WANCDR: Wasserstein adversarial Network for Cancer Drug Response

Hanjun Choi[1] and Mansu Kim[1,*]

Department of AI convergence, Gwangju Institute of Science and Technology, Korea
mansu.kim@gist.ac.kr

Abstract. Predicting patient-specific drug responses from preclinical cell-line data remains challenging due to significant heterogeneity between preclinical (cell-line) and clinical (patient) gene expression profiles. In this study, we propose WANCDR, a novel adversarial neural network framework designed to improve the generalization of drug-response predictions by aligning latent representations across preclinical and clinical domains. Specifically, we introduce a domain alignment module trained adversarially, which enforces the encoder to generate domain-invariant latent embeddings. Extensive experiments conducted on preclinical (GDSC) and clinical (TCGA) datasets demonstrate that WANCDR achieves robust predictive performance on preclinical data, while substantially outperforming existing approaches in clinical generalization, particularly when classifying responses for previously unseen drugs. Qualitative analyses via UMAP visualization further validate the superior domain alignment capability of WANCDR. Collectively, these results highlight the potential of WANCDR to bridge the translational gap from preclinical insights to clinical applications.

Keywords: Graph neural network · Cancer drug response · Wasserstein Adversarial Network

1 Introduction

Cancer drug response (CDR) prediction aims to identify how effectively a tumor will respond to a given anticancer treatment, and is a critical component of precision oncology [5]. Accurate CDR predictions help clinicians to select proper therapies based on individual patient profiles, improving clinical outcomes and reducing adverse side effects. However, substantial genetic and transcriptomic variability among tumors makes reliable predictions challenging [8]. To address this issue, it is critical to develop robust computational methods that capture complex biological variations.

Recent studies on CDR prediction have adopted deep learning-based approaches due to their effectiveness in handling complex biological data. For example, DeepCDR integrates multi-omics data with drug structural information using graph convolutional networks (GCNs). PANCDR employs adversarial learning techniques to reduce the domain gap between preclinical and clinical datasets

[11, 9]. Despite these advancements, existing methods still face limitations when applied to clinical patient data due to differences between training (preclinical) and testing (clinical) domains.

A significant limitation of current methods is that they are trained on preclinical datasets (e.g., Genomics of Drug Sensitivity in Cancer, GDSC), resulting in decreased prediction performance when evaluated on clinical cohorts (e.g., The Cancer Genome Atlas, TCGA) [4, 15]. Clinical data typically contain fewer samples, more diverse patient profiles, and distinct biological characteristics compared to preclinical datasets, leading to significant distributional shifts. For example, PANCDR employs a conventional adversarial framework to address these issues; however, it often suffers from instability during training and issues such as mode collapse, which limits its practical effectiveness [9]. Similarly, popular domain adaptation techniques borrowed from computer vision, including Gradient Reversal Layers (DANN) and correlation alignment (Deep CORAL), often struggle to effectively model the nuanced biological differences between preclinical and clinical data [4, 15].

To overcome these limitations, we propose a Wasserstein Adversarial Network-based model for CDR prediction (WANCDR), specifically designed to align latent feature distributions between preclinical and clinical data robustly. Our main scientific contributions are summarized as follows: (1) We propose a Wasserstein Generative Adversarial Network based CDR prediction model that effectively aligns latent feature distributions between preclinical and clinical data. (2) By employing a critic network to minimize the Wasserstein distance, our model ensures stable adversarial training, thus addressing common issues such as mode collapse and vanishing gradients found in conventional adversarial learning frameworks [6]. (3) We demonstrate the effectiveness of our domain alignment approach through extensive seen-unseen CDR experiments.

