The 33nd International Joint Conference on Artificial Intelligence
3th-9th August 2024, Jeju Island, South Korea

Paper ID:3261

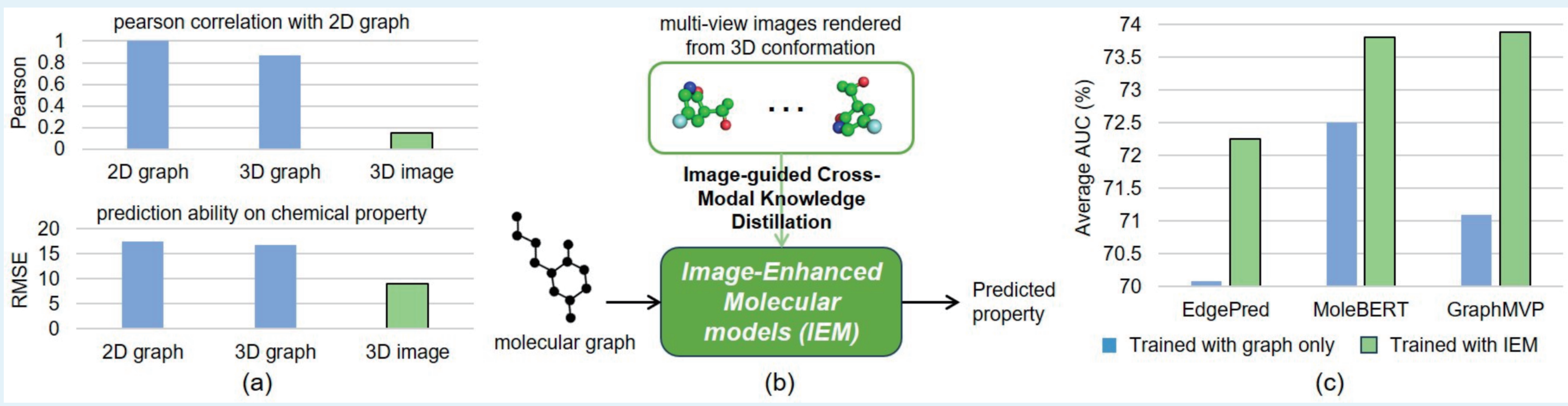
An Image-enhanced Molecular Graph Representation Learning Framework

Hongxin Xiang[1,2], Shuting Jin[3], Jun Xia[4], Man Zhou[5], Jianmin Wang[6], Li Zeng[2], Xiangxiang Zeng[1]

[1]HNU, [2]Yuyao Biotechnology Co., Ltd, [3]WUST, [4]Westlake University, [5]USTC, [6]Yonsei University

Introduction

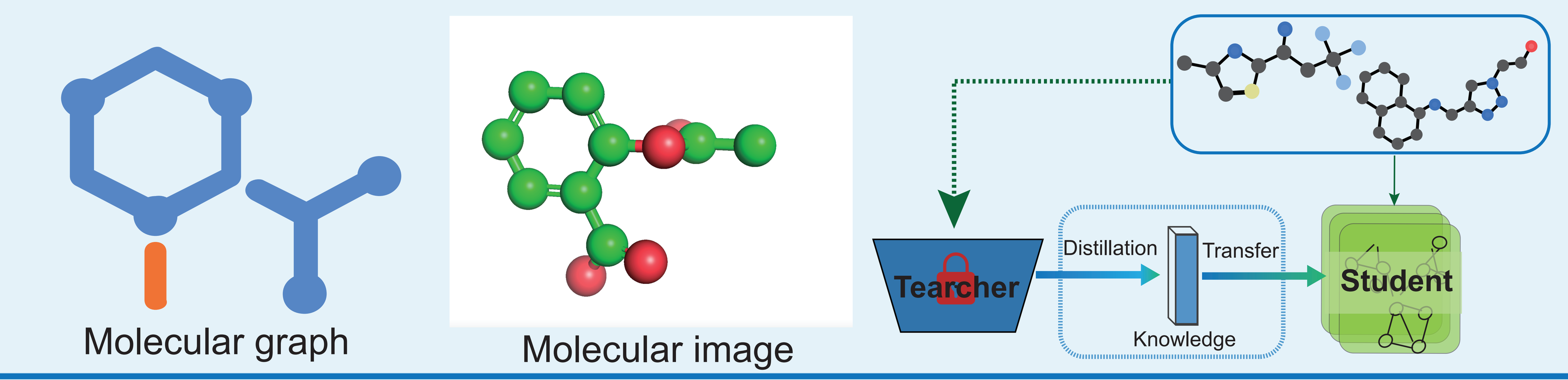
- **Molecular Representation Learning** plays an import role in high-precision drug discovery (such as molecular property prediction, target activity prediction).
  - **Limited by a single modality:** The paradigm of learning from a single modality gradually encounters the bottleneck of limited representation capabilities.
  - **Multimodal fusion has limited improvements:** 1) similar modalities and encoding ways. 2) weak feature extraction ability, resulting in insufficient complementary information between modalities.



