Multi-relational Contrastive Learning Graph Neural Network for Drug-drug Interaction Event Prediction

Zhankun Xiong[1*], Shichao Liu[1*], Feng Huang[1*], Ziyan Wang[1], Xuan Liu[1], Zhongfei Zhang[2], Wen Zhang[1†]

[1]College of Informatics, Huazhong Agricultural University
[2]Computer Science Department, Binghamton University
xiongzk@webmail.hzau.edu.cn, scliu@mail.hzau.edu.cn, {fhuang233, wangziyan, lx666}@webmail.hzau.edu.cn, zzhang@binghamton.edu, zhangwen@mail.hzau.edu.cn

AAAI-2023

Code: https://github.com/Zhankun-Xiong/MRCGNN

Reported by Dongdong Hu

Introduction

Most existing GNN-based models ignore either drug structural information or drug interactive information, but both aspects of information are important for DDI event prediction.

Furthermore, accurately predicting rare DDI events is hindered by their inadequate labeled instances.

Method

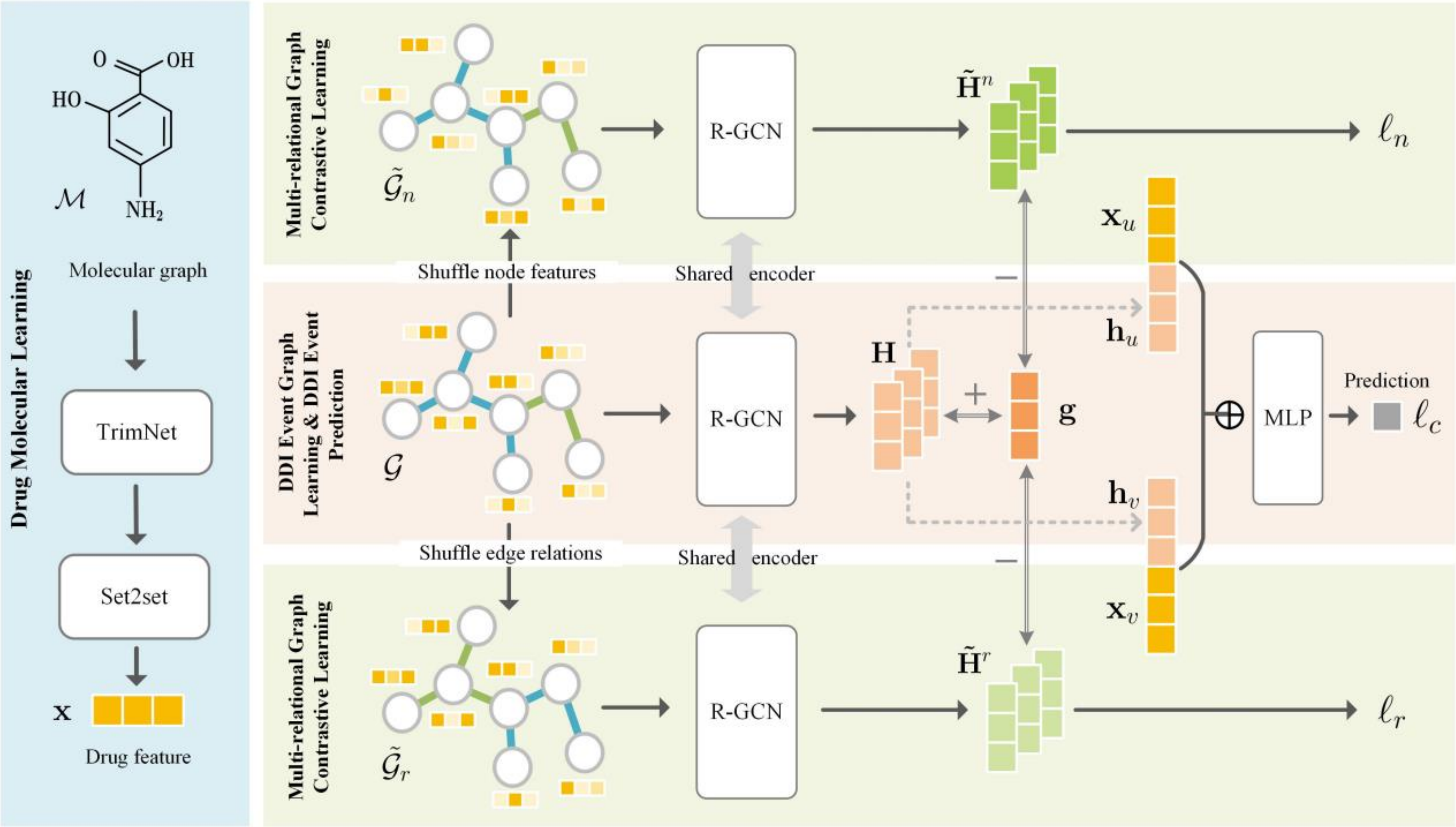

Figure 1: Overview of the proposed MRCGNN.

