COMPACT REPRESENTATION LEARNING FOR MULTIMODAL DRUG-DRUG INTERACTION EVENT PREDICTION

Zixuan Guo[1], *Hang Lv*[1], *Yanchao Tan*[1*], *Zhigang Lin*[2], *Hengyu Zhang*[3], *Xiping Chen*[4], *Carl Yang*[5]

[1]Fuzhou University, [2]The First Affiliated Hospital of Fujian Medical University,
[3]Macquarie University, [4]Hangzhou Bywin Technology Co., Ltd., [5]Emory University

ABSTRACT

Accurately predicting Drug-Drug Interaction Events (DDIEs) is essential for mitigating adverse reactions and improving patient safety. Despite the success of recent DDIE prediction methods in utilizing multimodal data, they often overlook redundancy both within and across modalities, leading to inconsistent and fragmented drug embeddings that miss informative pharmacological details. In this paper, we propose CompactDDI, a novel framework that learns compact intra- and inter-modality representations for improved DDIE prediction. Specifically, we design compact modality-specific extractors to generate consistent drug embeddings for intra-modality. Furthermore, we introduce a mutual information minimization regularization strategy that reduces redundancy and enhances complementarity across modalities. These compact representations are then fused to enable more accurate and generalized DDIE prediction. Extensive experiments on two commonly used datasets demonstrate the superior effectiveness and generalization of our proposed CompactDDI, particularly in predicting DDIEs between novel drugs. Our source code is publicly available at: https://github.com/Zixuan-Guo2005/CompactDDI.

***Index Terms*—** Drug-Drug Interaction Event, Multimodal Learning, Compact Representation Learning

1. INTRODUCTION

Drug-Drug Interaction Event (DDIE) prediction plays a critical role in ensuring medication safety and optimizing treatment plans, particularly as polypharmacy becomes increasingly prevalent in complex disease management [1, 2, 3]. Unexpected DDIEs can lead to adverse drug reactions, reduced therapeutic efficacy, or even fatal outcomes [4, 5]. Therefore, effectively identifying potential DDIEs can support public health security and clinical research, especially during novel drug development where clinical trials are limited [6].

Traditional approaches for DDIE prediction primarily relied on single-modality data, such as handcrafted features or molecular structures [7, 8]. For example, DeepDDI [7] utilized deep neural networks to infer interaction types based on chemical structure similarity. However, such models often struggle to capture the multifactorial nature of pharmacological signals. Recently, some studies have explored multimodal learning frameworks, integrating heterogeneous sources (e.g., biological profiles and Knowledge Graphs (KGs)) to enrich drug representations and improve predictive performance [9, 10, 11]. For instance, MDF-SA-DDI [10] introduced a transformer-based model that performs multi-source feature fusion with self-attention for drug embeddings.

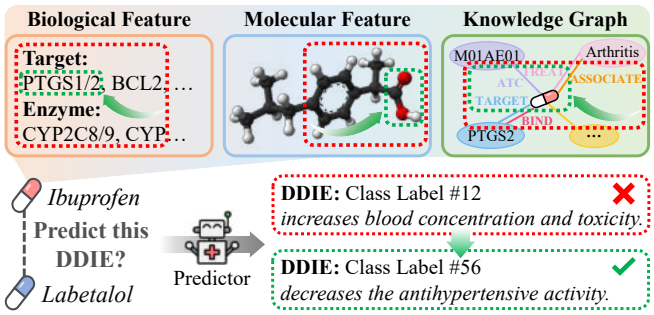


Fig. 1: An illustrative example of Drug-Drug Interaction Event (DDIE) prediction. Red dashed boxes contain redundant multimodal features that cause errors, while green dashed boxes indicate compact and informative features that guide accurate DDIE predictions.

However, these methods typically extract features from each modality independently while ignoring redundancy within and across modalities. As shown in Fig. 1, when all multimodal features of the drug *Ibuprofen* are considered without filtering (see red dashed boxes), redundant information across biological, molecular, and KG modalities (e.g., biological target BCL2 and molecular substructure benzene ring) causes a false toxicity event prediction with *Labetalol* [12]. In contrast, after reducing redundancy (see green dashed boxes), the multimodal information of *Ibuprofen* becomes compact, clearly reflecting its core pharmacological properties, such as PTGS pathway inhibition and the carboxyl group associated with anti-inflammatory effects [13]. This compact drug representation naturally leads to the correct DDIE prediction that *Labetalol* decreases the antihypertensive activity of *Ibuprofen*.