- **Our method:** We propose an image-enhanced molecular graph representation learning framework (called **IEM**) that leverages multi-view molecular images rendered from 3D conformations to boost molecular graph representations.

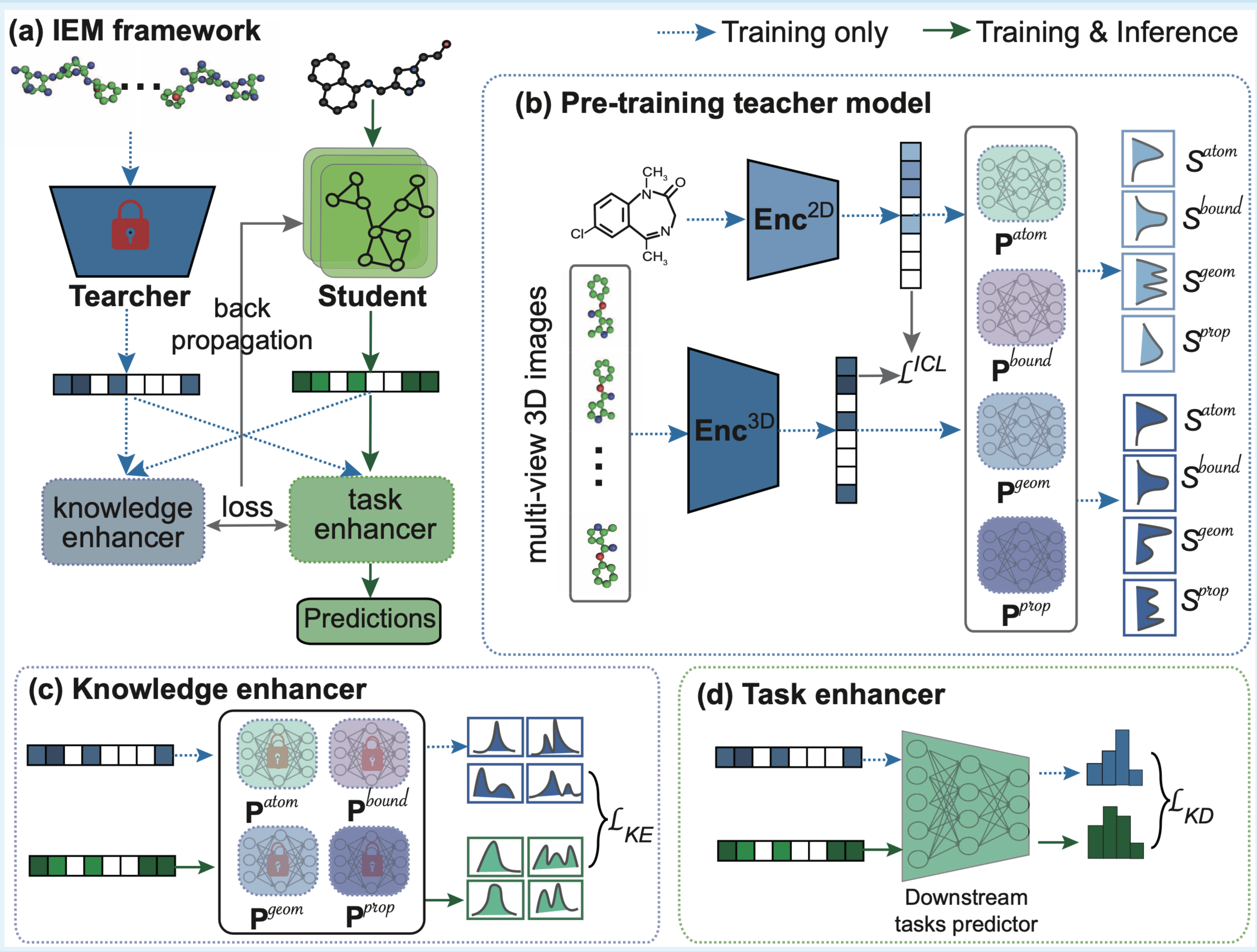
Related Work

- **Graph-based Molecular Representation Learning:** In view of the high cost of annotating molecules, recent studies mainly learn from large-scale label-free molecular databases by designing pre-training strategies.
- **Image-based Molecular Representation Learning:** Because graphs are discrete and unordered, some researchers consider representing molecules as images and utilizing mature computer vision techniques to extract features.
- **Cross-Modal Knowledge Distillation:** As an important branch of knowledge distillation, crossmodal knowledge distillation (CMKD) is still a relatively emerging field, which refers to using a teacher from another modality to supervise the learning model of the current modality and improve the performance of the student during inference.



