AFRL-AFOSR-JP-TR-2022-0002

Generic Multi-task Learning Framework for Structural Mixed Data using Deep Neural Network Feature Sharing Architecture -- Application to Predictive Tasks in Vietnamese Herbal Medicine

Ly, Le
INTERNATIONAL UNIVERSITY VIETNAM NATIONAL UNIVERSITY-HCM
QUARTER 6, LINH TRUNG, THU DUC DIST.
HO CHI MINH, , 700000
VN

11/09/2021
Final Technical Report

Air Force Research Laboratory
Air Force Office of Scientific Research
Asian Office of Aerospace Research and Development
Unit 45002, APO AP 96338-5002

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
09-11-2021	Final	08 Apr 2019 - 07 Apr 2021

4. TITLE AND SUBTITLE
Generic Multi-task Learning Framework for Structural Mixed Data using Deep Neural Network Feature Sharing Architecture -- Application to Predictive Tasks in Vietnamese Herbal Medicine

5a. CONTRACT NUMBER
FA2386-19-1-4032

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Le Ly

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
INTERNATIONAL UNIVERSITY VIETNAM NATIONAL UNIVERSITY-HCM
QUARTER 6, LINH TRUNG, THU DUC DIST.
HO CHI MINH, 700000
VN

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
AOARD
UNIT 45002
APO AP 96338-5002

10. SPONSOR/MONITOR'S ACRONYM(S)
AFRL/AFOSR IOA

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
AFRL-AFOSR-JP-TR-2022-0002

12. DISTRIBUTION/AVAILABILITY STATEMENT
A Distribution Unlimited: PB Public Release

13. SUPPLEMENTARY NOTES

14. ABSTRACT
While the development of new drugs is costly, time-consuming, and often accompanied by safety issues, drug repurposing, where old drugs with established safety are used for medical conditions other than originally developed, is an attractive alternative. In this project, our purpose is to use a Graph convolutional neural network to solve several tasks in drug discovery. First, drug-mediated toxicity is a heavy burden to the pharmaceutical industry, leading to safety-related failures in development and the high cost of drug discovery. Second, we consider the anti-cancer screening task which assesses the positive or negative of drug response to different types of cancer in humans. The success of deep learning on high-throughput chemical structure screening has offered unprecedented opportunities to detect toxicity compound candidates or negative responses to cancer cells. However, capturing graph structure and a very high number of toxicity/cancer type tasks of chemical structure has been challenging. We proposed to build a multitask graph neural network which is based on AdaShare capsule networks, graph convolutional neural network, and multi-task learning to predict the anti-cancer and toxic effect which represent in the chemical structure graph. We implemented two methods for toxicity testing prediction and anti-cancer screening by using the adaptive sharing features. The first method is adaptive sharing based on residual graph network and policy network and the second method is Capsule graph network with Task routing network. The graph-based AdaShare showed a potential performance on Tox21 datasets. Single task CapsGNN shows good performance. Multitask CapsGNN has been implemented and is ready for the training process, but the results have not been obtained yet. We continue to work on this and test on NCI and Tox21 datasets. We expect that the capsule mechanism of Multitask CapsGNN achieves better performance with a smaller number of parameters. A case study using multi-task solving approach on Drug-Induced Liver Injury (DILI) is also presented.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			ALAN LIN
U	U	U	SAR	12	**19b. TELEPHONE NUMBER** *(Include area code)* 227-7009

Final Report for AOARD Grant FA2386-19-1-4032

"Applying Graph Convolutional-based Multi-task Learning for Vietnamese Herbal Medicine Database Knowledge Discovery"

09 November 2021

Name of Principal Investigators (PI and Co-PIs): Ly Le

- e-mail address : ly.le@hcmiu.edu.vn
- Institution : International University, Vietnam National University
- Mailing Address : Ward 6, Linh Trung, Thu Duc City, Vietnam
- Phone : 0906578836
- Fax :

Period of Performance: 04/08/2019 –04/07/2021

Abstract:

