Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery

WILEY

WIREs DATA MINING AND KNOWLEDGE DISCOVERY

ADVANCED REVIEW OPEN ACCESS

Exploring Causal Learning Through Graph Neural Networks: An In-Depth Review

Simi Job[1] | Xiaohui Tao[1] | Taotao Cai[1] | Haoran Xie[2] | Lin Li[3] | Qing Li[4] | Jianming Yong[5]

[1]School of Mathematics, Physics, and Computing, University of Southern Queensland, Toowoomba, Queensland, Australia | [2]School of Data Science, Lingnan University, Tuen Mun, Hong Kong | [3]School of Computer Science and Artificial Intelligence, Wuhan University of Technology, Wuhan, China | [4]Department of Computing, Hong Kong Polytechnic University, Hung Hom, Hong Kong | [5]School of Business, University of Southern Queensland, Toowoomba, Queensland, Australia

Correspondence: Xiaohui Tao (xiaohui.tao@unisq.edu.au)

Received: 6 February 2025 | **Revised:** 15 April 2025 | **Accepted:** 3 May 2025

Associate Editor: Tianrui Li | **Editor-in-Chief:** Witold Pedrycz

Funding: This study was supported by the University of Southern Queensland, as part of the Wiley - University of Southern Queensland agreement via the Council of Australian University Librarians.

Keywords: causal Learning | causality | GNNs | graph neural networks

ABSTRACT

In machine learning, exploring data correlations to predict outcomes is a fundamental task. Recognizing causal relationships embedded within data is pivotal for a comprehensive understanding of system dynamics, the significance of which is paramount in data-driven decision-making processes. Beyond traditional methods, there has been a shift toward using graph neural networks (GNNs) for causal learning, given their capabilities as universal data approximators. Thus, a thorough review of the advancements in causal learning using GNNs is both relevant and timely. To structure this review, we introduce a novel taxonomy that encompasses various state-of-the-art GNN methods used in studying causality. GNNs are further categorized based on their applications in the causality domain. We further provide an exhaustive compilation of datasets integral to causal learning with GNNs to serve as a resource for practical study. This review also touches upon the application of causal learning across diverse sectors. We conclude the review with insights into potential challenges and promising avenues for future exploration in this rapidly evolving field of machine learning.

1 | Introduction

Exploration of causality forms the foundation of machine learning, aiming to uncover the intricate relationships embedded within data features. Predominantly, the focus has been on identifying correlations and associations, such as the links between lifestyle choices and health outcomes. Though traditional machine learning techniques, such as classification, are well-suited for predicting outcomes, they are inadequate for understanding causality. Moreover, there has been a growing recognition of the need to delve deeper into the actual causal elements behind these associations, especially when the objective extends beyond mere prediction to actual improvement or intervention. In building on this approach, it is essential to determine an alternative course of action that can lead to a more optimal outcome, inclusive of cause–effect relationships, which is where the concept of causality is significant. The process of causal analysis can be adopted to explore such causal effects, whereby the elements that contribute to a specific outcome can be determined.

Causal analysis is crucial in machine learning to ensure generalizability, explainability, and fairness of ML models (Zhao et al. 2022b). At its core, causality bifurcates into causal discovery and causal inference. The former unravels the structural relationships within data, while the latter investigates the consequences of various interventions. A significant challenge in causal inference is discerning the distinct effects of concurrent actions. A key deterrent to determining causation is the assumption that causal effects cannot be reliably evaluated statistically (Pearl 2000), on account of the impracticality of knowing possible outcomes in the event of an alternative course of action.

Causal learning has gained significance in several areas, including in the development of recommender systems (Si et al. 2023), language modeling (Vig et al. 2020), the medical field (Chai 2020), including medical imaging (Castro et al. 2020), and pharmacology (Zhao et al. 2022b), urban intelligence (Mannering et al. 2020), economics (Alam and Sumon 2020), and so on. In causal analysis, it is essential to address counterfactuals that allow replication of different approaches simultaneously. Causal machine learning provides a framework for including this aspect of causality and discovers the influence of confounders on an outcome. Incorporating causality into machine learning establishes causal relationships between features, which consequently lead to improved model accuracy (Carbo-Valverde et al. 2020). Causal forests, a variation of random forests (RF) (Carbo-Valverde et al. 2020) and ensemble methods that use logistic regression, RF, gradient boosting, and SVM (Decruyenaere et al. 2020) are examples of methods employed in causal machine learning. Propensity score (estimates treatment likelihood given covariates), covariate balancing (adjusts for characteristic differences between treatment groups), instrumental variable (IV-isolates causal effects by accounting for confounding), and Frontdoor criterion (uses a mediator for causal effects) are conventional methods for learning causal effects. Traditional causal discovery methods include constraint-based (PC (Spirtes et al. 2001), FCI (Spirtes et al. 2001)) and score-based (GES (Chickering 2002)) methods. Fundamentally, causal interactions can be represented in the form of graphs in most cases, rendering such traditional approaches ineffective for handling the dynamics of causal relationships.

Deep learning techniques such as MLPs (Multilayer Perceptron) have been used to learn causality (Si et al. 2023), with graph-based deep learning methods adopted in recent research studies (Chai 2020). Modeling causal relationships by means of graphs serves to establish causality dynamics. Causality learning using graph methods is gaining attention in recent times since GNNs have proven to be effective for causal analysis due to their potential as universal approximators of data (Zečević et al. 2021). Moreover, GNNs are adept at multi-modal deep learning and can also address the problem of feature sparsity through graph representations using feature interactions represented in the form of nodes (Zhai et al. 2023). GNN, with its neural network architecture, can handle complex graphs consisting of a multitude of nodes. Moreover, GNNs are capable of extracting information from unstructured data as well as modeling dynamic and evolving data structures. On these grounds, GNNs have an inherent potential to capture causality in data, especially when compared to traditional approaches as discussed by Feng et al. (2021) and Zhao et al. (2022a).

Existing surveys on graphical causal learning are predominantly centered on generalized graphical models. While Cheng et al. (2022) delved into graphical causal modeling techniques, they overlooked the role of GNNs. Other survey papers like Wu et al. (2020) and Zhou et al. (2020) have explored GNNs but disregarded their applications in causal learning. Yuan et al. (2022) focused on the explainability aspect of GNNs, with a brief reference to the causal screening method. Notably, causality-focused surveys such as (Nogueira et al. 2022) and (Yao et al. 2021) omitted GNNs for causality, while Kuang et al. (2020) provided only a cursory mention of GNNs in their discourse on causal inference. This survey comprehensively reviews GNNs for causal learning, addressing this identified gap in the literature. A visual summary of the survey is presented in Figure 1. To the best of our knowledge, this is the only study focusing on causal learning with GNNs. The key contributions of the article are the following:

- *Systematic taxonomy and in-depth review*: We introduce a meticulously designed taxonomy tailored for this survey, showcased in Figure 2. Every category undergoes a thorough review, summarization, and comparison against key works in the domain. Furthermore, we compile a detailed

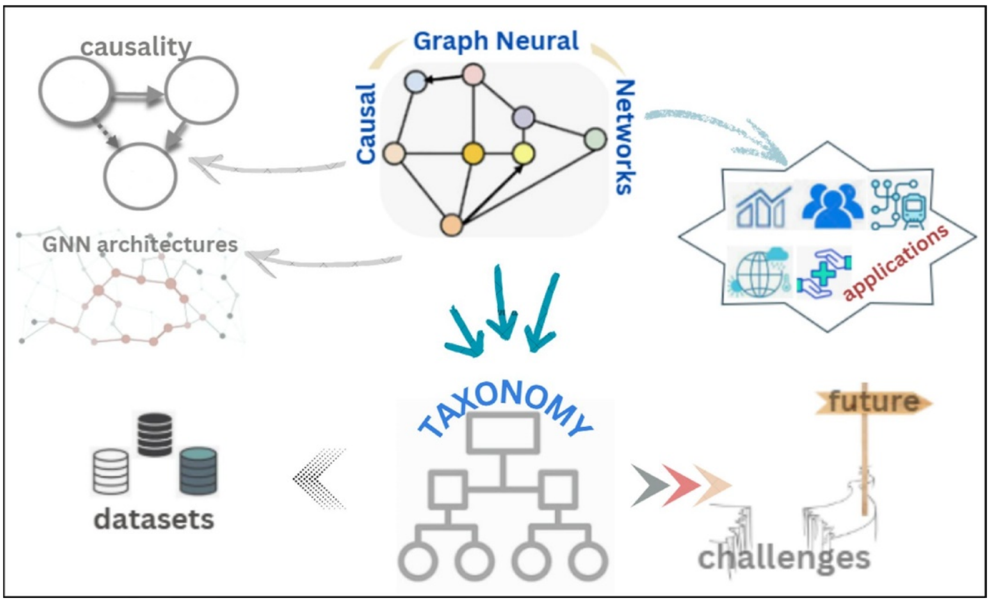


FIGURE 1 | Graphical abstract of the survey with a taxonomical approach to causal learning with graph neural networks.

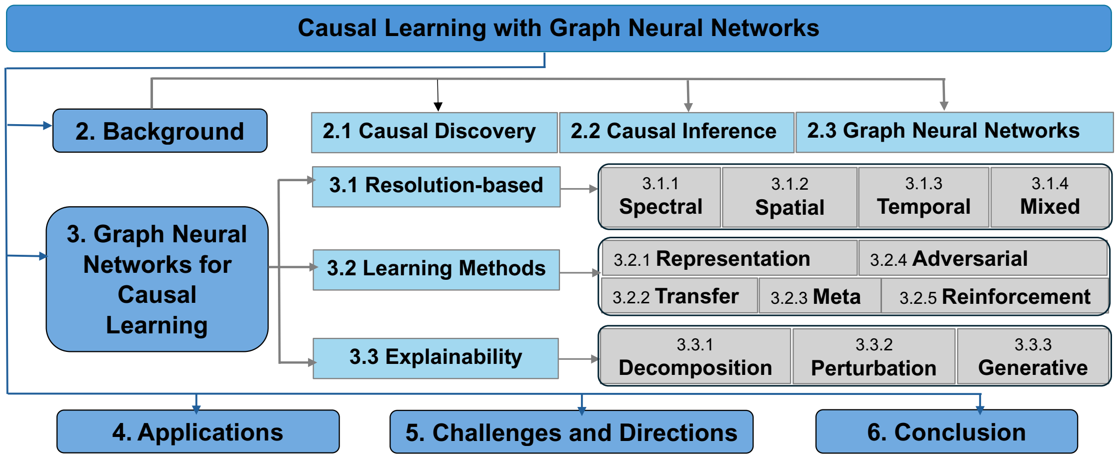


FIGURE 2 | Survey overview.

list of state-of-the-art methods employed in causal learning with GNNs, along with datasets pertinent to causality-focused GNN research (Section 3).

- *Rich resources and noteworthy references*: We present a curated collection of open resources, encompassing benchmark datasets, technical details, and practical applications. Our cited references are drawn from premier peer-reviewed journals covering areas such as data mining, AI, GNNs, and knowledge discovery (Section 3).
- *Future horizons*: We recognize the evolving nature of GNNs and its application in causality. Using this knowledge, we delve into its present constraints, challenges, and untapped opportunities, charting out potential avenues for future exploration in the field, and offer insights about future research in the domain (Section 5).

The remainder of the paper is organized as follows: Section 2: The background introduces causal learning concepts and GNNs, including their common variants, Section 3 presents the GNN approaches for causal learning, Section 4: Applications present the practical applications, and then Section 5 discusses challenges and future directions, and the final section concludes this survey.

2 | Background

Before deep learning methods, causal learning was studied using traditional methods. Figure 3 provides an outline of traditional methods used in graph learning and causal learning. GNNs, owing to their ability to act as functional approximators, have been researched in recent times due to the potential to extract causal relations and causal effects (Yu et al. 2019). This section provides a background on causal learning and GNNs.

2.1 | Definitions

2.1.1 | Definition 1 (Structural Causal Model)

Structural Causal Model (SCM; Pearl 2009) is a set of endogenous variables V and a set of exogenous variables U, mapped in terms of a set of functions F. The values of $V = V_0, V_1, \dots, V\{_i\}$ and $U = U_0, U_1, \dots, U\{_u\}$ are determined by F. A SCM is associated with a directed acyclic graph (DAG), with a node in U or V, and each edge in F.

2.1.2 | Definition 2 (Causal Modeling)

Causal modeling represents causal relationships mathematically through SCMs. *Potential outcome framework* (Rubin 2005), similar to the SCM model, is defined in terms of the treatment applied to a unit under study. *Potential outcome* is the outcome of a treatment. *Counterfactual outcome* forms the outcome if a different treatment were applied to the unit. Causal learning evaluates the change in outcome when an alternative treatment is applied, where control c is altered to treatment t, expressed as $E[y|\,t] - E[y|\,c]$.

2.1.3 | Definition 3 (Graphs)

A graph G is a set of nodes with edges and is represented as $G = (V, E)$, where V is the set of vertices (nodes) and E is the set of edges connecting the nodes. Graphs may be directed or undirected. Two nodes connecting an edge e, namely, $u \in V$ and $v \in V$ are neighbors and hence have an *adjacent* relationship. Adjacency matrix represents a graph as Boolean values for indicating the node connections.

A *causal graph* (Pearl 2009) is a directed acyclic graph (DAG) that represents causal relationships between variables. The nodes in a causal graph represent variables, which can be either observed (measured) or unobserved (latent), and the directed edges between the nodes indicate causal effects.

2.1.4 | Definition 4 (Back-Door Path)

A backdoor path between treatment and outcome variables (t,y) in a causal graph is a path that connects these two variables with an arrow pointing into t and does not follow a direct causal path from t to y (Pearl 2009). The treatment variable t corresponds to the intervention being studied, and the outcome variable y represents the result. A backdoor path can

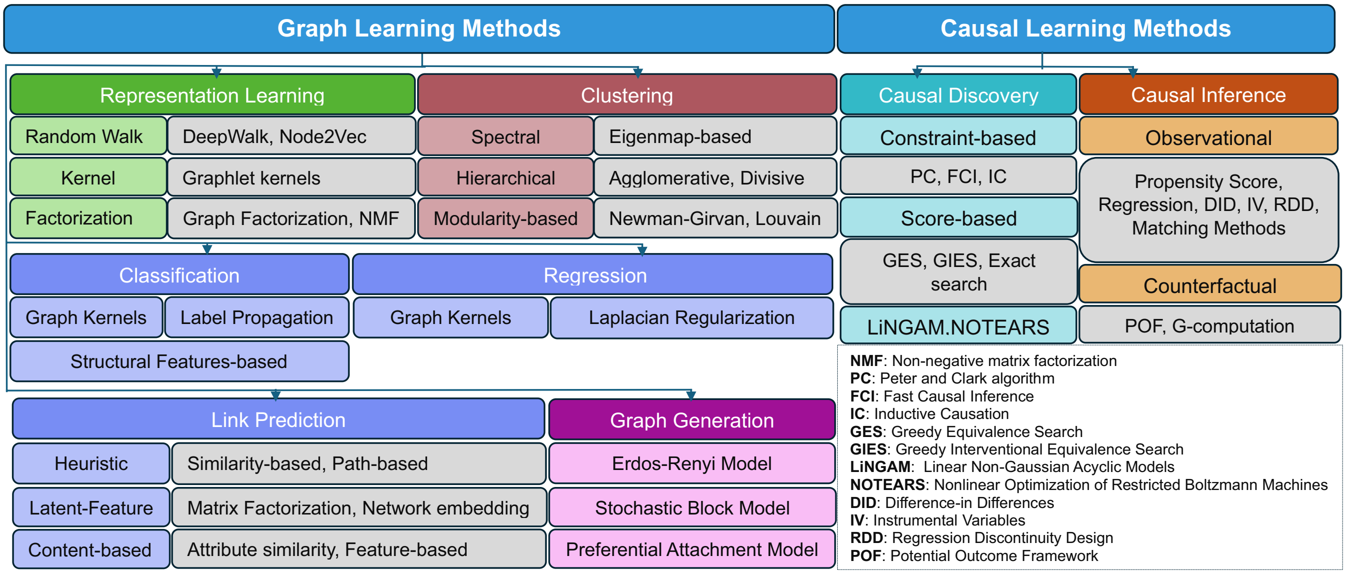


FIGURE 3 | Traditional graph learning and causal learning methods.

cause confounding bias unless blocked by controlling for the appropriate variables.

2.1.5 | Definition 5 (Confounder)

For a pair of variables (t,y), a variable $z \notin \{t, y\}$ is a confounder (Pearl 2009) of the causal effect $t \rightarrow y$, if z is the central node of a fork on a back-door path between t and y. z is a common cause of both t and y, possibly introducing confounding bias in the estimation of the causal effect of t on y. Conditioning on z blocks the back-door path and mitigates the confounding bias.

Causal learning involves causal discovery and causal inference. In causal discovery, the causal structure is derived from data to form a causal graph. Causal inference involves inferring causal effects from this causal graph. These concepts are briefly discussed in the following subsections.

2.2 | Causal Discovery

Causal discovery involves analyzing observational data to derive causal relations and is based on four assumptions (Spirtes et al. 2001):

- *Acyclicity*: A causal graph has no directed cycles, ensuring no variable is both a cause and an effect of itself.
- *Markov property*: A node in a causal graph is independent of nondescendant nodes, conditioned on its direct causes (parents), such that once the direct influences are accounted for, nondescendant nodes provide no additional information.
- *Faithfulness*: Causally connected variables in a graph are probabilistically dependent, and all observed dependencies are explained by the causal graph.

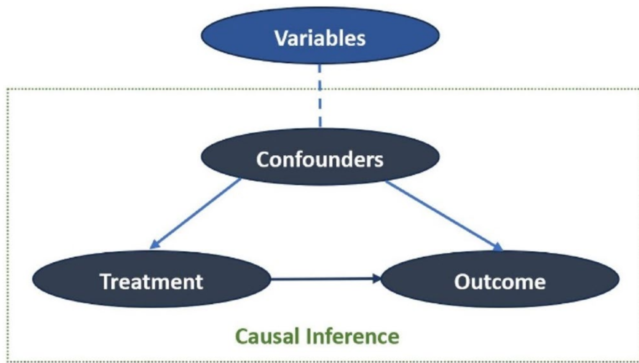


FIGURE 4 | Causal inference.

- *Sufficiency*: All confounders in the graph must be observed variables, ensuring that the causal discovery process is not biased by unobserved confounding factors.

2.3 | Causal Inference

Causal inference estimates the causal effect of a variable (treatment) over a variable of interest (outcome) as depicted in Figure 4. Here, a confounder is a variable that is causally associated with treatment and outcome. Causal inference derives the effect of an intervention on the outcome and requires the following conditions (Arboleda-Florez et al. 1998) to be satisfied:

- *Covariation*: A change in the causal variable should result in a change in the variable that is expected to be affected.
- *Temporal precedence*: If a causal variable causes an effect on the affected variable, then the occurrence of the causal variable must precede the effect on the affected variable.
- *Control extraneous variables*: Elimination of possible alternative causes.

Causal inference is based on the following assumptions:

- SUTVA (Stable Unit Treatment Value Assumption) (Fé et al. 2020): SUTVA assumes that there is no variation in the treatment within units and no interference between units' treatment statuses.
- Ignorability (Pearl 2010): Ignorability assumes that treatment assignment is independent of potential outcomes, with no unmeasured confounders.
- Consistency (Pearl 2010): Consistency assumes that the observed outcome for a unit is the same as the potential outcome under the treatment received.
- Positivity (Cole and Frangakis 2009): Positivity assumes that every unit has a positive chance of receiving each treatment level, ensuring that all treatment groups are represented in the data.

In causal learning, two approaches for computing causal effects from observational data are the Backdoor (Pearl 1995) and Frontdoor Adjustment criteria (Pearl 1995), given a causal graph. These concepts are briefly described in the following subsections.

2.3.1 | The Backdoor Adjustment

In a causal graph, a set of variables V satisfies the Backdoor criterion for a pair of variables (X,Y) if the following two conditions are met:

- Blocking all backdoor paths: V blocks all backdoor paths from X to Y, preventing bias in estimating the causal effect of X on Y.
- No descendants of X are in V: No node in V is a descendant of X, as this would reintroduce bias through confounding paths.

