JTE HCM-UTE
ISSN: 2615-9740

JOURNAL OF TECHNICAL EDUCATION SCIENCE
Ho Chi Minh City University of Technology and Engineering
Website: https://jte.edu.vn
Email: jte@hcmute.edu.vn

Graph Neural Network Model for Predicting Mitochondrial Toxicity via Complementary Molecular Features

Phat Dat Nguyen[1*], Trong Phuc Le[1], Dinh Khanh Ngan Nguyen[1], Thi Phuong Hoa Tran[1]
Ho Chi Minh City University of Technology and Engineering, Vietnam

[*]Corresponding author. *Email: datnp@hcmute.edu.vn*

ARTICLE INFO

Received: 22/06/2025
Revised: 29/12/2025
Accepted: 06/05/2026
Online First: 22/08/2026
Published:

KEYWORDS

Toxicity prediction;
Machine learning;
Deep learning;
Graph neural network;
Mitochondrial toxicity.

ABSTRACT

In this study, we introduce a hybrid Graph Neural Network model for predicting mitochondrial toxicity, a significant contributor to drug-induced liver injury. This model effectively captures both the structural and physicochemical properties of compounds by integrating graph embeddings and molecular fingerprints. After systematically optimizing the model architecture, we found that the combination of the GATv2Conv convolutional layer, 64 hidden neurons, and a single predictor layer was the most effective configuration for our purposes. When benchmarked against well-established 3D GNN models like DimeNet++ and ViSNet, our hybrid model consistently outperformed them, exhibiting improvements in AUC-ROC by up to 18.3% and 6.2%, respectively. These results emphasize the importance of achieving an optimal balance when integrating features and demonstrate the potential of our hybrid GNN model for computational toxicity prediction. This model could notably expedite drug discovery, reduce dependence on animal testing, and provide valuable insights into molecular toxicity.

Doi: https://doi.org/10.54644/jte.2026.1954

1. Introduction

Drug-induced toxicity continues to pose a considerable challenge in pharmaceutical development, contributing significantly to late-stage drug attrition and post-market withdrawals [1]. Among the various toxicity mechanisms, mitochondrial toxicity has been identified as a leading cause of adverse clinical outcomes, particularly in tissues demanding high energy like the heart, liver, and muscle [2]. In addition to their crucial role in ATP production, mitochondria are indispensable for detoxification, apoptosis regulation, fatty acid metabolism, and redox balance [3]. The disruption of these functions can give rise to hepatotoxicity, cardiotoxicity, and other severe toxicities.

An extensive range of drugs, such as antidiabetics, antivirals, statins, antibiotics, and NSAIDs, have been associated with mitochondrial toxicity. The diverse mechanisms include inhibition of respiratory complexes, uncoupling of oxidative phosphorylation, increased oxidative stress, impaired mitochondrial DNA replication, and disrupted biogenesis [4]. These effects can induce clinical symptoms like lactic acidosis, muscle weakness, and even organ failure, with individual susceptibility dependent upon genetic factors. Given the wide variety of implicated drug classes and mechanisms, accurate prediction of mitochondrial toxicity from chemical structure presents a significant challenge [2].

Addressing this, systematic mitochondrial toxicity screening has now become an integral part of drug development pipelines. The advancements in high-throughput assays, oxygen consumption measurements, and enhanced cell models have substantially improved the detection of mitochondrial toxicity [5]. However, translating these findings into accurately predicting clinical risk continues to be a challenge. This underscores the need for establishing more reliable, mechanistically informed, and universally applicable predictive models.

Recent advances in artificial intelligence (AI) and machine learning (ML) have revolutionized toxicity prediction, effectively integrating diverse data types—from chemical structures and omics profiles to clinical outcomes—into increasingly precise models [6]. In many instances, traditional

quantitative structure–activity relationship (QSAR) methods have been overtaken by deep learning techniques, particularly graph neural networks (GNNs), that excel in capturing the inherent graph-like nature of molecular structures. Specifically in complex endpoints like mitochondrial toxicity, these state-of-the-art approaches outperform conventional descriptor-based models when trained on comprehensive, well-curated datasets.

