



Transfer Learning-Assisted Survival Analysis of Breast Cancer Relying on the Spatial Interaction Between Tumor-Infiltrating Lymphocytes and Tumors

Yawen Wu¹, Yingli Zuo¹, Qi Zhu¹, Jianpeng Sheng², Daoqiang Zhang^{1(✉)},
and Wei Shao^{1(✉)}

¹ College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, MIIT Key Laboratory of Pattern Analysis and Machine Intelligence, Nanjing 211106, China

{dqzhang, shaowei20022005}@nuaa.edu.cn

² School of Medicine, Zhejiang University, Zhejiang 310058, China

Abstract. Whole-Slide Histopathology Image (WSI) is regarded as the gold standard for survival prediction of Breast Cancer (BC) across different subtypes. However, in cancer prognosis applications, the cost of acquiring patients' survival information is high and can be extremely difficult in practice. By considering that there exists a certain common mechanism for tumor progression among different subtypes of Breast Invasive Carcinoma (BRCA), it becomes critical to utilize data from a related subtype of BRCA to help predict the patients' survival in the target domain. To address this issue, we proposed a TILs-Tumor interactions guided unsupervised domain adaptation (T2UDA) algorithm to predict the patients' survival on the target BC subtype. Different from the existing feature-level or instance-level transfer learning strategy, our study considered the fact that the tumor-infiltrating lymphocytes (TILs) and its correlation with tumors reveal similar role in the prognosis of different BRCA subtypes. More specifically, T2UDA first employed the Graph Attention Network (GAT) to learn the node embeddings and the spatial interactions between tumor and TILs patches in WSI. Then, besides aligning the embeddings of different types of nodes across the source and target domains, we proposed a novel Tumor-TILs interaction alignment (TTIA) module to ensure that the distribution of interaction weights are similar in both domains. We evaluated the performance of our method on the BRCA cohort derived from the Cancer Genome Atlas (TCGA), and the experimental results indicated that T2UDA outperformed other domain adaption methods for predicting patients' clinical outcomes.

Y. Wu and Y. Zuo—Contribute equally to this work

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023
H. Greenspan et al. (Eds.): MICCAI 2023, LNCS 14225, pp. 612–621, 2023.
https://doi.org/10.1007/978-3-031-43987-2_59

Keywords: Tumor-infiltrating Lymphocytes · Unsupervised Domain Adaption · Prognosis Prediction · Graph Attention Network · Breast Cancer

1 Introduction

Breast cancer (BC) is the most common cancer diagnosed among females and the second leading cause of cancer death among women after lung cancer [1]. The BC differs greatly in clinical behavior, ranging from carcinoma in situ to aggressive metastatic disease [2, 3]. Thus, effective and accurate prognosis of BC as well as stratifying cancer patients into different subgroups for personalized cancer management has attracted more attention than ever before.

Among different types of imaging biomarkers, histopathological images are generally considered the golden standard for BC prognosis since they can confer important cell-level information that can reflect the aggressiveness of BC [4]. Recently, with the availability of digitalized whole-slide pathological images (WSIs), many computational models have been employed for the prognosis prediction of various subtypes of BC. For instance, Lu et al [5] presented a novel approach for predicting the prognosis of ER-Positive BC patients by quantifying nuclear shape and orientation from histopathological images. Liu et al [6] developed a gradient boosting algorithm to predict the disease progression for various subtypes of BC. However, due to the high-cost of collecting survival information from the patients, it is still a challenge to build effective machine learning models for specific BC subtypes with limited annotation data.

To deal with the above challenges, several researchers began to design domain adaption algorithms, which utilize the labeled data from a related cancer subtype to help predict the patients' survival in the target domain. Specifically, Alirezazadeh et al [7] presented a new representation learning-based unsupervised domain adaption method to predict the clinical outcome of cancer patients on the target domain. Zhang et al [8] proposed a collaborative unsupervised domain adaptation algorithm, which conducts transferability-aware adaptation and conquers label noise in a collaborative way. Other studies include Xu et al [9] developed graph neural networks for unsupervised domain adaptation in histopathological image analysis, based on a backbone for embedding input images into a feature space, and a graph neural layer for propagating the supervision signals of images with labels.