2.3.2 | The Frontdoor Adjustment

In a causal graph, a set of variables V satisfies the Frontdoor criterion for a pair of variables (X,Y) if:

- Blocking all directed paths from X to Y: V blocks all directed paths from X to Y.
- Blocking backdoor paths from V to Y by X: All backdoor paths from V to Y are blocked by X, preventing confounding through these paths.
- No backdoor paths from X to Y: There exist no backdoor paths from X to V, as this would introduce bias if V were included in the adjustment set.

2.4 | Graph Neural Networks

GNNs are deep learning techniques that are used for processing and analyzing graph-based data. With their neural architecture, they can learn node representation. The representation matrix is defined as $H \in R^{N \times F}$ where F is the node representation dimension. GNNs aggregate representations of node neighbors and update node representations in an iterative manner and are capable of learning causality with their functional approximation properties. A mathematical representation of the GNN framework is defined in (Equation (1); Wu et al. 2022), with *Aggregate* and *Combine* functions in each layer, where the node representation is initialized as $H^0 = X$. Here, $N(v)$ is the set of neighbors for node v. K is the total number of GNN layers, where $k = 1,2,\ldots,K$ and H^{K} is the finalized node representation. a_v^k is the aggregated node feature of the neighborhood H_u^{k-1}.

$$a_v^k = \mathrm{Aggregate}^k\left\{H_u^{k-1} : u \in N(v)\right\}$$

$$\boldsymbol{H_v^k} = \mathrm{Combine}^{\boldsymbol{k}}\left\{\boldsymbol{H_v^{k-1}}, \boldsymbol{a_v^k}\right\} \tag{1}$$

Graph learning includes the node-level, graph-level, or edge-level tasks briefly described here.

Node-level: *In node-level tasks, the property of each node in a graph is predicted. Node-level features may be importance-based or structure-based.*

Graph-level: *In graph-level tasks, the features of the entire graph structure are captured for computing graph similarities. Graphlet Kernel and Weisfeiler-Lehman Kernel are the two main approaches for feature learning at the graph level.*

Link or edge-level: *In link-level tasks, existing links are used to make predictions on new or missing links. Distance-based, local neighborhood overlap, and global neighborhood overlap are link-level features.*

There are four major graph learning tasks using machine learning, as outlined as follows:

- *Node Classification* is the task of determining the class of a node based on the classes of neighboring nodes.
- *Graph Classification* is the task of classifying an entire graph into various groups.
- *Node Regression* is the task of predicting a continuous value for a node.
- *Link Prediction* predicts potential relationships between two nodes in a graph.

For link prediction, either node-based or subgraph-based GNN methods may be used. A GNN layer builds node-level representations; hence, graph-level representations require graph pooling techniques (Wu et al. 2022) to reduce dimensionality while retaining maximum graph information. A general representation of the GNN classification pipeline is illustrated in Figure 5. Here, a GNN layer computes the input features from the input graph and aggregates these features to generate node embeddings. Node features are updated with an update function, and the transformed graph is passed through a classification layer to predict labels. The most important GNN architectures used in

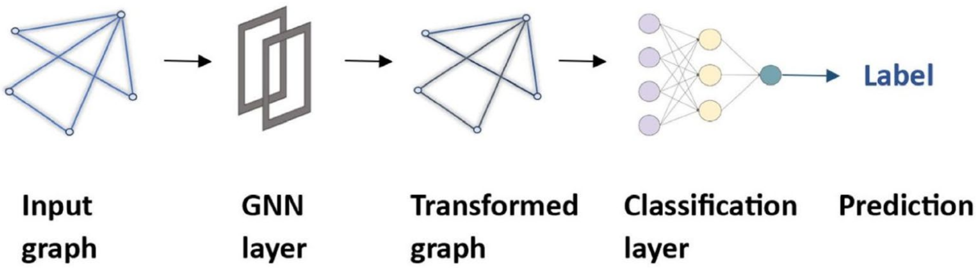


FIGURE 5 | Prediction with graph neural networks.

causal learning are GCN, GAT, GGNN, GAE, and GIN, which are briefly described as follows.

2.4.1 | Graph Convolutional Networks

Graph Convolutional Networks (GCNs; Kipf and Welling 2017) are GNNs that use convolutional aggregations and aggregate features, including neighbor features. A classic GCN framework consists of graph convolutional layer(s), a linear layer, and a nonlinear activation layer. The input graph undergoes a convolution operation that is applied to each node, following which feature information is aggregated and the node representations updated. An activation function such as ReLU is applied to the convolutional layer output.

GraphSAGE (Graph SAmple and aggreGatE; Hamilton et al. 2017) is an extension of GCN with an inductive learning framework, where sampling and aggregation of features from a node's neighborhood are used to build node representations. GraphSAGE samples a subset of neighbors for each node at various layers. The aggregator then aggregates the neighbors of preceding layers with an aggregating operation using mean aggregator, pooling aggregator, or LSTM aggregator.

2.4.2 | Graph Attention Networks

Graph Attention Networks (GATs; Velickovic et al. 2017) incorporate an attention layer into the GNNs for compiling neighborhood properties of nodes to form embeddings. Using the attention mechanism, the neighbors are assigned weights relative to their importance to a specific node. GATs attend to unseen nodes with their inductive learning, supporting both direct and indirect graphs. The graph attention layer applies a linear transformation to the node feature vectors with a weighted matrix. The activation function used is LeakyReLU, and the attention coefficients are computed based on the relative importance of neighbor features. To have a common scaling across the neighborhood, the coefficients are normalized. The process can be improvised with multi-head attention, wherein multiple attention maps are aggregated.

2.4.3 | Gated Graph Neural Networks

In gated graph neural networks (GGNNs; Li et al. 2016), gated recurrent units (GRU) are incorporated with GNNs as a recurrent function for the propagation step. Recurrence is performed for a fixed number of steps, and gradients are computed with backpropagation. With GGNN, node labels termed as node annotations may be added as inputs. Previous hidden states and neighboring hidden states are used for updating the hidden state of a node. GGNNs need to perform the recurrent function over all nodes multiple times, which is a drawback when processing large graphs.

2.4.4 | Graph Auto Encoder Networks

Graph Auto Encoder Networks (GAEs; Kipf and Welling 2016) utilize an encoder–decoder technique to encode nodes or graphs as a latent representation to construct network embeddings. The encoder generates embeddings for nodes using graph convolutional layers. The topological information of the nodes is incorporated in the representation. The decoder reconstructs the graph adjacency matrix after evaluating pairwise similarity of the network embeddings. The GAE aims to minimize the reconstruction loss of the decoded adjacency matrix as compared to the original matrix.

2.4.5 | Graph Isomorphism Networks

Graph Isomorphism Networks (GINs; Xu et al. 2019) are GNNs with high representational power as defined by the Weisfeiler–Lehman (WL) graph isomorphism test. The WL tests if two graphs are nonisomorphic by creating subtrees for the graph nodes, followed by color mapping each node based on the number of neighbors. The aggregate and combine functions are represented as the sum of the node features.

3 | Graph Neural Networks for Causal Learning

Causal learning focuses on uncovering cause-and-effect relationships in data, requiring specialized methods compared to traditional representation learning. GNNs generally model correlations between nodes, but their architecture must be adapted to explicitly capture causal structures (Job et al. 2025a; Sui et al. 2022a), which requires consideration of the following key factors:

- *Edge interpretation*: In causal graphs, edges represent direct causal relationships, with directionality indicating causality. GNNs, however, use edges to represent correlations. To adapt GNNs for causal learning, the causal structure must

FIGURE 6 | Causal learning graph neural networks (CLGNN) architecture.

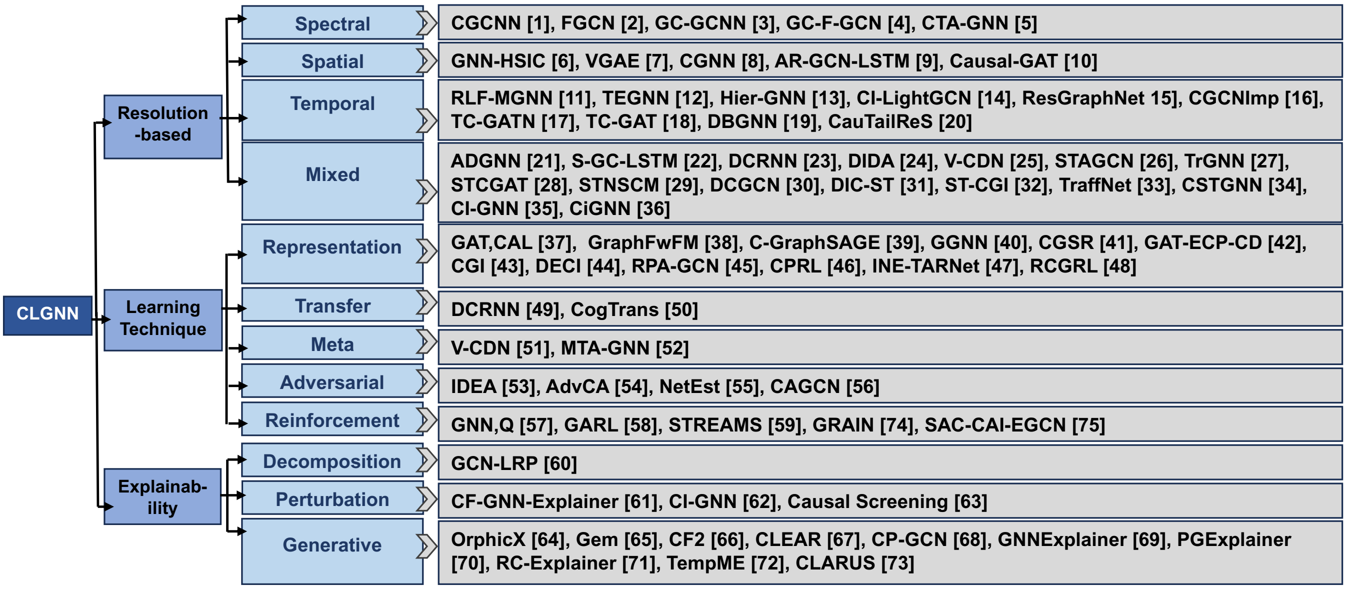
be determined, enabling the model to better capture causal relationships from data (Gao et al. 2024).

- *Establishing causal relationships using GNNs*: Adapting GNN architectures to infer causality involves incorporating causal constraints such as causal priors (Gao et al. 2024) or domain knowledge (Horiwaki 2023) into the learning process. These adjustments enable GNNs to propagate information based on both correlations and causal structures in the data. Techniques like feature emphasis and ignoring mechanisms also help to guide the model toward causal learning (Gao et al. 2024). Other methods, including causal weight masking, graph decomposition, and contrastive learning, have been applied to enhance causal learning in GNNs (Li et al. 2024b).
- *Linking correlation and causation in GNNs*: A key challenge in using GNNs for causal learning is distinguishing between correlation and causation. Correlation alone is insufficient to establish causality; it requires both correlation and the assumption of a directed cause–effect relationship. To address this, methods such as structural equation modeling (SEM), causal graph learning, and do-calculus-based interventions can be integrated into GNNs (Jiao et al. 2024), enabling them to move from correlation to causation and improve causal modeling.
- *Leveraging GNNs for causal learning*: The advantage of adapting GNNs for causal learning lies in their ability to combine the strengths of GNNs in learning causal representations with the capacity to infer cause and effect relationships. By modifying GNN architectures, such as adjusting edge weights (Fan et al. 2022) or incorporating causal constraints (Hu et al. 2025), GNNs can be trained to identify causal relationships from data.

A GCN framework integrated with mutual information was designed to analyze accident characteristics and learn causality in maritime vessel traffic accidents, aiming to improve decision making in accident investigations (Gan et al. 2025). A GNN enhanced with a causal attention-based sampling method was designed to learn geographic network representations (Wang et al. 2025b). This approach assigns weights to neighboring nodes based on similarity and causality, minimizing dependence on background representations. The model employs HSIC for sample reweighting to ensure independence between invariant and background representations, enhancing the GNN's ability to mitigate spurious correlations during predictions. An attention-based GNN was adapted for causal learning by incorporating the Hebbian principle and dynamic impact valuing, leading to the development of HebCGNN (Job et al. 2025b). This approach overcomes the limitations of attention mechanisms, which alone cannot capture causality, and improves graph classification performance by capturing complex feature interactions. In InsGNN (Fang et al. 2025), the GNN integrated with a mask generator enables causal interpretability by using learnable structural masking to identify invariant causal subgraphs, ensuring that predictions are causally relevant. CDA-GNN (Chang et al. 2025) is a causal discovery attention-based GNN that identifies key causal features from high-resolution breast cancer images. An attention mechanism incorporated into the GNN architecture is used to obtain causal features, and a backdoor adjustment strategy is employed to separate causal from noncausal features.

GNNs are also used in counterfactual causal learning to generate counterfactuals that consider causal relationships and confounders in the graph. The Authentic Graph Counterfactual Generator (AGCG) (Wang et al. 2025c) is a framework that addresses graph structure bias and enhances the fairness of GNN models. It improves counterfactual sample generation by identifying hidden confounders and ensures counterfactual fairness through concurrent learning from the original and corresponding true counterfactual sample. MMGCF (Zhang et al. 2025b), a GNN-based counterfactual explanation framework, generates counterfactuals through motif perturbations, identifying causally significant motifs and their impact on predictions. Counterfactual interventions are used to enhance graph contrastive learning through Counterfactual Intervention Enhancing Graph Contrastive Learning (CIEGCL), applying interventions during positive sampling and generating augmented information for contrast penalty to improve model generalization (Huang et al. 2025).

In this survey, the GNNs specifically employed for causality learning are given the term Causal Learning GNNs, abbreviated as CLGNNs, and the general architecture of a CLGNN is depicted in Figure 6. The causal discovery module infers causal structure from observational data. The causal graph is built to represent causal relationships between variables, from which causal effects are inferred using GNN. Neighborhood information in a graph is propagated through GNN layers, which is aggregated through an aggregation function. The final layer is then used to make the predictions. We group CLGNNs into three broad classes: (1) Resolution-based (pertaining to the nature of data under study), which are further categorized as Spectral, Spatial, Temporal, and Mixed methods; (2) Learning Methods (determined from the learning methodologies employed for causal learning), grouped into Representation, Transfer, Meta, Adversarial, and Reinforcement-based methods; and (3) Explainability (determined from the model explanation technique), comprised of Decomposition, Perturbation, and Generative methods. A summary of the datasets



[1]	Kong et al. (2022)	[26]	Gu and Deng (2022)	[51]	Li et al. (2020)
[2]	Li et al. (2023)	[27]	Li et al. (2021)	[52]	Meng et al. (2023)
[3]	Zhang et al. (2022)	[28]	Zhao et al. (2022)	[53]	Tao et al. (2023)
[4]	Zhang et al. (2022)	[29]	Deng et al. (2023)	[54]	Sui et al. (2022b)
[5]	Wang et al. (2023)	[30]	Lin et al. (2023)	[55]	Jiang and Sun (2022)
[6]	Ma and Tresp (2021)	[31]	Zhang et al. (2022)	[56]	Lee and Han (2023)
[7]	Zeˇcevi´c et al. (2021)	[32]	Zhang et al. (2022)	[57]	Amirinezhad et al. (2022)
[8]	Li et al. (2023)	[33]	Xu et al. (2023)	[58]	Yang et al. (2023)
[9]	Ntemi et al. (2022)	[34]	Jiang et al. (2024)	[59]	Sheth et al. (2023)
[10]	Liu et al. (2023)	[35]	Zheng et al. (2024)	[60]	Holzinger et al. (2021)
[11]	Wang et al. (2023)	[36]	Liu et al. (2024b)	[61]	Lucic et al. (2022)
[12]	Xu et al. (2020)	[37]	Sui et al. (2022a)	[62]	Zheng et al. (2023)
[13]	Wang et al. (2023)	[38]	Zhai et al. (2023)	[63]	Wang et al. (2021)
[14]	Ding et al. (2022)	[39]	Zhang et al. (2022)	[64]	Lin et al. (2022)
[15]	Chen et al. (2022)	[40]	Trust et al. (2022)	[65]	Lin et al. (2021)
[16]	Liu et al. (2022)	[41]	Wu et al. (2022)	[66]	Tan et al. (2022)
[17]	Li et al. (2022)	[42]	Cao et al. (2022)	[67]	Ma et al. (2022)
[18]	Yuan et al. (2023)	[43]	Feng et al. (2021)	[68]	Jin et al. (2022)
[19]	Qarkaxhija et al. (2022)	[44]	Phu and Nguyen (2021)	[69]	Ying et al. (2019)
[20]	Zeyu et al. (2023)	[45]	Chen et al. (2022)	[70]	Luo et al. (2020)
[21]	Wang et al. (2022)	[46]	Yang et al. (2022)	[71]	Wang et al. (2022)
[22]	Monken et al. (2021)	[47]	Adhikari, Zheleva(2023)	[72]	Chen and Ying (2024)
[23]	Wein et al. (2021)	[48]	Gao et al. (2023)	[73]	Metsch et al. (2024)
[24]	Zhang et al. (2022)	[49]	Wein et al. (2021)	[74]	Xiao et al. (2025)
[25]	Li et al. (2020)	[50]	Wu and Zhou (2023)	[75]	He et al. (2024)

FIGURE 7 | Taxonomy for causal learning graph neural networks (CLGNN).

discussed in the survey is tabulated in Table 1. The taxonomic structure is represented in Figure 7 and is discussed in this section.

3.1 | Resolution-Based Causal Learning GNNs

In this section, we present the CLGNNs in terms of the format of the modeled data. This aids in determining the specific GNN approaches that are capable of learning causality in the various data spaces. To this end, the resolution-based CLGNNs are further grouped into Spectral, Spatial, Temporal, and Mixed classes, each of which is described in this section.

3.1.1 | Spectral

Spectral data consists of information related to frequencies obtained from signals in machinery, brain monitoring, and so forth. For making effective use of signal data, it is pertinent to compile the interactive information between these signals. This process is simplified by a graphical approach that facilitates the capture of the topological structure of the system (Wang et al. 2023d). The signals are converted to adjacency matrices for representing graph connections and then transformed into causal graphs. The inference from the causal graph estimator is used for graph learning tasks. GNNs are adept at amassing

TABLE 1 | Datasets for causal learning graph neural networks (CLGNNs).