Method

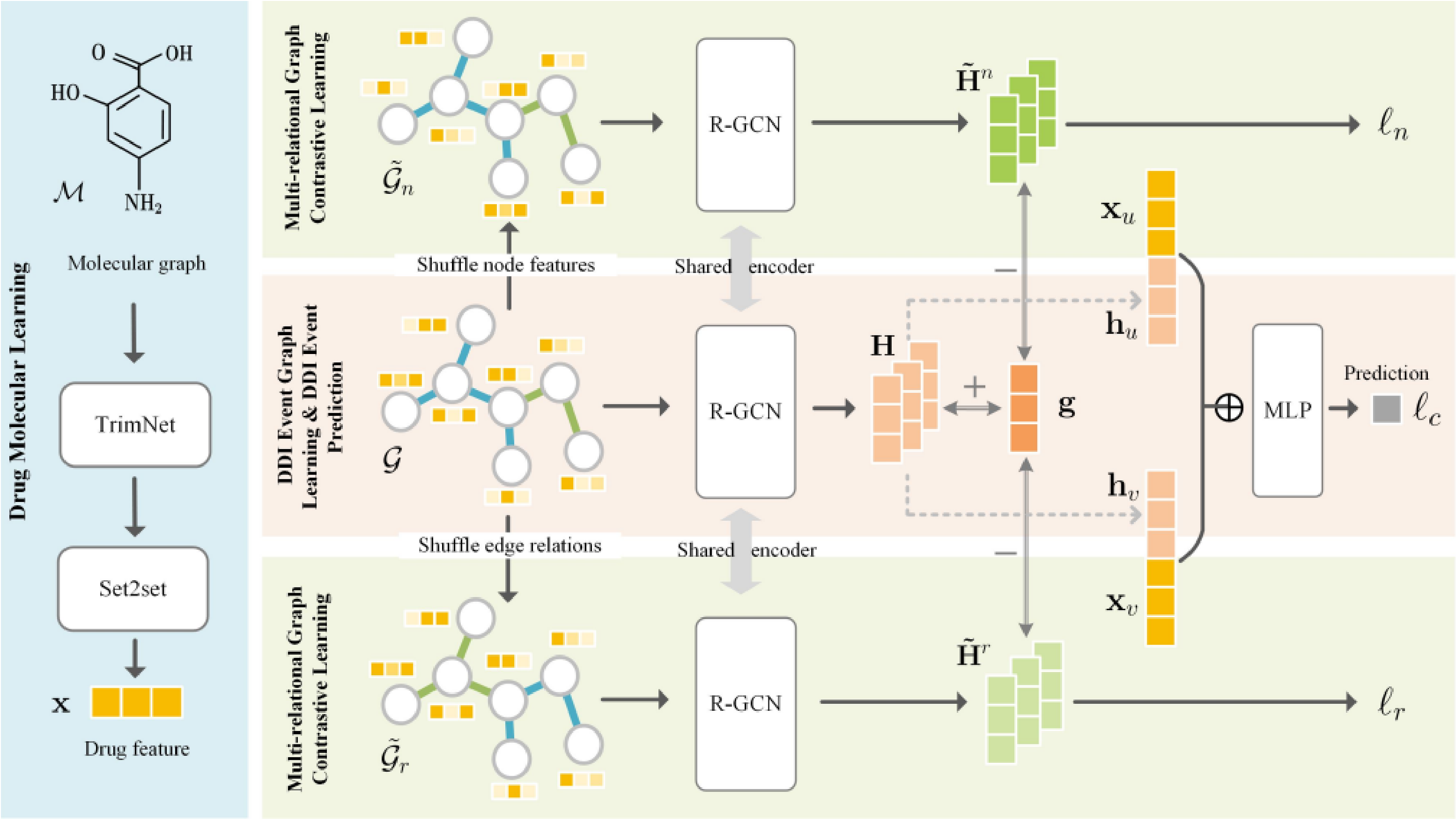

Figure 1: Overview of the proposed MRCGNN.

Drug Molecular Learning

$$\mathbf{s}_i^{(t+1)} = \mathrm{U}(\mathbf{s}_i^{(t)}, \sum\nolimits_{j\in\mathcal{N}_i} \mathbf{m}_{ij}^{(t+1)})$$
$$\mathbf{m}_{ij}^{(t+1)} = \mathrm{M}(\mathbf{s}_i^{(t)}, \mathbf{s}_j^{(t)}, \mathbf{e}_{ij}^{(t)}). \quad (1)$$

Note that $\mathcal{N}_i$ represents the neighbors of atom i, $\mathbf{m}_{ij}^{(t+1)}$ is the message from atom j to i, $\mathbf{e}_{ij}^{(t)}$ denotes the hidden state of the edge between i and j, and the update function U, a gated recurrent unit

$$\mathrm{M}(\mathbf{s}_i, \mathbf{s}_j, \mathbf{e}_{ij}) = ||_{k=1}^{K} \alpha_{ij}^{k} \odot \mathbf{W}_s^k \mathbf{s}_j \odot \mathbf{W}_e^k \mathbf{e}_{ij}$$
$$\alpha_{ij} = \mathrm{Softmax}(\sigma(\mathbf{u}^{\mathrm{T}}[\mathbf{W}_s \mathbf{s}_i || \mathbf{W}_e \mathbf{e}_{ij} || \mathbf{W}_s \mathbf{s}_j])), \quad (2)$$

where || represents vector concatenation operation, $\odot$ is the element-wise product,