Although much progress has been achieved, most of the existing studies applied the feature alignment strategy to reduce the distribution difference between source and target domains. However, such transfer learning methods neglected to take the interaction among different types of tissues into consideration. For example, it is widely recognized that tumor-infiltrating lymphocytes (TILs) and its correlation with tumors reveal a similar role in the prognosis of different BRCA subtypes. For instance, Kurozumi et al [10] revealed that high TILs expression was correlated with negative estrogen receptor (ER) expression and high histological grade ($P < 0.001$). Lu et al [11] utilized the TILs spatial pattern for survival analysis in different breast cancer subtypes including

ER-negative, ER-positive, and triple-negative. It can be expected that better prognosis performance can be achieved if we leveraged the TILs-Tumor interaction information to resolve the survival analysis task on the target domain.

Based on the above considerations, in this paper, we proposed a TILs-Tumor interactions guided unsupervised domain adaptation (T2UDA) algorithm to predict the patients’ survival on the target BC subtype. Specifically, T2UDA first applied the graph attention network (GATs) to learn node embeddings and the spatial interactions between tumor and TILs patches in WSI. In order to preserve the node-level and interaction-level similarities across different domains, we not only aligned the embedding for different types of nodes but also designed a novel Tumor-TILs interaction alignment (TTIA) module to ensure that the distribution of the interaction weights are similar in both domains. We evaluated the performance of our method on the Breast Invasive Carcinoma (BRCA) cohort derived from the Cancer Genome Atlas (TCGA), and the experimental results indicated that T2UDA outperforms other domain adaption methods for predicting patients’ clinical outcomes.

2 Method

We summarized the proposed T2UDA network in Fig. 1, which consists of three parts, *i.e.*, Graph Attention Network-based Framework, Feature Alignment(FA), and TILs-Tumor interaction alignment(TTIA). Next, we will introduce each part in detail.

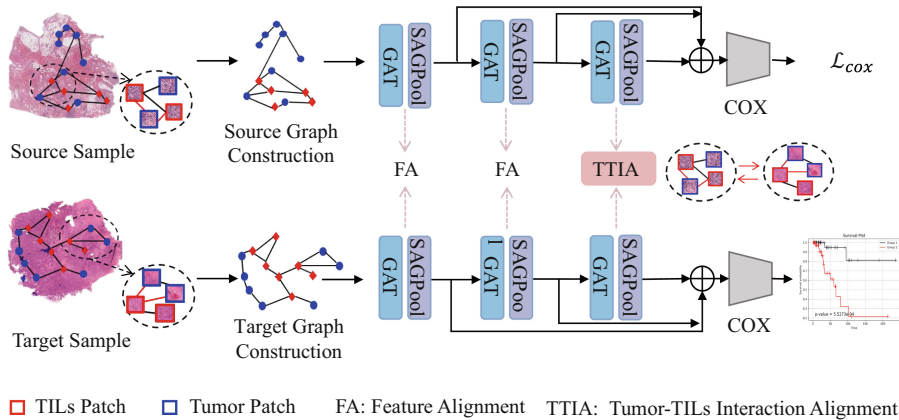


Fig. 1. The overall framework of the T2UDA network. T2UDA aligns TILs-Tumor edge interaction weights in the TTIA module and feature vectors in the FA module to reduce the discrepancy between different domains and achieving prognosis task on target BC patients.