Category	Dataset	CLGNN studies
Social networks	IMDB-B, IMDB-M (Yanardag and Vishwanathan 2015)	(Sui et al. 2022a; Ma et al. 2022)
	Yelp (Yelp 2022)	(Zhang et al. 2022d; Ding et al. 2022)
	COLLAB (Yanardag and Vishwanathan 2015)	(Sui et al. 2022a; Zhang et al. 2022d)
	Gowalla (Cho et al. 2011)	(Ding et al. 2022; Wu et al. 2023)
	Adressa (Gulla et al. 2017)	(Ding et al. 2022)
	Pokec (Takac and Zabovsky 2012)	(Ma and Tresp 2021)
	REDDIT-MULTI-5K (Yanardag and Vishwanathan 2015)	(Wang et al. 2022c; Xiang et al. 2021)
	Reddit (Hamilton et al. 2017)	(Tao et al. 2024)
	Flickr (Flickr 2020)	(Guo et al. 2020; Jiang and Sun 2022)
Citation networks	Cora (McCallum et al. 2000)	(Huang et al. 2022; Tao et al. 2024; Zhang et al. 2022f; Feng et al. 2021; Lee and Han 2023)
	Citeseer (Giles et al. 1998)	(Tan et al. 2022b; Tao et al. 2024; Zhang et al. 2022f; Feng et al. 2021; Lee and Han 2023)
	PubMed (Sen et al. 2008)	(Huang et al. 2022; Zhang et al. 2022f; Feng et al. 2021; Lee and Han 2023)
	MAG (Sinha et al. 2015)	(Cummings and Nassar 2020)
	ogbn-arxiv (Hu et al. 2020)	(Tao et al. 2024; Feng et al. 2021)
Purchase networks	Transaction (Zhang et al. 2022d)	(Zhang et al. 2022d)
	Amazon (Leskovec et al. 2007; Rakesh et al. 2018)	(Ma and Tresp 2021; Wu et al. 2023)
	ogbn-products (Hu et al. 2020)	(Tao et al. 2024)
	Diginetica (CIKM 2016)	(Wu et al. 2023)
Road networks	PeMS03/04/07/08 (Guo et al. 2021)	(Zhao et al. 2023b; Gu and Deng 2022)
	SG-TAXI (Land Transport Authority, SG-TAXI 2016)	(Li et al. 2021)
	NYC-Bike (Kaggle 2017)	(Deng et al. 2023)
	METR-LA (Jagadish et al. 2014)	(Lin et al. 2023; Wang et al. 2022c; Zhang et al. 2022c)
	PEMS-BAY (Li et al. 2018)	(Wang et al. 2022c; Zhang et al. 2022c)
Bioinformatics	MUTAG (Debnath et al. 1991)	(Zheng et al. 2024; Lin et al. 2022; Lin et al. 2021; Tan et al. 2022b; Sui et al. 2022a; Zhao et al. 2023a)
	Mutagenicity (Kazius et al. 2005)	(Wang et al. 2025a; Xiang et al. 2021)
	NCI1 (Wale et al. 2008)	(Zheng et al. 2024; Lin et al. 2022; Lin et al. 2021; Tan et al. 2022b; Sui et al. 2022a)
	PROTEINS (Borgwardt et al. 2005)	(Zheng et al. 2024; Sui et al. 2022a)
	Tox21 (Huang et al. 2016)	(Meng et al. 2023)
	ToxCast (EPA 2022)	(Meng et al. 2023)
	SIDER (Kuhn et al. 2016)	(Meng et al. 2023)
	MUV (Rohrer and Baumann 2009)	(Meng et al. 2023)
	Sachs (Sachs et al. 2005)	(Yang et al. 2023)
	Dream3 (Marbach et al. 2009)	(Amirinezhad et al. 2022)
Energy	UMass Smart (UMassTraceRepository 2017)	(Wang et al. 2025c)
	Energy (Candanedo 2017)	(Xu et al. 2020)

(Continues)

TABLE 1 | (Continued)

Category	Dataset	CLGNN studies
Image/video	MNIST (LeCun et al. 1998)	(Sui et al. 2022a)
	CIFAR-10 (Alex 2009)	(Sui et al. 2022a)
	CMNIST (Monti et al. 2017)	(Sui et al. 2022b)
	DAVIS (Pont-Tuset et al. 2017)	(Varga and Lrincz 2021)
	ADNI (ADNI 2017)	(Tang et al. 2023)
Medicine, Health	ABIDE (di Martino et al. 2014)	(Zheng et al. 2024)
	SRPBS (Tanaka et al. 2021)	(Zheng et al. 2024)
	REST-meta-MDD (Yan et al. 2019)	(Zheng et al. 2024)
	Infection (Faber et al. 2021)	(Zhao et al. 2023a)
	HCP (van Essen et al. 2013; Hodge et al. 2016)	(Wein et al. 2021)
	UR data (Wein et al. 2021)	(Wein et al. 2021)
	ASIA (Lauritzen and Spiegelhalter 1988)	(Zečević et al. 2021)
	Cancer (Korb 2003)	(Zečević et al. 2021)
	COVID-19 (Ritchie et al. 2020)	(Wang et al. 2022a; Ntemi et al. 2022)
	Wave1 (Chantala and Tabor 1999)	(Ma and Tresp 2021)
Psychology	SEED (Duan et al. 2013)	(Kong et al. 2022; Li et al. 2023a)
	SEED-IV (Zheng et al. 2018)	(Kong et al. 2022; Li et al. 2023a)
	DEAP (Koelstra et al. 2011)	(Zhang et al. 2022a)
Advertising/ sentiment	Criteo (Criteo 2014)	(Zhai et al. 2023)
	Avazu (Avazu 2015)	(Zhai et al. 2023)
	MovieLens-1 M (Grouplens 2003)	(Zhai et al. 2023)
	SST (Socher et al. 2013)	(Schnake et al. 2021)
Climate/geology	HadCRUT5 (Morice et al. 2021)	(Chen et al. 2022b)
	HadSST3 (Kennedy et al. 2011)	(Chen et al. 2022b)
	ERSSTv4/v3b (Smith et al. 2008)	(Chen et al. 2022b)
	ERA5 (Hersbach et al. 2020)	(Chen et al. 2022b)
	Berkeley-Earth	(Chen et al. 2022b)
	WADI (Ahmed et al. 2017)	(Wang et al. 2023c)
	Swat (Mathur and Tippenhauer 2016)	(Wang et al. 2023c)
Trade, finance	Nasdaq (Qin et al. 2017)	(Xu et al. 2020)
	UN Comtrade (2021)	(Monken et al. 2021)
	Exchangerate (Lai 2016)	(Xu et al. 2020)
Systems/Industrial	AIOps (Wang et al. 2023c)	(Wang et al. 2023c)
	TFF (Ruiz-Crcel et al. 2015)	(Wang et al. 2023d)
Causal datasets	Causal News Corpus (Tan et al. 2022a)	(Trust et al. 2022)
	ECPE (Xia and Ding 2019)	(Cao et al. 2022)
	EventStoryLine (Caselli and Vossen 2017)	(Phu and Nguyen 2021)
	Causal-TimeBank (Mirza 2014)	(Phu and Nguyen 2021)
	SemEval-2010 task 8 (Hendrickx et al. 2010)	(Gao et al. 2022; Yuan et al. 2023; Chen et al. 2022a)
	Altlex (Hidey and McKeown 2016)	(Chen et al. 2022a; Yuan et al. 2023)
	MEDCAUS (Moghimifar et al. 2020)	(Moghimifar et al. 2020)

(Continues)

TABLE 1 | (Continued)

Category	Dataset	CLGNN studies
General datasets	FB15K, WN18, WN18RR (Bordes et al. 2013)	(Wu and Zhou 2023)
	FB15K-237 (Toutanova et al. 2015)	(Wu and Zhou 2023)
	Motif (Ying et al. 2019)	(Sui et al. 2022b)
	Tree-Cycles (Ying et al. 2019)	(Lucic et al. 2022)
	Tree-Grids, BA-Shapes (Ying et al. 2019)	(Zhao et al. 2023a; Lucic et al. 2022)
	Graph-SST2/SST5 (Yuan et al. 2022)	(Zhao et al. 2023a; Gao et al. 2023)
	Graph-Twitter (Yuan et al. 2022)	(Gao et al. 2023)
	Earthquake (Korb 2003)	(Zečević et al. 2021)
	Visual genome (Krishna et al. 2017)	(Wang et al. 2022b; Xiang et al. 2021)
	MOL-BACE (Wu et al. 2018)	(Gao et al. 2023)
	MOL-BBBP (Wu et al. 2018)	(Sui et al. 2022b; Gao et al. 2023)
	MOL-HIV (Wu et al. 2018)	(Sui et al. 2022b; Ma et al. 2022)

global information with the graph Fourier transform, wherein the graph signals are projected in the eigenvector space.

Recent studies have shown that while emotion recognition traditionally relies on methods like CNNs (Huang et al. 2023), there is a growing interest in utilizing GNNs, particularly GCN and GAT, for causal learning in the spectral domain (Wang et al. 2023d; Kong et al. 2022; Li et al. 2023a). Kong et al. (2022) employed GCNs for causal emotion recognition from multichannel EEG signals. Each node was assigned as a channel, and Granger causality (GC) was calculated between each node for computing the adjacency matrix. The matrix and EEG information moved through convolution layers with Chebyshev polynomials (Hammond et al. 2011), and subsequently to a depth-wise separable convolution layer for extracting discriminative features for the final classification step.

Li et al. (2023a) employed a graph fusion strategy with GCN for emotion recognition. The approach encapsulated topological, functional, and causal features. A study on EEG and peripheral physiological signals by Zhang et al. (2022a) used GCN with GC analysis for emotion recognition. The differential entropy (DE) feature of the EEG signals formed the nodes of the graph, from which a matrix was computed to determine the edges with high causal values. GC-F-GCN (Zhang et al. 2022e), a GCN-based approach for emotion recognition, employed multi-frequency band EEG feature extraction. GC was used for computing GC matrices between EEG signals at each frequency band, with each converted to asymmetric binary GC matrices. These and DE features served as adjacency matrices and node values, respectively. The graph information from the EEG signals at different frequency bands for corresponding nodes was integrated based on GC-GCN features.

Industrial systems constitute large amounts of spectral information, making them another sector where causal learning is beneficial. Wang et al. (2023d) proposed a Causal Trivial Attention GNN (CTA-GNN) that used an attention mechanism and disentanglement-based causal learning with a backdoor criterion for fault diagnosis. The model estimated soft masks for obtaining node and edge representations. This strategy succeeded in weakening confounding effects by avoiding shortcut features. A summary of works in Spectral CLGNN is given in Table 2.

3.1.2 | Spatial

In spatial data, observations are spatially associated with each other. GCN is the most used type of GNN employed in the spatial domain, where the convolution operation is carried out on each node and the weights are shared across all locations. The node features are aggregated in layers by a permutation-invariant function, and the information is amassed within a localized boundary (Bo et al. 2023). Ma and Tresp (2021) studied causal effects under network interference and employed the Hilbert-Schmidt Independence Criterion (HSIC) for a feature space dependence test. HSIC is a statistical tool for measuring the dependency between two random variables by comparing the distance between their joint and marginal distributions in a Hilbert space, which is a vector space with an inner product. Zečević et al. (2021) also adopted an interventional approach for causal density and causal effect estimation with autoencoding. The GNN and VGAE models detected interventional change, including successive changes. Li et al. (2023c) employed multiple instance learning with GNNs for capturing spatial proximity and the similarity of features for cancer detection using the Contrastive Mechanism to segregate noncausal features.

GNN is effective in capturing spatially adjacent nodes in a graph and was efficiently implemented by Varga and Lrincz (2021) with label propagation for interactive video object segmentation using a Watershed algorithm. The graph was created from super pixel segments with dimensionality reduction and causality estimated based on optical flows. Causal-GAT with disentangled causal attention (DC-Attention) was proposed by Liu et al. (2023) for fault detection by incorporating disentangled representations with GAT. The causal graph was structured using monitoring

TABLE 2 | Spectral causal learning graph neural networks.

Model	Domain	Problem	Key functionality
CGCNN (Kong et al. 2022)	Psychology	Emotion classification	Emotion recognition from EEG signals with GC
FGCN (Li et al. 2023a)	Psychology	Emotion classification	EEG-based emotion recognition with GCN, graph fusion
GC-GCNN (Zhang et al. 2022a)	Psychology	Emotion classification	EEG-based emotion recognition using GCN and GC
GC-F-GCN (Zhang et al. 2022e)	Psychology	Emotion recognition	Emotion recognition & GCN-based EEG emotion recognition with GC
CTA-GNN (Wang et al. 2023d)	Industrial	Fault diagnosis	GAT with causal learning using disentanglement

variables, and DC-Attention was used to generate node representations from causal relations. Tang et al. (2023) proposed a causality-aware GCN framework for rigidity assessment in neuro-diseases. The GCN was built on node, structure, and representation levels and incorporated causal feature selection. A noncausal perturbation strategy with an invariance constraint ascertained the validity of the model under varying distributions. An autoregressive (AR) GCN framework with LSTM called AR–GCN–LSTM was developed by Ntemi et al. (2022) for epidemic case prediction. A GC adjacency matrix formed the input to GCN to extract causal information of locations that influenced each other's case numbers. AR modeling in AR-GCN-LSTM assisted in capturing the linear dependencies from time series, and feature representation was extracted with two LSTM layers. A summary of works related to CLGNNs in the spatial domain is given in Table 3.

3.1.3 | Temporal

Temporal data consists of observations related to time or date instances. Wang et al. (2023b) proposed an LSTM-based GNN model referred to as RLF-MGNN for load forecasting. Multiple temporal correlations of energy usage across households were captured along with transfer entropy to build a causality graph. The collective influence of household energy usage was employed for forecasting, with only positive correlations considered in building graphs. A transfer entropy-based causal analysis (TEGNN) was used in an experiment by Xu et al. (2020) for multivariate time series (MTS) forecasting, where causal prior information was captured using the transfer entropy matrix. Time series forecasting using GraphSAGE with residual neural networks (ResGraphNet) was studied by Chen et al. (2022b) for the causal prediction of global monthly mean temperature. Using the embedded ResNet, the connection weights of each node were used to build a learning framework. The training time for ResGraphNet was much higher than that of a GraphSAGE base model. Wang et al. (2023c) studied hierarchical GNNs for temporal root cause analysis in systems monitoring. The model integrated topological and individual causal discoveries by incorporating intra-level and inter-level system relationships while disregarding system logs. Ding et al. (2022) proposed a GCN-based model called CI-LightGCN for retraining recommender systems. Causal incremental graph convolution (IGC) and colliding effect distillation were employed, with IGC performing aggregation of only new neighbors. GCN was also employed for MTS imputation by Liu et al. (2022b), who proposed CGCNImp that incorporated GC alongside correlation and temporal dependencies. Attention mechanism and total variation reconstruction were used for retrieving latent temporal information.

A GAT framework called TC-GAT was proposed by Yuan et al. (2023) for temporal causality discovery from text. TC-GAT integrated temporal aspects with causality using graph attention by computing joint weights of temporal and causal features. A causal knowledge graph was used to obtain an adjacency matrix for causality extraction. GAT was also proposed by Li et al. (2022a) for industrial MTS forecasting, incorporating GC. Nonlinear interaction of node features was performed with parallel GRU encoders in the graph neighborhood space, and then the encoder hidden states were aggregated with an attention mechanism. In this approach, large-scale graphs introduce nonlinear interactive patterns that increase complexity. GGNNs are another causal learning approach proposed by Zeyu et al. (2023), who employed a framework called CauTailReS session recommendation with counterfactual reasoning. Popularity bias was eliminated using deconfounded training with causal intervention and do-calculus. GGNN was used to generate node embeddings, and both user interest and consistency embeddings were captured for learning causality. GNNs were also employed with De Bruijn graphs (Fredricksen 1992) using the DBGNN framework to learn causal topological patterns in dynamic graphs by Qarkaxhija et al. (2022). De Bruijn graphs were used to incorporate non-Markovian characteristics of causal walks, leading to a causality-aware node classification process. A summary of works related to CLGNNs in the temporal domain is given in Table 4.

3.1.4 | Mixed

In this section, we discuss causal graphs that handle mixed resolutions such as spatial–temporal data. In spatiotemporal graphs, node connections are formed as a function of time and space, and are applicable in instances where both time and location/space are of value in causal learning. Spatiotemporal CLGNNs have been beneficial for epidemic studies (Wang et al. 2022a), economic (Monken et al. 2021) and traffic (Gu and Deng 2022; Li et al. 2021; Zhao et al. 2023b; Deng et al. 2023) forecasting, and so forth. Wang et al. (2022a) proposed ADGNN, an attention-based dynamic GNN with causal learning for forecasting epidemic cases. The study included causal features and constraints with a minimal use of parameters. Causality also finds application in environmental forecasting, as demonstrated by Jiang

TABLE 3 | Spatial causal learning graph neural networks.

Model	Domain	Problem	Key functionality
GNN, HSIC (Ma and Tresp 2021)	Multiple	Network interference	Causal estimation with intervention policy optimization
GNN, VGAE (Zečević et al. 2021)	Multiple	Causal Inference	Estimate causal density and effect with AE intervention
CGNN (Li et al. 2023c)	Medicine	Cancer detection	Tumor proximity, similarity with multiple-instance learning
GCN (Tang et al. 2023)	Medicine	Neurology assessment	Causality-aware GCN for rigidity assessment in neuro-patients
AR-GCN-LSTM (Ntemi et al. 2022)	Medicine	Case prediction	Autoregressive GCN with LSTM for epidemic case prediction
GNN (Varga and Lrincz 2021)	Video	Segmentation	GNN for interactive video object segmentation
Causal-GAT (Liu et al. 2023)	Industry	Fault detection	Causal disentangled GAT for fault detection

et al. (2024), who developed CSTGNN to predict storm surges at various locations. In CSTGNN, temporal dependencies were modeled using GRU based on storm surge time series, while GNN was employed to capture spatial dependencies among observation centers. The Liang–Kleeman information flow theory facilitated causal inference for establishing causal connections between these centers.

Many traffic flow prediction studies incorporated attention mechanisms with GNN for causal learning, with both Gu and Deng (2022) and Zhao et al. (2023b) developing frameworks for capturing causal traffic dynamics using attention-based GNNs. The former proposed STAGCN, which considered both global and local traffic dynamics, assuming no interaction between the static and dynamic graph units. The latter work proposed the STCGAT framework, which captured overall spatiotemporal dependencies using local and global causal convolutions. Li et al. (2021) developed TrGNN, a trajectory-based model for predicting traffic flows using graphs that encapsulated temporal dependencies using attention mechanisms. In a similar domain, Deng et al. (2023) proposed STNSCM for building a causal model with the Frontdoor criterion for handling confounding bias in bike flow prediction. The framework integrated counterfactual reasoning and demonstrated effective long-term predictions. Lin et al. (2023) proposed Dynamic Causal GCN (DCGCN) for modeling spatiotemporal dependencies in the traffic prediction domain. Time-varying dynamic causal graphs were incorporated for constructing superior spatiotemporal topology representations. A similar approach was proposed by Wang et al. (2022c) for spatiotemporal forecasting with dynamic causality analysis. The causality adjacency matrix was fed to GCN for extracting dynamic correlations in road networks. GCN was also used for spatiotemporal cellular traffic forecasting by Zhang et al. (2022b) to develop a framework integrating GCN, referred to as DIC-ST. This framework was based on decomposition and integration with causal structure learning. An empirical mode decomposition (EMD) method was adopted for multi-scale decomposition, with the integration of various time series analyses of three different components. Following EMD, integration was performed based on KNN to subsequently employ the integration of prediction results. ST-CGI (Zhang et al. 2022c) employed GC with an autoregressive process to develop an interpretable traffic prediction model. A dilated causal convolution network was employed for encoding temporal information, with causal relationships extracted from the embeddings using GNNs. Xu et al. (2023) proposed TraffNet for causal learning of traffic volumes through path embedding, route learning, and road segment embedding. Bi-GRU was employed for path embedding, following which a meta-path-based GAT was used for route learning to determine the origin–destination demands for each route. Road segment embedding was obtained by aggregating all path embeddings using a position-aware message passing technique for capturing the road segment position. Finally, these embeddings were fed to a temporal GRU module for forecasting.

GAT is also useful in causal analysis of social and purchase networks that demand exploring both spatial and temporal elements. Disentangled Intervention-based Dynamic (DIDA) GAT, proposed by Zhang et al. (2022d), handles distribution shifts in dynamic graphs, using a disentangled attention-enabled causal inference framework to capture invariant patterns in graphs. For economic forecasting, Monken et al. (2021) proposed S-GC-LSTM for trade flow causal analysis of outlier events using Stateless Graph Convolutional LSTM to capture events that affect trading. A counterfactual study was performed, mainly focusing on highly disruptive events across geographical areas.

In the medical field, causality has been utilized for blood pressure (BP) estimation, with Liu et al. (2024b) employing a spatiotemporal GNN (STGNN) for cuffless continuous BP estimation with their model CiGNN. They first established a causal graph between BP and wearable features, followed by STGNN to learn from this causal graph. Spatiotemporal networks are also relevant in neuroscience, where brain functions need to be analyzed anatomically over time. The DCRNN model proposed by Wein et al. (2021) used GNNs with convolutional and recurrent networks for inference of causal relations in brain networks. Causal structural–functional interactions in brain regions were identified with the diffusion-convolution approach. Zheng et al. (2024) proposed CI-GNN, a brain network model for psychiatric

TABLE 4 | Temporal causal learning graph neural networks.