To this end, we propose a novel **Compact** representation learning framework for multimodal **D**rug-**D**rug **I**nteraction event prediction (CompactDDI). Specifically, we first integrate a compact intra-modality module with modality-specific extractors to generate consistent drug representations. We then employ mutual information minimization in the compact inter-modality module to reduce redundancy and capture cross-modality complementarity. Finally, we design a DDIE prediction module that fuses the compact embeddings into a unified representation for accurate DDIE prediction.

The main contributions of our work are summarized as follows: (1) We propose a novel compact multimodal framework that integrates intra- and inter-modality representation learning for DDIE prediction. (2) We design compact modality-specific extractors and a mutual information minimization regularization strategy, effectively reducing redundancy within and across modalities while capturing essential pharmacological signals of drugs. (3) We conduct extensive experiments on two benchmark datasets under two settings, demonstrating the superiority and generalization of our CompactDDI.

*Yanchao Tan is the corresponding author.

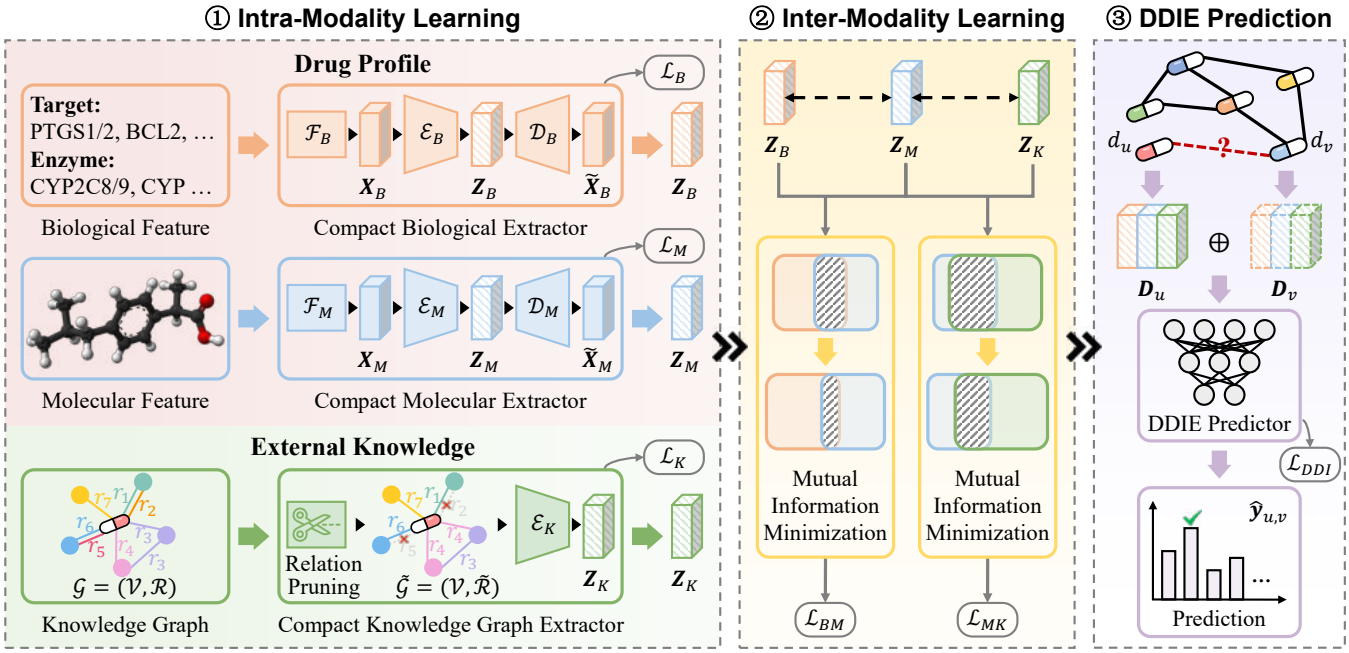


Fig. 2: The overall framework of our proposed CompactDDI, which captures compact multimodal drug representations through intra- and inter-modality learning to enhance Drug-Drug Interaction Event (DDIE) predictions.