Data Pre-processing. We obtained valid patches of 512×512 pixels from pathological images and segment the TILs and tumor tissues using a pre-trained U-Net++ model. Then we calculated the tumor and TILs area ratios in each patch and selected 300 patches with the largest ratios of each tissue type. Based on the selected patches, we constructed a graph $G = (V, E)$ for each WSI. Here, given patches as nodes V , we first calculated the pairwise distance among different nodes, and select the top 10 percent connections with the smallest distance values as edges E . For each node in V , we followed the study in [12], which applied the ResNet-101 model to extract node features. Then, the principal component analysis (PCA) is implemented to reduce dimensionalities of the node features to 128.

Calculating TILs-Tumor Interaction via Graph Attention Networks(GATs). To characterize the interaction between different TILs and tumor patches, we employed GAT [13], which has been proven to be useful in describing the spatial interaction between different tissues across WSIs. Our GAT-based framework consisted of 3 GAT layers interleaved with 3 graph pooling layers [14](shown in Fig. 1). The input of the GAT layer are $H = [H^L, H^T] \in R^{d \times 600}$, $H^w = [h_1^w, h_2^w, \dots, h_{300}^w] \in R^{d \times 300}$, $w \in \{L, T\}$ represent the output features for TILs and tumor nodes after PCA. The GAT layer generate a new group of node features $H' = [(H^L)', (H^T)']$ via a weight matrix $W \in R^{d \times d'}$ and $(H^w)' \in R^{d' \times 300}$, $w \in \{L, T\}$. Next, with a shared attentional mechanism $\mathbb{R}^{d'} \times \mathbb{R}^{d'} \rightarrow \mathbb{R}$, we calculated the attention coefficients among different nodes, which can be formulated as:

$$e_{ij} = a([Wh_i || Wh_j]), j \in \mathcal{N}_i. \quad (1)$$

Furthermore, a softmax function was then adopted to normalize the attention coefficients e_{ij} :

$$\alpha_{ij} = \text{softmax}_j(e_{ij}) = \frac{\exp(e_{ij})}{\sum_{k \in \mathcal{N}_i} \exp(e_{ik})}, \quad (2)$$

where \mathcal{N}_i represents all neighbors of node i . The new feature vector v_i for node i was calculated via a weighted sum:

$$v_i = \sigma \left(\sum_{v_j \in N(v_i)} \alpha_{ij} Wh_j \right). \quad (3)$$

Finally, the output features of each GAT layer were aggregated in the readout layer. We fed the generated output features from each readout layer into the Cox hazard proportional regression model for the final prognosis predictions.

Feature Alignment. In the proposed GAT-based transfer learning framework, the feature alignment component was employed on its first two layers. Then, for the node embeddings with different types (TILs and Tumor) in both the source and target domain, we performed a mean pooling operation to obtain their

aggregated features. Next, we aligned the aggregated tumor or TILs features from the two domains separately using Maximum Mean Discrepancy(MMD) [15]. Here, we adopted MMD for feature alignment due to its ability to measure the distance between two distributions without explicit assumptions on the data distribution, we showed the objective function of MMD in our method as follows:

$$L_{FA} = \sum_{r=1,2} \sum_{k \in L, T} \left\| \frac{1}{n} \sum_{i=1}^n (f_{i,k})^r - \frac{1}{m} \sum_{i=1}^m (f'_{i,k})^r \right\|_H^2 \quad (4)$$

where H is a Hilbert space, f represents the features from the source, f' represents the feature from the target, r represents the layer number, $k \in \{L, T\}$ referred to TILs or tumor node. In addition, n denotes the number of source samples, while m refers to the number of target samples.

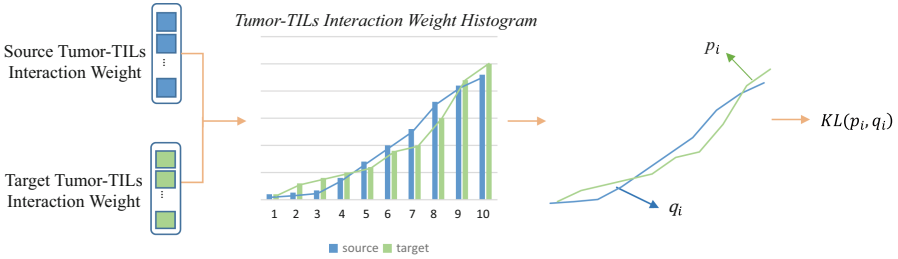


Fig. 2. The illustration of the proposed Interaction Weight Alignment module.

