out future research trends.

1. Introduction

In the era of big data, everyone is not only the center of data reception, but also the source of data generation. The data contains texts, images, videos, and graphs, etc., and its structure and probability distribution are also much different. Machine learning is one way that use the computer to find the underlying laws of data and obtain useful information in time. Traditional machine learning often assumes the test data and training data are independent and identically distributed (i.i.d). It means researchers need to re-collect new datasets and annotate them while studying a new field. It is time-consuming and labor-intensive and makes this hypothesis difficult to hold in practical applications. In other words, this hypothesis limits the use of models and datasets. Transfer learning can relax the limitation of the hypothesis, and help solve the problem of insufficient training data in machine learning. When researching the new field related to the previous research, the researchers can use the existing knowledge to study the new one and enhance the generalization ability and stability of the model.

In recent years, transfer learning has developed rapidly. It frequently appears in the top International Conferences, and is also widely used in various fields. Our contributions are as follows:

1) We introduce various classic transfer learning algorithms and recent research while pay more attention to feature-based transfer learning algorithms.

2) We introduce the application of transfer learning in image recognition, natural language processing, recommendation systems, collaborative filtering research, and related applications in other fields.

3) We try to point out the main problems of the current research, and the future research trends.
The rest of the paper is organized as follows: Sect. 2 gives the definition in the text, and introduces categories and related research of transfer learning. Sect. 3 introduces the applications. Sect. 4 briefly describes the current research directions and prospects of transfer learning. Sect. 5, the summary.

2. Transfer Learning

2.1. Transfer Learning: A Definition

The widely used definition of transfer learning[1-4] has two vital concepts: "domain" and "task". A domain \( \mathcal{D} \) consists of the feature space \( \mathcal{X} \) and the marginal probability distribution \( P(\mathcal{X}) \), identically equal(i.e., \( \mathcal{D} = \{ \mathcal{X}, P(\mathcal{X}) \} \), where sample \( \mathcal{X} = \{ x_1, ..., x_n \} \in \mathcal{X} \). Given the field \( \mathcal{D} \), the task \( \mathcal{T} \) is defined as a category space \( \mathcal{Y} \) and forecast model \( f(\mathcal{X}) \), i.e. \( \mathcal{T} = \{ \mathcal{Y}, f(\mathcal{X}) \} \). \( f(\mathcal{X}) \) can be interpreted statistically as a conditional probability distribution \( P(y|\mathcal{X}) \), i.e. \( f(\mathcal{X}) = P(y|\mathcal{X}) \), where \( y \in \mathcal{Y} \). The subscripts \( s \) and \( t \) in the following represent the source domain and the target domain respectively.

In current research, datasets collected from different sources are generally regarded as different domains where transfer learning can be applied for research across dataset.

2.2. Categories of Transfer Learning

In transfer learning, "when to transfer", "what to transfer" and "how to transfer" are three main research issues[1,5]. According to "how to transfer", it can be divided into four categories: instance-based[2-4], feature-based[6-12], model-based[13,9,12,14] and relation-based[15,16] transfer learning.

2.2.1. Instance-Based Transfer Learning

Instance-based transfer learning uses weighted combination or resampled source domain samples to study the target domain model, which is dependent on divergence for source domain and target domain. Because source domain instances always cannot be applied directly to the target domain when \( X_s = X_t \) and \( P(X_s) \neq P(X_t) \). Transfer AdaBoost learning Framework (TrAdaBoost)[2] is a classic instance-based transfer learning that addresses the problem of inductive transfer learning. It is suitable for transferring one source domain to one target domain when both domains are highly relevant. When the source and target domains are weakly correlated, transferring knowledge forcibly may lead to negative transfer. In response to this issue, YAO et al. [3] proposed the MultiSource-TrAdaBoost which was based on TrAdaBoost and transferred multiple source domains' knowledge to a target domain, that weakened the target domain classifier's requirement for domain relevance.

The above algorithms usually assume all training samples in the target domain are well known, however, in data flow situations training samples for the target area are often not available at once, so they are not applicable. The online transfer learning from multiple sources based on local classification accuracy algorithm(LC-MSOTL)[4] is suitable for data flow situations and can dynamically pick the most accurate local classifier from the source domain classifier. And the target domain classifier is weighted integrated by these source classifiers to categorize new target domain samples constantly coming.