Progress in ML-driven mitochondrial toxicity prediction has been characterized by both methodological innovation and improved interpretability. Early models, such as SVMs with genetic algorithm feature selection [7], not only demonstrated high accuracy but also enabled identification of toxicophores associated with mitochondrial dysfunction. Later studies incorporated mechanistic information, expanded HTS-derived datasets, and deployed random forest and deep learning algorithms to highlight structural alerts, known as aromatic rings and phenolics, implicated in mitochondrial impairment [8]–[10]. Most recently, ensemble and consensus modeling strategies, coupled with explainable AI methods, have further improved prediction accuracy and interpretability, identifying structural features such as nitro groups and phthalaldehyde derivatives as significant toxicity indicators [11]–[16]. The integration of these advances has produced highly predictive, transparent in silico platforms which support safer drug discovery and inform regulatory risk assessment for chemical safety evaluations.

Further improvements in predicting mitochondrial toxicity have been achieved through advanced deep learning. Studies have demonstrated that integrating chemical structure information with high-content phenotypic data, such as cell painting profiles, can create more robust and accurate models compared with using each modality alone, despite potential limitations regarding data size and quality [17]. The rise of graph neural networks (GNNs) has been especially impactful as their capability to directly model molecular graphs facilitates superior performance and mechanistic insight. For instance, combining GNNs with ensemble bagging has achieved high predictive accuracy and improved model interpretability [18], while hierarchical message-passing neural networks have outperformed several state-of-the-art models across various toxicity endpoints and effectively identified relevant structural toxicophores [19]. These developments highlight the potential of GNN-based approaches for both accuracy and interpretability, positioning them as vital tools for early toxicity screening and safer, more effective drug development.

Encouraged by these advancements, our research aims to develop a hybrid graph neural network (GNN) model combining both molecular graph structures and molecular fingerprints to predict mitochondrial toxicity. This study aims to merge the intuitive structural representation provided by GNNs with the detailed features derived from molecular fingerprints, harnessing complementary data for improved predictive accuracy and clarity. We designed this approach to not only leverage strengths from both graph-based and descriptor-based methods but also to address limitations seen in models relying on a single type of input. Ultimately, our goal is to present a robust, efficient tool for early-stage toxicity screening that can assist in promoting safer drug discovery and more informed risk assessment.

2. Computational Methods

2.1. Dataset Curation

The MitoTox database, introduced by Lin and coworkers in 2021, serves as the foundation for this study [20]. This database contains the mitochondrial toxicity information of organic molecules as binary labels (0 for non-toxic and 1 for toxic). The database was downloaded and processed according to a specific workflow to yield a refined dataset in a comma-separated values format, featuring 5894 molecules. For the construction of molecular graphs for GNN models, the 'SMILES' column, holding the Simplified Molecular Input Line-Entry System (SMILES) representations of the compounds, was utilized. The binary label for mitochondrial toxicity was recorded in the 'MTX' column, which served as the target for the GNN models.

Using common cheminformatics packages in Python, RDKit and DeepChem [21], we converted the SMILES representations into molecule objects, facilitating the extraction of various featurizers — molecular graphs and molecular fingerprints. Before the training process, the dataset was divided into

training, validation, and testing subsets in a 60:20:20 ratio for optimal model performance assessment. The extracted data is stored for each compound locally for future use in the training and evaluation phases of the GNN models.

2.2. GNN Models for Prediction of Mitochondrial Toxicity

The molecular graphs and fingerprints extracted from the training set served as inputs for the GNN models (refer to Figure 1). Our models are implemented using PyTorch-Geometric, a renowned open-source library for geometric deep learning. Our custom GNN architecture comprises three main components: a Graph Convolutional Network (GCN) block to learn molecular graphs, a multi-layer perceptron (MLP) block to process molecular fingerprints, and a second MLP block (the predictor block) that integrates outputs from the first two components to formulate the final predictions. We utilized various types of convolutional layers within the GCN architecture. Each block's parameters – the number of attention heads, the number of layers, and the number of neurons per layer – can be customized, allowing model design flexibility. Each linear layer is paired with a Rectified Linear Unit (ReLU) activation function, introducing non-linearity.