TILs-Tumor Interaction Alignment. To accurately characterize the interaction between TILs and tumors, we further analyzed the extracted interaction weights by dividing them into 10 intervals (i.e., bins). For each interval, we calculated the sum of all source domain interaction weights as i_k^s and the sum of all target domain interaction weights as i_k^t , where k represents the k -th interval. Consequently, we obtained two vectors and applied softmax on each of them for normalization that can be denoted as $p_i = [i_1^s, i_2^s, \dots, i_{10}^s]$ and $q_i = [i_1^t, i_2^t, \dots, i_{10}^t]$. In order to measure the dissimilarity between p_i and q_i , the Kullback-Leibler (KL) divergence is adapted on the third layer of GAT, which can be formulated as:

$$L_{TTIA} = KL(p_i, q_i) = \sum p_i \log(p_i/q_i). \quad (5)$$

According to Eq.(5), we can ensure that the weight distributions for the TIL-Tumor interaction are consistent in the source and target domain, which will benefit the following survival analysis. It is beneficial for the target domain.

Prognosis Prediction by the Cox Proportional Hazard Model. The Cox proportional hazard model was applied to predict the patients’ clinical outcome [16], and its negative log partial likelihood function can be formulated as:

$$l_{\text{prognosis}} = \sum_{i=1}^N \delta_i \left(\theta^T x_i - \log \sum_{j \in R(t_i)} \exp(\theta^T x_j) \right) \quad (6)$$

where x_i represents the output of the last layer for the prognosis task and $R(t_i)$ is the risk set at time t_i , which represents the set of patients that are still under risk before time t . In addition, δ_i is an indicator variable. Sample i refers to censored patient if $\delta_i = 0$, otherwise $\delta_i = 1$.

Overall Objective. To achieve domain-adaptive prognosis prediction, the final loss function included the Cox loss, FA loss, and TTIA loss as the following formula:

$$\mathbf{L}_t = \mathbf{L}_{\text{cox}} + \alpha \mathbf{L}_{\text{FA}} + \beta \mathbf{L}_{\text{TTIA}}, \quad (7)$$

where α and β represent the weights assigned to the importance of FA component and TTIA component respectively.

3 Experiments and Results

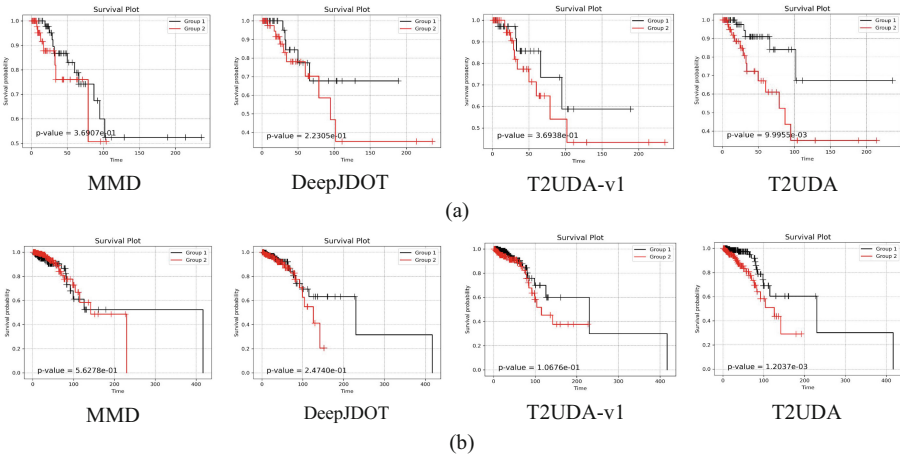
Datasets. We conducted our experiments on the breast invasive carcinoma (BRCA) dataset from The Cancer Genome Atlas (TCGA). Specifically, the BRCA dataset includes 661 patients with hematoxylin and eosin (HE)-stained pathological imaging and corresponding survival information. Among the collected BRCA patients in TCGA, the number of ER positive(ER+) and ER negative(ER−) patients are 515 and 146, respectively. We hope to investigate if the proposed T2UDA could be used to help improve the prognosis performance of (ER+) or (ER−) with the aid of the survival information on its counterpart.