2 Material and Methodology

2.1 Data preparation

Source Dataset: GDSC We use preclinical data obtained from the GDSC database [8]. Raw gene expression profiles for cancer cell lines are obtained from ArrayExpress (E-MTAB-3610), and drug response annotations (Sensitive or Resistant) are computed by binarizing IC_{50} values using LOBICO [10]. After removing samples without valid drug response annotations or drug identifiers (PubChem ID or SMILES), the final dataset contains 112,575 instances spanning 950 cell lines and 151 drugs, with an approximate Sensitive-to-Resistant ratio of 1:7.5. For drug representation, SMILES strings are converted into molecular graphs using RDKit (https://www.rdkit.org), and feature and adjacency matrices are computed using DeepChem [13].

Target Dataset: TCGA Gene expression profiles are downloaded from the GDAC Broad Institute (http://gdac.broadinstitute.org/), and clinical drug response annotations are acquired from Ding et al. [3]. Specifically, drug responses

are categorized as Sensitive (Complete Response/Partial Response) or Resistant (Progressive Disease/Stable Disease). Following MOLI's preprocessing strategy [14], cases involving multi-drug treatments are excluded, leaving only single-agent treatments. As a result, the final external test set comprises 666 instances from 569 patients treated with 69 distinct drugs, with a Sensitive-to-Resistant ratio of approximately 1:1. Additionally, we use 9,424 unlabeled primary solid tumor samples from TCGA for adversarial training, which helps to align feature distributions between preclinical and clinical datasets [9].

2.2 Wasserstein Adversarial Network for Cancer Drug Response (WANCDR)

The proposed WANCDR model comprises two primary components: a Wasserstein GAN (WGAN) for domain alignment and a Graph Convolutional Network (GCN)-based predictor for cancer drug response prediction. First, WANCDR utilizes a WGAN framework [1, 7] to robustly align latent genomic feature distributions between the source (cell-line) and target (patient) domains. Second, structural drug information is embedded through a GCN and integrated with the aligned genomic representation to accurately estimate drug response, quantified by IC_{50} values.

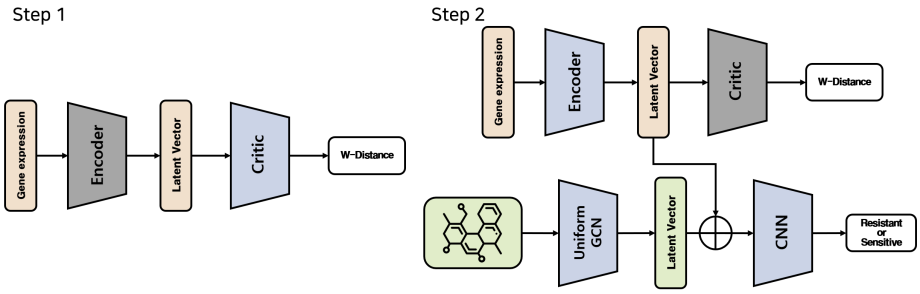

Fig. 1. WANCDR architecture: An MLP-based encoder generates latent representations of gene expression, while a GCN-based predictor captures drug features. A WGAN critic distinguishes between cell-line and patient-derived latent vectors. The training alternates between Step 1: freezing encoder/predictor and training the critic, and Step 2: freezing critic and jointly training encoder and predictor.

WGAN-Based Domain Alignment Module We employ a Wasserstein Generative Adversarial Network (WGAN) framework [1, 7] to effectively address the domain distribution mismatch between preclinical (cell-line) and clinical (patient) datasets. The WGAN framework is particularly suitable for our task, as it robustly aligns latent distributions by mitigating common adversarial training

challenges such as mode collapse and gradient instability through enforcing Lipschitz continuity constraints via gradient penalties. This approach provides enhanced training stability compared to traditional GAN-based methods.