Throughout training, we periodically validated the models to monitor their performance and make necessary parameter adjustments. The GNN models were trained in batches (with a batch size of 32) for 100 epochs using the Adam optimizer. We used binary cross-entropy as our loss function and set the learning rate at 0.001 and the exponential decay rate at 0.99. The training was carried out on a local computer equipped with an NVIDIA RTX 5070 Ti graphics card. To ensure reproducibility, we set all random seeds to 0. The test dataset was employed after the training to evaluate the model's predictive performance using binary classification metrics, including accuracy, F1-score, and the area under the receiver operating characteristic curve (AUC-ROC).

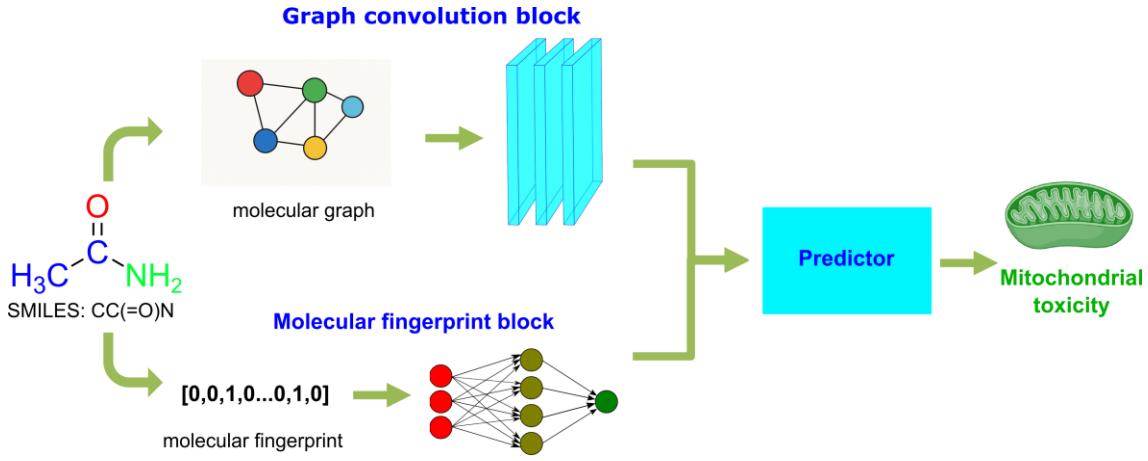


Figure 1. *Hybrid GNN model for prediction of mitochondrial toxicity.*

3. Results and Discussion

3.1. Combination of Molecular Graph and Molecular Fingerprint

In our study, we first evaluated an array of combinations of molecular graph featurizers and molecular fingerprints for their efficacy in predicting mitochondrial toxicity (Figure 2). Three graph featurizers were assessed: the Molecule Graph Convolution Featurizer (MGCF), the Path-Augmented Graph Transformer Network Molecule Graph Featurizer (PMGF), and the Direct Message Passing Neural Network Featurizer (DMPNNF). Additionally, we tested a variety of molecular fingerprints that included Mordred descriptors, as well as circular (MorganFP), Avalon (AvalonFP), atom-pair (APFP), topological torsion (TTFP), layered (LayeredFP), pattern (PatternFP), and RDKit (RDKitFP) fingerprints. For our initial model, we employed three GATv2 convolutional layers in the GCN block and three linear layers in the MLP block along with one linear layer in the predictor block, each layer comprising 128 neurons.