$$\mathbf{X} \in \mathbb{R}^{|\mathcal{V}|\times F}$$

Method

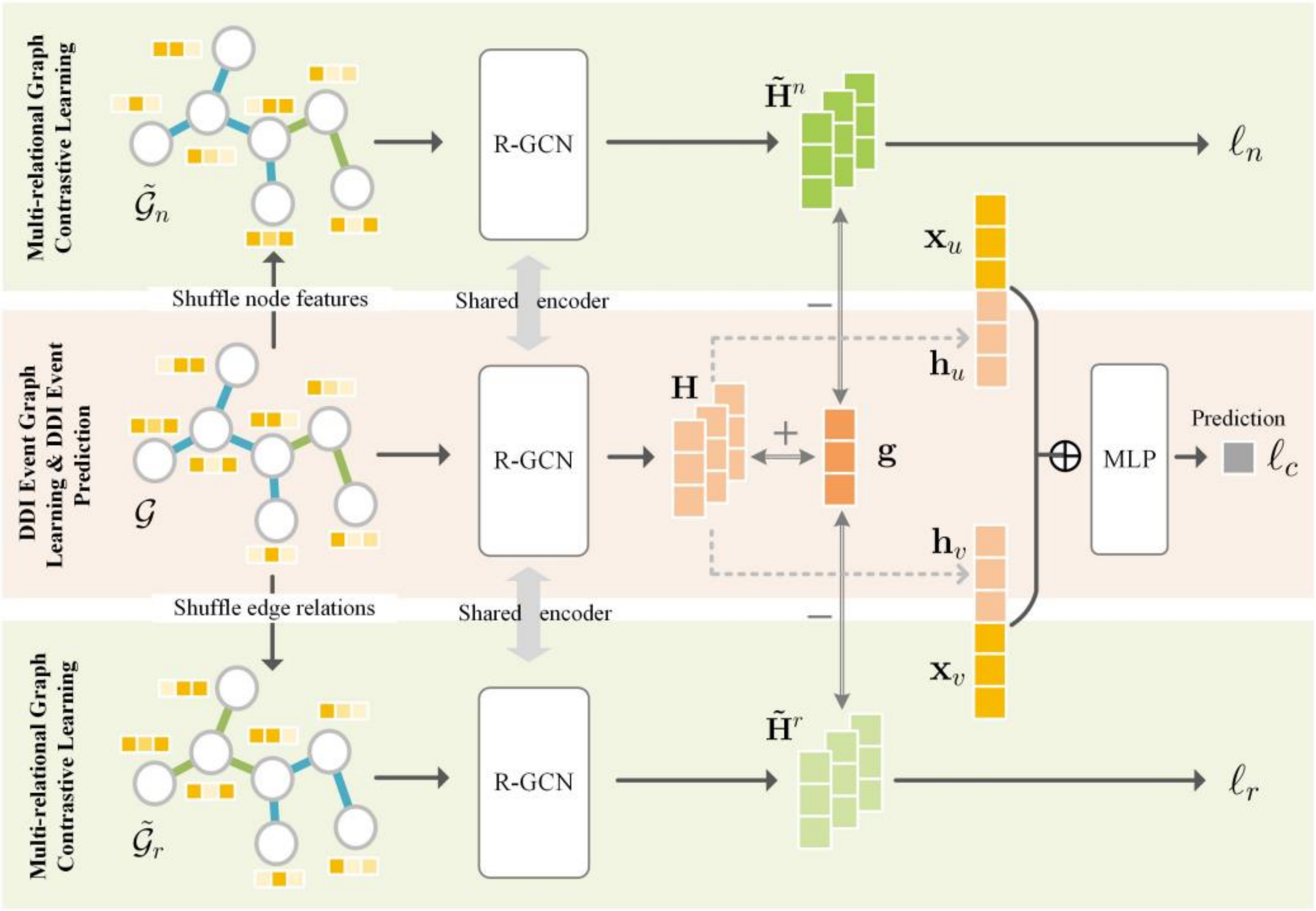

Figure 1: Overview of the proposed MRCGNN.

Drug-drug Interaction Event Graph Learning

$$\mathbf{h}_v^{(l+1)} = \sigma\left(\sum_{r\in\mathcal{R}}\sum_{u\in\mathcal{N}_v^r}\frac{1}{c_{vr}}\mathbf{W}_r^{(l)}\mathbf{h}_u^{(l)} + \mathbf{W}_o^{(l)}\mathbf{h}_v^{(l)}\right), \quad (3)$$

$$\mathbf{h}_v = \sum_{l=1}^{L}\alpha_l\mathbf{h}_v^{(l)}, \quad (4)$$

where α_l is a trainable parameter

$\mathbf{h}_v \in \mathbb{R}^Q$. $\mathbf{H} \in \mathbb{R}^{|\mathcal{V}|\times Q}$

Multi-Relational Contrastive Learning

shuffles the drug features $\mathbf{X}$

$$\mathcal{G} = (\mathcal{V},\mathcal{E},\mathcal{R}) \rightarrow \tilde{\mathcal{G}}_n = (\tilde{\mathcal{V}},\mathcal{E},\mathcal{R}) \qquad \tilde{\mathbf{H}}^n$$

shuffles the edge relations

$$\mathcal{G} = (\mathcal{V},\mathcal{E},\mathcal{R}) \rightarrow \tilde{\mathcal{G}}_r = (\mathcal{V},\mathcal{E},\tilde{\mathcal{R}}) \qquad \tilde{\mathbf{H}}^r$$