2.2.2. Feature-Based Transfer Learning

Feature-based transfer learning assumes the data of the source and target domain can be mapped into a common feature space through a certain feature representation[5]. It is crucial to learn a pair of mapping functions that project source and target domain data into the common feature space \( \mathcal{X} \). There are three main methods to find these functions: first, reducing domain differences; second, finding the universal feature between domains; third, augmenting feature across domains through latent relationships.

Deep transfer learning is a kind of feature-based transfer learning. TAN et al.[11] reviewed and summarized the research, and defined the concept. They divided deep transfer learning into four
categories: instance-based, mapping-based, network-based, and adversarial-based deep transfer learning.

Network-based or adversarial-based deep transfer learning stems from the transferability of deep neural networks, which Yosinski et al. [6] have already demonstrated first in 2014. Their experiments have shown with the network layer closer to the head layer, the transferability of the layer decreases, and the fine-tuning of the parameters helps to improve the performance. JANG et al. [8] proposed a novel heterogeneous neural network transfer learning. ResNet-32 is the source domain model and VGG-9 is the target model that the layers before the pool layer in ResNet can be transferred to VGG-9.

Deep Adaptation Network(DAN)[7] was proposed based on Reference[6] where \( Y_s = Y_t \) and \( Y_s \) are known. In a big data context, if the source domain data collected by researchers is large enough, its category space may be richer than the target domain, i.e. \( Y_s \supset Y_t \). Reference[9] called this application scenario as partial domain adaptation. In this case, the traditional transfer learning algorithm adapts the entire source domain to the target domain that will cause a negative transfer. In contrast, the Selective Adversarial Network (SAN)[9] is more suitable for this situation, formed by the combination of Generative adversarial Network (GAN) and multiple classifiers, as shown in Figure 1 (a). The SAN has the constraints mainly from the two aspects of sample and category to avoid the negative transfer caused by source domain private data. The Partial Adversarial Domain Adaptation(PADA) [13] is also based on GAN and it uses the weight \( Y \) to adjust the network and continuously reduces the non-target domain category data, as shown in Figure 1(b). The weight makes the negative transfer influence weaker. The PADA and the SAN are suitable for situations where there is prior information in the relationship between the source domain category space and the target domain category space, which also restricts their application scenarios. Universal Adaptation Network(UAN)[12] further generalizes the application background of transfer learning, which is suitable for the scenario without prior information about the category space between domains.

2.2.3. Model-Based Transfer Learning

For model-based transfer learning, the source domain and the target domain can share models or model parameters. It assumes there is the universal knowledge of source domain task and target domain task. Some algorithms of the network-based deep transfer learning mentioned in 2.2.2, such as SAN, PADA, and UDA algorithms, are also model-based migration algorithms.

According to the assumptions behind transfer learning, it can be divided into two categories. One is sharing the parameter between domains and the other is using regularization[5].

Some models or parameter transfer methods will take advantage of both of them. For example, reference[14] proposed a domain adaptation algorithm based on extreme learning machine (ELM) parameter transfer, which is a parameter-based transfer learning algorithm. The key idea is to project the ELM parameters of the target domain into the source domain parameter space, making them the same as the classifier parameter distribution of the source domain to the greatest extent. It adds a regular item to the objective function to solve the negative transfer, too.

![Figure 1](image-url)

Figure 1. (a) The architecture of SAN for partial transfer learning. An input x is fed to F to output a feature, which is then input to a classifier G for prediction \( \hat{y} \) as the probability to assign each data
point \( \mathbf{x}_i \) to the \( \mathbf{y}_i \) domain classifier \( G^k_d \), \( K=1,2\ldots \mathbf{y}_i \). The feature is also input to K classifier \( G^k_d \) for determining each data point \( \mathbf{x}_i \) come from the source domain or the target domain. \( L^k_d \) is the loss of the \( k \)th domain classifier. \( L_d \) is weighted integration of the loss \( L^k_d \) of \( K \) domain classifiers according to probability \( \mathbf{y}_i \). (b) The architecture of PADA, which is composed of \( F, G \) and \( D \) (domain discriminator, like \( G^k_d \)). GRL stands for Gradient Reversal Layer.

2.2.4. Relation-Based Transfer Learning
The relation-based transfer learning approach is to find out the implied relationship between the source and target domains and transfer the source domain knowledge to the target domain according to certain rules. It assumes that there is a common relationship between the source and target domain data.