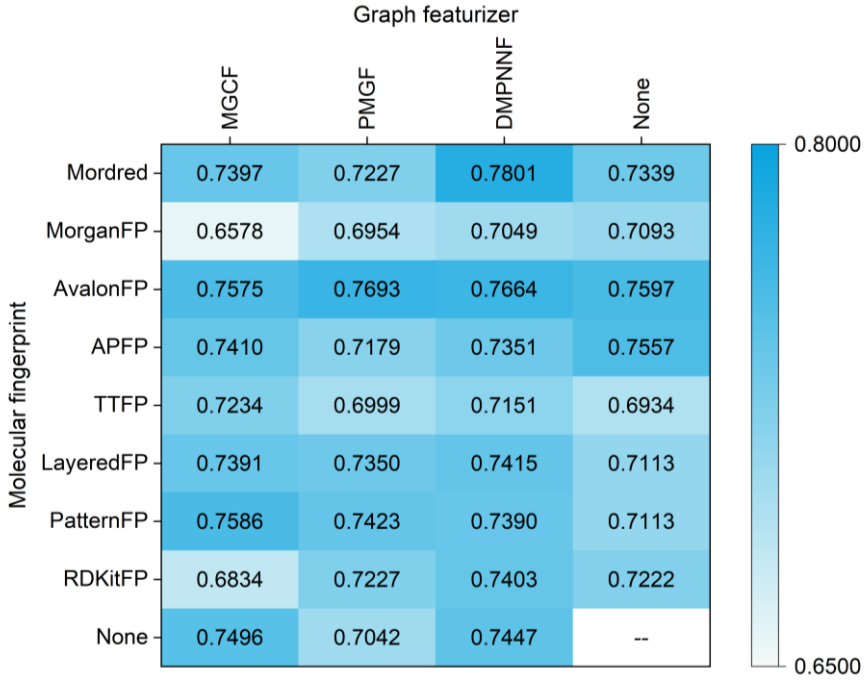


Figure 2. *AUC-ROC score of different combinations of graph featurizer and molecular fingerprint.*

As showcased by the heatmap in Figure 2, combining molecular graph features with molecular fingerprints consistently boosts the predictive performance for mitochondrial toxicity in comparison to using either feature type in isolation. When no graph featurizer is applied (as seen in the "None" section on the x-axis), the AUC-ROC scores tend to be lower than when any graph featurizer is combined with a fingerprint. Likewise, relying strictly on graph features without fingerprints, indicated as "None" on the y-axis, results in diminished performance. This pattern underscores the effectiveness of integrating both types of molecular representations, capturing complementary information leading to enhanced classification accuracy.

Of all the graph featurizers tested, DMPNNF consistently delivered the best results overall, particularly in combination with Mordred fingerprints, yielding the highest AUC-ROC score (0.7801). This suggests a synergistic effect, where the detailed structural encoding of DMPNNF could be complemented by the comprehensive physicochemical descriptors offered by Mordred fingerprints. Generally, MGCF outperformed PMGF across most fingerprints, implying that the graph convolutional approach exhibits more effectiveness than the path-augmented transformer network for this application.

As for fingerprints, AvalonFP and Mordred emerged as the top performers. AvalonFP showcased a strong predictive power when coupled with MGCF and PMGF, attaining AUC-ROC scores of 0.7575 and 0.7693, respectively. Conversely, Mordred clinched the top score when combined with DMPNNF and consistently delivered impressive results with other graph featurizers. On the other hand, MorganFP generated lower scores, suggestive of its encoding scheme potentially being less informative for predicting mitochondrial toxicity within this context. These outcomes highlight the importance of selecting feature representations that capably capture the relevant chemical information regarding the biological endpoint of interest.

We proceeded to evaluate the impact of the output size from both the graph convolution block and the molecular fingerprint block. These output embeddings were concatenated before passing through a final predictor block to derive the final prediction. As depicted in Figure 3, the alteration of the embedding size ratio between graph and fingerprint representations significantly impacted the performance of mitochondrial toxicity prediction, as measured by AUC-ROC.