$$\mathbf{g} = \Gamma(\mathbf{H}) \qquad \mathbf{g} \in \mathbb{R}^Q$$

Method

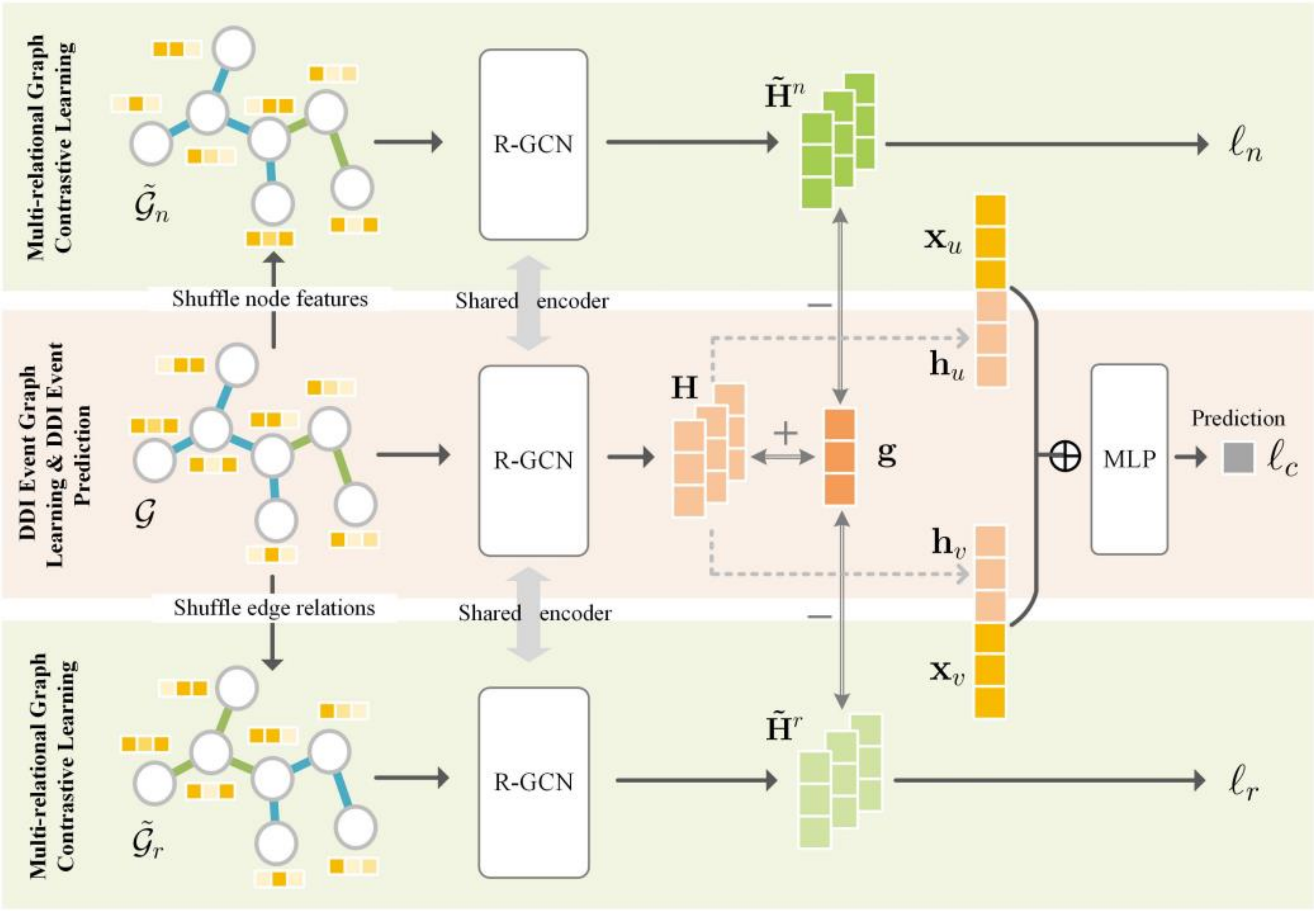

Figure 1: Overview of the proposed MRCGNN.

