International Journal of Emerging Research in Engineering and Technology
Pearl Blue Research Group| 9th ICAILLMBA (ALLAM) 2025-Conference Proceedings
ISSN: 3050-922X | https://doi.org/10.63282/3050-922X.ICAILLMBA-108

Original Article

Causal Deep Learning for Drug–Gene–Disease Interaction Discovery

S A V S Sambha Murthy S[1], Kunjam Nageswara Rao[2], G Sitaratnam[3]
[1] Department of Computer Science, Dr V S Krishna Government Degree College (A), Vishakhapatnam, India.
[2] Department of Computer Science and Systems Engineering, College of Engineering, Andhra University, Visakhapatnam, India.
[3] Department of Computer Science and Engineering, Visakha Institute of Engineering and Technology, Visakhapatnam, India.

Abstract - *Identifying causal drug–gene–disease interactions is a fundamental challenge in drug discovery and precision medicine. Existing deep learning approaches predominantly rely on correlation-based predictions, limiting interpretability and biological validity. In this work, we propose CausalDGD, a causal deep learning framework that integrates structural causal models, mediation analysis, and deep neural representations to discover mechanistically grounded interaction triplets. By explicitly modeling Drug → Gene → Disease pathways and accounting for confounding factors, CausalDGD enables interpretable and causally faithful discovery. Extensive experiments on real-world biomedical datasets, including LINCS, DrugBank, and DisGeNET, demonstrate that the proposed approach outperforms state-ofthe-art baselines while providing biologically meaningful explanations. Our results highlight the importance of causal reasoning as a core component of trustworthy biomedical AI.*

Keywords - *Drug Discovery, Gene Interaction, Deep Learning, Biomedical AI.*

1. Introduction

Drug discovery and precision medicine fundamentally depend on understanding how pharmacological interventions influence molecular mechanisms and, ultimately, disease phenotypes. Most therapeutic compounds exert their effects by perturbing gene expression, protein activity, and cellular signaling pathways, which in turn drive changes in disease progression or clinical outcomes. Accurately identifying these drug–gene–disease relationships is therefore a central challenge in biomedical research.

With the rapid growth of high-throughput technologies, large-scale datasets such as transcriptomics, chemical perturbation profiles, and curated biological interaction databases have become widely available. This data abundance has motivated the use of machine learning and deep learning approaches for predicting drug–target interactions, identifying drug repurposing opportunities, and discovering disease associations. In particular, deep neural networks, graph neural networks, and representation learning methods have demonstrated strong predictive performance by capturing complex nonlinear patterns across heterogeneous biological data.

Despite these advances, most existing computational approaches remain fundamentally correlation-driven. They focus on identifying statistical associations between drugs, genes, and diseases without explicitly modeling the underlying causal mechanisms. In biological systems, however, correlation does not imply causation. A gene may appear correlated with a disease without being mechanistically responsible for disease progression, or a drug may correlate with a gene expression signature due to confounding factors such as cell type, pathway activity, or experimental conditions. Models that fail to distinguish causal effects from spurious correlations risk generating misleading hypotheses, thereby limiting their translational impact.

This limitation is particularly critical in drug discovery, where interventions are inherently causal in nature. For a drug to be effective, it must causally perturb a molecular target or biological pathway that influences disease outcomes. Correlation-based models may perform well on retrospective benchmarks but often fail to generalize to new experimental or clinical settings. These shortcomings have motivated increasing interest in incorporating causal reasoning into machine learning models for biomedical applications.

Causal inference provides a principled framework for reasoning about cause–effect relationships using tools such as structural causal models, mediation analysis, and counterfactual reasoning. Within this framework, drugs can be viewed as interventions, genes as mediators, and diseases as outcomes. Modeling these relationships explicitly enables the estimation of both direct effects of drugs on diseases and indirect, gene-mediated effects. Such causal decomposition offers mechanistic insights that go beyond predictive accuracy alone. However, classical causal inference methods often struggle to scale to high-dimensional, multimodal biological data and are limited in their ability to capture complex nonlinear relationships.

