Supporting Information: AACDR: Integrating Graph Isomorphism Networks and Asymmetric Adversarial Domain Adaptation for Cancer Drug Response Prediction

Yi Zhang,[†,§] Tingfang Wu,[*,†,‡,¶,§] Liangpeng Nie,[†] Yelu Jiang,[†] Geng Li,[†] Jia Xu,[†] Zhenglong Zhou,[†] Yiwei Chen,[†] Lijun Quan,[*,†,‡,¶] and Qiang Lyu[*,†,‡,¶]

†*School of Computer Science and Technology, Soochow University, Ganjiang East Streat 333, 215006, Jiangsu, China*

‡*Province Key Lab for Information Processing Technologies, Soochow University, Ganjiang East Streat 333, 215006, Jiangsu, China*

¶*Collaborative Innovation Center of Novel Software Technology and Industrialization, Organization, Nanjing, 210000, Jiangsu, China*

§*Co-first authors*

E-mail: tfwu@suda.edu.cn; ljquan@suda.edu.cn; qiang@suda.edu.cn

The supplementary file consists of 15 Pages and 7 Tables in total:

Table S1: Kolmogorov-Smirnov and Mann-Whitney U Tests results for verifying the presence of the OOD (out-of-distribution) issue.

Table S2 & Table S3: Comparison of PANCDR and AACDR with Multiple Hyperparameter Configurations on the **PDX Dataset** (With and Without Domain Adaptation).

Table S4: Evaluation of AACDR interpretability through SHAP and cancer-type-specific

driver gene validation.

Table S5: Bootstrapped confidence interval analysis for AACDR and PANCDR.

Table S6 & Table S7: Reconstruction loss function and data preprocessing analysis.

Performance of AACDR under 100 random initializations

We trained and evaluated AACDR on the **Expr Dataset** with 100 runs using random initialization. The final results were divided into two parts: "Seen" and "Unseen", based on whether the drugs in the test entries appeared in the training set. The final results can be found on AACDR.

Comparison Between AACDR and PANCDR Across 28 Hyperparameter Configurations

We trained and tested 28 hyperparameter settings of AACDR and PANCDR on the **Expr Dataset**. For each hyperparameter combination, to mitigate the impact of randomness, we performed 10 random initializations and recorded the mean AUC on the test set. The detailed results can be found on AACDR.

Statistical Evidence of Distribution Shift

Through dual validation with the Kolmogorov-Smirnov test and the Mann-Whitney U test, we found that the gene expression profiles from PDX, cell lines, and cancer patients represent different distributions, as shown in Table S1.

The analysis was conducted on gene expression data standardized via z-score normalization. The table reports the number of genes (out of 702 selected) identified as exhibiting significantly different distributions by each respective statistical test.

Table S1: Distributional shift validation via Kolmogorov-Smirnov and Mann-Whitney U Tests.

$Domain_1$ - $Domain_2$	**Kolmogorov-Smirnov test**	**Mann-Whitney U test**
PDX - Cell Lines	701	646
PDX - Patients	702	657
PDX - Patients_Unlabeled	701	666
Cell Lines - Patients	699	665
Cell Lines - Patients_Unlabeled	701	684
Patients - Patients_Unlabeled	498	417

Processing and Splitting of the PDX Dataset

- Download the original data from Gao et al.[S1]
- Convert FPKM value to TPM value and apply log2 transformation.
- Apply z-score normalized for each sample.
- Select 702 genes from the COSMIC Cancer Gene Census https://cancer.sanger.ac.uk/census).[S2]
- Delete all recoreds which do not use single small molecule drugs.
- Treat "CR" and "PR" in the ResponseCategory as sensitive to the drug(labeled as 1), and "SD" and "PD" as resistant(labeled as 0).
- Use RDKit and DeepChem[S3] to convert Isomeric SMILES into atom-level feature matrices and adjacency matrices. Then pad virtual node.
- All positive samples (labeled as 1) were placed into the test set as (PDX_gene expr, drug, label) entries.
- For negative samples, if the corresponding PDX appeared in the positive samples, the entry was also included in the test set, resulting in a test set consisting of 331

instances made up of 24 PDX models and 42 drugs (68 positive samples and 263 negative samples).