$$\ell_n = -\frac{1}{|\mathcal{V}| + |\tilde{\mathcal{V}}|}\left(\sum_{v\in\mathcal{V}} \mathbb{E}_{(\mathcal{V},\mathcal{E},\mathcal{R})}[\log \mathrm{D}(\mathbf{h}_v, \mathbf{g})] + \sum_{u\in\tilde{\mathcal{V}}} \mathbb{E}_{(\tilde{\mathcal{V}},\mathcal{E},\mathcal{R})}[\log(1 - \mathrm{D}(\tilde{\mathbf{h}}^n_u, \mathbf{g}))]\right)$$

$$\ell_r = -\frac{1}{|\mathcal{V}| + |\mathcal{V}|}\left(\sum_{v\in\mathcal{V}} \mathbb{E}_{(\mathcal{V},\mathcal{E},\mathcal{R})}[\log \mathrm{D}(\mathbf{h}_v, \mathbf{g})] + \sum_{u\in\mathcal{V}} \mathbb{E}_{(\mathcal{V},\mathcal{E},\tilde{\mathcal{R}})}[\log(1 - \mathrm{D}(\tilde{\mathbf{h}}^r_u, \mathbf{g}))]\right), \quad (5)$$

where $D(\mathbf{h}_v, \mathbf{g}) = \sigma(\mathbf{h}_v^{\mathrm{T}}\mathbf{W}\mathbf{g})$ and $\mathbf{W}$ is a trainable parameter matrix.

DDI Event Prediction

$$\mathbf{h}_{(u,v)} = \mathbf{h}_u || \mathbf{x}_u || \mathbf{h}_v || \mathbf{x}_v$$

$$\hat{\mathbf{y}}_{(u,v)} = \mathrm{Softmax}(\mathrm{MLP}(\mathbf{h}_{(u,v)})), \quad (6)$$

$$\ell_c = -\sum_{(u,v)\in\Omega}\sum_{r\in\mathcal{R}} y^r_{(u,v)} \log \hat{y}^r_{(u,v)}, \quad (7)$$

$$\ell = \ell_c + \alpha\ell_r + \beta\ell_n, \quad (8)$$

Experiments

Table 1: Proportions of events in five groups to all events.

Datasets	Five groups				
	[1,10]	(10,50]	(50,100]	(100,300]	$(300,+\infty)$
Deng's	20.0%	21.5%	24.6%	15.4%	18.5%
Ryu's	5.8%	21.0%	11.6%	14.0%	47.6%

Table 2: Results of MRCGNN and baselines for DDI event prediction on two datasets.

Methods	Deng's dataset				Ryu's dataset			
	Acc.	Macro-F1	Macro-Rec.	Macro-Prec.	Acc.	Macro-F1	Macro-Rec.	Macro-Prec.
DeepDDI	0.7807	0.6055	0.5839	0.6611	0.9323	0.8643	0.8512	0.8928
SSI-DDI	0.7866	0.4216	0.3896	0.5139	0.9008	0.6663	0.6287	0.7507
TrimNet-DDI	0.8570	0.6548	0.6363	0.7046	0.9353	0.8288	0.8128	0.8627
MUFFIN	0.8269	0.5245	0.4844	0.6204	0.9510	0.8566	0.8339	0.8980
R-GCN	0.8695	0.7026	0.6878	0.7500	0.9284	0.8487	0.8291	0.8881
GoGNN	0.8766	0.6938	0.6841	0.7316	0.9424	0.8589	0.8451	0.8949
MRCGNN	**0.8979**	**0.7791**	**0.7688**	**0.8101**	**0.9566**	**0.8894**	**0.8727**	**0.9221**