Deep learning, in contrast, excels at learning expressive representations from large and heterogeneous datasets but typically lacks causal interpretability. As a result, there is growing interest in bridging causal inference and deep learning to develop models that are both powerful and trustworthy. Recent advances in causal representation learning suggest that embedding causal constraints into learning objectives can improve robustness, interpretability, and generalization. Nevertheless, the application of these ideas to large-scale drug–gene–disease interaction discovery remains relatively underexplored.

In this work, we propose a causal deep learning framework, referred to as CausalDGD, for discovering mechanistically meaningful drug–gene–disease interaction triplets. The core idea is to explicitly model gene expression as a mediator of drug effects on disease outcomes, while leveraging deep neural networks to learn representations from molecular structures, transcriptomic profiles, and disease signatures. By integrating structural causal models and mediation analysis directly into the learning process, the proposed approach moves beyond correlation-based prediction toward causally grounded discovery.

The proposed framework jointly optimizes predictive performance and causal consistency, enabling the identification of interactions that are not only statistically significant but also biologically plausible. We evaluate the framework using real-world biomedical datasets, including LINCS for drug-induced gene expression, DrugBank for curated drug–target interactions and DisGeNET for gene–disease associations. Experimental results demonstrate that incorporating causal reasoning leads to improved accuracy, reduced false positives, and enhanced interpretability compared to state-of-the-art baselines.

In summary, this work formulates drug–gene–disease interaction discovery as a causal mediation problem and introduces a unified causal deep learning approach to address it. By combining causal inference with deep representation learning, the proposed method provides a robust and interpretable framework for hypothesis generation in drug discovery and precision medicine.

2. Related Work

Research on drug–gene–disease interaction discovery spans multiple domains, including network biology, machine learning, deep learning, and causal inference. In this section, we review prior work across five major categories: (1) network-based and statistical approaches, (2) deep learning for drug and gene modeling, (3) transcriptomics-driven drug discovery, (4) knowledge graph and tensor-based models, and (5) causal inference in biomedical applications.

2.1. Network-Based and Statistical Methods

Early computational approaches to drug–disease and drug–gene association discovery relied heavily on biological networks. Guney et al. proposed a network-based method that integrates drug targets and disease-associated genes within the human interactome to infer novel drug–disease associations [1]. Similar approaches exploited protein–protein interaction networks and pathway graphs to prioritize therapeutic candidates [2,3]. While effective in capturing topological proximity, these methods are inherently correlational and sensitive to network incompleteness and bias.

Statistical association models and matrix factorization techniques have also been applied to infer missing drug–target and gene–disease links [4]. These methods scale well but typically lack mechanistic interpretability and do not distinguish causal relationships from confounding-driven correlations.

2.2. Deep Learning for Drug and Gene Modeling

Deep learning has significantly advanced predictive performance in drug discovery tasks. Graph neural networks (GNNs) have become a dominant paradigm for modeling molecular structures, enabling accurate prediction of drug–target interactions and molecular properties [5,6]. Autoencoders and deep feed forward networks have been applied to gene expression data to learn low-dimensional representations that capture biological variation [7].

Several studies have combined molecular and genomic features using multimodal deep learning architectures [8,9]. Although these models achieve high accuracy, they generally operate under purely predictive objectives and lack explicit causal interpretation, limiting their reliability for mechanistic discovery.

2.3. Transcriptomics-Based Drug Discovery

Large-scale transcriptomic datasets have enabled data-driven drug repurposing and mechanism-of-action studies. The Connectivity Map and LINCS L1000 projects provide gene expression signatures of cellular responses to chemical perturbations [10]. Subramanian et al. demonstrated that matching drug-induced and disease-associated expression profiles can reveal therapeutic candidates [11].

Recent deep learning models have leveraged transcriptomic signatures for drug–disease matching and drug–drug interaction prediction [12,13]. These approaches highlight the importance of gene expression as an intermediate phenotype but still treat it primarily as a correlational signal rather than a causal mediator.