Model	Domain	Problem	Key functionality
RLF-MGNN (Wang et al. 2023b)	Energy	Load forecast	Forecasting with LSTM and transfer entropy
TEGNN (Xu et al. 2020)	Multiple	MTS forecasting	MTS forecasting with transfer entropy-based causality
Hierarchical GNN (Wang et al. 2023c)	Systems	Root cause analysis	Systems causal analysis for identifying root cause of system issues
CI-LightGCN (Ding et al. 2022)	Social	System retraining	Re-train recommender with causal graph convolution
ResGraphNet (Chen et al. 2022b)	Climate	TS forecasting	Global mean temperature prediction using GraphSAGE
CGCNImp (Liu et al. 2022b)	Multiple	TS imputation	Inputting with GCN, correlation, and temporal dependency
TC-GATN (Li et al. 2022a)	Industry	MTS forecasting	MTS forecasting with GAT, GC, and GRU
TC-GAT (Yuan et al. 2023)	Text	Causality	Causality with GAT using temporal, causal relations
DBGNN (Qarkaxhija et al. 2022)	Multiple	Causality	De Bruijn graphs for causal learning in dynamic graphs
CauTailReS (Zeyu et al. 2023)	Multiple	Recommendation	GGNN-based session recommendation with counterfactuals

diagnosis, leveraging Granger causality. The model identified the most influential subgraph that demonstrated causal connections to either depressive or healthy patients through disentangled subgraph-level representation learning.

Spatial and time factors also play a role in predicting future movements in visual intelligence, as researched by Li et al. (2020). The authors proposed a V-CDN model with GNN as a spatial encoder, supplemented with counterfactual predictions. Among other domains, a GCN-based model employed by Cummings and Nassar (2020) for predicting papers with high future citation counts used percentile thresholds for ranking. A summary of works related to CLGNNs in the mixed domain is given in Table 5.

3.2 | Causal Learning GNNs Based on Learning Methods

In this section, we explore the utilization of CLGNNs across five learning domains: Representation, Transfer, Meta, Adversarial, and Reinforcement learning, summarizing related works in these different learning paradigms in Table 6.

3.2.1 | Representation Learning

Graph representation learning involves learning a model from graph-structured data by building features or embeddings that represent the structure. GraphSAGE is the most commonly used GNN approach for inductive representation learning and was integrated with causal inference to develop a framework called C-GraphSAGE (Zhang et al. 2022f) for classification. C-GraphSAGE incorporated causal sampling to control the influence of perturbations. C-GraphSAGE performed well under perturbation conditions, although GAT was superior with no perturbation. In the work by Trust et al. (2022), causal relationships between events in social-political news were extracted with GGNNs. The authors modeled event causality identification with contextualized language representations by building a graph representation of all documents in the Causal News Corpus dataset, along with their dependencies. GGNN layers were stacked to form an encoder, with information aggregated based on the edge type and direction, and GRU applied for node embedding updates.

Sui et al. (2022a) proposed Causal Attention Learning (CAL) with mitigation of confounding effects using softmask estimation from attention scores. The graph was decomposed into causal and trivial attended graphs with two GNN layers. The authors proposed the disentanglement of causal and trivial features, with GNNs filtering shortcut patterns for capturing causal features. An attention-based framework called Causality and Correlation Graph Modeling for Effective and Explainable Session-based Recommendation (CGSR) was proposed by Wu et al. (2023) for session-based recommendation with causality and correlation graph modeling. CGSR has four components, namely graph construction, item representation learner, session representation learner, and recommendation score generator. Effect graph, cause graph, and correlation graph were constructed, following which a weighted GAT was used for item representation learning on each of these graphs. The session representation learner forms a session representation by aggregating learned item representations in the session sequence using an attention layer. Cao et al. (2022) also proposed a GAT-based method for causality detection referred to as GAT-ECP-CD for textual emotion-cause pair causality. The study used BiLSTM to obtain a semantic representation. The sentence vector was fed to GAT for capturing dependencies

TABLE 5 | Mixed resolution causal learning graph neural networks.

Model	Domain	Problem	Key functionality
ADGNN (Wang et al. 2022a)	Health	Epidemic forecast	Attention-based dynamic GNN to predict COVID-19 cases
S-GC-LSTM (Monken et al. 2021)	Trade	Economic forecast	Trade flow causal analysis of outlier events with Graph Convolutional LSTM
DCRNN (Wein et al. 2021)	Medicine	Neuroscience	Causality of the brain with convolutional and recurrent networks
DIDA (Zhang et al. 2022d)	Multiple	Link prediction	Disentangled attention for invariant graph pattern extraction
V-CDN (Li et al. 2020)	Visual Intelligence	Prediction	Predict future movements with GNN as a spatial encoder
STAGCN (Gu and Deng 2022)	Urban Intelligence	Traffic forecasting	Traffic trends are captured with an attention-based GCN
TrGNN (Li et al. 2021)	Urban Intelligence	Traffic prediction	Trajectory-based model for predicting traffic flows with GAT
STCGAT (Zhao et al. 2023b)	Urban Intelligence	Traffic prediction	Causality-based GAT with causal convolutions for prediction
STNSCM (Deng et al. 2023)	Urban Intelligence	Bike prediction	Causality with the Frontdoor criterion to handle confounding
DCGCN (Lin et al. 2023)	Urban intelligence	Traffic prediction	Dynamic causal GCN to model spatiotemporal dependencies
GCN (Wang et al. 2022c)	Urban intelligence	Traffic prediction	Dynamic causal GCN for spatiotemporal forecasting
DIC-ST (Zhang et al. 2022b)	Urban intelligence	Traffic prediction	Spatiotemporal forecasting with decomposition, integration
ST-CGI (Zhang et al. 2022c)	Urban intelligence	Traffic prediction	Causal inference, autoregression for interpretable prediction
TraffNet (Xu et al. 2023)	Urban intelligence	Traffic prediction	Causality with route learning, path, and segment embeddings
GCN (Cummings and Nassar 2020)	Citation	Classification	Predict academic papers with high citation counts in the future
CSTGNN (Jiang et al. 2024)	Environment	Storm forecast	Predict storm surges at multiple locations with GRU and GNN
CI-GNN (Zheng et al. 2024)	Medicine	Psychiatric diagnosis	Psychiatric diagnosis from brain networks using GC
CiGNN (Liu et al. 2024b)	Medicine	BP estimation	Cuffless continuous BP estimation with STGNN

between clauses, followed by a joint prediction layer with three stages of predictions.

Zhai et al. (2023) developed a structured causality-based representation learning approach called GraphFwFM, incorporating a Field-weighted Factorization Machines (FWFM) mechanism for CTR prediction, with the GNN-based graph representation learning consolidating feature graphs, user graphs, and ad graphs. Gao et al. (2023) developed a Robust Causal Graph Representation Learning (RCGRL) framework to learn representations against confounding effects for improving causality prediction and generalization performance. A variant of the IV approach was used in the framework, where conditional moment restrictions for confounding elimination were transferred to unconditional restrictions. Causal GCN Inference (CGI) was proposed by Feng et al. (2021) for studying the causal effects of the local structure of a node when the labels of neighboring nodes vary in GCN. As a first step, intervention was done on predictions by blocking the graph structure, followed by a comparison with original predictions for determining the causal effects of the local structure. A GCN-based network deconfounder was proposed by Guo et al. (2020) for learning representations

TABLE 6 | Causal learning graph neural networks learning methods.

Type	Model	Domain	Key functionality
Representation	GAT, CAL (Sui, Wang, Wu, Lin et al. 2022)	Multiple	GAT with mitigation of confounding effects for graph classification
	GraphFwFM (Zhai et al. 2023)	Advertising	Causality-based CTR prediction using GraphSAGE with integrated representation learning
	C-GraphSAGE (Zhang et al. 2022f)	Citation	GraphSAGE with causal sampling for reducing perturbation influence
	GGNN (Trust et al. 2022)	ECI	Causal event detection to capture semantic and syntactic information
	CGSR (Wu et al. 2023)	Recommend	Session-based recommendation with causality, correlation graph modeling
	GAT-ECP-CD (Cao et al. 2022)	Text	Textual emotion-cause pair causal relationship detection with GAT
	CGI (Feng et al. 2021)	Citation	Causal GCN Inference model studying the causal effects of node local structure
	GCN (Guo et al. 2020)	Social Network	Network deconfounder for learning representations to uncover hidden confounders
	DECI (Phu and Nguyen 2021)	ECI	Document-level event causality identification with GCN
	GCN (Gao et al. 2022)	ECI	Extract event causality with GCN using textual, knowledge channels
	RPA-GCN (Chen et al. 2022a)	Text	Head-to-tail entity annotation approach for text with GAT and GCN
	CPRL (Yang et al. 2022)	Text	Causal representation learning using BERT, GCN, and attention mechanism
	INE-TARNet (Adhikari and Zheleva 2023)	General	Causality in the presence of heterogeneous peer influence with GNN estimator
	RCGRL (Gao et al. 2023)	General	Learn representations against confounding effects with IV
Transfer learning	DCRNN (Wein et al. 2021)	Medicine	Transfer learning enabled inference of causal relations in brain networks
	CogTrans (Wu and Zhou 2023)	KGR	Attention-based cognitive TL for knowledge graph reasoning
Meta learning	V-CDN (Li et al. 2020)	Visual	Long-term predictions from video sequences based on causal relationships
	MTA, GNN (Meng et al. 2023)	Drug discovery	Meta learning with motif-based task augmentation (MTA) technique
Adversarial learning	IDEA (Tao et al. 2024)	Multiple	Defense against graph attacks using invariance objectives from causal features
	AdvCA (Sui et al. 2022b)	Multiple	Graph augmentation to address covariate shift in OOD generalization
	NetEst (Jiang and Sun 2022)	Social Network	Causal inference using GNN for learning representations of confounders
	GCN (Moghimifar et al. 2020)	Text	Adaptive causality identification and localisation with GCN
	CAGCN (Lee and Han 2023)	Citation	Causal attention GCN with node and neighbor attention
Reinforcement learning	GNN (Amirinezhad et al. 2022)	General	GNN with Q-iteration to extract causality with minimal intervention
	GARL (Yang et al. 2023)	General	RL for causal discovery with graph attention
	STREAMS (Sheth et al. 2023)	Hydrology	RL, LSTM, and GCN for spatiotemporal causal discovery
	GRAIN (Xiao et al. 2025)	Network traffic	RL, GNN for causality discovery for multi-step attack scenario reconstruction
	SAC-CAI-EGCN (He et al. 2024)	Network topology	Actor-Critic, GCN with causality for efficient SDN routing

to uncover hidden confounders from network information. The representation learning function was parameterized using GCN towards learning causal effects. These authors customized the BlogCatalog dataset studied by Li et al. (2019) through a synthesis of outcomes and treatments.

A GCN-based framework called DECI was proposed by Phu and Nguyen (2021) for document-level event causality identification (ECI) using interaction graphs. The interaction graph nodes were formed from all the words, event mentions, and entity mentions in a document. Node connections were formed from discourse-based, syntax-based, and semantic-based information. The interaction graphs and representation vectors were regularized for improved representation learning. Gao et al. (2022) also proposed an approach for event causality extraction using GCN. The model was enhanced using a dual-channel approach with textual (TEC) and knowledge (KEC) enhancement channels. TEC learns significant intra-event features, and KEC uses GCN for assimilating external causality transition knowledge. GCN and GAT were employed by Chen et al. (2022a) developed relation position and attention-GCN (RPA-GCM) for establishing complex causal relations in text by marking entity boundaries. Relation features were extracted by integrating an attention network with a dependency tree. The interaction information between entities and relations was captured using a bi-directional GCN. GCN was employed for developing Causal Pattern Representation Learning (CPRL) (Yang et al. 2022) for extracting causality from text using an entity set of risk factors for various diseases. Input encoding was performed by BERT, followed by GCN for encoding dependencies, and causal features were extracted with an attention mechanism for constructing causal pattern representation. Adhikari and Zheleva (2023) devised INE-TARNet for estimating causal effects in the presence of heterogeneous peer influence using a GNN-based estimator. Arbitrary assumptions regarding network structure, interference conditions, and causal dependence were captured by the model. These assumptions were encoded by the Network Structural Causal Model (NSCM), which generated the Network Abstract Ground Graph (NAGG) for reasoning about treatment effects under arbitrary network interventions. Further, NSCM and NAGG were employed for extracting individual network effects (INE) using GNN.

3.2.2 | Transfer Learning

Transfer learning (TL) is a technique in which a machine learning model trained on a specific task is re-purposed to a related problem in a similar domain. Research on transfer learning with GNN is limited, and experiments so far demonstrate the transfer process to be plausible when source and target graphs are alike. This drawback, referred to as negative transfer, impacts target performance (Kooverjee et al. 2022).

A diffusion convolution recurrent neural network (DCRNN) based on GNN was employed by Wein et al. (2021) to infer causal relations in brain networks. Their model was pretrained on large volumes of fMRI sessions, ensuring better results for small datasets on account of diffusion-convolution used on massive graphs, leading to memory errors. Moreover, the authors experimented on the model using TL, with the results indicating a considerable reduction in error rates. Wu and Zhou et al. (2023) proposed CogTrans for cognitive transfer learning-based Knowledge Graph Reasoning (KGR) using GCN, where transferring was based on hierarchical structure similarity. The approach retained causal structure, with a hierarchical random walk being employed to acquire entity and relation characteristics for pretraining in the knowledge graph. A multi-head self-attention mechanism was used for the reasoning phase, in which an attention score is aggregated from entity and relation layers for encoding.

3.2.3 | Meta Learning

In machine learning, meta learning (ML) serves as a solution for data scarcity by exploiting previously learned experiences towards learning an algorithm that generalizes across various tasks. Model-Agnostic Meta-Learning (MAML) (Finn et al. 2017) is the most commonly adopted approach for training GNNs (Mandal et al. 2022) and trains model parameters for accelerated learning with minimal gradient updates. Meng et al. (2023) proposed ML with a motif-based task augmentation (MTA) technique for molecule property prediction, targeted at addressing the few-shot learning challenge associated with this task. Causal substructure was studied, with experiments conducted on molecule datasets. The model recovers the most relevant motifs from a previously defined motif vocabulary for label generation, for the purpose of task augmentation. Once augmented, classification was performed based on Euclidean distances. Li et al. (2020) experimented on custom video datasets to develop an ML framework called V-CDN for predicting future visuals. In V-CDN, Perception and Inference modules draw a graph inference, and the dynamics module is used to predict future visuals from the inferred graph. Inference was extracted from causal relationships, and the graph distribution was inferred based on the image representation for prediction.

3.2.4 | Adversarial Learning

Adversarial Learning (AL) is a machine learning technique that examines and devises defenses against adversarial attacks on models. Based on adversarial capability, attacks may be of the poisoning or evasion kind. Poisoning attacks involve training-time attacks that affect networks such as GNN through data poisoning. Evasion attacks are test-time attacks that add poisoned data at test time. Defense models for graph data serve to stabilize the model in the course of adversarial events.

Tao et al. (2024) proposed a framework called Invariant Causal Defense Method Against Adversarial Attacks (IDEA) to develop defense methods against adversarial attacks on GNNs. The defense was devised by developing invariant objectives from causal features and utilized a domain partitioner to manage mixed domains. IDEA handled evasion and poisoning attacks, with invariant defending plausible only with linear causality. A graph augmentation strategy called AdvCA (Adversarial Causal Augmentation) proposed by Sui et al. (2022b) addresses the problem of covariate shift in out-of-distribution (OOD) generalization. AdvCA performed adversarial data augmentation while preserving causal features across various environments. The covariate shift was devised based on the domain using graph size and type, or color. Jiang and Sun (2022) proposed Networked Causal Effects Estimation (NetEst) for inferring causality in network settings. The model used GNN for learning representations for confounders. Additionally, two AL modules were introduced to allow mismatched distributions to follow uniform distributions based on embeddings from confounders. NetEst presented a framework for addressing the issue of distribution gaps through representations. Moghimifar et al. (2020) proposed a GCN-based domain adaptive causality identification and localization with AL for extracting causal relationships in text. Distribution shifts were minimized using a gradient reversal approach. AL was applied to the training domain discriminator for differentiating the source and target domains. Moreover, feature representation was trained to overcome the domain discriminator. CAGCN (Lee and Han 2023) employed causal attention GCN with node and neighbor attention for enhanced robustness. AL was performed based on nettack (graph structure and attributes are attacked), metattack (global attack using meta learning), and random attack (random addition of edge to graph).

3.2.5 | Reinforcement Learning

Causal reinforcement learning explores causal mechanisms for the agent learning process towards improved decision making. The agent acts in the environment, and the effects of actions are observed, followed by counterfactual analysis. A general structure of the causal reinforcement learning process is illustrated in Figure 8. The environment and the agent are connected through the causal model and causal graph. Here, the action may be observational, interventional, or counterfactual.

Amirinezhad et al. (2022) researched learning causal structures from interventional data using GNN and RL. An active learning approach was employed, where the intervention results were used to derive causal relationships for formulating causal structures with minimal interventions. Training was performed using a Q-iteration algorithm, and the completed partially DAG (CPDAG) for each sub-network was derived from the causal network to serve as an input to another network, whose output in turn was used for intervention. A graph attention RL (GARL) framework was proposed by Yang et al. (2023) for causal discovery, which employed a GAT for embedding structural information and prior knowledge along with RL to generate variable orderings. RL was also employed by Sheth et al. (2023) to develop Spatio Temporal REinforcement learning and cAusal discovery for Streamflow prediction (STREAMS) to infer causality in hydrological models. The framework employed a spatiotemporal autoencoder consisting of LSTM and GCN, alongside RL to infer the causal structure of the process. Reinforcement learning, combined with causality discovery and GNN, was employed to develop GRAIN (Xiao et al. 2025), which reconstructs multistep attack scenarios without relying on external expertise, significantly enhancing accuracy and efficiency in identifying complex attacks. In the field of QoS-aware routing for Software-Defined Networking (SDN), He et al. (2024) proposed SAC-CAI-EGCN, a reinforcement learning method that includes an actor network, two critic networks, and two target critic networks, which quantifies the causal impact of agent actions on the environment and leverages GNN to embed node and link features, significantly improving performance in metrics like packet loss, latency, and throughput compared to baseline methods.

3.3 | Causal Learning GNNs Based on Explainability

Explainability of a machine learning model refers to the extent to which a model's output is meaningful. The terms explainability and interpretability are commonly interchanged, though interpretability focuses on the cause and effect of a model. With reference to GNN, interpretability is inherently designed in the GNN architecture, and explanations are post hoc. This review details explainability with reference to methods adopted for causal GNNs. The explainability techniques applied to CLGNNs are summarized in Table 7.

The standard methods used for GNN explanations are surrogate, decomposition, perturbation, generative, and gradient methods (Yuan et al. 2022). Surrogate models for graphs are more challenging to implement, owing to their topological design. Though surrogate approaches such as GraphLime and PGM-Explainer are used as graph explainers, ambiguity exists in defining the neighbors in the input graph. Similarly, gradient methods have been proven to be less advanced in GNN explanation compared to other explainers (Lin et al. 2021; Xiang et al. 2021). Hence, with a view to causal application, only decomposition, perturbation, and generative approaches are discussed in this study.

3.3.1 | Decomposition Methods

Decomposition methods decompose the original network into various elements to determine the most significant features, constituting the importance scores. Using score distribution rules, prediction scores are distributed from the output layer with a backpropagation approach towards the input layer. The importance scores for node features account for edge importance, as well as walk importance. Layer-wise relevance propagation (LRP) is a decomposition method that follows the law of total probability. Schnake et al. (2021) extended LRP to develop a technique adapted to GNNs referred to as GNN-LRP, which forms the score decomposition rule with high-order Taylor decomposition. A key difference from LRP is that GNN-LRP distributes scores to graph walks and not nodes or edges (Yuan et al. 2022). Holzinger et al. (2021) presented GNN-LRP with GCN for developing an explainable multi-modal causability framework for information fusion across various feature spaces. A joint multi-modal representation was proposed to be computed in a decentralized manner towards model scalability and security. This approach was proposed as an exploration-based, explainable method with counterfactual graphs for building an automated decision model.