Experiments

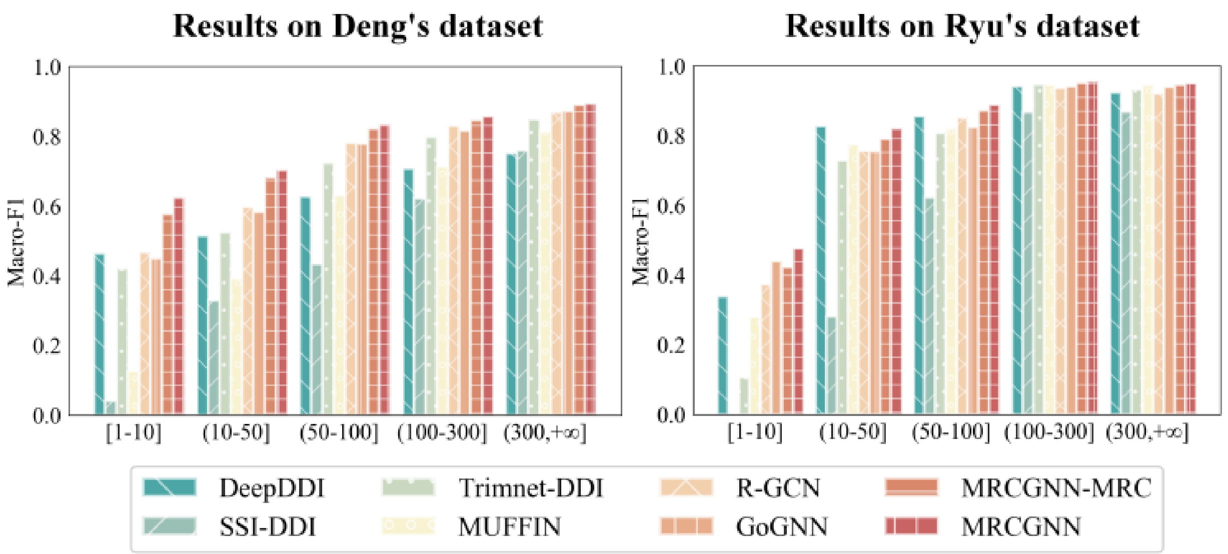

Figure 2: Results of MRCGNN and baselines on five groups of events.

Experiments

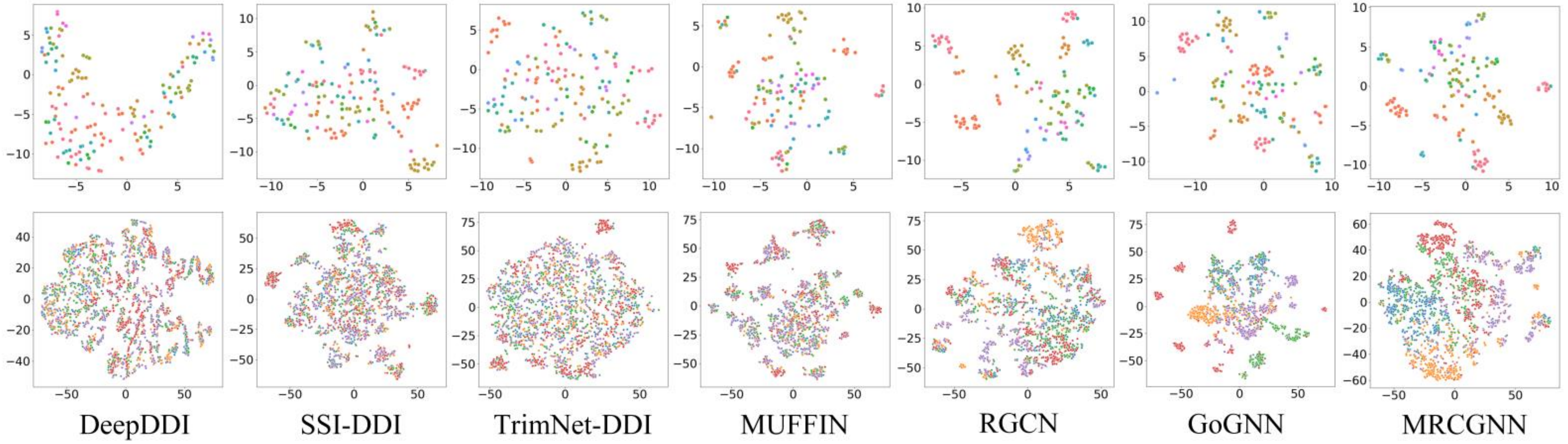

Figure 3: Visualization on Deng's dataset using the t-SNE. Each point represents a drug pair, and the color represents the DDI event. Upper: 20 events with the lowest frequency. Lower: 5 events with the highest frequency.