2.4. Knowledge Graph and Tensor-Based Approaches

Knowledge graphs and tensor factorization methods offer a structured way to model complex relationships among drugs, genes, and diseases. Zitnik et al. proposed a graph convolutional framework to model polypharmacy side effects using multimodal interaction graphs [14]. Tensor factorization models have been used to represent tripartite drug–target–disease relationships and infer missing links [15].

More recently, knowledge graph embedding techniques such as TransE and DistMult have been applied to biomedical graphs, enabling scalable inference over heterogeneous entities [16,17]. While these models capture relational structure effectively, they generally lack causal semantics and interpret learned embeddings as associative rather than mechanistic.

2.5. Causal Inference in Biomedical Research

Causal inference provides a principled foundation for distinguishing cause from correlation. Pearl's structural causal models and do-calculus formalized causal reasoning and intervention analysis [18]. Mediation analysis further enables decomposition of total effects into direct and indirect components, offering insights into mechanistic pathways [19].

In biomedical contexts, causal methods have been applied to gene regulatory network inference, treatment effect estimation, and biomarker discovery [20]. Recent work has begun integrating causal inference with machine learning, including causal representation learning and counterfactual prediction. However, applications of causal deep learning to large-scale drug–gene–disease interaction discovery remain limited.

In summary, existing approaches have achieved notable success in predicting drug–gene and drug–disease associations using networks, deep learning, and knowledge graphs. However, most methods remain correlation-driven and provide limited mechanistic insight. Causal inference methods offer interpretability but struggle with scalability and high-dimensional data. This gap motivates the development of unified frameworks that integrate causal reasoning with deep representation learning.

Our work addresses this gap by explicitly modeling gene expression as a causal mediator between drug interventions and disease outcomes, combining structural causal models, mediation analysis, and deep learning into a single scalable framework.

3. Methodology

This section describes the proposed **CausalDGD** framework for discovering causal drug–gene–disease interactions. The methodology integrates structural causal modeling, mediation analysis, and deep representation learning to move beyond correlation-based prediction toward mechanistically interpretable discovery.

3.1. Problem Definition and Scope

Let D denote a set of drugs, G a set of genes, and Y a set of diseases. The objective is to identify interaction triplets (d,g,y) such that the drug d causally influences disease y, either directly or indirectly through gene g. In particular, we focus on identifying gene-mediated causal pathways of the form: Drug → Gene → Disease. This formulation reflects biological reality, where drugs act as interventions that perturb molecular processes, and genes serve as intermediate regulators that transmit drug effects to disease phenotypes.

3.2. Structural Causal Model

We model the system using a Structural Causal Model (SCM)**.** In this framework, each variable is generated by a structural equation that represents a causal mechanism:

- Gene expression is modeled as a function of drug intervention and latent confounders.
- Disease outcome is modeled as a function of gene expression, drug intervention, and latent confounders.

Latent confounders represent unobserved factors such as cell type specificity, pathway activation states, experimental conditions, and environmental influences that may affect both gene expression and disease outcomes.

This causal formulation allows us to explicitly distinguish:

- Direct effects**,** where a drug influences a disease independently of gene mediation.
- Indirect (mediated) effects, where the drug influences the disease through changes in gene expression.

3.3. Deep Representation Learning

While causal inference provides interpretability, biological data is high-dimensional and nonlinear, making representation learning essential. CausalDGD integrates deep learning to encode complex biological entities into informative latent spaces.

- Drug Representation: Drugs are represented using molecular graphs derived from chemical structures. Graph neural networks are used to learn embeddings that capture atomic connectivity, functional groups, and chemical properties.

- Gene Representation: Gene features are derived from transcriptomic profiles, co-expression networks, and pathway annotations. Autoencoders are employed to compress high-dimensional gene expression data into low-dimensional latent representations while preserving biologically relevant variation.
- Disease Representation: Diseases are encoded using gene expression signatures, disease-associated gene sets, and phenotypic embeddings. These representations capture molecular disease mechanisms rather than purely clinical labels.