While the development of new drugs is costly, time-consuming, and often accompanied by safety issues, drug repurposing, where old drugs with established safety are used for medical conditions other than originally developed, is an attractive alternative. In this project, our purpose is to use a Graph convolutional neural network to solve several tasks in drug discovery. First, drug-mediated toxicity is a heavy burden to the pharmaceutical industry, leading to safety-related failures in development and the high cost of drug discovery. Second, we consider the anti-cancer screening task which assesses the positive or negative of drug response to different types of cancer in humans. The success of deep learning on high-throughput chemical structure screening has offered unprecedented opportunities to detect toxicity compound candidates or negative responses to cancer cells. However, capturing graph structure and a very high number of toxicity/cancer type tasks of chemical structure has been challenging. We proposed to build a multitask graph neural network which is based on AdaShare capsule networks, graph convolutional neural network, and multi-task learning to predict the anti-cancer and toxic effect which represent in the chemical structure graph. We implemented two methods for toxicity testing prediction and anti-cancer screening by using the adaptive sharing features. The first method is adaptive sharing based on residual graph network and policy network and the second method is Capsule graph network with Task routing network. The graph-based AdaShare showed a potential performance on Tox21 datasets. Single task CapsGNN shows good performance. Multitask CapsGNN has been implemented and is ready for the training process, but the results have not been obtained yet. We continue to work on this and test on NCI and Tox21 datasets. We expect that the capsule mechanism of Multitask CapsGNN achieves better performance with a smaller number of parameters. A case study using multi-task solving approach on Drug-Induced Liver Injury (DILI) is also presented.

Introduction:

I. Multi-task Graph Neural Network

Deep learning methods such as multi-task neural networks, singletask neural network or traditional machine learning have recently been applied to several applications, such as ligand-based virtual screening, reactivity to biological macromolecules,...[1][2]. Recently Graph Convolutional Neural Networks (GCNs) [3] and Multi-task Learning (MLT) have emerged as a powerful tool for computational drug discovery [4]. GCNs are generalizations of classical CNNs

to handle graph data such as molecular data. MLT is a subfield of machine learning in which multiple learning tasks are solved at the same time, while exploiting commonalities and differences across tasks. We have investigated the Graph Convolutional-based Multi-task Neural Network to solve Quantitative Structure-Activity Relationship to predict different aspects of drug-like molecule - metabolites in drug discovery field. As the graph data for different types of cancer may share common substructures, learning multiple related tasks together may potentially help improve the generalization performance of each single task [5].

Due to the above motivation, we propose to build a multitask graph neural network which is based on Adashare[6], CapsGNN [7], Taskrouting [8] and Multi-task Learning [5].

II. Investigating on drug-induced liver injury (DILI)

The human cytochrome P450 (CYP) superfamily holds responsibilities for the metabolism of both endogenous and exogenous compounds such as drugs, cellular metabolites, and toxins. Inhibition of CYP450 isoforms is closely associated with adverse drug reactions which may cause metabolic failures and even induce serious side effects. In modern drug discovery and development, identification of potential CYP isoforms' inhibitors is highly essential. Besides experimental approaches, numerous computational frameworks have been recently developed to address this biological issue.

In our study, we propose robust, stable, and effective prediction models for the virtual screening of ve CYP isoforms' inhibitors, including CYP1A2, CYP2C9, CYP2C19, CYP2D6, and CYP3A4. The method employs multitask learning combined with molecular fingerprint-embedded encoding to boost the predictive power. Our results showed that multitask learning had remarkably leveraged useful information from related tasks to promote global performance.

In comparison with several state-of-the-art methods, our proposed method is outperforming in all tasks with the highest area under the receiver operating characteristic (ROC) curve (ROC-AUC) of 0.93 and area under the precision-recall (PR) curve (PR-AUC) of 0.92. The evaluated performance once confirms our model's robustness, stability, and efficiency.

Method:

I. Multi-task Graph Neural Network

Given a set of K tasks T = {T1, T2, …, T3} defined a dataset for a graph classification task. We introduce two architectures: Adaptive Sharing based on Residual graph network and Policy network and Capsule graph network with Task routing network which performed the feature sharing mechanism that decides which network layers should be shared across which tasks and which layers should be task-specific.

a) Model architecture

i. The graph-based Adashare

AdaShare [6] is an approach for adaptively determining the feature sharing strategy across multiple tasks in deep multi-task learning. They learn the feature sharing policy and network weights jointly using standard backpropagation without adding any significant number of parameters.