Current research is mainly using Markov logic networks (MLNs) to find the relationship between different domains. For example, the language-bias based transfer learning algorithm(LTL)[16] uses MLNs to describe the first-order logical relationship of the source domain, builds a sequential search tree, and then migrates to the target domain; the two-order-deep transfer learning algorithm(TODTLER)[15]directly computes the posterior distributions of all second-order formulas given by the data in the source domain and then uses those posterior distributions as prior distributions over second-order formulas in the target domain to train the MLN in the target domain.

3. Applications of Transfer Learning
Transfer learning is now widely used in image recognition, natural language processing, text mining, speech classification, and other application fields.

In terms of image recognition, reference[17] has given a selective learning algorithm(SLA) to solve the problem of distance transfer learning, such as using face images to identify aircraft images. Reference [18] inspired by the principle of Okom shavers, has proposed a non-parametric simple transfer learning based on the linear structure of the domain. Reference [14] has proposed a convolutional neural network corn disease image recognition model based on transfer learning to realize the recognition of corn disease images under the complex field background of small data samples, which is intelligent recognition with rust images.

In natural language processing and text mining, Transfer Denoising Autoencoder (TDA) and a Transfer Deep Network(TDN) [19] have been proposed, both of which are more accurate in complex text classification, message classification, and news comment classification than SVM, DAN. Reference [20] has put forward a fine-grained entity classification method for the task of entity classification with a lack labeled dataset, which built a semantic mapping relationship between a labeled entity type and unlabeled entity type and combined attention mechanisms to realize the transfer learning. Transfer Capsule Network (TransCap)[21] not only can transfer text-level sentiment polarity knowledge to aspect-level sentiment classification tasks but also can transfer text-level and aspect-level knowledge to sentences level text classification tasks.

In terms of recommender systems and collaborative filtering, reference [22] proposed a novel cross-system recommendation system framework that are combined of active learning and transfer learning. They proposed Maximum-Margin Matrix Factorization, Regularized Low-rank Matrix Factorization, and Probabilistic Matrix Factorization based on this framework, which had achieved higher accuracy on the actual comment data of Netflix and Douban. Reference [23] proposed a cross-domain recommender system based on kernel-induced knowledge transfer, called KerKT. It is suitable for application scenarios where the source domain and target domain entities partially overlap. Reference [24] used transfer learning to solve the problem of data scarcity faced by the maximum margin matrix factorization algorithm in the application of recommendation systems.

4. Research Outlook
Transfer learning is one of the emerging fields of machine learning with great potential. Based on the above summary, we think the subsequent study may focus on the following areas.
One is on universal domain adaptation. Abstracting the similarities of different application scenarios and designing universal domain adaptive algorithms will help to further enhance the generalization of machine learning algorithms; the second is the study of transfer learning evaluation indicators. The current research always uses accuracy of multiple transfer tasks on the benchmark dataset to evaluate the algorithms' transferability. The future research may use the degree of similarity and the diversity of shared features space between domains, which will be more comprehensive; The third is the interpretability study of transfer learning about how to explain the feasibility and effectiveness of transfer learning from the perspectives of mathematical reasoning and visualization.

5. Conclusion
This paper summarizes and analyzes the transfer learning study. Currently there is more research on feature-based transfer learning and model-based transfer learning, especially deep transfer learning algorithms. Every algorithm has its advantages, disadvantages, and scope of application. Deep learning is naturally suitable for transfer learning, and its adversarial network may become a paradigm of transfer learning, which is one of the hot topics in transfer learning. How to avoid negative transfer is one of the key issues in this field. Transfer learning has been widely applied in many areas and the application is expanding with the development of computer network, information technology, and its industries develop rapidly. Transfer learning has great development potential and good development prospects.

Acknowledgments
This work was supported by the National Natural Science Foundation of China [71572125];