Experiments

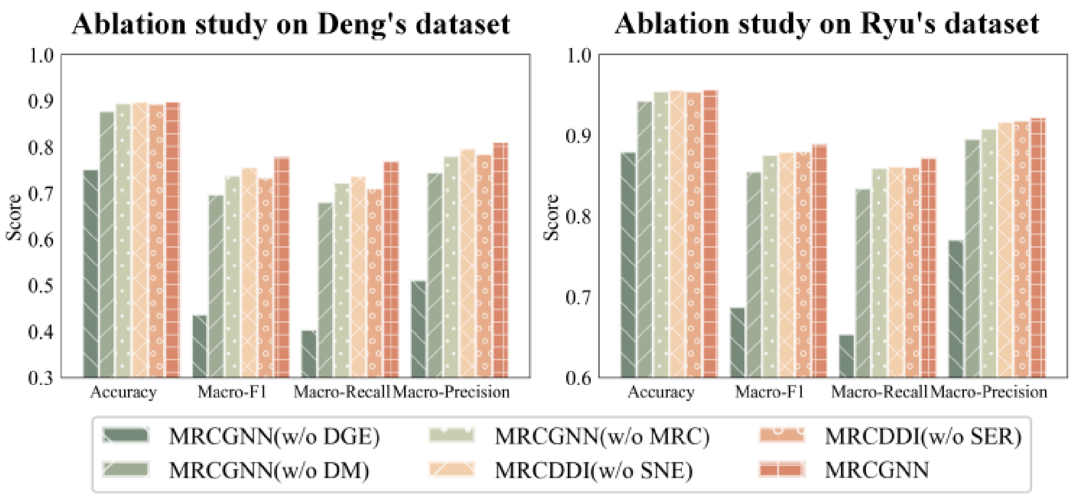

Figure 4: Results of MRCGNN and its variants in ablation study.

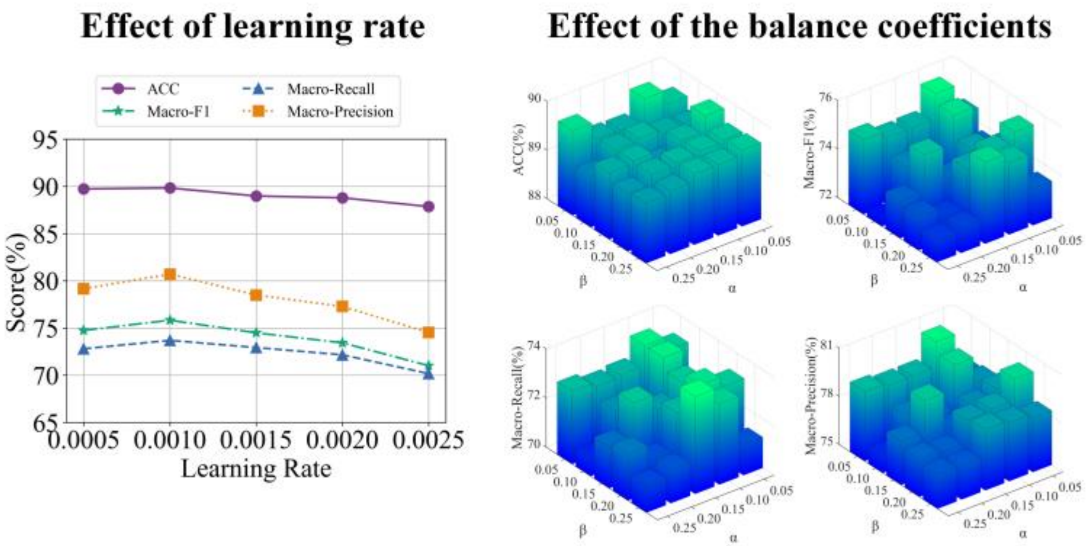

Figure 5: Hyper-parameter Sensitivity Analysis

Thanks