The main idea is to learn the sharing pattern through a task-specific policy that selectively

chooses which layers to execute for a given task in the multi-task network. This means that a single network for multi-task learning will separate execution paths for different tasks.

Adshare showed the effectiveness of our proposed approach on five standard datasets (including NYU v2, CityScapes, Tiny-Taskonomy, DomainNet, and Text Classification), outperforming several competing methods (Cross-Stitch Networks, Sluice Networks, NDDR-CNN…) [6]. The architecture not only achieves high accuracy on all tasks, but also restricts the number of new network parameters as much as possible as the number of tasks grows.

In this project, we transformed the architecture to graph-based network style with two main components which are Residual graph network as backbone and Policty network for the Task-specific Policy. we are aiming to solve the multi-task learning for the graph chemical structure problem.

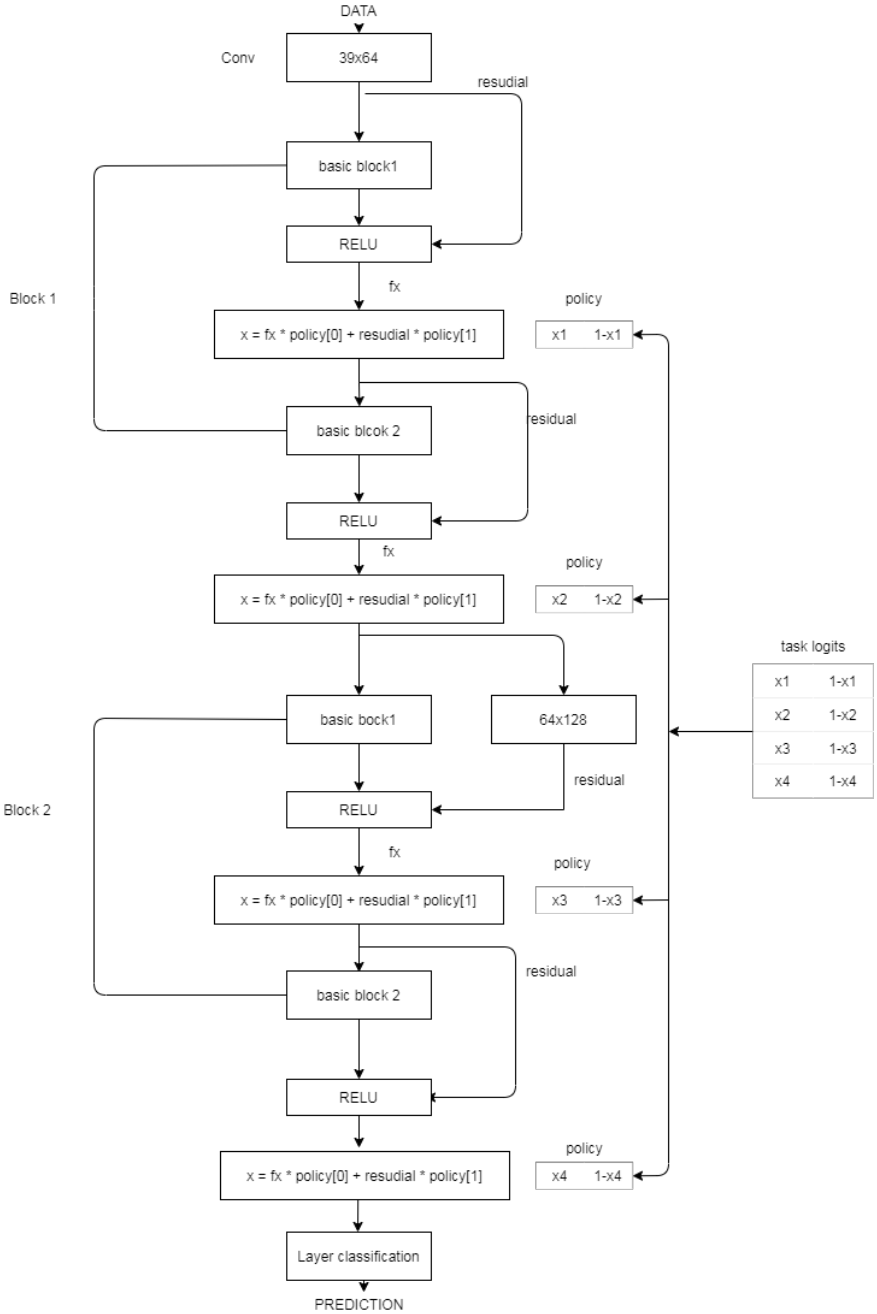

Figure 1. The graph-based Adashare model architecture (Example of 2 blocks of graph-based ResNet with Policy network)