Formally, let $E(\cdot)$ denote the MLP encoder network mapping gene expression profiles into latent feature representations. Specifically, given input profiles x_{GDSC} and x_{TCGA}, we have latent embeddings $z_g = E(x_{\mathrm{GDSC}})$ and $z_t = E(x_{\mathrm{TCGA}})$, which follow latent distributions $z_g \sim \mathbb{P}_z^{\mathrm{GDSC}}$ and $z_t \sim \mathbb{P}_z^{\mathrm{TCGA}}$, respectively. To quantify the discrepancy between these distributions, we introduce a critic network $C : z \mapsto \mathbb{R}$, which approximates the Wasserstein-1 distance:

$$W_1(\mathbb{P}_z^{\mathrm{TCGA}}, \mathbb{P}_z^{\mathrm{GDSC}}) = \sup_{\|C\|_L \leq 1} \left[\mathbb{E}_{z_t \sim \mathbb{P}_z^{\mathrm{TCGA}}}[C(z_t)] - \mathbb{E}_{z_g \sim \mathbb{P}_z^{\mathrm{GDSC}}}[C(z_g)]\right].$$

To enforce the critic's 1-Lipschitz constraint effectively, we adopt the gradient penalty regularization term from WGAN-GP [7], defined as:

$$\Omega_{\mathrm{GP}} = \mathbb{E}_{\hat{z} \sim P_{\hat{z}}}\left[(\|\nabla_{\hat{z}} C(\hat{z})\|_2 - 1)^2\right], \quad \hat{z} = \varepsilon \cdot z_t + (1 - \varepsilon)\, z_g,\ \varepsilon \sim U(0, 1).$$

Thus, the final objective for training the critic network becomes:

$$L_{\mathrm{critic}} = \mathbb{E}_{z_t \sim P_z^{\mathrm{TCGA}}}[C(z_t)] - \mathbb{E}_{z_g \sim P_z^{\mathrm{GDSC}}}[C(z_g)] + \lambda\, \Omega_{\mathrm{GP}} \tag{1}$$

where λ denotes a hyperparameter controlling the gradient penalty strength. By optimizing this objective, WANCDR effectively aligns the latent feature distributions of cell-line and patient data, thereby enhancing model generalization and predictive robustness across domains.

Graph Convolutional Network-based Prediction Module We employ a GCN-based prediction module that integrates genomic and structural drug information to accurately estimate CDR. Specifically, we adapt the GCN+MLP predictor architecture from PANCDR [9], which jointly models genomic features and drug structures. First, drug SMILES strings are converted into molecular graphs using RDKit, and atom-level features generated by DeepChem [13] are processed through a GCN to update node embeddings. A global average pooling layer is then applied to aggregate these node embeddings into a fixed-length drug representation h_{drug}.

Formally, let $GCN(\cdot)$ represent the GCN-based embedding network for structural drug information. Given a molecular graph (G), constructed from SMILES strings using RDKit, GCN generates a fixed-length embedding vector:

$$h_{\mathrm{drug}} = GCN(\mathrm{G}).$$

We then combine two latent embeddings into a unified representation:

$$h_{\mathrm{combined}} = h_{\mathrm{drug}} \oplus z_{\mathrm{i}}, i \in \{t, g\}$$

which is further processed by a fully-connected layer followed by a 1D convolutional layer, as described previously in PAN-CDR [9], to capture high-level

interactions between drug and genomic features. Finally, a sigmoid activation function predicts the probability of drug sensitivity ($\hat{y}$):

$$\hat{y} = \sigma(\text{Conv1D}(\text{FC}(h_{\text{combined}}))).$$

The prediction module is trained by minimizing the binary cross-entropy loss:

$$\mathcal{L}_{\text{pred}} = -\left[y \log \hat{y} + (1-y) \log(1-\hat{y})\right]. \tag{2}$$

By integrating the domain-aligned genomic representations, z_{expr}, with structural drug embeddings, h_{drug}, we accurately predict drug responses and enhance the generalization performance across both clinical and preclinical populations.