- Due to the limited size of the dataset, all PDX gene expression data were included in the training set as an unlabeled target domain dataset for transfer learning, totaling 399 gene expression values.

Here are all durgs and its Isomeric SMILES we found in Gao's data:

- **paclitaxel**:

  CC1=C2[C@H](C(=O)[C@@]3([C@H](C[C@@H]4[C@]([C@H]3[C@@H]([C@@](C2(C)C)(C[C@@H]1OC(=O)[C@@H]([C@H](C5=CC=CC=C5)NC(=O)C6=CC=CC=C6)O)O)OC(=O)C7=CC=CC=C7)(CO4)OC(=O)C)O)C)OC(=O)C

- **INC280**:

  CNC(=O)C1=C(F)C=C(C=C1)C1=NN2C(CC3=CC4=C(C=C3)N=CC=C4)=CN=C2N=C1

- **BGJ398**:

  CCN1CCN(CC1)C2=CC=C(C=C2)NC3=CC(=NC=N3)N(C)C(=O)NC4=C(C(=CC(=C4Cl)OC)OC)Cl

- **LEE011**:

  CN(C)C(=O)C1=CC2=CN=C(N=C2N1C3CCCC3)NC4=NC=C(C=C4)N5CCNCC5

- **CGM097**:

  CC(C)OC1=C(C=C2CC(=O)N([C@H](C2=C1)C3=CC=C(C=C3)Cl)C4=CC=C(C=C4)N(C)CC5CCC(CC5)N6CCN(C(=O)C6)C)OC

- **LGH447**:

  C[C@H]1C[C@H](C[C@H](C1)N)C2=C(C=NC=C2)NC(=O)C3=NC(=C(C=C3)F)C4=C(C=CC=C4F)F

- **LDE225**:

  C[C@@H]1CN(C[C@@H](O1)C)C2=NC=C(C=C2)NC(=O)C3=CC=CC(=C3C)C4=CC=C(C=C4)OC(F)(F)F

- **dacarbazine**:

  CN(C)/N=N/C1=C(NC=N1)C(=O)N

- **abraxane**:

  CC1=C2[C@H](C(=O)[C@@]3([C@H](C[C@@H]4[C@]([C@H]3[C@@H]([C@@](C2(C)C)(C[C@@H]1OC(=O)[C@@H]([C@H](C5=CC=CC=C5)NC(=O)C6=CC=CC=C6)O)O)OC(=O)C7=CC=CC=C7)(CO4)OC(=O)C)O)C)OC(=O)C

- **trametinib**:

  CC1=C2C(=C(N(C1=O)C)NC3=C(C=C(C=C3)I)F)C(=O)N(C(=O)N2C4=CC=CC(=C4)NC(=O)C)C5CC5

- **gemcitabine-50mpk**:

  C1=CN(C(=O)N=C1N)[C@H]2C([C@@H]([C@H](O2)CO)O)(F)FS

- **CLR457**:

  C[C@@H]1[C@@H](N(C(=O)O1)C2=NC(=NC(=C2)C3=CN=C(N=C3C(F)(F)F)N)N4CCOCC4)CO

- **erlotinib**:

  COCCOC1=C(C=C2C(=C1)C(=NC=N2)NC3=CC=CC(=C3)C#C)OCCOC

- **binimetinib**:

  CN1C=NC2=C1C=C(C(=C2F)NC3=C(C=C(C=C3)Br)F)C(=O)NOCCO

- **INC424**:

  C1CCC(C1)[C@@H](CC#N)N2C=C(C=N2)C3=C4C=CNC4=NC=N3

- **BKM120**:

  C1COCCN1C2=NC(=NC(=C2)C3=CN=C(C=C3C(F)(F)F)N)N4CCOCC4

- **LDK378**:

  CC1=CC(=C(C=C1C2CCNCC2)OC(C)C)NC3=NC=C(C(=N3)NC4=CC=CC=C4S(=O)(=O)C(C)C)Cl

- **BYL719**:

  CC1=C(SC(=N1)NC(=O)N2CCC[C@H]2C(=O)N)C3=CC(=NC=C3)C(C)(C)C(F)(F)F

- **5FU**:

  C1=C(C(=O)NC(=O)N1)F

- **encorafenib**:

  C[C@@H](CNC1=NC=CC(=N1)C2=CN(N=C2C3=C(C(=CC(=C3)Cl)NS(=O)(=O)C)F)C(C)C)NC(=O)OC

- **tamoxifen**:

  CC/C(=C(\C1=CC=CC=C1)/C2=CC=C(C=C2)OCCN(C)C)/C3=CC=CC=C3

- **HSP990**:

  CC1=C2C(=NC(=N1)N)C[C@@H](NC2=O)C3=C(C=C(C=C3)F)C4=NC(=CC=C4)OC

- **HDM201**:

  CC(C)N1C2=C(C(=O)N([C@H]2C3=CC=C(C=C3)Cl)C4=CC(=CN(C4=O)C)Cl)N=C1C5=CN=C(N=C5OC)OC

- **WNT974**:

  CC1=CC(=CN=C1C2=CC(=NC=C2)C)CC(=O)NC3=NC=C(C=C3)C4=NC=CN=C4

Assessing the Effectiveness of Domain Adaptation in AACDR and PANCDR

For fairness, we trained both the basic and domain adaptation models of PANCDR and AACDR on the constructed **PDX dataset**. We explored multiple combinations of hyperparameters for the basic models of each method (consisting of the feature extractor and predictor, excluding the discriminator and its associated loss). We observed that the two models achieve similar levels of performance. After applying their respective domain adaptation strategies, AACDR demonstrated a clear advantage. This indicates that its asymmetric adversarial domain adaptation remains effective even under a different distribution shift. The specific results are shown in Tables S2 and S3.

Specifically, the basic model of PANCDR involves hyperparameters {**nz**; **d_dim**, **lr**; **gdsc_size**}, while the additional hyperparameters related to its domain adaptation component include {**lr_adv**; **lam**; **pdx_size**}, as is mentioned in.[S4] The hyperparameter **nz** means the dimension of cancer features, **d_dim** means the dimension of drug embeddings, **lr** means the learning rate of the prediction model(uniform graph convolution networks, variational autoencoder and the 1D convolutional layers for cancer drug response prediction), **lr_adv** means the learning rate of the discriminator, **lam** is a regularization coefficient, **gdsc_size** is the batch size of gdsc data and **pdx_size** is that of pdx data.

For our model, the basic model includes hyperparameters {**lr**; **gdsc_size**; $\mathbf{gin}_\epsilon$}, while the domain adaptation component involves additional hyperparameters {**r**; **m**; **margin**; **pdx_size**}. The **lr** is the learning rate of the model, **gdsc_size** and **pdx_size** correspond to the batch sizes, $\mathbf{gin}_\epsilon$ is the parameter in GIN mlp layer. **r**, **m** and **margin** control the strength of domain adaptation.

Table S2: The results of training on PDX dataset without domain adaptation.

Model	**Hyperparameters**	**AUC**	**ACC**	**Precision**	**Recall**	**F1**
PANCDR-	128;100;0.0001;128	0.5026	0.6737	0.2619	0.3235	0.2895
PANCDR-	128;100;0.0001;256	0.5389	0.7583	0.3571	0.2206	0.2727
PANCDR-	128;100;0.0003;256	0.4903	0.7372	0.3061	0.2206	0.2564
PANCDR-	128;100;0.0003;128	0.4947	0.6314	0.2404	0.3676	0.2907
PANCDR-	128;128;0.0001;256	0.5374	0.4048	0.2346	0.8382	0.3666
PANCDR-	128;128;0.0001;128	0.5330	0.4532	0.2347	0.7353	0.3559
PANCDR-	128;128;0.0003;128	0.4530	0.2749	0.2152	0.9559	0.3514
PANCDR-	128;128;0.0003;256	0.5413	0.7372	0.3273	0.2647	0.2927
PANCDR-	128;100;0.0001;128	0.5029	0.6737	0.2619	0.3235	0.3895
PANCDR-	128;100;0.0001;256	0.5389	0.7583	0.3571	0.2206	0.2727
AACDR-	0.0003;768;0	0.5029	0.5377	0.2293	0.5294	0.3200
AACDR-	0.0003;512;0	0.5417	0.5045	0.2391	0.6471	0.3492
AACDR-	0.0001;512;0	0.5219	0.3353	0.2246	0.9112	0.3605
AACDR-	0.0001;768;0	0.4908	0.4532	0.2244	0.6765	0.3370
AACDR-	0.0003;512;0.2	0.4934	0.4985	0.2216	0.5735	0.3197
AACDR-	0.0001;512;0.2	0.5042	0.5650	0.2467	0.5441	0.3394
AACDR-	0.0001;768;0.2	0.4825	0.5347	0.2208	0.5	0.3063