Description of main components:

- **Residual graph block:** we consider using graph-based ResNet which means ResNet [7] architecture but replaced convolution layers by graph convolution layer. In particular, a residual block is said to be shared across two tasks if it is being used by both of them, or task-specific if it is being used by only one task for predicting the output. In this way, the select-or-skip policy of all blocks and tasks determines the adaptive feature sharing mechanism over the given task set T.

- **Learning a Task-Specific Policy:** we learn the select-or-skip policy U and network weights jointly through standard backpropagation from our designed loss functions. However, each select-or-skip policy is discrete and non-differentiable, and this makes direct optimization difficult.

ii. CapsGNN + Task routing

CapsGNN architecture considers each graph as multiple embeddings and each embedding reflects the graph properties from different aspects [8]. It means we can capture the important information at the graph level based on the extracted node features in the form of capsules through GNN. The high-level graph capsules and class capsules are also achieved due to the routing mechanism.

Besides, the task routing layer works as masks that enable task routing being generated randomly when the model is instantiated and is kept constant through the training process. By applying the task routing to the network, we are able to reuse units between tasks and scale up the number of tasks that can be performed with a single model [8].

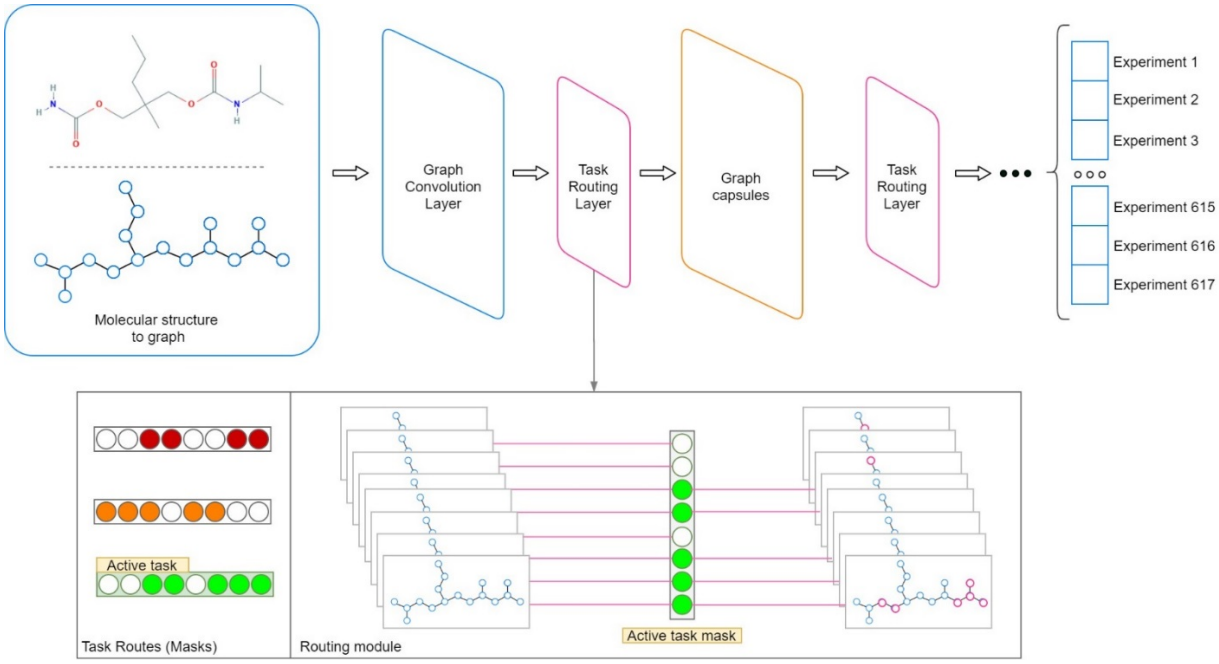

Figure 2. The multi-task CapsGNN architecture + Task Routing layer

Description of main components:

- **CapsGNN** [8]: In general, the network contains three main blocks including:
  - ✓ "Basic node capsules extraction block" which is applied to extract local vertices features with different receptive-fields and then primary node capsules are built in this block.
  - ✓ "High level graph capsules extraction block": Attention Module and Dynamic Routing are fused to generate multiple capsules for graphs.
  - ✓ "Graph classification block" which again uses Dynamic Routing to generate class capsules for graph classification

- **Task routing** [9]: We applied a channel-wise task-specific binary mask over the convolutional activations, restricting the input to the following layer to contain only activations assigned to the task. The flow of activations does not follow its conventional route, but is rerouted to an alternate one, the corresponding Task Routing Layer. The masks that enable task routing are generated randomly when the model is instantiated and is kept constant through the training process. These masks are created using a sharing ratio hyper-parameter σ, which is pre-defined and indicates how many units are task specific, and how many are shared between tasks.

b) Dataset

i. Tox21

The Tox21 comprised 12,000 environmental chemicals and drugs which were measured for 12 different toxic effects by specifically designed assays [10]. This physiology data collection provides toxicology testing for a large library of compounds based on in vitro high-throughput screening. The chemical structure is represented by graph with binary target variable for each experiment result. Each compound is represented by SMILES code which is transformed into graph data structure. The toxic effects are defined as graph classification tasks in the multitask learning model.

ii. NCI

NCI Anti-cancer activity prediction data are benchmark for predicting biological activities of small molecules for different types of cancers [11]. Each molecule is represented as a graph, with atoms representing nodes and bonds denoting edges. A molecule is positive if it is active against a certain type of cancer, or negative otherwise.

Collections	ID	#Pos	#Total	Dataset Description
NCI	1	1793	37349	Non-Small Cell Lung
	33	1467	37022	Melanoma
	41	1350	25336	Prostate
	47	1735	37298	Central Nerv Sys
	81	2081	37549	Colon
	83	1959	25550	Breast
	109	1773	37518	Ovarian
	123	2715	36903	Leukemia
	145	1641	37043	Renal

Table 1. Description of NCI Graph Datasets

Table 1 shows nine different cancer targets which can be used as graph classification tasks in our experiments. The balanced dataset will be randomly built from the original graph set. Since all these tasks are relevant in cancer prediction and some common substructures may exist for all types of cancers, NCI could be considered as an ideal benchmark for multi-task graph classification.

II. Investigating on drug-induced liver injury (DILI)

a) Model architecture

The model was designed with an embedding (Embed) layer, a convolutional (Conv) block, and a fully-connected (FC) block. The FC block consisted of three FC layers. Except for the final layer of the FC block, the others were associated with batch normalization layers, an essential regularization layers for weight standardization to lessen the effect of weight initialization. Adam optimizer was utilized to iteratively update network weights with the learning rate of 0.0025. The leaky rectified linear unit (LeakyRELU) was used as an activation function with a slope of 0.01 for the Conv layer and the next two FC layers. The sigmoid function was the activation function of the last layer (Figure 3).

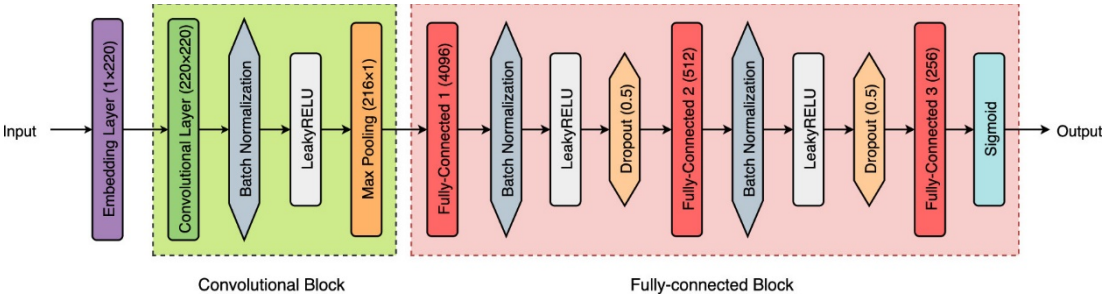

Figure 3. Model architecture

b) Dataset

Compounds collected from Zhang et al. [12], Ai et al. [13] Liew et al. [14], and Kotsampasakou et al. [15], and Chen et al. [16] were merged and curated. The detailed curation steps were discussed in our publication [17].