Loss Function We alternately minimize two losses to optimize our model. First, we update the *Domain Alignment Module* by minimizing the critic loss $\mathcal{L}_{\text{critic}}$, defined in Eq. 1. Subsequently, we optimize the *prediction module* by minimizing the joint predictive and adversarial loss:

$$\mathcal{L}_{\text{joint}} = \mathcal{L}_{\text{pred}} + \lambda \mathcal{L}_{\text{adv}}, \tag{3}$$

where the predictive loss $\mathcal{L}_{\text{pred}}$ is defined in Eq. 2, and λ is a hyperparameter balancing predictive accuracy and domain alignment. The adversarial loss is given by:

$$\mathcal{L}_{\text{adv}} = -\mathbb{E}_{x_{GDSC} \sim \mathbb{P}_{\text{GDSC}}}[C(E(x_{GDSC}))],$$

which encourages the encoder $E(\cdot)$ to generate domain-invariant latent representations by restricting the critic's capability to discriminate between preclinical (cell-line) and clinical (patient) data. This joint optimization ensures robust predictive performance and effective alignment of latent features across domains.

3 Experiments and results

3.1 Experiments setup

To assess the classification performance of WANCDR, we conduct experiments on the preclinical GDSC dataset using a stratified 5-fold cross-validation approach. Specifically, among the 5 folds, four are assigned as the training set, from which 5% is set as an internal validation set for early stopping; the remaining fold serves as the test set. The classification performance is evaluated in terms of Area Under the Curve (AUC), F1-score, Precision, and Recall. To ensure the robustness of the results, we repeat this entire cross-validation process 10 times and report the average outcomes.

Furthermore, to assess the model's generalization to clinical data, we perform additional evaluations using the clinical TCGA dataset. Briefly, we train WANCDR on 95% of the GDSC preclinical dataset, reserving 5% as an internal validation set to determine optimal hyperparameters via random search [2]. The model trained on GDSC is then directly applied to the external TCGA clinical

dataset. Similar to the preclinical experiment setup, we measure the model's ability to classify Sensitive and Resistant patient-drug pairs using AUC, F1-score, Precision, and Recall. This procedure is also repeated 10 times to ensure robust and reliable estimates of clinical generalization performance.

3.2 Results on Preclinical (GDSC) Dataset

We evaluate the predictive performance of our proposed model on the preclinical GDSC dataset. The comparative results are summarized in Table 1. In detail, WANCDR achieve an AUC of 0.815 ± 0.010, accuracy (ACC) of 0.753 ± 0.023, precision of 0.284 ± 0.019, recall of 0.723 ± 0.032, and F1-score of 0.407 ± 0.016. Although DeepCDR slightly outperform WANCDR in terms of AUC (0.832 ± 0.002), ACC (0.762 ± 0.017), precision (0.295 ± 0.014), recall (0.737 ± 0.022), and F1-score (0.421 ± 0.011), WANCDR shows competitive performance on preclinical data, demonstrating the effectiveness of the proposed approach.

Table 1. Comparison of classification performance between WANCDR and DeepCDR on preclinical (GDSC) and clinical (TCGA) datasets. Results are reported in the format of mean (standard deviation).

Dataset	Model	AUC	ACC	Precision	Recall	F1
GDSC	WANCDR	0.815(0.010)	0.753(0.023)	0.284(0.019)	0.723(0.032)	0.407(0.016)
	DeepCDR	**0.832(0.002)**	**0.762(0.017)**	**0.295(0.014)**	**0.737(0.022)**	**0.421(0.011)**
TCGA	WANCDR	**0.672(0.032)**	**0.644(0.024)**	**0.656(0.015)**	**0.561(0.096)**	**0.601(0.059)**
	DeepCDR	0.550(0.030)	0.574(0.028)	0.609(0.070)	0.416(0.143)	0.496(0.054)

3.3 Results on Clinical (TCGA) Dataset

To evaluate the clinical generalization capability of WANCDR, we train the model on the preclinical GDSC dataset and assess its predictive performance directly on the external clinical TCGA dataset. As summarized in Table 2, WANCDR significantly outperforms DeepCDR across all performance metrics, achieving an AUC of 0.672 ± 0.032, accuracy (ACC) of 0.644 ± 0.024, precision of 0.656 ± 0.015, recall of 0.561 ± 0.096, and F1-score of 0.601 ± 0.059.