2. METHODOLOGY

2.1. Problem Formulation and CompactDDI Overview

Drug-Drug Interaction Event (DDIE) prediction aims to classify the specific interaction events $\boldsymbol{y}_{u,v} \in \mathcal{Y}$ between a drug pair (d_u, d_v), where $\mathcal{Y}$ is the DDIE matrix. Each drug is characterized by two primary multimodal sources: drug profile (e.g., biological and molecular features) and external knowledge (e.g., knowledge graph). To achieve this, we propose a CompactDDI framework, which learns compact multimodal drug representations for improved DDIE predictions (shown in Fig. 2). Specifically, we first design compact modality extractors to transform biological, molecular, and knowledge graph inputs into latent embeddings $\boldsymbol{Z}_B, \boldsymbol{Z}_M, \boldsymbol{Z}_K$. Then, we minimize inter-modality mutual information to reduce redundancy and enhance complementarity across modalities. Finally, the multimodal representations of a drug pair $(\boldsymbol{D}_u, \boldsymbol{D}_v)$ are fused and fed into a DDIE predictor to generate the predicted interactions $\hat{\boldsymbol{y}}_{u,v}$.

2.2. Compact Representation Learning for Intra-Modality

To effectively capture informative and compact representations for each modality, we model drugs from the drug profile and external knowledge via modality extractors. The former reflects intrinsic biomedical properties, including biological features and molecular structures, while the latter encodes extrinsic contextual information, such as Knowledge Graphs (KGs). Notably, the two perspectives can flexibly integrate other modalities (e.g., textual descriptions).

Compact Drug Profile Representation. We leverage biological and molecular modalities to represent each drug profile. For the biological feature of a drug, we compute Jaccard similarity over three commonly adopted descriptors (i.e., chemical substructures, targets, and enzymes) [14, 10], thereby obtaining the initial biological representation $\boldsymbol{X}_B$. For the molecular feature, we first obtain the SMILES string from DrugBank [15] and parse it with RDKit [16] to construct a molecular graph with atomic and bond information. The graph is then processed by a Heterogeneous Graph Neural Network (HGNN) [17] to generate the molecular representation $\boldsymbol{X}_M$. Subsequently, the biological and molecular representations are fed into modality-specific autoencoders to produce latent embeddings $\boldsymbol{Z}_B, \boldsymbol{Z}_M \in \mathbb{R}^{1\times dim}$. The encoder projects each input into a low-dimensional representation, while the decoder reconstructs the original feature space. The reconstruction loss is defined as:

$$\mathcal{L}_{\{B,M\}} = \frac{1}{N}\sum_{i=1}^{N}\left\|\boldsymbol{X}^i_{\{B,M\}} - \widetilde{\boldsymbol{X}}^i_{\{B,M\}}\right\|^2, \tag{1}$$

where $\boldsymbol{X}^i_{\{B,M\}}$ and $\widetilde{\boldsymbol{X}}^i_{\{B,M\}}$ denote the original and reconstructed multimodal representations of drug d_i, and N is the number of drugs. By minimizing this loss, the autoencoders are guided to preserve expressive modality-specific information while alleviating redundancy, yielding compact and informative embeddings.

Compact External Knowledge Representation. KGs provide abundant semantic context, but the large number of relations often leads to overlapping or inconsistent expressions [18], which may capture inaccurate drug characteristics. To achieve a more concise and semantically consistent knowledge representation, we introduce relation pruning strategy on the given KG $\mathcal{G} = (\mathcal{V}, \mathcal{R})$, transforming the original relation set $\mathcal{R}$ into a refined set $\widetilde{\mathcal{R}}$ as follows:

$$\widetilde{\mathcal{R}} = \text{softmax}\left(\frac{\log(\boldsymbol{r}) + \boldsymbol{G}}{\tau}\right), \tag{2}$$

where $\boldsymbol{r} = \text{MLP}(\boldsymbol{R})$ represents learnable pruning weights derived from the original relation embeddings $\boldsymbol{R} \in \mathbb{R}^{|\mathcal{R}|\times dim}$. $\boldsymbol{G}$ is noise sampled from a Gumbel distribution, and τ is a temperature hyperparameter. We then obtain the refined relations embeddings $\widetilde{\boldsymbol{R}} = \widetilde{\mathcal{R}}\boldsymbol{R}$. This strategy is regularized by a Kullback-Leibler (KL) divergence, $\mathcal{L}_K = \text{KL}(\widetilde{\boldsymbol{R}}\|\boldsymbol{R})$, which encourages relations with consistent semantics in the context of $\mathcal{G}$ to be categorized into the same index, while preventing excessive pruning and preserving core information. Ultimately, the compact KG $\widetilde{\mathcal{G}} = (\mathcal{V}, \widetilde{\mathcal{R}})$ is encoded by an HGNN to generate the final knowledge embeddings $\boldsymbol{Z}_K \in \mathbb{R}^{1\times dim}$. We

Table 1: Statistics of the datasets used in our experiments.