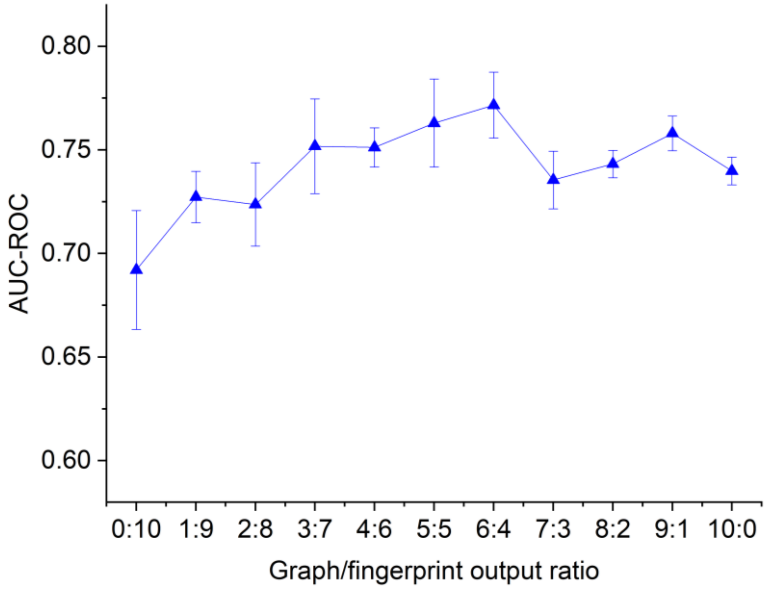


Figure 3. *Evaluation of GCN block/MLP block output size ratio.*

When using strictly graph embeddings (ratio 10:0) or solely fingerprint embeddings (ratio 0:10), there was a noticeable decline in predictive accuracy (0.7399 and 0.6921, respectively). This reinforces the importance of integrating both molecular feature types. The performance progressively escalated as the ratio gradually skewed towards a balanced combination, achieving peak performance at a graph-to-fingerprint embedding ratio of about 6:4. This combination resulted in an AUC-ROC score of 0.7716, indicating an optimal balance where enriched structural information from graph embeddings effectively complements the physicochemical detail from fingerprint embeddings. However, the results plateaued beyond this point, suggesting diminishing returns and emphasizing the criticality of preserving complementary contributions from both feature sets for maximal predictive accuracy.

3.2. Optimization of GNN Model Architecture

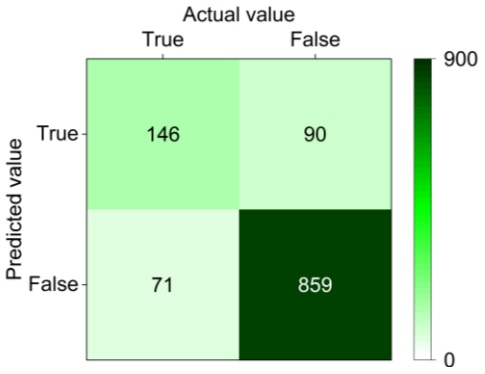
Capitalizing on the best graph embedding/fingerprint embedding ratio of 6:4, we further honed our model architecture to enhance performance. Table 1 presents a systematic comparison of different graph convolutional architectures, hidden neuron sizes, and predictor layer depths for mitochondrial toxicity prediction. Among the varying types of convolutional layers evaluated in the GCN block (Entries 1–8), GATv2Conv consistently delivered superior results with a hidden size of 128 and a single predictor layer (Entry 6). This suggests that GATv2Conv's enhanced attention mechanism provides a more effective and expressively rich molecular representation for this classification task compared to alternatives such as GCNConv, SAGEConv, or GraphConv. When paired with 64 hidden neurons and a single predictor layer (Entry 9), the model's performance was augmented, registering the highest F1-score (0.6446) and AUC-ROC (0.7890).

The results demonstrate that merely ramping up the model complexity does not necessarily lead to performance gains. Rather, higher increases to 256 or 512 hidden neurons actually decreased the model's performance (Entries 10 and 11). These results suggest that an optimal balance between model capacity and overfitting risk might be achieved by using a moderate embedding dimension. Additionally, an increase in the number of layers in the predictor block from one to two or three did not yield further improvements but instead caused a small decline in F1-score and AUC-ROC (Entries 12 and 13).