IEM Framework

- **Overview of the IEM:** The image-enhanced molecular graph representation learning framework (IEM), which **equips knowledgeable teachers and distillation strategies** to prevent negative transfer
- **The process of pre-training the teacher:** Use **5 pre-training tasks** to train a knowledgeable teacher.
- **Execution process of the knowledge enhancer:** Exploit **image-based teacher** to enhance graph-based student by using the knowledge enhancer and task enhancer.
- **Execution process of the task enhancer:** Train IEM and inference in downstream tasks.



- **IEM has the following advantages:** (1) **Universality**: IEM can be integrated with any graph-based method. (2) **Effectiveness**: IEM significantly improves the performance of several graph-based baselines. (3) **Efficiency**: As low as 5% of training images can still improve performance; (4) **Compatibility**: IEM is compatible with both 2D and 3D molecular images and different rendering strategies.

Results

- **Datasets:** 1) 2 millions unlabeled molecules with 3D conformations from PCQM4Mv2 database. 2) 8 binary classification datasets from MoleculeNet. 3) 4 regression datasets included in GraphMVP.
- **Comparsion with other methods in classification tasks:** The ROC-AUC (%) performance of different methods on 8 classification datasets of MPP.

	Tox21	ToxCast	Sider	ClinTox	MUV	HIV	BBBP	BACE	Average
#Molecules	7831	8576	1427	1478	93087	41127	2039	1513	-
#Task	12	617	27	2	17	1	1	1	-
GIN [Xu *et al.*, 2018]	74.3(0.9)	61.5(0.8)	57.3(1.2)	57.2(4.1)	71.6(2.8)	75.2(2.0)	66.7(1.8)	69.6(5.5)	66.68
IEM-GIN	74.5(0.4)	62.5(0.8)	59.1(1.7)	62.6(4.1)	77.7(2.9)	77.9(1.3)	69.3(1.9)	77.7(3.5)	70.16
Δ	↑ 0.2	↑ 1.0	↑ 1.8	↑ 5.4	↑ 6.1	↑ 2.7	↑ 2.6	↑ 8.1	↑ 3.5
EdgePred [Hu *et al.*, 2020a]	76.0(0.6)	64.1(0.6)	60.4(0.7)	64.1(3.7)	75.1(1.2)	76.3(1.0)	67.3(2.4)	77.3(3.5)	70.08
IEM-EdgePred	76.3(0.6)	64.6(0.6)	61.2(0.6)	67.5(2.3)	78.3(1.3)	78.3(1.3)	67.8(2.2)	**84.1(0.8)**	72.26
Δ	↑ 0.3	↑ 0.5	↑ 0.8	↑ 3.4	↑ 3.2	↑ 2.0	↑ 0.5	↑ 6.8	↑ 2.2
GraphMVP [Liu *et al.*, 2021]	74.5(0.7)	63.4(0.5)	60.7(1.4)	78.4(6.4)	73.0(2.3)	75.6(1.6)	67.4(2.4)	75.8(3.0)	71.10
IEM-GraphMVP	75.9(0.7)	64.4(0.6)	61.9(1.7)	**80.8(3.1)**	77.3(1.2)	**78.8(1.1)**	68.7(1.0)	83.3(1.4)	**73.89**
Δ	↑ 1.4	↑ 1.0	↑ 1.2	↑ 2.4	↑ 4.3	↑ 3.2	↑ 1.3	↑ 7.5	↑ 2.8
GraphMVP-C [Liu *et al.*, 2021]	74.6(0.4)	63.4(0.6)	60.6(1.3)	76.9(3.7)	72.8(2.4)	77.1(2.1)	69.9(1.4)	79.6(1.7)	71.86
IEM-GraphMVP-C	75.6(0.6)	64.8(0.5)	62.0(0.9)	79.2(2.9)	77.0(1.7)	78.2(1.0)	**71.4(1.4)**	81.9(1.6)	73.76
Δ	↑ 1.0	↑ 1.4	↑ 1.4	↑ 2.3	↑ 4.2	↑ 1.1	↑ 1.5	↑ 2.3	↑ 1.9
Mole-BERT [Xia *et al.*, 2023]	77.0(0.3)	64.4(0.2)	63.2(0.7)	72.7(2.7)	79.2(2.0)	77.7(0.7)	65.7(2.3)	80.2(0.9)	72.51
IEM-Mole-BERT	**77.8(0.4)**	**65.6(0.3)**	**65.3(0.8)**	72.2(1.4)	**79.7(1.8)**	**78.8(0.6)**	68.1(1.0)	83.0(0.9)	73.81
Δ	↑ 0.8	↑ 1.2	↑ 2.1	-0.5	↑ 0.5	↑ 1.1	↑ 2.4	↑ 2.8	↑ 1.3

- **Comparsion with other methods in regression tasks:** The ROC-AUC (%) performance of different methods on 4 regression datasets of MPP.