Dataset	No. of compounds			Sources
	DILI	Non-DILI	Total	
Development	946	651	1597	Liew et al. Zhang et al. Kotsampasakou et al. Ai et al.
Test	128	194	322	Chen et al.

Table 2. Curated data for model training and evaluation

c) Model training

The model was trained over 20 epochs. The optimal network was selected at the epoch displaying minimal validation loss. The loss function used is binary cross-entropy expressed as:

$$Loss = \sum_{i=1}^{n} \hat{y}_i \times log(\hat{y}_i) + (1 - \hat{y}_i) \times log(1 - \hat{y}_i)$$

The details on input and output sizes of data at each layer were described in our publication [17].

Result:

I. Multi-task Graph Neural Network

1. The Graph-based Adashare

We report the results obtained so far. With our architecture, in epoch 109, we obtained ROC-AUC of 12 tasks is 0.85 which is in the fifth place in the benchmarking (https://paperswithcode.com/sota/drug-discovery-on-tox21). We continue to work on model improvement, e.g., trying different base models and using other datasets.

Table 3. Epoch 109 Avenger ROC_AUC of 12 task: 0.85

Task	**Performance (ROC-AUC)**
To NR-AR	0.8596
NR-AR-LBD	0.9032
NR-AhR	0.8252
NR-Aromatase	0.8564
NR-ER	0.7251
NR-ER-LBD	0.8312
NR-PPAR-gamma	0.8687
SR-ARE	0.8092
SR-ATAD5	0.8827
SR-HSE	0.8624
SR-MMP	0.8734
SR-p53	0.8982

2. Multi task CapsGNN with Task routing

Table 2 shows us the results of the performance of single task CapsGNN on biological datasets [8]. The baseline methods including both kernel-based and deep-learning-based algorithms.

Algorithm	**Mutag**	**NCI1**	**PROTEINS**	**D&D**	**ENZYMES**
WL	82.05±0.36	**82.19±0.18**	74.68±0.49	**79.78±0.36**	52.22±1.26
GK	81.58±2.11	62.49±0.27	71.67±0.55	78.45±0.26	32.70±1.20
RW	79.17±2.07	>3days	74.22±0.42	>3days	24.16±1.64
Graph2vec	83.15±9.25	73.22±1.81	73.30±2.05	-	-
AWE	**87.87±9.76**	-	-	71.51±4.02	35.77±5.93
DGK	87.44±2.72	80.31±0.46	75.68±0.54	73.50±1.01	53.43±0.91
PSCN	**88.95±4.37**	76.34±1.68	75.00±2.51	76.27±2.64	-
DGCNN	85.83±1.66	74.44±0.47	75.54±0.94	7**9.37±0.94**	51.00±7.29
ECC	76.11	76.82	-	72.54	45.67
GCAPS-CNN	-	**82.72±2.38**	**76.40±4.17**	77.62±4.99	**61.83±5.39**
CapsGNN	86.67±6.88	78.35±1.55	**76.28±3.63**	75.38±4.17	**54.67±5.67**

Table 4. Single Task Result of Biology Dataset [8]

Kernel-based methods: Weisfeiler-Lehman subtree kernel (WL), graphlet count kernel (GK), and Random Walk (RW)

Deep-learning-based methods: Graph2vec, Deep Graph Kernel (DGK), AWE, PATCHY-SAN (PSCN), GCAPS-CNN, Dynamic Edge CNN (ECC), and Deep Graph CNN (DGCNN)

Within these methods, the performance of CapsGNN in single task archives top 2 on 2 out of 5 biological datasets which indicate its high potential of large graph dataset analysis. Based on this result, we have completed building a multi-task CapsGNN architecture which we hope improve the performance by sharing the information between graph classification tasks, but the results are not yet available. We continue to work on this.

II. Investigating on drug-induced liver injury (DILI)

Three evaluation metrics including area under the receiver operating characteristic curve (AUC), accuracy (ACC), and Matthews's correlation coefficient (MCC) were used to evaluate the model performance.