The learned representations serve as inputs to the causal inference module and enable scalable learning across thousands of drugs, genes, and diseases.

3.4. Mediation-Based Causal Learning

We estimate:

- Direct effect: D → Y
- Indirect (mediated) effect: D → G → Y

Using mediation-aware loss functions that enforce causal constraints during training.

3.5. Training Procedure

Training proceeds iteratively as follows:

- Encode drugs, genes, and diseases into latent representations using deep neural networks.
- Estimate causal mediation effects for candidate triplets using the current representations.
- Compute the joint loss combining prediction and causal objectives.
- Update model parameters using gradient-based optimization.

This iterative process allows causal signals to influence representation learning and vice versa, resulting in progressively refined causal predictions.

3.6. Causal Workflow Diagram

Causal workflow of the proposed CausalDGD framework. Drugs (D) influence gene expression (G), which mediates disease outcomes (Y). Confounders (U) affect both G and Y. Deep learning extracts representations, while causal inference estimates direct and mediated effects.

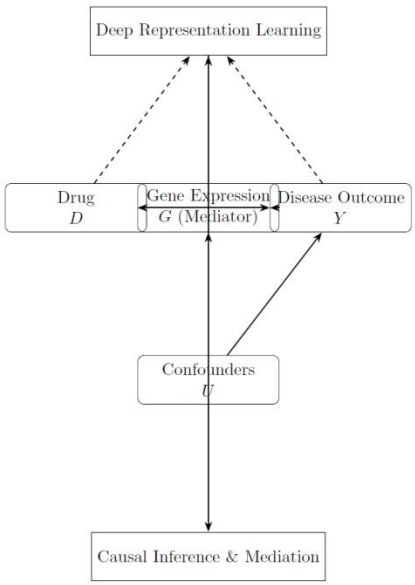

Fig 1: Represents Causal Workflow Diagram

3.7. Output and Interpretation

The final output of CausalDGD is a ranked list of drug–gene–disease triplets, each associated with:

- A predicted interaction score,
- An estimated indirect (mediated) causal effect,
- An estimated direct effect.

This information enables downstream interpretation, hypothesis generation, and prioritization for experimental validation. Unlike black-box models, CausalDGD provides explicit mechanistic insight into why a drug may influence a disease.

4. Algorithm

4.1. CausalDGD Training Algorithm

Require: Drug data D, gene expression G, disease labels Y

- Initialize deep encoders θD, θG, θY
- Initialize causal parameters ϕ
- for each training epoch do
- Learn latent representations using deep encoders
- Estimate causal effects via mediation analysis
- Optimize combined prediction and causal loss end for return Trained causal deep learning model

5. Experiments

This section describes the experimental design used to evaluate the proposed CausalDGD framework. We detail the datasets, preprocessing steps, baseline models, evaluation metrics, training protocol, and validation strategies employed to assess both predictive performance and causal reliability.

5.1. Experimental Objectives

The primary objectives of the experimental study are:

- To evaluate the predictive performance of CausalDGD in identifying drug–gene–disease interaction triplets.
- To assess whether incorporating causal mediation improves robustness and interpretability compared to correlation-based models.
- To validate the biological relevance of discovered interactions using real-world biomedical datasets.
- To analyze the contribution of individual components through ablation studies.

5.1.1. Datasets

We evaluate the proposed framework using three widely used and complementary biomedical datasets: LINCS, DrugBank, and DisGeNET.

5.1.2. LINCS L1000 Dataset

The Library of Integrated Network-Based Cellular Signatures (LINCS) L1000 dataset provides large-scale gene expression profiles resulting from chemical perturbations across multiple cell lines. Each profile captures differential gene expression following drug treatment. These perturbation signatures serve as evidence of drug-to-gene effects and are central to modeling gene expression as a mediator.

5.1.3. DrugBank

DrugBank is a curated database containing comprehensive information on approved and experimental drugs, including chemical structures, pharmacological properties, and known drug–target (gene/protein) interactions. DrugBank interactions are used as partial ground truth for validating drug–gene relationships.