Dataset	Deng et al.	Lin et al.
# of Drugs	572	1,258
# of DDIs	74,528	323,539
# of DDI Events	65	100
# of KG Entities	97,243	
# of KG Relations	107	

define the overall intra-modality compact loss as follows:

$$\mathcal{L}_{Intra} = \mathcal{L}_B + \mathcal{L}_M + \mathcal{L}_K. \tag{3}$$

2.3. Compact Representation Learning for Inter-Modality

While each modality extractor produces compact intra-modality embeddings directly combining them can still introduce redundancy due to overlapping semantics across modalities. To capture compact inter-modality signals, we introduce a Mutual Information (MI) minimization regularization [19]. A high MI indicates overlapping and redundant signals across modalities. Such redundancy causes the model to focus on shared interaction mechanisms and consequently misses informative pharmacological details. Formally, the MI between two embeddings $\boldsymbol{Z}_i$ and $\boldsymbol{Z}_j$ is defined as:

$$\text{MI}(\boldsymbol{Z}_i, \boldsymbol{Z}_j) = H(\boldsymbol{Z}_i) + H(\boldsymbol{Z}_j) - H(\boldsymbol{Z}_i, \boldsymbol{Z}_j), \tag{4}$$

where $H(\cdot)$ denotes the entropy. A high MI value indicates large information overlap and redundancy between modalities. To enable efficient computation and optimization, we reformulate MI with conditional entropy and KL divergence as follows:

$$\text{MI}(\boldsymbol{Z}_i, \boldsymbol{Z}_j) = H_{\boldsymbol{Z}_j}(\boldsymbol{Z}_i) + H_{\boldsymbol{Z}_i}(\boldsymbol{Z}_j) - \text{KL}(\boldsymbol{Z}_i \| \boldsymbol{Z}_j) - \text{KL}(\boldsymbol{Z}_j \| \boldsymbol{Z}_i), \tag{5}$$

where $H_{\boldsymbol{Z}_j}(\boldsymbol{Z}_i) = -\mathbb{E}[\log p(\boldsymbol{Z}_i | \boldsymbol{Z}_j)]$. Based on this formulation, the overall inter-modality compact loss is computed as:

$$\mathcal{L}_{Inter} = \mathcal{L}_{BM} + \mathcal{L}_{MK}, \tag{6}$$

$$\mathcal{L}_{BM} = \mathbb{E}[\text{MI}(\boldsymbol{Z}_B, \boldsymbol{Z}_M)], \quad \mathcal{L}_{MK} = \mathbb{E}[\text{MI}(\boldsymbol{Z}_M, \boldsymbol{Z}_K)]. \tag{7}$$

By aligning biological and knowledge embeddings with the molecular anchor, the inter-modality loss reduces redundancy and yields more compact and distinctive representations for DDIE prediction.

2.4. Drug-Drug Interaction Event Prediction

For each drug, we construct a fused representation by concatenating its compact biological, molecular, and knowledge graph embeddings, i.e., $\boldsymbol{D} = [\boldsymbol{Z}_B || \boldsymbol{Z}_M || \boldsymbol{Z}_K] \in \mathbb{R}^{1 \times 3dim}$. Given a drug pair (d_u, d_v), we then predict their DDIE as follows:

$$\hat{\boldsymbol{y}}_{u,v} = \text{MLP}([\boldsymbol{D}_u || \boldsymbol{D}_v]), \tag{8}$$

where $\text{MLP} : \mathbb{R}^{1 \times 6dim} \to \mathbb{R}^{1 \times C}$ is a multi-layer perceptron and C is the number of DDIE classes. We treat DDIE prediction as a multi-class classification task and optimize via the cross-entropy loss:

$$\mathcal{L}_{\text{DDI}} = -\sum_{(u,v) \in \mathcal{D}} \boldsymbol{y}_{u,v} \log \hat{\boldsymbol{y}}_{u,v}, \tag{9}$$

where $\mathcal{D}$ is the set of training pairs, and $\boldsymbol{y}_{u,v}$ is the ground-truth. Finally, the objective function of CompactDDI is calculated as:

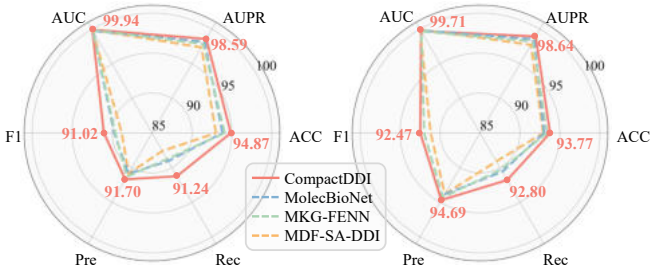
$$\mathcal{L} = \mathcal{L}_{\text{DDI}} + \alpha \mathcal{L}_{Intra} + \beta \mathcal{L}_{Inter}, \tag{10}$$

where α and β are hyperparameters balancing the intra- and inter-modality compact losses.

3. EXPERIMENTS

3.1. Experimental Settings

Datasets and Evaluation Protocols. We verify the effectiveness of CompactDDI on two commonly used datasets, i.e., Deng et al. [14] Lin et al. [10]. Additionally, we use DRKG [21] as the external biomedical knowledge graph, which deletes DDIs overlapping with both datasets to avoid data leakage. The detailed statistics are summarized in Table 1. For evaluation metrics, we adopt accuracy (ACC), area under the precision-recall curve (AUPR), the area under the receiver operating characteristic curve (AUC), macro-F1 score (F1), precision (Pre), and recall (Rec), consistent with [10, 22].



(a) Deng et al. (b) Lin et al.

Fig. 3: Performance (%) comparisons of CompactDDI on the two datasets under the Existing Drug-Existing Drug setting.

Baselines. To comprehensively evaluate CompactDDI, we compare it with ten representative state-of-the-art baselines, including traditional classification methods (RF [23] and DNN [24]), a single-modality DDIE method (DeepDDI [7]), and multimodal DDIE methods (Lee et al. [6], DDIMDL [14], MDNN [9], MDF-SA-DDI [10], HetDDI [17], MKG-FENN [22], and MolecBioNet [11]).

Implementation Details. Following [10, 17, 11], we evaluate DDIE prediction under two settings: Novel Drug-Novel Drug and Existing Drug-Existing Drug, where novel drugs appear only in the test set to assess generalization. Each model is evaluated using 5-fold cross-validation. We train for 120 epochs with a learning rate of 0.01, batch size of 1024, embedding dimension $dim = 128$, and temperature $\tau = 0.1$. The HGNN encoder is configured with three hidden layers, and the hyperparameters are set to $\alpha = 10$ and $\beta = 1$. All experiments are conducted on two NVIDIA RTX 3090 Ti GPUs.

3.2. Overall Performance Comparison

We evaluate CompactDDI on the Deng et al. and Lin et al. datasets under both the Novel Drug-Novel Drug and Existing Drug-Existing Drug settings. The results are shown in Table 2 and Fig. 3. Overall, our CompactDDI consistently outperforms all baselines across both settings, demonstrating the effectiveness of compact representation learning for intra- and inter-modality. In particular, in the more challenging Novel Drug-Novel Drug setting, CompactDDI achieves significant average improvements of 24.16%, which highlights its strong generalization ability to unseen drugs.

Traditional classification and single-modality approaches (e.g., DNN and DeepDDI), which rely on handcrafted features or single molecular inputs, fail to capture comprehensive drug information and thus yield limited performance. By contrast, our proposed CompactDDI integrates multimodal sources and achieves substantial gains up to 200.91% with Rec on the Deng et al. dataset under the Novel Drug-Novel Drug setting.

Among multimodal methods, MDF-SA-DDI leverages multiple biological features, HetDDI combines molecular structures with KG, and MKG-FENN integrates multimodal knowledge graphs. However, these methods simply concatenate multimodal drug features without capturing semantic dependencies or complementary information across modalities. While MolecBioNet alleviates inter-modality overlap between biomedical knowledge graphs and molec-

Table 2: Experimental results (%) on both datasets under the Novel Drug-Novel Drug setting, where * denotes a significant improvement according to the Wilcoxon signed-rank test [20]. The best performances are highlighted in **boldface** and the second runners are underlined.