To train the model, 10% of the development set was randomly selected to construct the validation set for model monitoring while the rest of the development data were used for training. The experiment was repeated 30 times (trials) with random sampling of validation data. All the models were selected at the epochs where the validation losses were minimum. For each trial, the best model was reloaded to train with one, two, and three additional epochs using the entire development set with the learning rate of 0.001 to observe any possible improvements. The outcomes indicate that additional training for several epochs can significantly boost the model performance on the independent test set. Besides, the small variation in AUC, ACC, and MCC over 30 experimental trials suggests the model stability and robustness. The AUC value of about 0.96 is therefore highly meaningful to address the problem of screening DILI compounds

Metric	**Model (i)**[a]	**Model (ii)**[b]	**Model (iii)**[c]	**Model (vi)**[d]
ACC	0.95 ± 0.01	0.96 ± 0.01	0.96 ± 0.01	0.96 ± 0.01
MCC	0.84 ± 0.06	0.89 ± 0.03	0.91 ± 0.01	0.92 ± 0.01
AUC	0.72 ± 0.09	0.80 ± 0.05	0.83 ± 0.02	0.83 ± 0.01

[a]*Model at the converged epoch (setup (i))*
[b]*Model with one additional training epoch (setup (ii))*
[c]*Model with two additional training epochs (setup (iii))*
[d]*Model with three additional training epochs (setup (iv))*

Table 5: Model performance

Although there are many computational frameworks designed for hepatotoxicity prediction, only methods that have been cross-validated were selected for comparison. The experimental results confirm that our proposed computational frameworks to predict DILI compounds perform significantly better than other state-of-the-art methods based on all evaluated metrics.

Method	**No. of Samples**	**Test Method**	**Test AUC**
Bayesi[a]	295	10-fold CV	0.62
Stacking Ensemble with Naïve Bayes[b]	1087	5-fold CV	0.74
Decision Forest[c]	197	10-fold CV	-
SVM[d]	1317	Test set	0.65
Naïve Bayes[e]	420	Test set	-
Ensemble-Top5[f]	1241	5-fold CV	0.76
		Test set	0.90
Our methods	1919	Test set	0.96

[a]Ekins et al.; [b]Liew et al.; [c]Chen et al.; [d]Zhang et al.; [e]Zhang et al.; [f]Ai et al.

Discussion:
Feature sharing strategy across multiple tasks are the best way to solve the graph classification multitask with deep learning. We presented two solutions for tasks for toxicity testing prediction and anti-cancer screening by using adaptive sharing feature. The first solution is Adaptive Sharing based on Residual graph network and Policy network and the second solution is Capsule graph network with Task routing network.

The graph-based Adashare shows a potential performance on Tox21 datasets which could be further improved in the near future. Single task GapsGNN shows good performance. Multitask CapsGNN has been implemented and is ready for the training process, but the results have not been obtained yet. We continue to work on this and test on NCI and Tox21. We expect that the capsule mechanism of Multitask CapsGNN achieves better performance with smaller number of parameters growing in the architecture.

Finally, according to all evaluated metrics, applying the multitask learning method to predict human cytochrome P450 inhibition can significantly boost the predictive power to effectively recognize inhibitors of all five CYP isoforms. The combination of the multitask learning and molecular fingerprint-embedded encoding creates a robust, stable, and efficient computational framework to virtually screen for potential CYP450 inhibitors which may cause severe adverse drug reactions.

List of Publications and Significant Collaborations that resulted from your AOARD supported project:
b) papers published in peer-reviewed conference proceedings:

1. Nguyen, L., Nguyen-Vo, T. H., Trinh, Q. H., Nguyen, Bach., Nguyen, P. U., Le, L., Nguyen, B. P. (2021) iANP-EC: Identifying Anticancer Natural Products using Ensemble Learning Incorporated with Evolutionary Computing. *Journal of chemical information and modeling (**In revision**)*

2. Nguyen-Vo, T. H., Trinh, Q. H., Nguyen, L., Nguyen, P. U., Nguyen, T. N., Nguyen, D., Nguyen, Nguyen, P. B., Le, L. (2021) iCYP-MFE: Identifying Human Cytochrome P450 Inhibitors using Multi-task Learning and Molecular Fingerprint-embedded Encoding. *Journal of chemical information and modeling (**In revision**)*

3. Nguyen-Vo, T. H., Nguyen, L., Do, N., Le, P. H., Nguyen, T. N., Nguyen, B. P., & Le, L. (2020). Predicting Drug-Induced Liver Injury Using Convolutional Neural Network and Molecular Fingerprint-Embedded Features. *ACS omega*, *5*(39), 25432-25439.