Hyparameters of PANCDR: **nz**; **d_dim**; **lr**; **gdsc_size**
Hyparameters of AACDR: **lr**; **gdsc_size**; $\mathbf{gin}_{\epsilon}$

Table S3: The results of training on PDX dataset with domain adaptation.

Model	Hyperparameters	AUC	ACC	Precision	Recall	F1
PANCDR	128;100;0.0003;0.0001;0.05;256;128	0.5923	0.6858	0.3200	0.4706	0.3810
PANCDR	128;100;0.0001;0.0001;0.01;128;14	0.5518	0.5801	0.2552	0.5441	0.3474
PANCDR	128;128;0.0001;0.0001;0.01;128;14	0.5481	0.5861	0.2518	0.5147	0.3882
PANCDR	128;128;0.0001;0.0001;0.01;128;32	0.5528	0.4713	0.2535	0.8088	0.3819
PANCDR	128;100;0.0001;0.0001;0.01;256;128	0.5733	0.7523	0.3704	0.2941	0.3279
PANCDR	128;128;0.0001;0.0001;0.03;256;128	0.5795	0.5196	0.2486	0.6618	0.3614
PANCDR	128;128;0.0001;0.0001;0.05;256;128	0.5914	0.4622	0.2500	0.8088	0.3819
PANCDR	128;100;0.0003;0.0003;0.01;256;128	0.5992	0.6405	0.2893	0.5147	0.3704
PANCDR	128;128;0.0003;0.0003;0.01;256;128	0.5614	0.5891	0.2671	0.5735	0.3645
PANCDR	128;128;0.0003;0.0003;0.03;256;128	0.5686	0.5438	0.2573	0.6471	0.3682
PANCDR	128;128;0.0003;0.0003;0.05;256;128	0.5954	0.6193	0.2836	0.5589	0.3762
PANCDR	128;100;0.0001;0.0001;0.05;256;128	0.5536	0.5378	0.2455	0.6029	0.3489
PANCDR	128;100;0.0003;0.0001;0.05;256;256	0.5814	0.6858	0.2907	0.3676	0.3247
PANCDR	128;100;0.0001;0.0001;0.01;128;256	0.5420	0.4260	0.2371	0.8088	0.3667
AACDR	0.1;0.02;0.3;768;256;0.0001	0.5138	0.7538	0.3421	0.1912	0.2453
AACDR	0.1;0.02;0.3;768;256;0.0003	0.5338	0.5952	0.2462	0.4706	0.3232
AACDR	0.5;0.04;0.6;768;128;0.0001	0.5712	0.5710	0.2566	0.5735	0.3545
AACDR	0.5;0.04;0.7;768;128;0.0001	0.6280	0.6375	0.3143	0.6471	0.4231
AACDR	0.6;0.05;0.7;512;128;0.0001	0.5925	0.7311	0.3433	0.3382	0.3407
AACDR	0.6;0.05;0.8;512;128;0.0001	0.5818	0.6284	0.2800	0.5147	0.3627
AACDR	0.65;0.06;0.9;512;128;0.0001	0.5815	0.5710	0.2716	0.6471	0.3826
AACDR	0.7;0.06;0.7;768;128;0.0001	0.5973	0.6828	0.3093	0.4412	0.3636
AACDR	0.7;0.05;0.7;512;128;0.0001	0.6024	0.7674	0.4211	0.3529	0.3840
AACDR	0.7;0.04;0.7;512;128;0.0001	0.6215	0.7039	0.3404	0.4706	0.3951
AACDR	0.7;0.03;0.7;512;128;0.0001	0.6134	0.7402	0.3714	0.3824	0.3768
AACDR	0.7;0.045;0.7;512;128;0.0001	0.6276	0.7160	0.3617	0.5000	0.4198
AACDR	0.5;0.04;0.6;768;256;0.0001	0.5203	0.7976	0.5263	0.1471	0.2299