3.3.2 | Perturbation Methods

Perturbation methods investigate variations in output on changing inputs to a model, whereby the changes in output specify the input elements relevant for inference. GNNExplainer, PGExplainer, SubgraphX, Causal Screening, and CF-GNNExplainer are a few

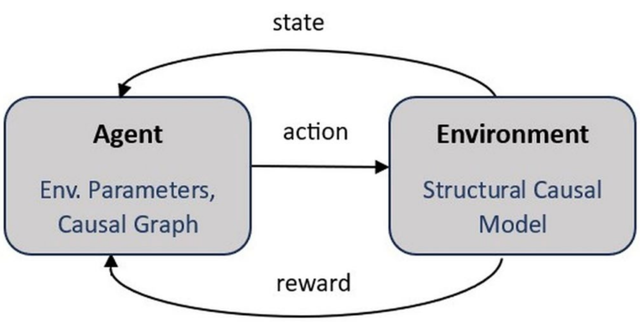


FIGURE 8 | Causal reinforcement learning.

of the perturbation-based explainer methods used for explaining graph models, with the latter two specifically designed for causality. CF-GNNExplainer, a counterfactual explainer for GNNs, proposed by Lucic et al. (2022), used minimal perturbation to the input graph for varying predictions. The framework incorporated edge deletions for generating counterfactual explanations. Xiang et al. (2021) developed causal screening for explainable GNN for classification by selecting a graph feature with large causal attribution. The approach involved identifying predominant edges for explainable classification. CI-GNN, another explainable GNN with a focus on causality, was developed by Zheng et al. (2024) to extract functionality connections in brain networks using GraphVAE to learn disentangled latent representations.

3.3.3 | Generative Methods

In the generative approach, graph generation methods are employed to provide explanations by devising a search strategy to obtain the most explanatory subgraph. GNN generative explainable models include approaches such as mask generation, VGAE, GAN, Diffusion, and RL (Cheng et al. 2023a). CLEAR (Ma et al. 2022) is a VGAE approach for optimizing graph data for generalization on unseen graphs with counterfactual explanation generation. A GCN was used as the encoder, and the decoder was an MLP, which also learned the mean and covariance of the latent variables' prior distribution. The input graph and counterfactuals were matched using graph matching. OrphicX (Lin et al. 2022) is also a VGAE approach that provides explanations from latent causal factors using an information flow mechanism and backdoor adjustment. In the model, linear independence of explained features was not assumed. Gem, another VGAE approach proposed by Lin et al. (2021), provided explanations with GC. A rapid explanation process with no requirement for prior knowledge of GNN structure was employed. The framework received the input graph into the explainer to output a compact explanation graph without prior structural knowledge of the GNN. Tan et al. (2022b) developed a framework called CF2 for explaining node classification with counterfactual and factual reasoning, with evaluation done based on the probability of necessity and sufficiency. RL-based GNN explanation called RC-Explainer was proposed by Wang et al. (2022b), where the explainable model was defined in a sequential decision-making process that uncovered edge attributions for causal screening. In the framework, edge dependencies were considered for deriving causal effects.

Similar to Gem, PGExplainer also provided local and global views for explanations with parameterized networks. A distinction from the Gem framework is that PGExplainer, a mask generation method, learns MLP from the target GNN's node embeddings (Lin et al. 2021). PGExplainer learned discrete edge masks by training a parameterized mask predictor. Edge embeddings were obtained from an input graph by combining node embeddings. These edge embeddings are used for predictions, following which the discrete masks were sampled with reparameterization. The mask predictor training was performed by assimilating the original and model predictions. GNNExplainer, which is also a mask generation method, learned both edge and node soft masks for explaining predictions. Similar to PGExplainer, the masks are combined with the original graph, and the original and new predictions are assimilated. Zhao et al. (2023a) extended PGExplainer and GNNEXplainer with a three-layer GCN by aligning internal embeddings of the raw graph and the explainable subgraph. A three-layer GCN was employed in these two approaches for evaluating their faithfulness and consistency in comparison to other explanatory approaches such as GRAD, Gem, and RG-Explainer. The proposed approach, mainly PGExplainer, demonstrated commendable performance compared to other explainers, along with the base PGExplainer and GNNEXplainer. Metsch et al. (2024) developed CLARUS, an explainable platform for visualizing patient-specific networks and gene interactions. The platform aimed to develop an interactive, explainable GNN model by manually posing counterfactual questions and analyzing their effects on GNN predictions.

Explainability is also important in text and temporal data analysis to improve the transparency of the factors influencing predictions. Jin et al. (2022) proposed a Causality-Pruned semantic dependency forest Graph Convolutional Network (CP-GCN) for relation extraction from text using causality-pruned dependency forests. Semantic and syntactic information was introduced to the dependency tree for constructing forests. A task-specific causal explainer was then trained for pruning dependency forests, which was subsequently fed to GCN for learning representations. The interpretability of predictive models constructed from temporal data is enhanced by explainable temporal graphs. Chen and Ying (2024) proposed a framework called TempME to improve the interpretability of temporal GNNs. This model uncovers essential temporal motifs guiding the prediction process. The approach involves capturing temporal causality, which includes sequencing events based on their temporal order to maintain causality between interactions.

3.4 | Summary

While various GNN architectures have been employed for causal learning across different domains, it is important to recognize that only a few are specifically designed for this purpose, particularly attention-based models, as emphasized in several studies (Sui et al. 2022a; Liu et al. 2024c; Sui et al. 2024). Models like GATs leverage attention mechanisms to prioritize the significance of neighboring nodes during aggregation, allowing them to focus on the most relevant connections, which enhances the detection of causal relationships. GATs are especially effective for heterogeneous graphs, where the importance of relationships varies, thus improving their capacity to capture diverse causal

TABLE 7 | Explainable causal graph neural networks.

Method	Model	Domain	Key functionality
Decomposition	GCN, LRP (Holzinger et al. 2021)	Medicine	Explainable multi-modal causality framework for medical decision modeling
Perturbation	CF-GNN-Explainer (Lucic et al. 2022)	General	Counterfactual explainer for GNN with minimal perturbation
	CI-GNN (Zheng et al. 2024)	Medicine	Explainable causality to extract brain connections
	Causal Screening (Xiang et al. 2021)	General	Explainable GNN by selecting a graph feature with large causal attribution
Generative	OrphicX (Lin et al. 2022)	Bioinformatics	Explanations from latent causal factors with VGAE
	Gem (Lin et al. 2021)	Bioinformatics	Auto-encoder-based explanations with Granger Causality
	CF2 (Tan et al. 2022b)	Multiple	Explain classification with counterfactuals, factuals
	CLEAR (Ma et al. 2022)	Multiple	VGAE classification with counterfactual explanation
	CP-GCN (Jin et al. 2022)	Text	Relation extraction with causality-pruned dependency forests
	GNNExplainer (Ying et al. 2019)	Multiple	Explainable graph classification with masks (Zhao et al. 2023a)
	PGExplainer (Luo et al. 2020)	Multiple	Explainable graph classification with masks (Zhao et al. 2023b)
	RC-Explainer (Wang et al. 2022b)	Multiple	Explainability in a sequential decision process uncovering edge attributions
	TempME (Chen and Ying 2024)	Multiple	Explainability of temporal graphs with motif discovery
	CLARUS (Metsch et al. 2024)	Medicine	Explainability through manual counterfactuals

structures. However, a drawback is their computational complexity, which may affect performance in large-scale graphs.

In contrast, the simpler GCN model, while computationally efficient, relies on localized convolution operations that aggregate information from a node's neighbors, effectively capturing local graph structures. Nevertheless, GCNs use fixed weights for neighbor aggregation, which can restrict their ability to adapt to the varying importance of different neighbors, potentially missing significant causal relationships. Additionally, the issue of over-smoothing can impede the model's ability to differentiate between distinct causal effects in complex graphs. In the case of GINs, they are highly expressive, which can make them more effective at capturing complex causal relationships where subtle structural differences are important. However, the increased expressiveness and complexity of GINs can also lead to higher computational costs and longer training times compared to GCNs.

4 | Applications

Causality finds extensive applications in various domains such as medicine, social sciences, trade and commerce, and so forth, which we discuss in detail in the following sections. A list of code repositories for these application domains is provided in Table 8.

4.1 | Medicine and Healthcare

Identifying the causal factors influencing patient health facilitates personalized care and assists in determining the contributors to specific diseases. This approach enables healthcare professionals to make informed decisions regarding disease diagnosis and patient care, thereby improving treatment effectiveness. Hızlı et al. (2024) proposed a nonparametric temporal causal mediation framework aimed at estimating the direct and indirect effects of healthcare interventions. In this methodology, the effects are modeled as stochastic processes, with the mediator interacting with the outcome processes to capture future trajectories. Causality also plays a significant role in medical imaging, aiding in disease diagnosis, monitoring, and treatment planning. Image segmentation, viewed as a form of causal prediction by Castro et al. (2020) in medical image analysis, was addressed using causal reasoning by the authors. This approach was suggested as a potential strategy to address the scarcity of labeled data in medical imaging. Within the same domain, Miao et al. (2023) employed a causality-based convolutional network framework for segmentation, relying on statistical quantification.

Hızlı et al. (2024) analyzed the impact of healthcare interventions over time using a nonparametric, stochastic approach to capture the temporal dynamics. In contrast, Castro et al. (2020) focused on addressing the scarcity of labeled data by applying causal reasoning to medical image segmentation tasks. Similarly, Miao et al. (2023) worked on image segmentation tasks, but their approach aimed to address the complexity of these tasks. The integration of causality into these different subdomains highlights the versatility of causal GNNs in improving medical outcomes.

4.2 | Social and Environmental Sciences

Causality is crucial in policy making, behavioral studies, climate science, economics, and other fields, aiding in the development of evidence-based interventions to address social and environmental issues. Causal modeling was utilized to examine how adolescent brain processes criticism (Chen et al. 2024). The analysis combined causality with parametric empirical Bayesian analysis to map interactions among brain regions. Lindsay (2024) also explored changes in behavioral patterns using causality to identify the degree to which neural changes contribute to changes in behaviors.

Iglesias-Suarez et al. (2024) used a causal discovery approach in climate modeling to construct a causally informed model that uncovers causal drivers across various climate regimes. Through the exclusion of spurious links, this model identifies essential causal drivers, thus facilitating the enhancement of climate modeling systems. Causality was also employed for analyzing the impact of geopolitical risks on energy prices using variational mode decomposition and multiresolution causality with autoregressive models (Sadaoui and Jabeur 2023).

In these domains, Causal modeling with parametric empirical Bayesian analysis was used to map brain interactions (Chen et al. 2024), with causal analysis also employed by Lindsay (2024) in the same task to assess the neural-behavioral link. Causal techniques were also applied in climate modeling, where Iglesias-Suarez et al. (2024) focused on genuine causal drivers, as well as in geopolitics with a multi-method causal approach (Sadaoui and Jabeur 2023). In all these applications, causality is incorporated using varying approaches, including Bayesian analysis, multiresolution methods, and others.

4.3 | Trade and Commerce

Building brand loyalty is a key objective for businesses aiming to gain a competitive edge in the market. Understanding the causal connections among factors such as marketing efforts and customer experiences aids in comprehending the drivers of customer loyalty. This understanding facilitates predictive modeling and the implementation of targeted interventions to improve loyalty outcomes. Singh (2021) conducted a causal study to predict airline loyalty, identifying the most critical loyalty predictor as well as the least influential predictors. Similarly, causality also plays a significant role in recommendation systems, where recommendations are customized by understanding causal connections between user preferences and attributes. Additionally, causal analysis aids in dynamically updating recommendations by identifying changes in preferences and behaviors over time. Yu et al. (2023) and Wu et al. (2023) utilized causality in session-based recommendation systems. The former blocked shortcut paths on the session graph to capture causal relations, while the latter developed an attention-based framework capturing both causal and correlation relationships.

Price forecasting holds a significant place in business planning and risk management. Causality plays an important role in understanding the underlying factors that drive price movements in markets by uncovering causal relationships among factors such as economic indicators and geopolitical events. Identifying causal links among market variables enhances predictive accuracy by disregarding spurious correlations. Cheng et al. (2023b) integrated causality with stochastic frontier analysis for crude oil price forecasting, conducting causality assessments across eight categories, including demand, supply, economic policy, geopolitical risk, and others. Wind power forecasting is vital for energy trading, and understanding causal relationships among factors such as wind speed, atmospheric conditions, and geographic features is essential in this domain. Li et al. (2023b) created a framework for wind power forecasting by employing causal convolutions combined with transformers to improve feature extraction.

Singh (2021) identifies key predictors of customer loyalty using a causal study, focusing on their relative influence. In contrast, Yu et al. (2023) used the blocking of shortcut paths in session graphs to capture causal relationships in recommendation systems, while Wu et al. (2023) employed attention methods to model user preferences in the same task. Similarly, Cheng et al. (2023b) used stochastic frontier analysis, and Li et al. (2023b) applied transformer models to incorporate causality in the economic domain and wind power forecasting, respectively.

4.4 | Urban Intelligence

Causality in the field of urban intelligence aids in infrastructure planning and management, particularly when considering factors such as transportation patterns and population growth. A study of traffic flow focused on how contextual conditions causally influence spatial dependencies, with a specific emphasis on triggering effects to enhance important features resulting from these conditions (Xiong and Wang 2024). Liu et al. (2024a) approached the modeling of traffic trajectories as a causal task, intending to explore dependencies within these trajectories. By employing the front-door criterion based on causal interventions, they conducted feature engineering. Additionally, through counterfactual reasoning representation, future contextual information was inferred.

The CaST framework (Xia et al. 2023) enhances spatiotemporal graph forecasting by addressing temporal OOD issues and dynamic spatial causation through causal treatments. It employs backdoor and frontdoor adjustments to separate environment

TABLE 8 | Code repositories for causality applications across various tasks.

Application	Task	Model	Code Link
Medicine and healthcare	Psychiatric diagnosis	CI-GNN (Zheng et al. 2024)	https://github.com/ZKZ-Brain/CI-GNN
	Healthcare intervention	Nonparametric mediator-outcome (Hızlı et al. 2024)	https://github.com/caglar-hizli/dynamic-causal-mediation
	Medical imaging	CauSSL (Miao et al. 2023)	https://github.com/JuzhengMiao/CauSSL
Social and environmental sciences	Atmospheric processes	Benchmarking Causality (Huang et al. 2021)	https://github.com/big-data-lab-umbc/cybertraining/tree/master/year-3-projects/team-6
	Weather forecasting	DGFormer (Xu et al. 2024)	https://github.com/xzwbsz/DGFormer
	Social influence	CDRSB (Wang et al. 2025a)	https://github.com/Lili1013/CDRSB
Trade and commerce	Trading	AINET (Monken et al. 2021)	https://github.com/AndersonMonken/AINET-GNN-Trade21
	Prediction uncertainty	GraphDK (Wen et al. 2023)	https://github.com/uqhwen2/GraphDKL
Urban intelligence	Traffic forecasting	CaST (Xia et al. 2023)	https://github.com/yutong-xia/CaST
	Spatiotemporal learning	CauSTG (Zhou et al. 2023)	https://github.com/zzyy0929/KDD23-CauSTG
	Bike flow prediction	STNSCM (Deng et al. 2023)	https://github.com/EternityZY/STNSCM
Engineering	Speech enhancement	MHANet (Nicolson and Paliwal 2020)	https://github.com/anicolson/DeepXi
	Bioprocess forecasting	CEGLo-GNN (Sun et al. 2024)	https://github.com/YueYueXia/CEGLo-GNN
	fMRI contrastive learning	CIIGCL (Wei et al. 2025)	https://github.com/wby920920/CIIGCL
Natural language processing	Text classification	DAS (Wu et al. 2024b)	https://github.com/6666ev/DAS
	Hate speech detection	CATCH (Sheth et al. 2024)	https://github.com/paras2612/CATCH
	Causal sentence extraction	Text mining (Norouzi et al. 2024)	https://github.com/rasoulnorouzi/cessc
	Causal LLMs	CLADDER (Jin et al. 2023)	https://github.com/causalNLP/cladder

and entity features, using edge-level convolution with the Hodge–Laplacian operator to capture spatial causality. OOD challenges in spatiotemporal learning for urban digitization using CauSTG (Zhou et al. 2023) leveraged causal spatiotemporal relations and invariant learning techniques. This approach included temporal environmental partitioning, spatiotemporal consistency learning, and hierarchical invariance exploration, focusing on disentangling seasonal trend patterns and local–global invariance filtering to improve generalization in dynamic environments. Similarly, Deng et al. (2023) proposed a spatiotemporal neural structure causal model (STNSCM) with counterfactual reasoning to address contextual conditions and inter-regional time-varying causality in bike-sharing flow prediction. Self-supervised learning is leveraged in the Spatial–Temporal sElf-superVised confoundEr learning (STEVE) model to enhance confounder representation using a basis vector approach (Ji et al. 2025). This approach enables better decoupling of confounder effects from direct traffic relations and adapts well to unseen confounders. Each of these approaches addresses causal-based learning in urban intelligence using methods ranging from contextual condition analysis to spatiotemporal forecasting, aiming towards improving urban systems through a causal approach.

4.5 | Engineering

Causality in engineering encompasses various activities, including root cause analysis, process improvement, risk mitigation, design optimization, predictive modeling, and more. Speech enhancement systems are essential for maintaining speech quality in various settings such as videoconferencing, broadcasting, and public address systems. Nicolson and Paliwal (2020) developed a causal speech enhancement system named *MHANet*, utilizing DNN and an attention mechanism. In *MHANet*, positional encoding was omitted, and the multi-heads from the attention unit attended to specific speech regions, including both target and noisy signals. *MHANet* depended on the current and preceding time frames, whereas equivalent noncausal models utilize multiple past frames, leading to a performance delay.

Fault diagnosis is a critical requirement in industrial processes for uncovering the underlying reasons behind process faults, and causal discovery is emerging as a key tool in this task. Wang et al. (2023a) introduced a multisensor time series causality discovery method utilizing convolutional neural networks for diagnosing root causes of process faults, validated through permutation importance causality assessment. Likewise, attention-based causal methodologies were employed by Wang et al. (2023d) and Liu et al. (2023) for fault detection through disentangled representations. *MHANet* focuses on real-time processing with limited context. On the other hand, the multisensory approach by Wang et al. (2023a) is a more comprehensive, though computationally intensive, method.

4.6 | Natural Language Processing

The use of Natural Language Processing in social media is important for understanding user behavior, conducting sentiment analysis, social network analysis, and other related tasks. Li et al. (2024c) employed a causal public opinion framework to construct a price prediction model by leveraging Granger causal testing and linear regression. This was achieved through the mining of public opinion text using sentiment analysis. Sheth et al. (2024) employed causality to understand invariant representations in hate speech by disentangling platform-invariant and platform-dependent components from texts. Causal text mining is also relevant in academic papers and medical texts for effective and targeted information retrieval, as well as for identifying causal pathways, risk factors, and other important factors. Norouzi et al. (2024) designed a text mining framework for extracting causal sentences from social science papers. This framework facilitates the analysis of causal claims, enabling the discovery of mechanisms underlying social phenomena. Wu et al. (2024b) implemented causality-based text mining using a debiased attention supervision method to mitigate label-based and word-based biases through causal techniques. The model was employed on medical and legal texts for text classification, achieving high performance through coherent attention distributions. The first two works focus on sentiment and representation analysis; the latter two specialize in causal sentence extraction and bias mitigation in academic texts. The main difference lies in their application domain and complexity, with the latter using more refined approaches for specialized tasks.

4.6.1 | Large Language Models

Large language models (LLMs) have promoted natural language understanding by enabling the generation, processing, and interpretation of text with a high level of contextual accuracy across a wide range of applications, contributing to content creation and problem solving. In recent research, causality has become an important factor in advancing the capabilities of LLMs. The integration of causal reasoning helps LLMs to model cause and effect relationships in text, contributing to better contextual understanding. This enables LLMs to produce more accurate outputs based on complex reasoning and deeper insights.