3.1 Implementation Details and Evaluation Metrics

The dimension of intermediate layers in GAT was 256. The pooling ratio in SagPool was set to 0.7. α and β were tuned from $\{0.01, 0.1\}$. During training, the model was trained for 150 epochs for both the main experiment and all comparative experiments. We used the Adam optimizer with a learning rate tuned from $\{1e-5, 1e-4\}$. We evaluated the performance of our model using the Concordance Index (CI) and Area Under the Curve (AUC) as performance metrics. Both CI and AUC range from 0 to 1, with larger values indicating better prediction performance and vice versa [1].

Table 1. Quantitative Performance Comparison between Different Unsupervised Domain Adaptation Methods and Our Method.

	ER+ \rightarrow ER-		ER+ \rightarrow ER-	
	CI	AUC	CI	AUC
Source only	0.6187	0.6135	0.5771	0.6054
DDC	0.6118	0.6202	0.5854	0.6382
DANN	0.6414	0.6456	0.6274	0.6206
DeepJDOT	0.6535	0.6633	0.6357	0.6428
MDD	0.6498	0.6651	0.6261	0.6367
T2UDA-v1	0.6314	0.6450	0.6266	0.6427
T2UDA	0.6933	0.7097	0.6803	0.6903

**Fig. 3.** The survival curves by applying different methods on two benchmark settings: (1) source domain is ER+ BC and Target domain is ER- BC in (a); (2) source domain is ER- BC and Target domain is ER+ BC in (b).

3.2 Result and Discussion

In this study, we compared the performance of our proposed model with several existing domain adaptation methods, including 1) DDC [17]: Utilize the Maximum Mean Discrepancy (MMD) to calculate the domain difference loss between source and target data and optimize both classification loss and disparity loss. 2) DANN [18]: An adversarial learning method that used gradient backpropagation to extract domain-independent features. 3) MDD [19]: An adversarial training method that combined metric learning and domain adaptation. 4) DeepJDOT [20]: An Unsupervised Domain Adaptation method based on optimal transport that simultaneously learns features and optimizes classifiers by measuring joint feature/label differences. 5) Source only: it was trained on

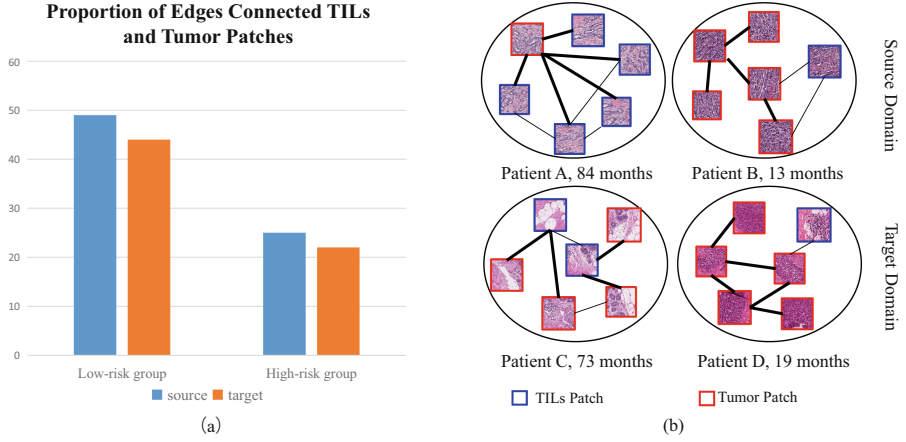


Fig. 4. (a): Compare proportion of Edges Connected TILs and Tumor patches of source and target domain (b): Compare interactions between TILs patches and tumor patches for long survival patients and short survival patients of source and target domain. The thick black line indicates the edges with higher weight.