Table 2. Comparison of classification performance for seen and unseen drug-gene expression pairs on the clinical TCGA dataset. Results are reported in the format of mean (standard deviation).

Dataset	Model	AUC	ACC	Precision	Recall	F1
Seen	WANCDR	**0.690(0.048)**	**0.660(0.029)**	**0.671(0.027)**	**0.569(0.144)**	**0.604(0.087)**
	DeepCDR	0.554(0.044)	0.595(0.045)	0.639(0.097)	0.431(0.121)	0.496(0.054)
Unseen	WANCDR	**0.632(0.017)**	**0.622(0.015)**	**0.641(0.031)**	**0.623(0.143)**	**0.622(0.059)**
	DeepCDR	0.534(0.040)	0.568(0.022)	0.597(0.058)	0.614(0.239)	0.573(0.107)

Furthermore, we evaluate the model's classification performance specifically on unseen drug-gene expression pairs. As presented in Table 2, WANCDR consistently demonstrates superior results for both previously seen and unseen drug classifications. Specifically, we observe the robust predictive performance achieved on unseen drugs (AUC: 0.632 ± 0.017, ACC: 0.622 ± 0.015, and F1-score: 0.622 ± 0.059), highlighting WANCDR's promising potential for generalization to novel therapeutic agents.

These findings clearly demonstrate that WANCDR provides robust predictive performance and enhanced clinical generalization, effectively bridging the domain gap between preclinical (cell-line) and clinical (patient) data.

3.4 Domain Alignment Analysis

To qualitatively assess the effectiveness of domain alignment achieved by WANCDR, we visualize latent embeddings of the GDSC (blue) and TCGA (orange) datasets using UMAP [12] (Fig. 2). As shown in Fig. 2-(a), latent embeddings from the two domains form clearly separated clusters before training, indicating a considerable domain gap between preclinical (GDSC) and clinical (TCGA) data. After training with DeepCDR (Fig. 2-(b)), the embeddings within each domain become more structured; however, they still remain distinctly separated, highlighting DeepCDR's limited capacity for cross-domain generalization. In contrast, WANCDR (Fig. 2-(c)) generates a cohesive latent space wherein the two domains substantially overlap, forming a continuous manifold that effectively bridges GDSC and TCGA samples. This significant intermingling demonstrates the successful domain-invariant representation learning capability of WANCDR. The qualitative observation aligns closely with the quantitative evaluations, where WANCDR shows comparable performance with DeepCDR on the preclinical GDSC dataset and significantly superior performance on the clinical TCGA dataset.

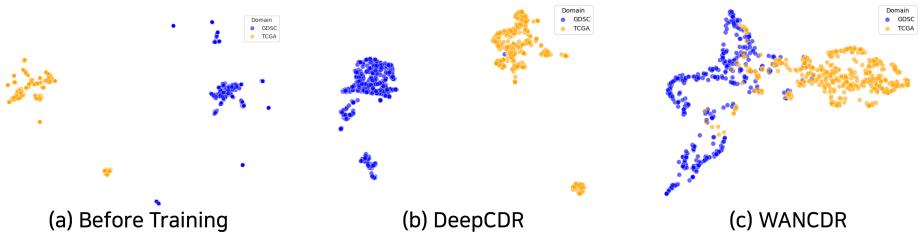

Fig. 2. UMAP of latent gene-expression embeddings for GDSC (blue) and TCGA (orange): (a) Before training, (b) DeepCDR, (c) WANCDR.