Ban et al. (2025) proposed a causal LLM framework that focuses on individual causal aspects to create a harmonized prior using a set of structural constraints. The integration of this prior enhances model performance in structure learning. Similarly, a causality-based LLM model that integrates commonsense knowledge and causal reasoning for emotion analysis was used to develop an LLM-driven sentiment analysis framework with example retrieval (Zhang et al. 2025a). Causal reasoning in LLMs is also explored using a chain-of-thought prompting strategy called CAUSALCOT, with the goal of analyzing their causal reasoning abilities (Jin et al. 2023). LLMs are also used in pharmacovigilance to identify causal relationships between drugs and adverse effects using a causal temporal GNN with motif embedding and similarity search to address issues related to NER approaches (Kalla et al. 2023). LLM Reasoning Graphs (LLMRG) create tailored reasoning graphs that link user profiles and behaviors through causal analysis, with GNNs encoding the graphs to improve recommender systems without requiring further user or item information (Wang et al. 2024). Thus, recent advancements in causal reasoning with LLMs, including frameworks for sentiment analysis, pharmacovigilance, and recommender systems, demonstrate their potential to improve model performance through causal analysis.

5 | Challenges and Directions

There are several challenges associated with causal learning and GNNs, which we discuss in this section.

5.1 | Data Quality

Observational data, often marred by incompleteness and biases such as selection bias and confounding (Bareinboim and Pearl 2012; Hammerton and Munafò 2021), presents a unique set of challenges in the context of causal learning. Traditional modeling assumptions that address these challenges might not always encapsulate the system's intricate dynamics. For instance, deriving causality from observational data can be hindered by issues such as deficient anomaly detection approaches (Qiu et al. 2020; Wu and Liu 2021) and the complexity of certain treatments, such as those related to images (Liu et al. 2022).

Opportunities: Constructed data, derived from scenarios similar to observational studies, emerges as a promising solution (Keith et al. 2020), especially when ground truth is elusive. This approach aids in evaluating causality and can be complemented by weak supervision for causal feature selection in cases with sparse labeled data. Furthermore, the integration of interventional and observational data can enhance causal learning. However, there remain challenges such as dataset imbalances and discrepancies in training and test set distributions (Castro et al. 2020). Future endeavors should prioritize exploring methods to address these challenges, such as techniques to handle dataset shifts caused by factors like irregular data collection.

Delving into distributional disparities, aspects like population, annotation, and manifestation shifts, along with considerations for anticausal tasks, need thorough investigation. Additionally, the development of models that can seamlessly adapt to these shifts and biases will be pivotal in advancing the field of causal learning.

5.2 | Causal Assumptions

Causal discovery makes strong assumptions such as the Markov property (Spirtes et al. 2001), and causal inference makes assumptions such as SUTVA and consistency, among others (Fé et al. 2020; Pearl 2010). Both require ongoing investigation in scenarios that deviate from these premises. Similarly, assumptions such as positivity (Cole and Frangakis 2009) are not adhered to when certain elements in a system remain idle or untouched. Furthermore, causal studies generally represent treatments as discrete events (Peters et al. 2022), which is inadequate for exploring causality in continuous data.

Opportunities: Sensitivity of causal study outcomes to violations of these assumptions (Igelstrm et al. 2022) must be investigated to design a flexible framework for validating causal models. A concrete validation approach needs to be designed for validating causal estimates in the presence of potential confounding, leading to unbiased estimators. Estimating causal effects in a continuous environment requires further research to model treatments as continuous events.

5.3 | Research Scalability

Many causal datasets are small-scale datasets, which are typically scaled-down graph datasets that do not align with real-world dynamics. Synthetic or semi-synthetic datasets are the core of many causal experiments (Zhao et al. 2024; Rao et al. 2024).

Opportunities: Translating new approaches in the context of high-dimensional data needs to be further investigated using larger causal data. This is particularly significant, since the primary step in preparing high-dimensional data for causality learning is encoding. This is adequate for nonsemantic data; however, the viability of encoding semantic data to a low-dimensional representation is largely domain-dependent (Farahani et al. 2021). Low-dimensional features must encompass treatment, outcomes, and confounding information to lead towards representation learning. Few areas requiring investigation in representation learning are adequate management of confounding information and noisy outcomes.

5.4 | Dynamic Causality

Causality research is largely based on static observational data studies. Causal inference for modeling dynamic data would pave the way for capturing causality in dynamic environments. Moreover, causal inference in the spatial domain requires the inclusion of spatial heterogeneity, spatial interactions, and spatial lag effects for accurate causal estimation. This will open a path for designing dedicated spatial causal learning methods (Akbari et al. 2023).

Opportunities: If the mechanisms associated with causal structure vary to the extent of affecting causal links, distribution shift must be incorporated in modeling causality (Glymour et al. 2019). Causality-based feature selection can handle distribution shift data, but this requires interventional information for extracting causal variables. Moreover, it is challenging to capture causal dynamics in a continuous environment, on account of the difficulty in gathering time series data rapidly enough to account for the swift changes occurring in the system. Additionally, for spatial causal analysis, the development of appropriate tools competent enough to capture spatial complexities would pave the way for more significant research in the domain.

5.5 | Multiple Tasks and Multi-Modal Causality

There is a dearth of studies exploring causal learning with multimodal data. Publicly available multi-modal causal datasets can motivate research in this area and enhance causal applicability in multiple fields. Causal representations for multiple tasks can be extracted by clustering trajectories from the various domains (Kurutach et al. 2018), though determining the granularity of causal variables is largely dependent on the individual task.

Opportunities: Diligent curation of multi-modal datasets would uplift multi-modal research to some degree. Additionally, taking into consideration the impracticality of validating causal models with several tasks and interventions, meta learning needs to be further researched alongside reinforcement learning in the field of causality. GNNs can improve causal inference in multiagent systems by modeling complex interactions and handling multi-modal data (Zhang et al. 2022g; Sui et al. 2022a; Ektefaie et al. 2023). They also enhance causal understanding and support meta-learning and reinforcement learning for more adaptive and accurate modeling.

5.6 | Training Data

While GNNs are adept at learning complex patterns in data, their effectiveness is largely dependent on the volume of training data. GNNs reach high performance and generalizability with large datasets, which presently have limited availability. Furthermore, such studies with limited data tend to result in biased outcomes. Conversely, when large-scale data is available, experiments are often restricted to data subsets (Long et al. 2021) on account of the computational requirements of GNNs.

Opportunities: Transfer learning can be utilized in circumstances where data is limited. This approach is also beneficial for learning networks with constant changes. Both instances are made possible by model transferability through the process of training a model with data-rich historical or topical subsets (Jiang and Luo 2022). In scenarios with large-scale data, graph partitioning is one possible approach, in which networks are split into smaller units (Jiang and Luo 2022). Generative adversarial networks (GAN) can also enhance training data for

GNN models with a generator–discriminator approach, with Wasserstein GAN (WGAN; Arjovsky et al. 2017) able to generate data for graph modeling purposes.

5.7 | Temporal Data

A major challenge in temporal graphs is managing nonstationary data, which involves adapting to evolving relationships and patterns over time (Li et al. 2024a). As nodes and edges evolve, traditional models often struggle to capture these shifts accurately, leading to outdated or incorrect predictions.

Opportunities: To address nonstationary data in temporal graphs, GNNs can employ dynamic architectures and causal inference methods to adapt to changing relationships. Given that temporal nonstationarity can introduce causal calculation bias, directly applying existing causal effects to predict delays may reduce their accuracy (Zhu et al. 2024). Since discovering causal relations helps in handling the nonstationarity of temporal data (Faruque et al. 2024), incorporating causal models improves the GNN's capacity to interpret and predict the impacts of temporal shifts. Meanwhile, adaptive learning strategies ensure the model remains accurate and relevant over time. This approach provides more precise and timely insights into the evolving dynamics of the graph.

5.8 | Graph Structures and Scalability

Graph modeling in a multi-tasking framework remains a challenge, with problems such as forecasting in various task areas demanding multiple graph structures. Most studies address this problem by employing GNN on the same graph using feedforward layers, leading to multiple results (Jiang and Luo 2022). Furthermore, scalability is an open problem in processing massive graphs. Development of GNN architectures that are scalable for continually evolving graphs is critical for speeding up the graph training process.

Opportunities: Neighbor sampling and mini-batch processing are a few approaches currently used by researchers (Hamilton et al. 2017), but require further study toward overcoming the overfitting and oversmoothing problems of GNNs. Meta learning is another approach that can equip GNNs with the ability to learn on a task for generalizing to multiple tasks. When a model learned on a graph is applied to other graph structures, interactions between multiple graphs are not captured. This includes the incorporation of cross-network node similarities, anchor links, and so forth. Similarly, graph correlations are also not accounted for; hence, information from multiple graphs must be consolidated in representation learning for effective modeling.

5.9 | Complexity of Graph Structure and Graph Layers

GNN models are mostly studied on homogeneous graphs on account of the complexity of implementing GNNs as heterogeneous structures. This is particularly important for capturing dynamic spatial information. Moreover, this aspect also limits GNN's capability in handling multi-modal data. Multi-modal GNN research primarily uses balanced data, and thus requires further exploration in the case of imbalanced datasets. Besides, adding unlimited layers in GNN models can contribute to a performance decline (Zhang et al. 2019), making this a possible research direction for developing deep, robust GNNs that are computationally efficient.

Opportunities: Modeling in complex environments requires research into higher-order structures such as graphlets. Laplacian matrix estimator and data adaptive graph generation are some existing approaches proposed to capture dependencies in a dynamic environment (Jiang and Luo 2022). Incorporating embedding propagation operations is a strategy adopted in this area. Further research into dynamic and heterogeneous graph structures would contribute to the stability and adaptability of GNNs in varying structures. Furthermore, the challenge associated with increasing GNN depth can be alleviated to some degree by adopting skip connection-based structures (Zhang et al. 2019); however, more flexible architectures are required for building deeper models. In complex networks, nodes and edge attributes may be dynamic with a multilayered structure. Each layer must be defined as a separate dimension, necessitating encoding at these various network layers. Studies must be expanded on encoding dynamic networks, incorporating node attributes for predicting topology dynamics.

5.10 | Adversarial Attacks on Graphs

A major drawback of the GNN architecture is that it is susceptible to adversarial attacks (Zhang and Zitnik 2020), leading to performance decline. A benchmark approach for building defenses against adversarial attacks on GNNs is required to develop a suitable framework for defense against graph attacks. The framework must incorporate structure as well as attribute perturbations for building robust GNNs. In addition to formulating defense approaches, more research should be conducted on identifying and cleaning contaminated graphs.

Opportunities: Measuring perturbations on graphs is undetectable at the human level and hence requires robust perturbation evaluation measures to address this problem. Adversarial attacks on graphs are mostly researched on static graphs with node attributes (Jin et al. 2021). An important research direction in adversarial studies is a focus on complex graphs with edge attributes, as well as dynamic graphs. Transferability of graph adversarial examples is another area that requires further research for building efficient and robust graph models.

5.11 | Explainability of Graph Data

The complex nature of graph data leads to abstract explanations, which can be further compounded by a lack of domain knowledge (Agarwal et al. 2023). This problem can be alleviated to some extent by developing standardized datasets for explanation tasks. Moreover, heterogeneous data contributes to complex structures, which can complicate the explainability process (Wu et al. 2024a). Most explainable GNN models are instance-level and not at the model level. Explainable GNNs are critical for identifying the subgraphs (Agarwal et al. 2023) that can significantly contribute to model outcomes. Lack of locality information and varying node neighbors (Yuan et al. 2022) make explaining graph structures a challenging task. Extending existing explainable methods to GNNs is not very dependable on account of the topology information in adjacency matrices being presented as discrete values (Yuan et al. 2022). For the same reason, input optimization approaches commonly utilized in explaining models cannot be extended to GNNs. Learning soft masks is another approach used as an explainability technique, but if translated to an adjacency matrix, this will interfere with its discrete nature (Yuan et al. 2022).

Opportunities: Explaining GNNs involves understanding the structural information of graphs, which is not directly explainable by current methods. Evaluation of explainable methods in the graph domain is not straightforward, since graph visualizations are not easy to render for direct human understanding (Li et al. 2022b). Both task-specific and task-agnostic evaluation metrics must be researched and developed with a view towards quantitative explanation of GNN models.

6 | Conclusion

In recent years, GNNs have gained recognition as trustworthy and effective tools for a variety of tasks related to graphs. This survey provides a comprehensive review of the latest advancements in GNNs in the domain of causal learning. There has been a notable increase in the application of GNN-based methods for causal learning, significantly transforming the field. Based on our extensive analysis of peer-reviewed publications, it is clear that these GNN-based approaches have enhanced the efficiency, robustness, and versatility of causal learning. Contemporary methods can adapt more readily to various situations or datasets than their traditional counterparts, and they also leverage a broader range of data types and sources during the initial analysis, leading to more comprehensive and effective results. This adaptability underscores the need for a clear framework to categorize and understand the diverse approaches in causal GNN methods, prompting us to develop a new taxonomic system. Consequently, we have organized these methods into three distinct categories: resolution-based, learning-based, and explainability-based. This classification allows us to better address their unique characteristics, functionalities, and applications in various contexts. We explored the advancements in each category, highlighting the unique contributions and development paths for each of these approaches. Furthermore, we compiled a concise guide to the resources commonly cited in the literature, including both datasets and the overall framework of resources used in researching CGNNs. Additionally, we provided an overview of the real-world applications of causality across various domains, illustrating its practical significance. To conclude, we have identified several challenges and open research avenues that warrant attention in this rapidly evolving field.

Author Contributions

Simi Job: conceptualization (equal), data curation (equal), formal analysis (equal), investigation (equal), methodology (equal), resources (equal), writing – original draft (equal), writing – review and editing (equal). **Xiaohui Tao:** conceptualization (equal), formal analysis (equal), investigation (equal), methodology (equal), project administration (equal), supervision (equal), validation (equal), writing – original draft (equal), writing – review and editing (equal). **Taotao Cai:** formal analysis (equal), investigation (equal), methodology (equal), supervision (equal), writing – original draft (equal), writing – review and editing (equal). **Haoran Xie:** conceptualization (equal), investigation (equal), methodology (equal), writing – review and editing (equal). **Lin Li:** conceptualization (equal), formal analysis (equal), investigation (equal), methodology (equal), writing – review and editing (equal). **Qing Li:** conceptualization (equal), investigation (equal), writing – review and editing (equal). **Jianming Yong:** conceptualization (equal), investigation (equal), project administration (equal), supervision (equal), writing – review and editing (equal).

Acknowledgments

This work is supported by the SAGE Athena Swan Scholarship. Open access publishing facilitated by University of Southern Queensland, as part of the Wiley - University of Southern Queensland agreement via the Council of Australian University Librarians.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

Related WIREs Articles

Methods and tools for causal discovery and causal inference

References

Adhikari, S., and E. Zheleva. 2023. "Inferring Causal Effects Under Heterogeneous Peer Influence." *arXiv Preprint* arXiv:2305.17479.

ADNI. 2017. "Alzheimer's Disease Neuroimaging Initiative" [Dataset]. https://adni.loni.usc.edu/ [Accessed February 3, 2025].

Agarwal, C., O. Queen, H. Lakkaraju, and M. Zitnik. 2023. "Evaluating Explainability for Graph Neural Networks." *Scientific Data* 10, no. 1: 144.

Ahmed, C. M., V. R. Palleti, and A. P. Mathur. 2017. "Wadi: A Water Distribution Testbed for Research in the Design of Secure Cyber Physical Systems," in *Proceedings of the 3rd International Workshop on Cyber-Physical Systems for Smart Water Networks*, Pittsburgh, Pennsylvania, pp. 25–28. https://doi.org/10.1145/3055366.3055375.

Akbari, K., S. Winter, and M. Tomko. 2023. "Spatial Causality: A Systematic Review on Spatial Causal Inference." *Geographical Analysis* 55, no. 1: 56–89.

Alam, K. J., and K. K. Sumon. 2020. "Causal Relationship Between Trade Openness and Economic Growth: A Panel Data Analysis of Asian Countries." *International Journal of Economics and Financial Issues* 10, no. 1: 118–126.

Alex, K. 2009. Learning Multiple Layers of Features From Tiny Images [Thesis]. https://www.cs.toronto.edu/kriz/learning-features-2009-TR.pdf.

Amirinezhad, A., S. Salehkaleybar, and M. Hashemi. 2022. "Active Learning of Causal Structures With Deep Reinforcement Learning." *Neural Networks* 154: 22–30.

Arboleda-Florez, J., H. Holley, and A. Crisanti. 1998. "Understanding Causal Paths Between Mental Illness and Violence." *Social Psychiatry and Psychiatric Epidemiology* 33: S38–S46.

Arjovsky, M., S. Chintala, and L. Bottou. 2017. “Wasserstein Generative Adversarial Networks,” in *International Conference on Machine Learning* (PMLR), pp. 214–223.

Avazu. 2015. Kaggle Click-Through Rate Prediction [Dataset]. https://www.kaggle.com/c/avazu-ctr-prediction/ [Accessed March 2, 2025].

Ban, T., L. Chen, D. Lyu, X. Wang, Q. Zhu, and H. Chen. 2025. “Llm-Driven Causal Discovery via Harmonized Prior.” *IEEE Transactions on Knowledge and Data Engineering* 37: 1943–1960.

Bareinboim, E., and J. Pearl. 2012. “Controlling Selection Bias in Causal Inference.” In *Artificial Intelligence and Statistics*, 100–108. PMLR.

Bo, D., X. Wang, Y. Liu, Y. Fang, Y. Li, and C. Shi. 2023. “A Survey on Spectral Graph Neural Networks.” *arXiv preprint* arXiv:2302.05631.

Bordes, A., N. Usunier, A. Garcia-Duran, J. Weston, and O. Yakhnenko. 2013. “Translating Embeddings for Modeling Multi-Relational Data,” in *Advances in Neural Information Processing Systems*, 26.

Borgwardt, K. M., C. S. Ong, S. Schnauer, S. V. N. Vishwanathan, A. J. Smola, and H.-P. Kriegel. 2005. “Protein Function Prediction via Graph Kernels.” *Bioinformatics* 21, no. suppl_1: i47–i56.

Candanedo, L. 2017. Appliances Energy Prediction [Database]. https://doi.org/10.24432/C5VC8G [Accessed March 2, 2025].

Cao, Q., X. Hao, H. Ren, W. Xu, S. Xu, and C. J. Asiedu. 2022. “Graph Attention Network Based Detection of Causality for Textual Emotion-Cause Pair.” *World Wide Web* 26: 1–15.

Carbo-Valverde, S., P. Cuadros-Solas, and F. Rodrguez-Fernndez. 2020. “A Machine Learning Approach to the Digitalization of Bank Customers: Evidence From Random and Causal Forests.” *PLoS One* 15, no. 10: e0240362.

Caselli, T., and P. Vossen. 2017. “The Event Storyline Corpus: A New Benchmark for Causal and Temporal Relation Extraction.” In *Proceedings of the Events and Stories in the News Workshop*, 77–86. Association for Computational Linguistics. https://doi.org/10.18653/v1/W17-2711.

Castro, D. C., I. Walker, and B. Glocker. 2020. “Causality Matters in Medical Imaging.” *Nature Communications* 11, no. 1: 3673.

Chai, X. 2020. “Diagnosis Method of Thyroid Disease Combining Knowledge Graph and Deep Learning.” *IEEE Access* 8: 149787–149795.

Chang, X., Z. Zhang, J. Sun, K. Lin, and P. Song. 2025. “Breast Cancer Image Classification Based on H&E Staining Using a Causal Attention Graph Neural Network Model.” *Medical & Biological Engineering & Computing* 63, no. 2: 1–15. https://doi.org/10.1007/s11517-025-03303-3.

Chantala, K., and J. Tabor. 1999. *National Longitudinal Study of Adolescent Health:Strategies to Perform a Design-Based Analysis Using the Add Health Data*. Carolina Population Center.

Chen, J., and R. Ying. 2024. “Tempme: Towards the Explainability of Temporal Graph Neural Networks via Motif Discovery.” *Advances in Neural Information Processing Systems* 36: 29005–29028.

Chen, Q., S. L. B. Bonduelle, G.-R. Wu, M.-A. Vanderhasselt, R. de Raedt, and C. Baeken. 2024. “Unraveling How the Adolescent Brain Deals With Criticism Using Dynamic Causal Modeling.” *NeuroImage* 286: 120510.