Collectively, these findings underscore the essential role of meticulous architectural selection and hyperparameter tuning. The configuration pairing GATv2Conv with 64 hidden neurons and one predictor layer emerged as the optimal choice for this application.

Table 1. *Optimization of GNN model for prediction of mitochondrial toxicity.*

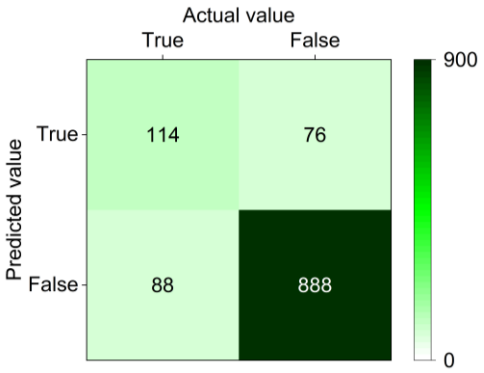
Entry	Type of convolutional layer	Hidden neurons	Predictor layers	F1-score	AUC-ROC
1	GCNConv	128	1	0.6200	0.7546
2	SAGEConv	128	1	0.5959	0.7546
3	GraphConv	128	1	0.5831	0.7382
4	ClusterGCNConv	128	1	0.6000	0.7498
5	GATConv	128	1	0.5825	0.7298
6	GATv2Conv	128	1	0.6366	0.7801
7	GeneralConv	128	1	0.6318	0.7621
8	TransformerConv	128	1	0.5874	0.7450
9	GATv2Conv	64	1	0.6446	0.7890
10	GATv2Conv	256	1	0.5859	0.7217
11	GATv2Conv	512	1	0.5682	0.7139
12	GATv2Conv	64	2	0.6389	0.7869
13	GATv2Conv	64	3	0.6224	0.7532



Accuracy	0.8619
Precision	0.6186
Recall	0.6728
Specificity	0.9052
F1-score	0.6446
AUC-ROC	0.7890

(a) Hybrid GNN Model

(b) DimeNet++



(c) ViSNet

Figure 4. *Comparison of hybrid GNN model and 3D GNN models.*

We compared our model architecture with state-of-the-art 3D GNN models such as DimeNet++ [22] and ViSNet [23]. The results affirm the notable effectiveness of our hybrid GNN model (Figure 4). It consistently mirrors or outperforms its counterparts across most performance metrics. Our optimized model delivers a standout accuracy of 0.8619, superior to DimeNet++ and ViSNet (with 0.8419 and 0.8593 respectively), emphasizing its prowess in correctly classifying toxic and non-toxic compounds. Moreover, the hybrid model provides optimized precision (0.6186), recall (0.6728), and F1-score (0.6446), showcasing its improved ability to correctly identify true positives while minimizing false predictions.

The hybrid GNN model also excels in the Area Under the Receiver Operating Characteristic (AUC-ROC) with a score of 0.7890. This outperforms the considerably lower AUC-ROC scores of 0.6670 and 0.7427 for DimeNet++ and ViSNet respectively. This suggests that our hybrid model possesses superior discriminative power in distinguishing between toxic and non-toxic compounds. Though DimeNet++ and ViSNet report marginally higher specificity (0.9215 and 0.9212 respectively compared to 0.9052 from the hybrid model), the hybrid GNN model's overall performance demonstrates that it is a more comprehensive and balanced solution for predicting mitochondrial toxicity.

4. Conclusions

In conclusion, this study presents a novel hybrid Graph Neural Network (GNN) approach integrating graph and molecular fingerprint embeddings for enhancing the predictive accuracy of mitochondrial toxicity. By systematically varying their combination and performing comprehensive hyperparameter and architectural optimization, we highlight the importance of a balanced feature integration for optimal performance. Specifically, our results suggest that an approximately 6:4 graph-to-fingerprint embedding ratio provides the richest and most informative molecular representation. Furthermore, within the GNN architecture, the GATv2Conv convolutional layer paired with 64 hidden neurons and a single predictor layer emerged as the most effective configuration.