4. Nguyen-Vo, T. H., Nguyen, L., Do, N., Nguyen, T. N., Trinh, K., Cao, H., & Le, L. (2019). Plant metabolite databases: from herbal medicines to modern drug discovery. *Journal of chemical information and modeling*, *60*(3), 1101-1110.

f) provide a list any interactions with industry or with Air Force Research Laboratory scientists or significant collaborations that resulted from this work.

Vietherb website: http://vietherb.com.vn/

REFERENCES

1. Goh, G. B., Hodas, N. O. & Vishnu, A. Deep learning for computational chemistry. Journal of Computational Chemistry (2017) doi:10.1002/jcc.24764.
2. Ma, J., Sheridan, R. P., Liaw, A., Dahl, G. E. & Svetnik, V. Deep neural nets as a method for quantitative structure-activity relationships. J. Chem. Inf. Model. (2015) doi:10.1021/ci500747n.
3. Ragoza, M., Hochuli, J., Idrobo, E., Sunseri, J. & Koes, D. R. Protein-Ligand Scoring

with Convolutional Neural Networks. J. Chem. Inf. Model. (2017) doi:10.1021/acs.jcim.6b00740.
4. Kipf, T. N. & Welling, M. Semi-supervised classification with graph convolutional networks. in 5th International Conference on Learning Representations, ICLR 2017 - Conference Track Proceedings (2019).
5. Gawehn, E., Hiss, J. A. & Schneider, G. Deep Learning in Drug Discovery. Molecular Informatics (2016) doi:10.1002/minf.201501008.
6. Argyriou, A., Evgeniou, T. & Pontil, M. Multi-task feature learning. in Advances in Neural Information Processing Systems (2007).
7. Targ, S., Almeida, D., & Lyman, K. (2016). Resnet in resnet: Generalizing residual architectures. arXiv preprint arXiv:1603.08029.
8. Xinyi, Z. & Chen, L. Capsule graph neural network. in 7th International Conference on Learning Representations, ICLR 2019 (2019).
9. Strezoski, G., Noord, N. V., & Worring, M. (2019). Many task learning with task routing. In Proceedings of the IEEE/CVF International Conference on Computer Vision (pp. 1375-1384).
10. Richard, A. M., Huang, R., Waidyanatha, S., Shinn, P., Collins, B. J., Thillainadarajah, I., ... & Tice, R. R. (2020). The Tox21 10K Compound Library: Collaborative Chemistry Advancing Toxicology. Chemical Research in Toxicology.
11. Shirui Pan, Xingquan Zhu, Chengqi Zhang, and Philip S. Yu. "Graph Stream Classification using Labeled and Unlabeled Graphs", International Conference on Data Engineering (ICDE), pages 398-409, 2013.
12. Zhang, C. et al. In silico Prediction of Drug Induced Liver Toxicity Using Substructure Pattern Recognition Method. Mol. Inform. (2016) doi:10.1002/minf.201500055.
13. Ai, H. et al. Predicting Drug-Induced Liver Injury Using Ensemble Learning Methods and Molecular Fingerprints. Toxicol. Sci. (2018) doi:10.1093/toxsci/kfy121.
14. Liew, C. Y., Lim, Y. C. & Yap, C. W. Mixed learning algorithms and features ensemble in hepatotoxicity prediction. J. Comput. Aided. Mol. Des. (2011) doi:10.1007/s10822-011-9468-3.
15. Kotsampasakou, E., Montanari, F. & Ecker, G. F. Predicting drug-induced liver injury: The importance of data curation. Toxicology (2017) doi:10.1016/j.tox.2017.06.003.
16. Chen, M. et al. DILIrank: The largest reference drug list ranked by the risk for developing drug-induced liver injury in humans. Drug Discovery Today (2016) doi:10.1016/j.drudis.2016.02.015.
17. Nguyen-Vo, T. H. et al. Predicting Drug-Induced Liver Injury Using Convolutional Neural Network and Molecular Fingerprint-Embedded Features. ACS Omega (2020) doi:10.1021/acsomega.0c03866.