Chen, Y., W. Wan, J. Hu, Y. Wang, and B. Huang. 2022a. “Complex Causal Extraction of Fusion of Entity Location Sensing and Graph Attention Networks.” *Information* 13, no. 8: 364.

Chen, Z., Z. Wang, Y. Yang, and J. Gao. 2022b. “Resgraphnet: Graphsage With Embedded Residual Module for Prediction of Global Monthly Mean Temperature.” *Artificial Intelligence in Geosciences* 3: 148–156.

Cheng, D., J. Li, L. Liu, J. Liu, and T. D. Le. 2022. “Data-Driven Causal Effect Estimation Based on Graphical Causal Modelling: A Survey.” *ACM Computing Surveys* 56, no. 5: 127.

Cheng, J., K. Amara, Y. Junchi, and R. Ying. 2023a. “Generative Explanations for Graph Neural Network: Methods and Evaluations.” *IEEE Data Engineering Bulletin* 47, no. 2: 64–79.

Cheng, X., P. Wu, S. S. Liao, and X. Wang. 2023b. “An Integrated Model for Crude Oil Forecasting: Causality Assessment and Technical Efficiency.” *Energy Economics* 117: 106467.

Chickering, D. M. 2002. “Optimal Structure Identification With Greedy Search.” *Journal of Machine Learning Research* 3: 507–554.

Cho, E., S. A. Myers, and J. Leskovec. 2011. “Friendship and Mobility: User Movement in Location-Based Social Networks,” in *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Diego, CA, pp. 1082–1090. https://doi.org/10.1145/2020408.2020579.

CIKM. 2016. CIKM Cup 2016 [Database]. https://competitions.codalab.org/competitions/11161.

Cole, S. R., and C. E. Frangakis. 2009. “The Consistency Statement in Causal Inference: A Definition or an Assumption?” *Epidemiology* 20, no. 1: 3–5.

UN Comtrade. 2021. United Nations Department of Economic and Social Affairs/Statistics Division. United Nations Commodity Trade Statistics Database [Database]. https://comtradeplus.un.org/ [Accessed March 2, 2025].

Criteo. 2014. Kaggle Display Advertising Challenge Dataset [Database]. https://www.kaggle.com/c/criteo-display-ad-challenge/data [Accessed April 2, 2025].

Cummings, D., and M. Nassar. 2020. “Structured Citation Trend Prediction Using Graph Neural Networks.” In *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 3897–3901. IEEE.

Debnath, A. K., R. L. Lopez de Compadre, G. Debnath, A. J. Shusterman, and C. Hansch. 1991. “Structure-Activity Relationship of Mutagenic Aromatic and Heteroaromatic Nitro Compounds. Correlation With Molecular Orbital Energies and Hydrophobicity.” *Journal of Medicinal Chemistry* 34, no. 2: 786–797.

Decruyenaere, A., J. Steen, K. Colpaert, D. D. Benoit, J. Decruyenaere, and S. Vansteelandt. 2020. “The Obesity Paradox in Critically Ill Patients: A Causal Learning Approach to a Casual Finding.” *Critical Care* 24, no. 1: 1–11.

Deng, P., Y. Zhao, J. Liu, X. Jia, and M. Wang. 2023. “Spatio-Temporal Neural Structural Causal Models for Bike Flow Prediction.” In *Proceedings of the Thirty-Seventh AAAI Conference on Artificial Intelligence and Thirty-Fifth Conference on Innovative Applications of Artificial Intelligence and Thirteenth Symposium on Educational Advances in Artificial Intelligence*, vol. 37, 4242–4249. AAAI Press.

di Martino, A., C.-G. Yan, Q. Li, et al. 2014. “The Autism Brain Imaging Data Exchange: Towards a Large-Scale Evaluation of the Intrinsic Brain Architecture in Autism.” *Molecular Psychiatry* 19, no. 6: 659–667.

Ding, S., F. Feng, X. He, Y. Liao, J. Shi, and Y. Zhang. 2022. “Causal Incremental Graph Convolution for Recommender System Retraining.” *IEEE Transactions on Neural Networks and Learning Systems* 35, no. 4: 4718–4728.

Duan, R.-N., J.-Y. Zhu, and B.-L. Lu. 2013. “Differential Entropy Feature for Eeg-Based Emotion Classification.” In *2013 6th International IEEE/EMBS Conference on Neural Engineering (NER)*, 81–84. IEEE.

Ektefaie, Y., G. Dasoulas, A. Noori, M. Farhat, and M. Zitnik. 2023. “Multimodal Learning With Graphs.” *Nature Machine Intelligence* 5, no. 4: 340–350.

EPA. 2022. Toxcast Database [Database]. https://www.epa.gov/comptox-tools/exploring-toxcast-data [Accessed April 2, 2025].

Faber, L. K., A. Moghaddam, and R. Wattenhofer. 2021. “When Comparing to Ground Truth Is Wrong: On Evaluating GNN Explanation

Methods," in *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*, 332–341.

Fan, S., X. Wang, Y. Mo, C. Shi, and J. Tang. 2022. "Debiasing Graph Neural Networks via Learning Disentangled Causal Substructure." *Advances in Neural Information Processing Systems* 35: 24934–24946.

Fang, H., H. Wang, Y. Gao, et al. 2025. "Insgnn: Interpretable Spatio–Temporal Graph Neural Networks via Information Bottleneck." *Information Fusion* 119: 102997.

Farahani, A., S. Voghoei, K. Rasheed, and H. R. Arabnia. 2021. "A Brief Review of Domain Adaptation," in *Advances in Data Science and Information Engineering*: *Proceedings From ICDATA 2020 and IKE 2020*, pp. 877–894.

Faruque, O., S. Ali, X. Zheng, and J. Wang. 2024. Ts-causalnn: Learning Temporal Causal Relations From Non-Linear Non-Stationary Time Series Data. *arXiv preprint* arXiv:2404.01466.

Fé, E., P. Atkinson, S. Delamont, A. Cernat, J. W. Sakshaug, and R. A. Williams. 2020. *Causal Estimation and Causal Inference*. Sage.

Feng, F., W. Huang, X. He, X. Xin, Q. Wang, and T.-S. Chua. 2021. "Should Graph Convolution Trust Neighbors? A Simple Causal Inference Method," in *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 1208–1218.

Finn, C., P. Abbeel, and S. Levine. 2017. "Model-Agnostic Meta-Learning for Fast Adaptation of Deep Networks." In *International Conference on Machine Learning*, 1126–1135. PMLR.

Flickr. 2020. Flickr Social Network [Database]. https://www.flickr.com/.

Fredricksen, H. 1992. "A New Look at the de Bruijn Graph." *Discrete Applied Mathematics* 37: 193–203.

Gan, L., Z. Gao, X. Zhang, et al. 2025. "Graph Neural Networks Enabled Accident Causation Prediction for Maritime Vessel Traffic." *Reliability Engineering & System Safety* 257: 110804.

Gao, H., J. Li, W. Qiang, et al. 2023. "Robust Causal Graph Representation Learning Against Confounding Effects." *Proceedings of the AAAI Conference on Artificial Intelligence* 37, no. 6: 7624–7632.

Gao, H., C. Yao, J. Li, et al. 2024. "Rethinking Causal Relationships Learning in Graph Neural Networks." *Proceedings of the AAAI Conference on Artificial Intelligence* 38, no. 11: 12145–12154.

Gao, J., H. Yu, and S. Zhang. 2022. "Joint Event Causality Extraction Using Dual-Channel Enhanced Neural Network." *Knowledge-Based Systems* 258: 109935.

Giles, C. L., K. D. Bollacker, and S. Lawrence. 1998. "Citeseer: An Automatic Citation Indexing System." In *Proceedings of the Third ACM Conference on Digital Libraries*, 89–98. ACM. https://doi.org/10.1145/276675.276685.

Glymour, C., K. Zhang, and P. Spirtes. 2019. "Review of Causal Discovery Methods Based on Graphical Models." *Frontiers in Genetics* 10: 524.

Grouplens. 2003. Movielens 1m Dataset. https://grouplens.org/datasets/movielens/1m/.

Gu, Y., and L. Deng. 2022. "Stagcn: Spatial–Temporal Attention Graph Convolution Network for Traffic Forecasting." *Mathematics* 10, no. 9: 1599.

Gulla, J. A., L. Zhang, P. Liu, Ö. Özgöbek, and X. Su. 2017. "The Adressa Dataset for News Recommendation," in *Proceedings of the International Conference on Web Intelligence*, pp. 1042–1048.

Guo, R., J. Li, and H. Liu. 2020. "Learning Individual Causal Effects From Networked Observational Data," in *Proceedings of the 13th International Conference on Web Search and Data Mining*, pp. 232–240.

Guo, S., Y. Lin, H. Wan, X. Li, and G. Cong. 2021. "Learning Dynamics and Heterogeneity of Spatial-Temporal Graph Data for Traffic Forecasting." *IEEE Transactions on Knowledge and Data Engineering* 34, no. 11: 5415–5428.

Hamilton, W., Z. Ying, and J. Leskovec. 2017. Inductive Representation Learning on Large Graphs. Advances in Neural Information Processing Systems, 30.

Hammerton, G., and M. R. Munafò. 2021. "Causal Inference With Observational Data: The Need for Triangulation of Evidence." *Psychological Medicine* 51, no. 4: 563–578.

Hammond, D. K., P. Vandergheynst, and R. Gribonval. 2011. "Wavelets on Graphs via Spectral Graph Theory." *Applied and Computational Harmonic Analysis* 30, no. 2: 129–150.

He, Y., G. Xiao, J. Zhu, T. Zou, and Y. Liang. 2024. "Reinforcement Learning-Based Sdn Routing Scheme Empowered by Causality Detection and Gnn." *Frontiers in Computational Neuroscience* 18: 1393025.

Hendrickx, I., S. N. Kim, Z. Kozareva, et al. 2010. "SemEval-2010 Task 8: Multi-Way Classification of Semantic Relations between Pairs of Nominals." In *Proceedings of the 5th International Workshop on Semantic Evaluation*, 33–38. ACL.

Hersbach, H., B. Bell, P. Berrisford, et al. 2020. "The era5 Global Reanalysis." *Quarterly Journal of the Royal Meteorological Society* 146, no. 730: 1999–2049.

Hidey, C., and K. McKeown. 2016. "Identifying Causal Relations Using Parallel Wikipedia Articles." In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics, Vol. 1: Long Papers*, 1424–1433. Association for Computational Linguistics.

Hızlı, Ç., S. John, A. Juuti, T. Saarinen, K. Pietiläinen, and P. Marttinen. 2024. "Temporal Causal Mediation Through a Point Process: Direct and Indirect Effects of Healthcare Interventions," in *Advances in Neural Information Processing Systems*, 36.

Hodge, M. R., W. Horton, T. Brown, et al. 2016. "Connectomedb—Sharing Human Brain Connectivity Data." *NeuroImage* 124: 1102–1107.

Holzinger, A., B. Malle, A. Saranti, and B. Pfeifer. 2021. "Towards Multi-Modal Causability With Graph Neural Networks Enabling Information Fusion for Explainable Ai." *Information Fusion* 71: 28–37.

Horiwaki, K. 2023. "Domain-Based Graph Neural Network: Evaluations of the Effectiveness of Graph Representation Based on Domain Knowledge." In *2023 IEEE 19th International Conference on Automation Science and Engineering (CASE)*, 1–8. IEEE.

Hu, W., M. Fey, M. Zitnik, et al. 2020. "Open Graph Benchmark: Datasets for Machine Learning on Graphs." *Advances in Neural Information Processing Systems* 33: 22118–22133.

Hu, W., J. Wu, and Q. Qian. 2025. "Cirlexplainer: Causality-Inspired Explainer for Graph Neural Networks via Reinforcement Learning." *IEEE Transactions on Neural Networks and Learning Systems* PP: 1–15.

Huang, Q., M. Yamada, Y. Tian, D. Singh, and Y. Chang. 2022. "Graphlime: Local Interpretable Model Explanations for Graph Neural Networks." *IEEE Transactions on Knowledge and Data Engineering* 35, no. 7: 6968–6972.

Huang, R., M. Xia, D.-T. Nguyen, et al. 2016. "Tox21 Challenge to Build Predictive Models of Nuclear Receptor and Stress Response Pathways as Mediated by Exposure to Environmental Chemicals and Drugs." *Frontiers in Environmental Science* 3: 85.

Huang, W., X. Gao, G. Zhao, et al. 2023. "A Novel Brain Inception Neural Network Model Using Eeg Graphic Structure for Emotion Recognition." *Brain-Apparatus Communication: A Journal of Bacomics* 2, no. 1: 2222159.

Huang, Y., M. Kleindessner, A. Munishkin, D. Varshney, P. Guo, and J. Wang. 2021. "Benchmarking of Data-Driven Causality Discovery Approaches in the Interactions of Arctic Sea Ice and Atmosphere." *Frontiers in Big Data* 4: 642182.

Huang, Z., T. Jiang, Y. Feng, Z. Wen, and X. Cui. 2025. "Ciegcl: Counterfactual Intervention Enhancing Graph Contrastive Learning in Implicit Feedback." In *ICASSP 2025-2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 1–5. IEEE.

Igelstrm, E., P. Craig, J. Lewsey, J. Lynch, A. Pearce, and S. V. Katikireddi. 2022. "Causal Inference and Effect Estimation Using Observational Data." *Journal of Epidemiology and Community Health* 76, no. 11: 960–966.

Iglesias-Suarez, F., P. Gentine, B. Solino-Fernandez, et al. 2024. "Causally-Informed Deep Learning to Improve Climate Models and Projections." *Journal of Geophysical Research: Atmospheres* 129, no. 4: e2023JD039202.

Jagadish, H. V., J. Gehrke, A. Labrinidis, et al. 2014. "Big Data and Its Technical Challenges." *Communications of the ACM* 57, no. 7: 86–94.

Ji, J., W. Zhang, J. Wang, and C. Huang. 2025. "Seeing the Unseen: Learning Basis Confounder Representations for Robust Traffic Prediction." *arXiv Preprint* arXiv:2311.12472.

Jiang, S., and Y. Sun. 2022. "Estimating Causal Effects on Networked Observational Data via Representation Learning," in *Proceedings of the 31st ACM International Conference on Information & Knowledge Management*, pp. 852–861.

Jiang, W., and J. Luo. 2022. "Graph Neural Network for Traffic Forecasting: A Survey." *Expert Systems With Applications* 207: 117921.

Jiang, W., J. Zhang, Y. Li, et al. 2024. "Advancing Storm Surge Forecasting From Scarce Observation Data: A Causal-Inference Based Spatio-Temporal Graph Neural Network Approach." *Coastal Engineering* 190: 104512.

Jiao, L., Y. Wang, X. Liu, et al. 2024. "Causal Inference Meets Deep Learning: A Comprehensive Survey." *Research* 7: 467.

Jin, W., Y. Li, H. Xu, et al. 2021. "Adversarial Attacks and Defenses on Graphs." *ACM SIGKDD Explorations Newsletter* 22, no. 2: 19–34.

Jin, Y., J. Li, Z. Lian, C. Jiao, and X. Hu. 2022. "Supporting Medical Relation Extraction via Causality-Pruned Semantic Dependency Forest," in *Proceedings of the 29th International Conference on Computational Linguistics*, pp. 2450–2460.

Jin, Z., Y. Chen, F. Leeb, et al. 2023. "Cladder: Assessing Causal Reasoning in Language Models." *Advances in Neural Information Processing Systems* 36: 31038–31065.

Job, S., X. Tao, T. Cai, et al. 2025a. "Causal Integration in Graph Neural Networks Toward Enhanced Classification: Benchmarking and Advancements for Robust Performance." *World Wide Web* 28, no. 3: 30.

Job, S., X. Tao, T. Cai, et al. 2025b. "Hebcgnn: Hebbian-Enabled Causal Classification Integrating Dynamic Impact Valuing." *Knowledge-Based Systems* 311: 113094.

Kaggle. 2017. New York City Bike Share Dataset [Database]. https://www.kaggle.com/datasets/akkithetechie/new-york-city-bike-share-dataset/ [Accessed April 2, 2025].

Kalla, A., S. Mukhopadhyay, Z. Ralte, and I. Kar. 2023. "Exploring the Impact of Motif-Driven Causal Temporal Analysis Using Graph Neural Network in Improving Large Language Model Performance for Pharmacovigilance," in *Proceedings of the 2023 9th International Conference on Advanced Computing and Communication Systems* (*ICACCS*), IEEE, pp.1769–1776.

Kazius, J., R. McGuire, and R. Bursi. 2005. "Derivation and Validation of Toxicophores for Mutagenicity Prediction." *Journal of Medicinal Chemistry* 48, no. 1: 312–320.

Keith, K. A., D. Jensen, and B. O'Connor. 2020. "Text and Causal Inference: A Review of Using Text to Remove Confounding From Causal Estimates." *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pp. 5332–5344.

Kennedy, J. J., N. A. Rayner, R. O. Smith, D. E. Parker, and M. Saunby. 2011. "Reassessing Biases and Other Uncertainties in Sea Surface Temperature Observations Measured In Situ Since 1850: 1. Measurement and Sampling Uncertainties." *Journal of Geophysical Research* 116, no. D14: D14103.

Kipf, T. N., and M. Welling. 2016. "Variational Graph Auto-Encoders." *Stat* 21: 1050.

Kipf, T. N., and M. Welling. 2017. "Semi-Supervised Classification With Graph Convolutional Networks," in *Proceedings of the 5th International Conference on Learning Representations* (ICLR 2017).

Koelstra, S., C. Muhl, M. Soleymani, et al. 2011. "Deap: A Database for Emotion Analysis; Using Physiological Signals." *IEEE Transactions on Affective Computing* 3, no. 1: 18–31.

Kong, W., M. Qiu, M. Li, X. Jin, and L. Zhu. 2022. "Causal Graph Convolutional Neural Network for Emotion Recognition." *IEEE Transactions on Cognitive and Developmental Systems* 15, no. 4: 1686–1693.

Kooverjee, N., S. James, and T. Van Zyl. 2022. "Investigating Transfer Learning in Graph Neural Networks." *Electronics* 11, no. 8: 1202.

Korb, K. B. 2003. *Bayesian Artificial Intelligence*. Chapman Hall/CRC.

Krishna, R., Y. Zhu, O. Groth, et al. 2017. "Visual Genome: Connecting Language and Vision Using Crowdsourced Dense Image Annotations." *International Journal of Computer Vision* 123: 32–73.

Kuang, K., L. Li, Z. Geng, et al. 2020. "Causal Inference." *Engineering* 6, no. 3: 253–263.

Kuhn, M., I. Letunic, L. J. Jensen, and P. Bork. 2016. "The Sider Database of Drugs and Side Effects." *Nucleic Acids Research* 44, no. D1: D1075–D1079.

Kurutach, T., A. Tamar, G. Yang, S. J. Russell, and P. Abbeel. 2018. "Learning Plannable Representations With Causal Infogan." *Advances in Neural Information Processing Systems*, 31.

Lai, G. 2016. Exchangerate Dataset [Database]. https://www.kaggle.com/datasets/federalreserve/exchange-rates [Accessed February 4, 2025].

Land Transport Authority, SG-TAXI. 2016. SG-TAXI Dataset [Database]. https://datamall.lta.gov.sg/ [Accessed January 27, 2025].

Lauritzen, S. L., and D. J. Spiegelhalter. 1988. "Local Computations With Probabilities on Graphical Structures and Their Application to Expert Systems." *Journal of the Royal Statistical Society: Series B: Methodological* 50, no. 2: 157–194.

LeCun, Y., L. Bottou, Y. Bengio, and P. Haffner. 1998. "Gradient-Based Learning Applied to Document Recognition." *Proceedings of the IEEE* 86, no. 11: 2278–2324.

Lee, Y., and S. W. Han. 2023. "Cagcn: Causal Attention Graph Convolutional Network Against Adversarial Attacks." *Neurocomputing* 538: 126187.

Leskovec, J., L. A. Adamic, and B. A. Huberman. 2007. "The Dynamics of Viral Marketing." *ACM Transactions on the Web* 1, no. 1: 5.