Dataset	Deng et al.						Lin et al.					
Metric	ACC	AUPR	AUC	F1	Pre	Rec	ACC	AUPR	AUC	F1	Pre	Rec
RF	36.26	33.21	86.01	15.57	23.59	13.11	40.07	38.16	96.54	19.64	31.86	15.67
DNN	40.87	37.76	95.50	11.52	18.36	10.93	45.70	41.29	95.65	29.97	43.45	25.08
DeepDDI	36.02	27.81	90.59	13.73	15.86	14.50	36.11	28.20	92.64	18.68	23.01	17.11
Lee et al.	40.97	31.84	83.02	20.22	22.16	20.27	48.67	43.49	90.93	30.82	33.55	30.66
DDIMDL	40.75	36.35	95.12	15.90	24.08	14.52	46.99	43.86	96.85	30.32	37.73	27.29
MDNN	45.75	42.15	92.03	16.97	25.64	17.09	50.56	45.43	95.29	21.40	34.63	20.41
MDF-SA-DDI	43.38	38.73	86.30	23.29	27.15	22.26	47.94	44.50	96.86	29.37	36.67	26.59
HetDDI	44.86	40.12	88.45	22.04	25.98	21.73	49.58	46.10	96.27	28.79	35.09	25.96
MKG-FENN	45.52	41.62	91.49	21.86	27.54	21.31	50.30	46.32	96.69	29.57	37.20	25.46
MolecBioNet	46.54	42.69	90.07	24.50	28.41	22.85	51.43	46.75	95.09	24.59	37.86	27.31
CompactDDI	**53.19***	**53.75***	**96.95***	**29.95***	**32.30***	**32.89***	**68.92***	**73.36***	**98.16***	**39.47***	**45.26***	**39.38***

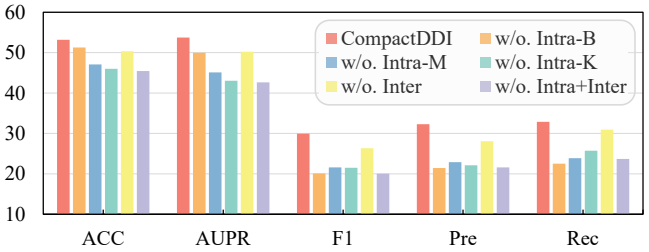
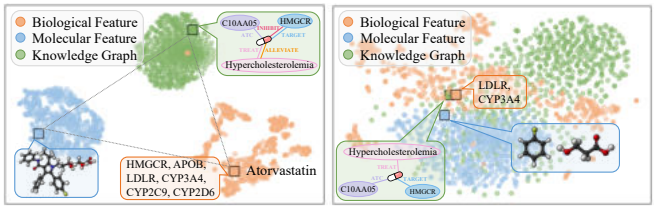


Fig. 4: Performance (%) comparisons of CompactDDI and its five variants on Deng et al. under the Novel Drug-Novel Drug setting.



(a) CompactDDI w/o. Intra+Inter (b) CompactDDI

Fig. 5: Visualization of drug embeddings from three modalities on Deng et al., learned by different models. Best viewed in color.

ular features, it still ignores redundancy within individual modalities. In contrast, CompactDDI explicitly learns more compact and informative drug embeddings for intra- and inter-modality, achieving up to 60.51% performance gains over MolecBioNet.

3.3. Ablation Studies

To better understand the contribution of each component in CompactDDI, we conduct ablation studies on the Deng et al. dataset under the Novel Drug-Novel Drug setting as follows: w/o. Intra-B and w/o. Intra-M replace the biological and molecular autoencoders with simple MLPs, respectively; w/o. Intra-K removes the relation pruning strategy in the knowledge graph; w/o. Inter removes the inter-modality compact loss; and w/o. Intra+Inter removes both intra- and inter-modality compact mechanisms.

As shown in Fig. 4, compared with w/o. Intra+Inter, our CompactDDI achieves up to 49.54% improvement in Pre. Moreover, CompactDDI consistently outperforms all intra-modality ablations (i.e., w/o. Intra-B, w/o. Intra-M, w/o. Intra-K), with an average improvement of 25.73%. This indicates that compact modeling within each modality is essential, while also revealing that the influence of redundancy on overall performance differs across modalities. Incorporating the inter-modality compact loss further boosts performance up to 15.07% in Pre, highlighting the importance of capturing complementary and reducing redundancy across modalities. These results also confirm that our design extracts more concise and informative drug embeddings for enhanced DDIE predictions.

3.4. Case Studies

To demonstrate the advantages of CompactDDI in learning compact multimodal drug representations, we visualize the embeddings from three modalities generated by CompactDDI w/o. Intra+Inter (which removes both intra- and inter-modality compact mechanisms) and by CompactDDI. Additionally, we highlight the drug Atorvastatin (□) with its used multimodal information in Fig. 5.