the source domain and applied directly to the target domain. 6) T2UDA-v1: it was a variant of T2UDA which didn't use TTIA. The experimental results were presented in Table 1.

The results presented in Table 1 revealed several key observations. First, our proposed method outperformed feature alignment-based methods such as DDC and DeepJDOT in terms of both CI and AUC values. The reason lies in that these methods only transferred the knowledge at the feature level and neglected the inter-relationship between TILs and tumors. Second, our method outperformed adversarial-based methods such as DANN and MDD, as the high heterogeneity between the target and source domains results in negative transfer through adversarial training. Instead of directly aligning regions, our proposed method focused on similar TILs-Tumor interactions and aligning patches of the same tissue.

We also evaluated the contributions of the key components of our framework and found that T2UDA performed better than Source only and T2UDA-v1, which shows the advantage of minimizing differences in TILs-Tumor interaction weights.

In addition, we also evaluated the patient stratification performance of different methods. As shown in Fig. 3, our proposed T2UDA outperformed feature alignment-based methods (such as DDC and DeepJDOT), adversarial-based methods (such as DANN and MDD), and T2UDA-v1 in stratification performance, proving that considering the interaction between TILs and tumors as migration knowledge leads to better prognostic results.

We also examined the consistency of important edges in each group of stratified patients based on the TILs-Tumor interaction weights calculated by the

GAT-based framework in the source and target domains. As seen in Fig. 4(a), for both the source and target domains, the proportion of edges that connect TILs and tumor regions in the low-risk group was higher than that in the high-risk group, showing that the interaction between TILs and tumors played a critical role in prognostic prediction in different BC subtypes. Furthermore, as shown in Fig. 4(b), the weights of the edges connecting tumor and TILs regions were higher for patients in the low survival risk group in both source and target domains. This was consistent with our knowledge that brisk interaction between TILs and tumor regions indicates a better clinical outcome and demonstrates the transferability of this knowledge.

4 Conclusion

In this paper, we presented an unsupervised domain adaptation algorithm that leverages TILs-Tumor interactions to predict patients' survival in a target BC subtype(T2UDA). Our results demonstrated that the relationship between TILs and tumors is transferable and can be effectively used to improve the accuracy of survival prediction models. To the best of our knowledge, T2UDA was the first method to successfully achieve interrelationship transfer between TILs and tumors across different cancer subtypes for prognosis tasks.

Acknowledgements. This work was supported by the National Natural Science Foundation of China (Nos.62136004, 61902183, 61876082, 61732006, 62076129), the National Key R&D Program of China (Grant Nos.2018YFC2001600, 2018YFC2001602).

References

1. Shao, W., Wang, T., Huang, Z., Han, Z., Zhang, J., Huang, K.: Weakly supervised deep ordinal cox model for survival prediction from whole-slide pathological images. *IEEE Trans. Med. Imaging* **40**(12), 3739–3747 (2021)
2. Okabe, M., et al.: Predictive factors of the tumor immunological microenvironment for long-term follow-up in early stage breast cancer. *Cancer Sci.* **108**(1), 81–90 (2017)
3. Mizukami, Y., et al.: Detection of novel cancer-testis antigen-specific t-cell responses in til, regional lymph nodes, and pbl in patients with esophageal squamous cell carcinoma. *Cancer Sci.* **99**(7), 1448–1454 (2008)
4. Yawen, W., et al.: Recent advances of deep learning for computational histopathology: principles and applications. *Cancers* **14**(5), 1199 (2022)
5. Cheng, L., et al.: Nuclear shape and orientation features from h&e images predict survival in early-stage estrogen receptor-positive breast cancers. *Lab. Invest.* **98**(11), 1438–1448 (2018)
6. Liu, P., Fu, B., Yang, S.X., Deng, L., Zhong, X., Zheng, H.: Optimizing survival analysis of xgboost for ties to predict disease progression of breast cancer. *IEEE Trans. Biomedical Eng.* **68**(1), 148–160 (2020)