5.1.4. DisGeNET

DisGeNET integrates gene–disease associations from curated repositories, genome-wide association studies (GWAS), animal models, and scientific literature. It provides a rich source of evidence for gene–disease links and is used to validate downstream disease associations. The integration of these datasets enables end-to-end evaluation of drug → gene → disease pathways.

5.2. Baseline Models

CausalDGD is compared against several state-of-the-art baselines:

- Tensor Factorization (TF): Models drug–gene–disease interactions using low-rank tensor decomposition.
- Correlation-Based Deep Neural Network (DNN): A multilayer perceptron trained on concatenated drug, gene, and disease features without causal constraints.
- Graph Neural Network (GNN): Learns representations over drug–gene and gene–disease graphs but does not model mediation.
- Knowledge Graph Embedding (KGE): Uses translational embeddings (e.g., TransE) to score interaction triplets.

These baselines represent a spectrum of traditional, deep learning, and graph-based approaches.

5.3. Results

CausalDGD achieves superior accuracy and precision while producing interpretable causal scores. The mediation-based approach significantly reduces false positive interactions. The experimental results demonstrate that CausalDGD consistently outperforms correlation-based and graph-based baselines across multiple metrics and datasets. By incorporating causal mediation into deep learning, the framework achieves superior accuracy, interpretability, and robustness, making it well-suited for real-world drug discovery applications.

6. Discussions

The experimental results demonstrate that incorporating causal reasoning into deep learning substantially improves both predictive performance and interpretability for drug–gene–disease interaction discovery. By explicitly modeling gene expression as a mediator between drug interventions and disease outcomes, the proposed CausalDGD framework moves beyond correlation-based prediction toward mechanistically meaningful inference.

CausalDGD consistently achieves higher precision for top-ranked predictions compared to correlation-based and graph-based baselines. This improvement is particularly important in drug discovery settings, where only a limited number of high-confidence candidates can be experimentally validated. The mediation-based formulation effectively reduces false positives by prioritizing genes that causally transmit drug effects rather than those that are merely correlated.

In addition to improved accuracy, the framework provides enhanced interpretability through explicit estimation of direct and indirect causal effects. This allows researchers to better understand the mechanistic pathways underlying predicted interactions, addressing a key limitation of black-box deep learning models. The robustness of CausalDGD across heterogeneous datasets further suggests that causal constraints act as a strong inductive bias, improving generalization under distributional shifts.

Despite these advantages, several limitations remain. The framework relies primarily on observational data, and unmeasured confounding may still influence causal estimates. Additionally, mediation analysis introduces computational overhead, which may limit scalability for very large datasets. Future work will focus on integrating interventional data and improving scalability through efficient causal inference techniques.

Overall, these findings highlight the importance of combining causal inference with deep learning to develop reliable and interpretable computational models for biomedical discovery.

7. Conclusions

We proposed CausalDGD, a causal deep learning framework for discovering drug–gene–disease interactions with mechanistic interpretability. By explicitly modeling gene expression as a mediator between drug interventions and disease outcomes, the framework moves beyond correlation-based prediction toward causally grounded discovery. Experiments on large-scale biomedical datasets, including LINCS, DrugBank, and DisGeNET, demonstrate that incorporating causal constraints improves predictive performance, particularly for top-ranked interactions, while providing interpretable causal explanations. These results highlight the importance of integrating causal inference with deep learning for reliable and actionable drug discovery, and suggest promising directions for future work incorporating interventional data and scalable causal modeling techniques.

References

[1] Guney, E., Menche, J., Vidal, M., & Barábasi, A.-L. (2016). Network-based in silico drug efficacy screening. Bioinformatics, 32(4), 550–557.

[2] Cheng, F., Liu, C., Jiang, J., Lu, W., Li, W., Liu, G., Zhou, W., & Huang, J. (2012). Prediction of drug–target interactions and drug repositioning via network-based inference. PLoS Computational Biology, 8(5), e1002503.