As shown in Fig. 5(a), without compact mechanisms, the distributions of biological, molecular, and KG embeddings are fragmented, generating inconsistent drug representations. For example, Atorvastatin's biological targets, such as APOB, introduce off-target lipid metabolism signals. Meanwhile, repeated aromatic nuclei and pyrroline motifs in the molecular structure interfere with the modeling of its hydroxyl-acid group's cholesterol-lowering role. These redundancies scatter the embeddings and hinder the capturing of its lipid-lowering mechanism. In contrast, Fig. 5(b) shows that CompactDDI produces compact distributions that preserve informative pharmacological details. For example, filtering out redundant peripheral proteins and irrelevant KG relations enables Atorvastatin's biological and KG embeddings to consistently emphasize key targets. The resulting partial overlap between the two modalities aligns with Atorvastatin's lipid-regulation mechanism, providing a unified representation for accurate DDIE prediction.

4. CONCLUSION

In this paper, we propose CompactDDI, a novel compact representation learning framework for multimodal Drug-Drug Interaction Event (DDIE) prediction. Particularly, we integrate the compact mechanisms for intra- and inter-modality, which effectively reduce redundancy and enhance complementarity within and across modalities. This design generates consistent and compact drug representations, capturing the essential pharmacological details for improved DDIE prediction. Extensive experiments on two datasets demonstrate that CompactDDI consistently outperforms state-of-the-art baselines, showing strong accuracy and generalization under Novel Drug-Novel Drug and Existing Drug-Existing Drug settings.

5. ACKNOWLEDGEMENTS

This work was supported in part by the Fujian Provincial Artificial Intelligence Industry Development Technology Project under Grant (2025H0042), Fujian Provincial Natural Science Foundation of China under Grants (2025J01540), National Natural Science Foundation of China under Grants (62302098). Carl Yang was not supported by any fund from China.

6. REFERENCES

[1] Chaoqi Yang, Cao Xiao, Fenglong Ma, Lucas Glass, and Jimeng Sun, "Safedrug: Dual molecular graph encoders for recommending effective and safe drug combinations," in *Proceedings of the 30th International Joint Conference on Artificial Intelligence, IJCAI 2021*, 2021, pp. 3735–3741.

[2] Chuang Zhao, Hongke Zhao, Xiaofang Zhou, and Xiaomeng Li, "Enhancing precision drug recommendations via in-depth exploration of motif relationships," *IEEE Transactions on Knowledge and Data Engineering*, vol. 36, no. 12, pp. 8164–8178, 2024.

[3] Jie Tan, Yu Rong, Kangfei Zhao, Tian Bian, Tingyang Xu, Junzhou Huang, Hong Cheng, and Helen Meng, "Natural language-assisted multi-modal medication recommendation," in *Proceedings of the 33rd ACM International Conference on Information and Knowledge Management*, 2024, pp. 2200–2209.

[4] Jia Jia, Feng Zhu, Xiaohua Ma, Zhiwei W Cao, Yixue X Li, and Yu Zong Chen, "Mechanisms of drug combinations: interaction and network perspectives," *Nature Reviews Drug Discovery*, vol. 8, no. 2, pp. 111–128, 2009.

[5] Michael Dougan, Adrienne M Luoma, Stephanie K Dougan, and Kai W Wucherpfennig, "Understanding and treating the inflammatory adverse events of cancer immunotherapy," *Cell*, vol. 184, no. 6, pp. 1575–1588, 2021.

[6] Geonhee Lee, Chihyun Park, and Jaegyoon Ahn, "Novel deep learning model for more accurate prediction of drug-drug interaction effects," *BMC Bioinformatics*, vol. 20, no. 1, pp. 415, 2019.

[7] Jae Yong Ryu, Hyun Uk Kim, and Sang Yup Lee, "Deep learning improves prediction of drug–drug and drug–food interactions," *Proceedings of the National Academy of Sciences*, vol. 115, no. 18, pp. E4304–E4311, 2018.

[8] Kexin Huang, Cao Xiao, Trong Hoang, Lucas Glass, and Jimeng Sun, "Caster: Predicting drug interactions with chemical substructure representation," in *Proceedings of the AAAI conference on Artificial Intelligence*, 2020, vol. 34, pp. 702–709.