Hyparameters of PANCDR: **nz**; **d_dim**; **lr**; **lr_adv**; **lam**; **gdsc_size**; **pdx_size**
Hyparameters of AACDR: **r**; **m**; **margin**; **gdsc_size**; **pdx_size**; **lr**

Evaluation for model interpretability

To evaluate the interpretability of AACDR and enhance its potential for clinical application, we conducted a SHAP (SHapley Additive exPlanations) analysis on the test set. First, the test samples were reclassified according to the cancer types of the patients, based on annotations from The Cancer Genome Atlas (TCGA). For each cancer type, we identified the corresponding driver genes from the Network of Cancer Genes (NCG), which includes genes validated through real clinical cases. Subsequently, SHAP analysis was performed separately for each cancer type to determine the gene–level feature importance within the model's predictions. The final results are summarized in Table S4.

The second column of the table indicates, for each cancer type, how many of the top 10 most important genes identified by AACDR through SHAP analysis are known driver genes for that cancer, based on clinical evidence. This demonstrates that AACDR is capable of effectively identifying representative and biologically meaningful genes from gene expression data, thereby enhancing the model's acceptability and credibility in therapeutic prediction.

Table S4: SHAP-based interpretability analysis by cancer type. For each cancer type in the test set, this table reports the number of clinically validated driver genes (from NCG) among the top 10 most influential features identified by SHAP, along with the number of corresponding patient-drug entries.

Cancer Type	Driver Genes in SHAP Top 10 Genes	Corresponding Patient-Drug Pair Counts
ACC	5	1
BLCA	7	15
BRCA	6	15
CESE	8	56
COAD	7	14
HNSC	7	56
KICH	5	1
KIRC	5	11
KIRP	5	7
LGG	7	97
LIHC	6	17
LUAD	7	13
LUSC	7	7
MESO	5	14
PAAD	5	57
PCPG	6	2
PRAD	5	18
READ	5	1
SARC	6	24
SKCM	7	8
STAD	7	57
TGCT	5	6
THCA	5	10
UCEC	8	14
UCS	5	9

Bootstrapped confidence intervals for AACDR and PANCDR

We performed 1000-iteration bootstrapped sampling with replacement on the test set to estimate confidence intervals for AACDR and PANCDR across multiple evaluation metrics. The resulting 95% confidence intervals are shown it Table S5.

Table S5: The results of bootstrapped confidence intervals.

Method	AUC	ACC	Precision	Recall	F1
AACDR	0.6957-0.7710	0.6656-0.7328	0.6187-0.7415	0.6224-0.8278	0.6491-0.7433
PANCDR	0.6642-0.7364	0.6411-0.7102	0.6039-0.7226	0.5987-0.8083	0.6273-0.7279

Additional analyses on loss function design and preprocessing strategy

To further address concerns regarding the custom asymmetric reconstruction loss function $\mathcal{L}_{AE}$ (as shown in Equation (1)) and the model's reliance on specific preprocessing protocols, we conducted complementary ablation studies.