When benchmarked against well-established 3D GNN models, DimeNet++ and ViSNet, our optimized hybrid model consistently demonstrated superior performance across multiple evaluation metrics. Its enhanced performance serves as evidence of the value in hybridizing different molecular representations to capture their complementary information. This study highlights the potential of our hybrid GNN model as a powerful tool in computational toxicology, pointing to promising applications not only in reducing animal testing and accelerating drug discovery but also in providing insights into how different molecular features can be effectively integrated within GNNs to enhance model prediction.

Acknowledgments

This work is part of a 2025 project funded by the Ho Chi Minh City University of Technology and Engineering, Vietnam.

Conflict of Interest

The authors declare no conflict of interest.

REFERENCES

[1] J. A. Dykens and Y. Will, "The significance of mitochondrial toxicity testing in drug development," *Drug Discovery Today,* vol. 12, no. 17, pp. 777–785, 2007.
[2] K. Chan, T. Don, S. Nandita, and P. J. O'Brien, "Drug-induced mitochondrial toxicity," *Expert Opinion on Drug Metabolism & Toxicology,* vol. 1, no. 4, pp. 655–669, 2005.
[3] Y. Will and J. Dykens, "Mitochondrial toxicity assessment in industry – a decade of technology development and insight," *Expert Opinion on Drug Metabolism & Toxicology,* vol. 10, no. 8, pp. 1061–1067, 2014.
[4] S. Nadanaciva and Y. Will, "New insights in drug-induced mitochondrial toxicity," *Current Pharmaceutical Design,* vol. 17, no. 20, pp. 2100–2112, 2011.
[5] M. Porceddu, N. Buron, P. Rustin, B. Fromenty, and A. Borgne-Sanchez, "In vitro assessment of mitochondrial toxicity to predict drug-induced liver injury," *Drug-induced liver toxicity,* pp. 283–300, 2018.
[6] C. Bai, L. Wu, R. Li, Y. Cao, S. He, and X. Bo, "Machine learning-enabled drug-induced toxicity prediction," *Advanced Science,* vol. 12, no. 16, p. 2413405, 2025.
[7] H. Zhang, Q. Y. Chen, M. L. Xiang, C. Y. Ma, Q. Huang, and S. Y. Yang, "In silico prediction of mitochondrial toxicity by using GA-CG-SVM approach," *Toxicology in Vitro,* vol. 23, no. 1, pp. 134–140, 2009.