7. Alirezazadeh, P., Hejrati, B., Monsef-Esfahani, A., Fathi, A.: Representation learning-based unsupervised domain adaptation for classification of breast cancer histopathology images. *Biocybern. Biomed. Eng.* **38**(3), 671–683 (2018)
8. Zhang, Y., et al.: Collaborative unsupervised domain adaptation for medical image diagnosis. *IEEE Trans. Image Process.* **29**, 7834–7844 (2020)
9. Xu, D., Cai, C., Fang, C., Kong, B., Zhu, J., Li, Z.: Graph neural networks for unsupervised domain adaptation of histopathological image analytics. *arXiv preprint arXiv:2008.09304* (2020)
10. Kurozumi, S., et al.: Prognostic significance of tumour-infiltrating lymphocytes for oestrogen receptor-negative breast cancer without lymph node metastasis. *Oncol. Lett.* **17**(3), 2647–2656 (2019)
11. Zixiao, L., et al.: Deep-learning-based characterization of tumor-infiltrating lymphocytes in breast cancers from histopathology images and multiomics data. *JCO Clin. Cancer Informat.* **4**, 480–490 (2020)
12. Zuo, Y.: Identify consistent imaging genomic biomarkers for characterizing the survival-associated interactions between tumor-infiltrating lymphocytes and tumors. In: *Medical Image Computing and Computer Assisted Intervention-MICCAI 2022: 25th International Conference, Singapore, 18–22 September 2022, Proceedings, Part II*, pp. 222–231. Springer (2022). https://doi.org/10.1007/978-3-031-16434-7_22
13. Veličković, P., Cucurull, G., Casanova, A., Romero, A., Lio, P., Bengio, Y.: Graph attention networks. *arXiv preprint arXiv:1710.10903* (2017)
14. Lee, J., Lee, I., Kang, J.: Self-attention graph pooling. In: *International Conference on Machine Learning*, pp. 3734–3743. PMLR (2019)
15. Borgwardt, K.M., Gretton, A., Rasch, M.J., Kriegel, H.-P., Schölkopf, B., Smola, A.J.: Integrating structured biological data by kernel maximum mean discrepancy. *Bioinformatics* **22**(14), e49–e57 (2006)
16. Shao, W., et al.: Integrative analysis of pathological images and multi-dimensional genomic data for early-stage cancer prognosis. *IEEE Trans. Med. Imaging* **39**(1), 99–110 (2019)
17. Tzeng, E., Hoffman, J., Zhang, N., Saenko, K., Darrell, T.: Deep domain confusion: Maximizing for domain invariance. *arXiv preprint arXiv:1412.3474* (2014)
18. 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), 2030–2096 (2016)
19. Zhang, Y., Liu, T., Long, M., Jordan, M.: Bridging theory and algorithm for domain adaptation. In: *International Conference on Machine Learning*, pp. 7404–7413. PMLR (2019)
20. Damodaran, B.B., Kellenberger, B., Flamary, R., Tuia, D., Courty, N.: DeepJDOT: deep joint distribution optimal transport for unsupervised domain adaptation. In: Ferrari, V., Hebert, M., Sminchisescu, C., Weiss, Y. (eds.) *ECCV 2018. LNCS*, vol. 11208, pp. 467–483. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-01225-0_28