$$\mathcal{L}_{AE} = \begin{cases} \frac{1}{2\beta}\left(\mathbf{x} - \mathbf{x}_{rec}\right)^2, & \text{if } |\mathbf{x} - \mathbf{x}_{rec}| < \beta, \\ |\mathbf{x} - \mathbf{x}_{rec}| - \frac{\beta}{2}, & \text{otherwise.} \end{cases} \tag{1}$$

$$\mathcal{L} = \begin{cases} \frac{1}{2}\left(\mathbf{x} - \mathbf{x}_{rec}\right)^2, & \text{if } |\mathbf{x} - \mathbf{x}_{rec}| < \beta, \\ \beta\,|\mathbf{x} - \mathbf{x}_{rec}| - \frac{\beta}{2}, & \text{otherwise.} \end{cases} \tag{2}$$

Specifically, we compared our loss formulation (a specific version of Huber loss shown in Equation (2), where $\beta = 1.0$) against standard alternatives, including Mean Squared Error (MSE), Mean Absolute Error (MAE) and Huber loss with other parameter values

(e.g., $\beta = 0.1$ and $\beta = 10$). In addition, we evaluated the impact of removing z-score normalization from cancer representation data.

The following tables (Table S6 and Table S7)summarize the comparative results of these ablations. It can be observed that the Huber loss clearly outperforms both MSE and MAE. Moreover, both the settings of $\beta = 0.1$ and $\beta = 1.0$ yield satisfactory results, indicating that the hyperparameter has a relatively wide effective range. These findings also support the effectiveness of the configuration used in our method. In addition, the results demonstrate that normalization contributes positively to model stability and accuracy. These findings underscore the rationality and robustness of our design choices.

Table S6: The ablation results of reconstruction loss.

Loss	**AUC(std)**	**ACC(std)**	**Precision(std)**	**Recall(std)**	**F1(std)**
$\mathcal{L}_{AE}$ (Huber with $\beta = 1.0$)	**0.7338(0.0230)**	**0.6947**(0.0193)	0.6749(**0.0245**)	**0.7218(0.0811)**	**0.6943(0.0349)**
Huber with $\beta = 0.1$	0.7328(0.0230)	0.6925(**0.0170**)	**0.6760**(0.0280)	0.7123(0.0819)	0.6898(0.0354)
Huber with $\beta = 10$	0.7244(0.0277)	0.6744(0.0214)	0.6508(0.0298)	0.7225(0.0916)	0.6804(0.0379)
MSE	0.7116(0.0355)	0.6674(0.0223)	0.6488(0.0299)	0.6985(0.1073)	0.6669(0.0483)
MAE	0.7261(0.0341)	0.6785(0.0256)	0.6613(0.0360)	0.7048(0.1075)	0.6762(0.0519)

std: standard deviation

Table S7: The ablation results of Z-score normalization.

Method	**AUC**	**ACC**	**Precision**	**Recall**	**F1**
Z-score	**0.7338(0.0230)**	**0.6947(0.0193)**	**0.6749(0.0245)**	**0.7218(0.0811)**	**0.6943(0.0349)**
Without Z-score	0.7182(0.0287)	0.6854(0.0261)	0.6714(0.0366)	0.6725(0.1107)	0.6698(0.0524)

Notably, when z-score normalization was not applied, the model exhibited a decline in performance after a comparable number of training iterations. This highlights that appropriate standardization is indeed beneficial for deep learning tasks, contributing to improved training effectiveness and model performance.

References

(S1) Gao, H.; Korn, J. M.; Ferretti, S.; Monahan, J. E.; Wang, Y.; Singh, M.; Zhang, C.; Schnell, C.; Yang, G.; Zhang, Y.; others High-throughput screening using patient-derived tumor xenografts to predict clinical trial drug response. *Nature medicine* **2015**, *21*, 1318–1325.

(S2) Tate, J. G.; Bamford, S.; Jubb, H. C.; Sondka, Z.; Beare, D. M.; Bindal, N.; Boutselakis, H.; Cole, C. G.; Creatore, C.; Dawson, E.; others COSMIC: the catalogue of somatic mutations in cancer. *Nucleic acids research* **2019**, *47*, D941–D947.

(S3) 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, Inc.", 2019.

(S4) Kim, J.; Park, S.-H.; Lee, H. PANCDR: precise medicine prediction using an adversarial network for cancer drug response. *Briefings in Bioinformatics* **2024**, *25*, bbae088.