	ESOL	Lipo	Malaria	CEP
#Molecules	1,128	4,200	9,999	29,978
#Task	1	1	1	1
GIN w/o pre-train	1.472(0.038)	0.832(0.025)	1.113(0.011)	1.340(0.018)
IEM-GIN	1.346(0.045)	0.817(0.019)	1.084(0.003)	1.329(0.021)
Δ	↑ 8.56%	↑ 1.80%	↑ 2.61%	↑ 0.82%
EdgePred	1.367(0.041)	0.778(0.013)	1.110(0.011)	1.362(0.025)
IEM-EdgePred	1.350(0.027)	0.769(0.006)	1.088(0.005)	1.345(0.016)
Δ	↑ 1.24%	↑ 1.16%	↑ 1.98%	↑ 1.25%
GraphMVP	1.322(0.062)	0.773(0.016)	1.128(0.019)	1.308(0.024)
IEM-GraphMVP	1.281(0.044)	0.754(0.015)	1.089(0.005)	**1.294**(0.020)
Δ	↑ 3.10%	↑ 2.46%	↑ 3.46%	↑ 1.07%
GraphMVP-C	1.333(0.055)	0.768(0.013)	1.114(0.008)	1.304(0.020)
IEM-GraphMVP-C	1.274(0.037)	0.761(0.017)	1.090(0.004)	1.296(0.012)
Δ	↑ 4.43%	↑ 0.91%	↑ 2.15%	↑ 0.61%
MoleBERT	1.115(0.017)	0.727(0.006)	1.137(0.021)	1.350(0.015)
IEM-MoleBERT	1.090(0.031)	**0.716**(0.003)	**1.080**(0.003)	1.343(0.013)
Δ	↑ 2.24%	↑ 1.51%	↑ 5.01%	↑ 0.52%
GraphMVP-F	1.094(0.037)	0.724(0.009)	1.106(0.013)	1.397(0.040)
IEM-GraphMVP-F	**1.067**(0.039)	**0.716**(0.010)	1.093(0.012)	1.392(0.026)
Δ	↑ 2.47%	↑ 1.10%	↑ 1.18%	↑ 0.36%

- **Different GNN Architectures:** The average ROC-AUC (%) performance on 8 classification datasets with different GNN architectures.

	GCN	GIN	GAT	GraphSAGE
w/o IEM	66.88	66.68	66.53	66.99
w/ IEM	69.81	70.16	69.76	69.61
Δ	↑ 4.39%	↑ 5.23%	↑ 4.87%	↑ 3.92%

- **Different Image Rendering Strategies:** The average ROC-AUC (%) performance on 8 classification datasets with different image rendering methods.

Image rendering		Method	
Image type	Rendering strategy	EdgePred	GraphMVP
×	×	70.08	71.1
2D	RDKit	72.21 (↑ 3.04%)	73.34 (↑ 3.15%)
2D	PyMol	72.00 (↑ 2.74%)	73.41 (↑ 3.25%)
3D	PyMol	**72.26** (↑ 3.11%)	**73.89** (↑ 3.92%)

- **Image Efficiency:** The average ROC-AUC (%) performance on 8 classification datasets with different number of images.

	image size					
	0%	5%	10%	20%	50%	100%
IEM	71.10	72.20	72.26	72.95	73.38	73.89
Δ	-	↑ 1.55%	↑ 1.64%	↑ 2.60%	↑ 3.20%	↑ 3.92%

- **Ablation Study:** Ablation results on knowledge enhancer (KE) and task enhancer (TE).

Enhancer		Method			
KE	TE	GCN	GIN	GAT	GraphSAGE
×	×	66.88	66.68	66.53	66.99
×	✓	68.07 (1.19)	68.16 (1.48)	68.48 (1.95)	68.44 (1.45)
✓	×	68.26 (1.38)	68.60 (1.92)	68.59 (2.06)	68.58 (1.59)
✓	✓	69.81 (2.93)	70.16 (3.48)	69.76 (3.23)	69.61 (2.62)

Code