4 Conclusion

In this study, we propose WANCDR improving cross-domain generalization in cancer drug response prediction through latent representation alignment between preclinical (GDSC) and clinical (TCGA) gene expression profiles. Experimental results demonstrate that WANCDR achieves competitive predictive performance on the preclinical dataset while substantially outperforming baseline methods in clinical generalization, particularly for unseen therapeutic agents. Qualitative analysis using UMAP visualizations further confirms the effective alignment of latent features across domains, highlighting WANCDR's capability to bridge the domain gap between preclinical cell-line data and clinical patient data. Our findings collectively indicate that WANCDR represents a promising approach towards translating preclinical insights into clinical practice.

References

1. Arjovsky, M., Chintala, S., Bottou, L.: Wasserstein gan. In: International Conference on Machine Learning (ICML) (2017)
2. Bergstra, J., Bengio, Y.: Random search for hyper-parameter optimization. The journal of machine learning research **13**(1), 281–305 (2012)
3. Ding, Z., Zu, S., Gu, J.: Evaluating the molecule-based prediction of clinical drug responses in cancer. Bioinformatics **32**(19), 2891–2895 (2016)
4. Ganin, Y., Ustinova, E., Ajakan, H., Germain, P., Larochelle, H., Laviolette, F., March, M., Lempitsky, V.: Domain-adversarial training of neural networks. Journal of machine learning research **17**(59), 1–35 (2016)
5. Garnett, M.J., Edelman, E.J., Heidorn, S.J., Greenman, C.D., Dastur, A., Lau, K.W., Greninger, P., Thompson, I.R., Luo, X., Soares, J., et al.: Systematic identification of genomic markers of drug sensitivity in cancer cells. Nature **483**(7391), 570–575 (2012)
6. Goodfellow, I.J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y.: Generative adversarial nets. Advances in neural information processing systems **27** (2014)
7. Gulrajani, I., Ahmed, F., Arjovsky, M., Dumoulin, V., Courville, A.C.: Improved training of wasserstein gans. In: Advances in Neural Information Processing Systems (NeurIPS) (2017)
8. Iorio, F., Knijnenburg, T.A., Vis, D.J., Bignell, G.R., Menden, M.P., Schubert, M., Aben, N., Gonçalves, E., Barthorpe, S., Lightfoot, H., et al.: A landscape of pharmacogenomic interactions in cancer. Cell **166**(3), 740–754 (2016)
9. Kim, J., Park, S.H., Lee, H.: Pancdr: precise medicine prediction using an adversarial network for cancer drug response. Briefings in Bioinformatics **25**(2), bbae088 (2024)
10. Knijnenburg, T.A., Klau, G.W., Iorio, F., Garnett, M.J., McDermott, U., Shmulevich, I., Wessels, L.F.: Logic models to predict continuous outputs based on binary inputs with an application to personalized cancer therapy. Scientific reports **6**(1), 36812 (2016)
11. Liu, Q., Hu, Z., Jiang, R., Zhou, M.: Deepcdr: a hybrid graph convolutional network for predicting cancer drug response. Bioinformatics **36**(Supplement_2), i911–i918 (2020)

12. McInnes, L., Healy, J., Melville, J.: Umap: Uniform manifold approximation and projection for dimension reduction. arXiv preprint arXiv:1802.03426 (2018)
13. Ramsundar, B., Eastman, P., Walters, P., Pande, V.: Deep learning for the life sciences: applying deep learning to genomics, microscopy, drug discovery, and more. O'Reilly Media (2019)
14. Sharifi-Noghabi, H., Zolotareva, O., Collins, C.C., Ester, M.: Moli: multi-omics late integration with deep neural networks for drug response prediction. Bioinformatics **35**(14), i501–i509 (2019)
15. Sun, B., Saenko, K.: Deep coral: Correlation alignment for deep domain adaptation. In: European conference on computer vision. pp. 443–450. Springer (2016)