[9] Tengfei Lyu, Jianliang Gao, Ling Tian, Zhao Li, Peng Zhang, and Ji Zhang, "Mdnn: A multimodal deep neural network for predicting drug-drug interaction events," in *Proceedings of the 30th International Joint Conference on Artificial Intelligence, IJCAI 2021*, 2021, pp. 3536–3542.

[10] Shenggeng Lin, Yanjing Wang, Lingfeng Zhang, Yanyi Chu, Yatong Liu, Yitian Fang, Mingming Jiang, Qiankun Wang, Bowen Zhao, Yi Xiong, et al., "Mdf-sa-ddi: predicting drug–drug interaction events based on multi-source drug fusion, multi-source feature fusion and transformer self-attention mechanism," *Briefings in Bioinformatics*, vol. 23, no. 1, pp. bbab421, 2022.

[11] Mengjie Chen, Ming Zhang, and Cunquan Qu, "Towards interpretable drug-drug interaction prediction: A graph-based approach with molecular and network-level explanations," in *Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining V. 2*, 2025, pp. 203–214.

[12] KD Rainsford, "Ibuprofen: pharmacology, efficacy and safety," *Inflammopharmacology*, vol. 17, no. 6, pp. 275–342, 2009.

[13] Randall E Harris, Joanne Beebe-Donk, Hani Doss, and Deborah Burr Doss, "Aspirin, ibuprofen, and other non-steroidal anti-inflammatory drugs in cancer prevention: a critical review of non-selective cox-2 blockade," *Oncology Reports*, vol. 13, no. 4, pp. 559–583, 2005.

[14] Yifan Deng, Xinran Xu, Yang Qiu, Jingbo Xia, Wen Zhang, and Shichao Liu, "A multimodal deep learning framework for predicting drug–drug interaction events," *Bioinformatics*, vol. 36, no. 15, pp. 4316–4322, 2020.

[15] David S Wishart, Yannick D Feunang, An C Guo, Elvis J Lo, Ana Marcu, Jason R Grant, Tanvir Sajed, Daniel Johnson, Carin Li, Zinat Sayeeda, et al., "Drugbank 5.0: a major update to the drugbank database for 2018," *Nucleic Acids Research*, vol. 46, no. D1, pp. D1074–D1082, 2018.

[16] Greg Landrum et al., "Rdkit: A software suite for cheminformatics, computational chemistry, and predictive modeling," *Greg Landrum*, vol. 8, no. 31.10, pp. 5281, 2013.

[17] Zhe Li, Xinyi Tu, Yuping Chen, and Wenbin Lin, "Hetddi: a pre-trained heterogeneous graph neural network model for drug–drug interaction prediction," *Briefings in Bioinformatics*, vol. 24, no. 6, pp. bbad385, 2023.

[18] Tengfei Ma, Yujie Chen, Liang Wang, Xuan Lin, Bosheng Song, and Xiangxiang Zeng, "S^2dn: Learning to denoise unconvincing knowledge for inductive knowledge graph completion," in *Proceedings of the AAAI Conference on Artificial Intelligence*, 2025, vol. 39, pp. 12355–12363.

[19] Jing Zhang, Deng-Ping Fan, Yuchao Dai, Xin Yu, Yiran Zhong, Nick Barnes, and Ling Shao, "Rgb-d saliency detection via cascaded mutual information minimization," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021, pp. 4338–4347.

[20] Robert F Woolson, "Wilcoxon signed-rank test," *Wiley Encyclopedia of Clinical Trials*, pp. 1–3, 2007.

[21] Vassilis N. Ioannidis, Xiang Song, Saurav Manchanda, Mufei Li, Xiaoqin Pan, Da Zheng, Xia Ning, Xiangxiang Zeng, and George Karypis, "Drkg - drug repurposing knowledge graph for covid-19," `https://github.com/gnn4dr/DRKG/`, 2020.

[22] Di Wu, Wu Sun, Yi He, Zhong Chen, and Xin Luo, "Mkgfenn: A multimodal knowledge graph fused end-to-end neural network for accurate drug–drug interaction prediction," in *Proceedings of the AAAI Conference on Artificial Intelligence*, 2024, vol. 38, pp. 10216–10224.

[23] Weiwei Cheng and Eyke Hüllermeier, "Combining instance-based learning and logistic regression for multilabel classification," *Machine Learning*, vol. 76, no. 2, pp. 211–225, 2009.

[24] Geoffrey E Hinton and Ruslan R Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, no. 5786, pp. 504–507, 2006.