References
[1] Pan, S.J., Yang, Q.(2010) A Survey on Transfer Learning. IEEE Transactions on Knowledge and Data Engineering,22(10): 1345-1359.
[2] Dai, W. ,Yang, Q. ,Xue, G.-R., Yu, Y. (2007) Boosting for transfer learning. In: Proceedings of the 24th international conference on Machine learning. Corvalis Oregon. pp. 193-200
[3] Yao, Y., Doretto, G.(2010) Boosting for transfer learning with multiple sources. In: the Proceedings of the 24th international conference on Machine learning. Corvalis Oregon. pp. 193-200 .
[4] Tang, S., Wen, Y., Qin, Y.(2019) Online Transfer Learning from Multiple Sources Based on Local Classification Accuracy. Journal of Software, 28(11): 2940-2960.
[5] Yang, Q. ,Zhang, Y. ,Dai, W., Pan, S.J.(2020) Transfer learning. Cambridge University Press,
[6] Yosinski, J. ,Clune, J. ,Bengio, Y., Lipson, H. (2014) How transferable are features in deep neural networks? In: Advances in neural information processing systems. British Columbia. pp. 3320-3328
[7] Long, M. ,Cao, Y. ,Wang, J., Jordan, M. (2015) Learning transferable features with deep adaptation networks. In: International conference on machine learning. Lille. pp. 97-105
[8] Jang, Y. ,Lee, H. ,Hwang, S.J., Shin, J. (2019) Learning What and Where to Transfer. In: International Conference on Machine Learning. Long Beach. pp. 3030-3039.
[9] Cao, Z. ,Long, M. ,Wang, J., Jordan, M.I. (2018) Partial transfer learning with selective adversarial networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake. pp. 2724-2732.
[10] Sun, B. ,Yang, L. ,Lin, M., Young, C. ,Dong, J. ,Zhang, W., Dong, P. (2019) SuperChat: dialogue generation by transfer learning from vision to language using two-dimensional word embedding. In: Proceedings of the 1st International Workshop on Deep Learning Practice for High-Dimensional Sparse Data. Anchorage Alaska. pp. 1-4.
[11] Tan, C. ,Sun, F. ,Kong, T. ,Zhang, W. ,Yang, C., Liu, C. (2018) A survey on deep transfer learning. In: International conference on artificial neural networks. Rhodes. pp. 270-279.
[12] You, K., Long, M., Cao, Z., Wang, J., Jordan, M.I. (2019) Universal domain adaptation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Long Beach. pp. 2720-2729.

[13] Cao, Z., Ma, L., Long, M., Wang, J. (2018) Partial adversarial domain adaptation. In: Proceedings of the European Conference on Computer Vision (ECCV). Munich. pp. 135-150.

[14] Xu, S., Cai D., Luo C. (2018) Domain adaptation algorithm with ELM parameter transfer. Acta Automatica Sinica, 44(02): 311-317.

[15] Van Haaren, J., Kolobov, A., Davis, J. (2016) Todtler: Two-order-deep transfer learning. In: Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence. Austin. pp. 3007-3015.

[16] Kumaraswamy, R., Odom, P., Kersting, K., Leake, D., Natarajan, S. (2015) Transfer Learning via Relational Type Matching. In: the 2015 IEEE International Conference on Data Mining. Atlantic. pp. 811-816.

[17] Tan, B., Zhang, Y., Pan, S.J., Yang, Q. (2017) Distant domain transfer learning. In: Thirty-first AAAI conference on artificial intelligence. San Francisco. pp. 2604-2610.

[18] Wang, J., Chen, Y., Yu, H., Huang, M., Yang, Q. (2019) Easy transfer learning by exploiting intra-domain structures. In: 2019 IEEE International Conference on Multimedia and Expo (ICME). Shanghai. pp. 1210-1215.

[19] Long, M., Cao, Y., Cao, Z., Wang, J., Jordan, M.I. (2019) Transferable Representation Learning with Deep Adaptation Networks. IEEE Trans Pattern Anal Mach Intell, 41(12): 3071-3085.

[20] Feng, J., Guo, Z., Liu, Y., Ma X. (2019) Fine-grained Entity Type Classification Based on Transfer Learning. Acta Automatica Sinica. http://kns.cnki.net/kcms/detail/11.2109.TP.20190917.1704.007.html (2019).

[21] Chen, Z., Qian, T. (2019) Transfer capsule network for aspect level sentiment classification. In: Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics. Florence. pp. 547-556.

[22] Zhao, L., Pan, S.J., Yang, Q. (2017) A unified framework of active transfer learning for cross-system recommendation. Artificial Intelligence, 245: 38-55.

[23] Zhang, Q., Lu, J., Wu, D., Zhang, G. (2019) A Cross-Domain Recommender System With Kernel-Induced Knowledge Transfer for Overlapping Entities. IEEE Transactions on Neural Networks and Learning Systems, 30(7): 1998-2012.

[24] Veeramachaneni, S.D., Pujari, A.K., Padmanabhan, V., Kumar, V. (2019) A Maximum Margin Matrix Factorization based Transfer Learning Approach for Cross-Domain Recommendation. Applied Soft Computing, 85: 105751.