[8] M. D. Nelms, C. L. Mellor, M. T. D. Cronin, J. C. Madden, and S. J. Enoch, "Development of an in silico profiler for mitochondrial toxicity," *Chemical Research in Toxicology,* vol. 28, no. 10, pp. 1891–1902, 2015.
[9] H. Zhang *et al.*, "Development of novel prediction model for drug-induced mitochondrial toxicity by using naïve Bayes classifier method," *Food and Chemical Toxicology,* vol. 110, pp. 122–129, 2017.
[10] D. A. Dreier, N. D. Denslow, and C. J. Martyniuk, "Computational in vitro toxicology uncovers chemical structures impairing mitochondrial membrane potential," *Journal of Chemical Information and Modeling,* vol. 59, no. 2, pp. 702–712, 2019.
[11] J. Hemmerich, F. Troger, B. Füzi, and G. F. Ecker, "Using machine learning methods and structural alerts for prediction of mitochondrial toxicity," *Molecular Informatics,* vol. 39, no. 5, p. 2000005, 2020.
[12] W. Tang, J. Chen, and H. Hong, "Discriminant models on mitochondrial toxicity improved by consensus modeling and resolving imbalance in training," *Chemosphere,* vol. 253, p. 126768, 2020.
[13] F. Bringezu, J. Carlos Gómez-Tamayo, and M. Pastor, "Ensemble prediction of mitochondrial toxicity using machine learning technology," *Computational Toxicology,* vol. 20, p. 100189, 2021.
[14] P. Zhao *et al.*, "In silico prediction of mitochondrial toxicity of chemicals using machine learning methods," *Journal of Applied Toxicology,* vol. 41, no. 10, pp. 1518–1526, 2021.
[15] K. Jaganathan, M. U. Rehman, H. Tayara, and K. T. Chong, "XML-CIMT: Explainable machine learning (XML) model for predicting chemical-induced mitochondrial toxicity," *International Journal of Molecular Sciences,* vol. 23, no. 24, p. 15655, 2022.
[16] R. Zhang, Z. Chen, B. Wang, Y. Li, Y. Mu, and X. Li, "Modeling and insights into the structural characteristics of chemical mitochondrial toxicity," *ACS Omega,* vol. 8, no. 35, pp. 31675–31682, 2023.
[17] M. G. De Lomana, P. M. Zapata, and F. Montanari, "Predicting the mitochondrial toxicity of small molecules: insights from mechanistic assays and cell painting data," *Chemical Research in Toxicology,* vol. 36, pp. 1107–1120, 2023.
[18] Y. Igarashi, R. Kojima, S. Matsumoto, H. Iwata, Y. Okuno, and H. Yamada, "Developing a GNN-based AI model to predict mitochondrial toxicity using the bagging method," *The Journal of Toxicological Sciences,* vol. 49, no. 3, pp. 117–126, 2024.
[19] Y. Zhou *et al.*, "ToxMPNN: A deep learning model for small molecule toxicity prediction," *Journal of Applied Toxicology,* vol. 44, no. 7, pp. 953–964, 2024.
[20] Y. T. Lin, K. H. Lin, C. J. Huang, and A. C. Wei, "MitoTox: a comprehensive mitochondrial toxicity database," *BMC Bioinformatics,* vol. 22, no. 10, p. 369, 2021.
[21] B. Ramsundar, P. Eastman, P. Walters, and V. Pande, "Deep learning for the life sciences: applying deep learning to genomics, microscopy, drug discovery, and more." O'Reilly Media, Inc., 2019.
[22] J. Gasteiger, S. Giri, J. T. Margraf, and S. Günnemann, "Fast and uncertainty-aware directional message passing for non-equilibrium molecules," *arXiv preprint arXiv:2011.14115,* 2020.
[23] Y. Wang *et al.*, "Enhancing geometric representations for molecules with equivariant vector-scalar interactive message passing," *Nature Communications,* vol. 15, no. 1, p. 313, 2024.

Phat Dat Nguyen earned B.Eng. in Chemical Engineering from the Ho Chi Minh City University of Technology in 2015. In 2016, he relocated to the USA and started doing photochemical research under the guidance of Prof. Oleg V. Larionov at the University of Texas at San Antonio. He then joined the research group of Prof. Loi H. Do at the University of Houston in 2017. His doctoral work encompassed studies on metal-catalyzed transfer hydrogenation and single-molecule detection using super-resolution fluorescence imaging (SRFM) within living cells. After earning the Ph.D degree in 2022, he joined the Faculty of Chemistry and Life Sciences at the HCMC University of Technology and Engineering, Vietnam.
Email: datnp@hcmute.edu.vn. ORCID: https://orcid.org/0000-0002-0634-8994

Trong Phuc Le is an undergraduate student at the HCMC University of Technology and Engineering, Vietnam. His major is organic chemical engineering and technology.
Email: 22128166@student.hcmute.edu.vn. ORCID: https://orcid.org/0009-0007-6088-4767

Dinh Khanh Ngan Nguyen is an undergraduate student at the HCMC University of Technology and Engineering, Vietnam. Her major is pharmaceutical chemical engineering and technology.
Email: 22128046@student.hcmute.edu.vn. ORCID: https://orcid.org/0009-0004-2219-9724

Thi Phuong Hoa Tran is an undergraduate student at the HCMC University of Technology and Engineering, Vietnam. Her major is organic chemical engineering and technology.
Email: 21128151@student.hcmute.edu.vn. ORCID: https://orcid.org/0009-0005-6305-6743