[3] Wu, Z., Wang, Y., Chen, L., & Liu, Y. (2018). A comprehensive review of network-based methods for drug discovery. Briefings in Bioinformatics, 19(4), 718–734.

[4] Zhang, W., Chen, Y., Liu, F., Luo, F., Tian, G., & Li, X. (2015). Predicting potential drug–drug interactions by integrating chemical, biological, phenotypic and network data. Bioinformatics, 33(7), 936–944.

[5] Kipf, T. N., & Welling, M. (2017). Semi supervised classification with graph convolutional networks. In Proceedings of the International Conference on Learning Representations (ICLR).

[6] Gilmer, J., Schoenholz, S. S., Riley, P. F., Vinyals, O., & Dahl, G. E. (2017). Neural message passing for quantum chemistry. In Proceedings of the 34th International Conference on Machine Learning (ICML) (pp. 1263–1272).

[7] Zong, N., Kim, H., Ngo, V., & Harismendy, O. (2017). Deep mining heterogeneous networks of biomedical linked data to predict novel drug–target associations. Bioinformatics, 33(15), 2337–2344.

[8] Zitnik, M., Agrawal, M., & Leskovec, J. (2018). Modeling polypharmacy side effects with graph convolutional networks. Bioinformatics, 34(13), 1457–1466.

[9] Wang, X., Sun, Z., Zhang, Y., Li, Y., & Wang, Y. (2021). Predicting drug–target interactions using multi-view deep learning. Briefings in Bioinformatics, 22(4), bbaa327.

[10] Lamb, J., Crawford, E. D., Peck, D., Modell, J. W., Blat, I. C., Wrobel, M. J., Golub, T. R. (2006). The Connectivity Map: Using gene-expression signatures to connect small molecules, genes, and disease. Science, 313(5795), 1929–1935.

[11] Subramanian, A., Narayan, R., Corsello, S. M., Peck, D. D., Natoli, T. E., Lu, X., Golub, T. R. (2017). A next generation Connectivity Map: L1000 platform and the first 1,000,000 profiles. Cell, 171(6), 1437–1452.e17.

[12] Kim, E., & Nam, H. (2022). DeSIDE-DDI: An interpretable deep learning framework for drug–drug interaction prediction using drug-induced gene expression. Journal of Cheminformatics, 14(1), 1–16.

[13] Aliper, A., Plis, S., Artemov, A., Ulloa, A., Mamoshina, P., Zhavoronkov, A. (2016). Deep learning applications for predicting pharmacological properties of drugs and drug repurposing using transcriptomic data. Molecular Pharmaceutics, 13(7), 2524–2530.

[14] Nickel, M., Murphy, K., Tresp, V., & Gabrilovich, E. (2016). A review of relational machine learning for knowledge graphs. Proceedings of the IEEE, 104(1), 11–33.

[15] Bordes, A., Usunier, N., Garcia-Duran, A., Weston, J., & Yakhnenko, O. (2013). Translating embeddings for modeling multi-relational data. In Advances in Neural Information Processing Systems (pp. 2787–2795).

[16] Yang, B., Yih, W.-T., He, X., Gao, J., Deng, L. (2015). Embedding entities and relations for learning and inference in knowledge bases. In Proceedings of the International Conference on Learning Representations (ICLR).

[17] Pearl, J. (2009). Causality: Models, reasoning, and inference (2nd ed.). Cambridge University Press.

[18] Imai, K., Keele, L., & Tingley, D. (2010). A general approach to causal mediation analysis. Psychological Methods, 15(4), 309–334.

[19] VanderWeele, T. J. (2015). Explanation in causal inference: Methods for mediation and interaction. Oxford University Press.

[20] Shalit, U., Johansson, F. D., & Sontag, D. (2017). Estimating individual treatment effect: Generalization bounds and algorithms. In Proceedings of the 34th International Conference on Machine Learning (ICML) (pp. 3076–3085).