Li, J., R. Guo, C. Liu, and H. Liu. 2019. "Adaptive Unsupervised Feature Selection on Attributed Networks," in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pp. 92–100.

Li, J., Y. Shi, H. Li, and B. Yang. 2022a. "Tc-Gatn: Temporal Causal Graph Attention Networks With Nonlinear Paradigm for Multivariate Time-Series Forecasting in Industrial Processes." *IEEE Transactions on Industrial Informatics* 19, no. 6: 7592–7601.

Li, L., H. Wang, W. Zhang, and A. Coster. 2024a. "Stg-Mamba: Spatial-Temporal Graph Learning Via Selective State Space Model." *arXiv Preprint* arXiv:2403.12418.

Li, M., M. Qiu, W. Kong, L. Zhu, and Y. Ding. 2023a. "Fusion Graph Representation of Eeg for Emotion Recognition." *Sensors* 23, no. 3: 1404.

Li, M., P. Tong, M. Li, Z. Jin, J. Huang, and X.-S. Hua. 2021. "Traffic Flow Prediction With Vehicle Trajectories." *Proceedings of the AAAI Conference on Artificial Intelligence* 35, no. 1: 294–302.

Li, N., J. Dong, L. Liu, H. Li, and J. Yan. 2023b. "A Novel Emd and Causal Convolutional Network Integrated With Transformer for Ultra Short-Term Wind Power Forecasting." *International Journal of Electrical Power & Energy Systems* 154: 109470.

Li, P., Y. Yang, M. Pagnucco, and Y. Song. 2022b. "Explainability in Graph Neural Networks: An Experimental Survey." *arXiv Preprint* arXiv:2203.09258.

Li, X., R. Guo, J. Lu, T. Chen, and X. Qian. 2023c. "Causality-Driven Graph Neural Network for Early Diagnosis of Pancreatic Cancer in Non-Contrast Computerized Tomography." *IEEE Transactions on Medical Imaging* 42, no. 6: 1656–1667.

Li, X., R. Guo, H. Zhu, T. Chen, and X. Qian. 2024b. "A Causality-Informed Graph Intervention Model for Pancreatic Cancer Early Diagnosis." *IEEE Transactions on Artificial Intelligence* 1, no. 1: 1–11.

Li, Y., A. Torralba, A. Anandkumar, D. Fox, and A. Garg. 2020. "Causal Discovery in Physical Systems From Videos." *Advances in Neural Information Processing Systems* 33: 9180–9192.

Li, Y., J. Yao, J. Song, et al. 2024c. "Investigation of Causal Public Opinion Indexes for Price Fluctuation in Vegetable Marketing." *Computers and Electrical Engineering* 116: 109227.

Li, Y., R. Yu, C. Shahabi, and Y. Liu. 2018. "Diffusion Convolutional Recurrent Neural Network: Data-Driven Traffic Forecasting." International Conference on Learning Representations. https://openreview.net/forum?id=SJiHXGWAZ.

Li, Y., R. Zemel, M. Brockschmidt, and D. Tarlow. 2016. "Gated Graph Sequence Neural Networks," in *4th International Conference on Learning Representations* (*ICLR2016*).

Lin, J., Z. Li, Z. Li, L. Bai, R. Zhao, and C. Zhang. 2023. "Dynamic Causal Graph Convolutional Network for Traffic Prediction," in *Proceedings of the 2023 IEEE 19th International Conference on Automation Science and Engineering* (*CASE*), pp. 1–8.

Lin, W., H. Lan, and B. Li. 2021. "Generative Causal Explanations for Graph Neural Networks," in *Proceedings of the International Conference on Machine Learning* (PMLR), pp. 6666–6679.

Lin, W., H. Lan, H. Wang, and B. Li. 2022. "Orphicx: A Causality-Inspired Latent Variable Model for Interpreting Graph Neural Networks," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 13729–13738.

Lindsay, G. W. 2024. "Grounding Neuroscience in Behavioral Changes Using Artificial Neural Networks." *Current Opinion in Neurobiology* 84: 102816.

Liu, B., D. Wang, X. Yang, et al. 2022. "Show, Deconfound and Tell: Image Captioning With Causal Inference," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 18041–18050.

Liu, C., G. Cui, and S. Liu. 2022b. "Cgcnimp: A Causal Graph Convolutional Network for Multivariate Time Series Imputation." *PeerJ Computer Science* 8: e966.

Liu, J., H. Lin, X. Wang, L. Wu, S. Garg, and M. M. Hassan. 2024a. "Reliable Trajectory Prediction in Scene Fusion Based on Spatio-Temporal Structure Causal Model." *Information Fusion* 107: 102309.

Liu, J., S. Zheng, and C. Wang. 2023. "Causal Graph Attention Network With Disentangled Representations for Complex Systems Fault Detection." *Reliability Engineering & System Safety* 235: 109232.

Liu, L., H. Lu, M. Whelan, Y. Chen, and X. Ding. 2024b. "Cignn: A Causality-Informed and Graph Neural Network Based Framework for Cuffless Continuous Blood Pressure Estimation." *IEEE Journal of Biomedical and Health Informatics* 28, no. 5: 2674–2686.

Liu, R., Q. Zhang, D. Lin, W. Zhang, and S. X. Ding. 2024c. "Causal Intervention Graph Neural Network for Fault Diagnosis of Complex Industrial Processes." *Reliability Engineering & System Safety* 251: 110328.

Long, Q., L. Xu, Z. Fang, and G. Song. 2021. "HGK-GNN: Heterogeneous Graph Kernel Based Graph Neural Networks," in *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*, pp. 1129–1138.

Lucic, A., M. A. ter Hoeve, G. Tolomei, M. de Rijke, and F. Silvestri. 2022. "Cf-Gnnexplainer: Counterfactual Explanations for Graph Neural Networks," in *International Conference on Artificial Intelligence and Statistics* (PMLR), pp. 4499–4511.

Luo, D., W. Cheng, D. Xu, et al. 2020. "Parameterized Explainer for Graph Neural Network." *Advances in Neural Information Processing Systems* 33: 19620–19631.

Ma, J., R. Guo, S. Mishra, A. Zhang, and J. Li. 2022. "Clear: Generative Counterfactual Explanations on Graphs." *Advances in Neural Information Processing Systems* 35: 25895–25907.

Ma, Y., and V. Tresp. 2021. "Causal Inference Under Networked Interference and Intervention Policy Enhancement," in *Proceedings of the International Conference on Artificial Intelligence and Statistics* (PMLR), pp. 3700–3708.

Mandal, D., S. Medya, B. Uzzi, and C. Aggarwal. 2022. "Metalearning With Graph Neural Networks: Methods and Applications." *ACM SIGKDD Explorations Newsletter* 23, no. 2: 13–22.

Mannering, F., C. R. Bhat, V. Shankar, and M. Abdel-Aty. 2020. "Big Data, Traditional Data and the Tradeoffs Between Prediction and Causality in Highway-Safety Analysis." *Analytic Methods in Accident Research* 25: 100113.

Marbach, D., T. Schaffter, C. Mattiussi, and D. Floreano. 2009. "Generating Realistic In Silico Gene Networks for Performance Assessment of Reverse Engineering Methods." *Journal of Computational Biology* 16, no. 2: 229–239.

Mathur, A. P., and N. O. Tippenhauer. 2016. "Swat: A Water Treatment Testbed for Research and Training on Ics Security." In *2016 International Workshop on Cyber-Physical Systems for Smart Water Networks (CySWater)*, 31–36. IEEE.

McCallum, A. K., K. Nigam, J. Rennie, and K. Seymore. 2000. "Automating the Construction of Internet Portals With Machine Learning." *Information Retrieval* 3, no. 1: 127–163.

Meng, Z., Y. Li, P. Zhao, Y. Yu, and I. King. 2023. "Meta-Learning With Motif-Based Task Augmentation for Few-Shot Molecular Property Prediction." In *Proceedings of the 2023 SIAM International Conference on Data Mining (SDM)*, 811–819. SIAM.

Metsch, J. M., A. Saranti, A. Angerschmid, et al. 2024. "Clarus: An Interactive Explainable Ai Platform for Manual Counterfactuals in Graph Neural Networks." *Journal of Biomedical Informatics* 150: 104600.

Miao, J., C. Chen, F. Liu, H. Wei, and P.-A. Heng. 2023. "Caussl: Causality-Inspired Semi-Supervised Learning for Medical Image Segmentation," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 21426–21437.

Mirza, P. 2014. "Extracting Temporal and Causal Relations Between Events," in *Proceedings of the ACL 2014 Student Research Workshop*, pp. 10–17.

Moghimifar, F., G. Haffari, and M. Baktashmotlagh. 2020. "Domain Adaptative Causality Encoder." In *Proceedings of the 18th Annual Workshop of the Australasian Language Technology Association*, 1–10. Australasian Language Technology Association. aclanthology.org/2020.alta-1.1/.

Monken, A., F. Haberkorn, M. Gopinath, L. Freeman, and F. A. Batarseh. 2021. Graph Neural Networks for Modeling Causality in International Trade. The International FLAIRS Conference Proceedings, 34. https://doi.org/10.32473/flairs.v34i1.128485.

Monti, F., D. Boscaini, J. Masci, E. Rodolà, J. Svoboda, and M. M. Bronstein. 2017. "Geometric Deep Learning on Graphs and Manifolds Using Mixture Model CNNs," in *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition* (*CVPR*), 21–26 July 2017, pp. 5425–5434. https://doi.org/10.1109/CVPR.2017.576.

Morice, C. P., J. J. Kennedy, N. A. Rayner, et al. 2021. "An Updated Assessment of Near-Surface Temperature Change From 1850: The hadcrut5 Data Set." *Journal of Geophysical Research: Atmospheres* 126, no. 3: e2019JD032361.

Nicolson, A., and K. K. Paliwal. 2020. "Masked Multi-Head Self-Attention for Causal Speech Enhancement." *Speech Communication* 125: 80–96.

Nogueira, A. R., A. Pugnana, S. Ruggieri, D. Pedreschi, and J. Gama. 2022. "Methods and Tools for Causal Discovery and Causal Inference." *WIREs Data Mining and Knowledge Discovery* 12, no. 2: e1449.

Norouzi, R., B. Kleinberg, J. Vermunt, et al. 2024. Capturing Causal Claims: A Fine Tuned Text Mining Model for Extracting Causal Sentences From Social Science Papers. https://doi.org/10.31234/osf.io/kwtpm.

Ntemi, M., I. Sarridis, and C. Kotropoulos. 2022. "An Autoregressive Graph Convolutional Long Short-Term Memory Hybrid Neural Network for Accurate Prediction of Covid-19 Cases." *IEEE Transactions on Computational Social Systems* 10, no. 2: 724–735.

Pearl, J. 1995. "Causal Diagrams for Empirical Research." *Biometrika* 82, no. 4: 669–688.

Pearl, J. 2000. *Causality: Models, Reasoning, and Inference*. 2nd ed. Cambridge University Press.

Pearl, J. 2009. *Causality*. 2nd ed. Cambridge University Press.

Pearl, J. 2010. "An Introduction to Causal Inference." *International Journal of Biostatistics* 6, no. 2: 7.

Peters, J., S. Bauer, and N. Pfister. 2022. "Causal Models for Dynamical Systems," in *Probabilistic and Causal Inference*: *The Works of Judea Pearl*. pp. 671–690.

Phu, M. T., and T. H. Nguyen. 2021. Graph Convolutional Networks for Event Causality Identification With Rich Document-Level Structures, in *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics*: *Human Language Technologies*, pp. 3480–3490.

Pont-Tuset, J., F. Perazzi, S. Caelles, P. Arbeláez, A. Sorkine-Hornung, and L. van Gool. 2017. "The 2017 Davis Challenge on Video Object Segmentation." *arXiv Preprint* arXiv:1704.00675.

Qarkaxhija, L., V. Perri, and I. Scholtes. 2022. De Bruijn Goes Neural: Causality-Aware Graph Neural Networks for Time Series Data on Dynamic Graphs. Learning on Graphs Conference (PMLR), pp. 51:1–51:21.

Qin, Y., D. Song, H. Chen, W. Cheng, G. Jiang, and G. Cottrell. 2017. "A Dual-Stage Attention-Based Recurrent Neural Network for Time Series Prediction," in *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence* (IJCAI-17), pp. 2627–2633. https://doi.org/10.24963/ijcai.2017/366.

Qiu, J., Q. Du, K. Yin, S.-L. Zhang, and C. Qian. 2020. "A Causality Mining and Knowledge Graph Based Method of Root Cause Diagnosis for Performance Anomaly in Cloud Applications." *Applied Sciences* 10, no. 6: 2166.

Rakesh, V., R. Guo, R. Moraffah, N. Agarwal, and H. Liu. 2018. "Linked Causal Variational Autoencoder for Inferring Paired Spillover Effects" in *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*, pp. 1679–1682.

Rao, J., J. Xie, H. Lin, S. Zheng, Z. Wang, and Y. Yang. 2024. "Incorporating Retrieval-Based Causal Learning With Information Bottlenecks for Interpretable Graph Neural Networks." arXiv preprint arXiv:2402.04710.

Ritchie, H., E. Mathieu, L. Rodés-Guirao, et al. 2020. Coronavirus Pandemic (Covid-19). Our World in Data, 2022.

Rohrer, S. G., and K. Baumann. 2009. "Maximum Unbiased Validation (Muv) Data Sets for Virtual Screening Based on Pubchem Bioactivity Data." *Journal of Chemical Information and Modeling* 49, no. 2: 169–184.

Rubin, D. B. 2005. "Causal Inference Using Potential Outcomes: Design, Modeling, Decisions." *Journal of the American Statistical Association* 100, no. 469: 322–331.

Ruiz-Crcel, C., Y. Cao, D. Mba, L. Lao, and R. T. Samuel. 2015. "Statistical Process Monitoring of a Multiphase Flow Facility." *Control Engineering Practice* 42: 74–88.

Sachs, K., O. Perez, D. Pe'er, D. A. Lauffenburger, and G. P. Nolan. 2005. "Causal Protein-Signaling Networks Derived From Multiparameter Single-Cell Data." *Science* 308, no. 5721: 523–529.

Sadaoui, F., and S. B. Jabeur. 2023. "Analyzing the Influence of Geopolitical Risks on European Power Prices Using a Multiresolution Causal Neural Network." *Energy Economics* 124: 106793.

Schnake, T., O. Eberle, J. Lederer, et al. 2021. "Higher-Order Explanations of Graph Neural Networks via Relevant Walks." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 44, no. 11: 7581–7596.

Sen, P., G. Namata, M. Bilgic, L. Getoor, B. Galligher, and T. Eliassi-Rad. 2008. "Collective Classification in Network Data." *AI Magazine* 29, no. 3: 93.

Sheth, P., R. Moraffah, T. S. Kumarage, A. Chadha, and H. Liu. 2024. "Causality Guided Disentanglement for Cross-Platform Hate Speech Detection," in *Proceedings of the 17th ACM International Conference on Web Search and Data Mining*, pp. 626–635.

Sheth, P., A. Mosallanezhad, K. Ding, et al. 2023. "Streams: Towards Spatio-Temporal Causal Discovery With Reinforcement Learning for Streamflow Rate Prediction," *in Proceedings of the 32nd ACM International Conference on Information and Knowledge Management* (*CIKM 2023*), pp. 4815–4821.

Si, Z., Z. Sun, X. Zhang, et al. 2023. "Enhancing Recommendation With Search Data in a Causal Learning Manner." *ACM Transactions on Information Systems* 41, no. 4: 1–31.

Singh, B. 2021. "Predicting Airline Passengers' Loyalty Using Artificial Neural Network Theory." *Journal of Air Transport Management* 94: 102080.

Sinha, A., Z. Shen, Y. Song, et al. 2015. "An Overview of Microsoft Academic Service (MAS) and Applications," in *Proceedings of the 24th International Conference on World Wide Web*, pp. 243–246. https://doi.org/10.1145/2740908.2742839.

Smith, T. M., R. W. Reynolds, T. C. Peterson, and J. Lawrimore. 2008. "Improvements to NOAA's Historical Merged Land–Ocean Surface Temperature Analysis (1880–2006)." *Journal of Climate* 21, no. 10: 2283–2296.

Socher, R., A. Perelygin, J. Wu, et al. 2013. "Recursive Deep Models for Semantic Compositionality Over a Sentiment Treebank," in Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing, pp. 1631–1642.

Spirtes, P., C. Glymour, and R. Scheines. 2001. *Causation, Prediction, and Search*. MIT Press. https://doi.org/10.7551/mitpress/1754.001.0001.

Sui, Y., W. Mao, S. Wang, et al. 2024. "Enhancing Out-Of-Distribution Generalization on Graphs via Causal Attention Learning." *ACM Transactions on Knowledge Discovery From Data* 18, no. 5: 1–24.

Sui, Y., X. Wang, J. Wu, M. Lin, X. He, and T.-S. Chua. 2022a. "Causal Attention for Interpretable and Generalizable Graph Classification," in *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pp. 1696–1705.

Sui, Y., X. Wang, J. Wu, A. Zhang, and X. He. 2022b. "Adversarial Causal Augmentation for Graph Covariate Shift." *arXiv preprint* arXiv:2211.02843.

Sun, Z., Y. Li, Q. He, H. Xu, W. Wang, and X. Liu. 2024. "Causality Enhanced Global-Local Graph Neural Network for Bioprocess Factor Forecasting." *IEEE Transactions on Industrial Informatics* 20: 12428–12438.

Takac, L., and M. Zabovsky. 2012. "Data Analysis in Public Social Networks," in *International Scientific Conference and International Workshop Present Day Trends of Innovations*, pp. 1–6.

Tan, F. A., A. Hrriyetolu, T. Caselli, et al. 2022a. "The Causal News Corpus: Annotating Causal Relations in Event Sentences From News." In *Proceedings of the Thirteenth Language Resources and Evaluation Conference*, 2298–2310. European Language Resources Association. aclanthology.org/2022.lrec-1.246.

Tan, J., S. Geng, Z. Fu, et al. 2022b. "Learning and Evaluating Graph Neural Network Explanations Based on Counterfactual and Factual Reasoning," in *Proceedings of the ACM Web Conference* 2022, pp. 1018–1027.

Tanaka, S. C., A. Yamashita, N. Yahata, et al. 2021. "A Multi-Site, Multi-Disorder Resting-State Magnetic Resonance Image Database." *Scientific Data* 8, no. 1: 227.

Tang, X., C. Zhang, R. Guo, X. Yang, and X. Qian. 2023. "A Causality-Aware Graph Convolutional Network Framework for Rigidity Assessment in Parkinsonians." *IEEE Transactions on Medical Imaging* 43, no. 1: 229–240. https://doi.org/10.1109/TMI.2023.3294182.

Tao, S., Q. Cao, H. Shen, Y. Wu, B. Xu, and X. Cheng. 2024. "Idea: Invariant Causal Defense for Graph Adversarial Robustness." *Information Sciences* 680: 121171.

Toutanova, K., D. Chen, P. Pantel, H. Poon, P. Choudhury, and M. Gamon. 2015. "Representing Text for Joint Embedding of Text and Knowledge Bases," in *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, pp. 1499–1509.

Trust, P., R. Minghim, E. Milios, and K. Provia. 2022. "Ggnn@ Causal News Corpus 2022: Gated Graph Neural Networks for Causal Event Classification From Social-Political News Articles," in *Proceedings of the 5th Workshop on Challenges and Applications of Automated Extraction of Socio-Political Events From Text* (*CASE*), pp. 85–90.

UMassTraceRepository. 2017. Umass Smart Dataset [Database]. https://traces.cs.umass.edu/docs/traces/smartstar/ [Accessed February 4, 2025].

van Essen, D. C., S. M. Smith, D. M. Barch, et al. 2013. "The Wu-Minn Human Connectome Project: An Overview." *NeuroImage* 80: 62–79.

Varga, V., and A. Lrincz. 2021. "Fast Interactive Video Object Segmentation With Graph Neural Networks," in *Proceedings of the 2021 International Joint Conference on Neural Networks* (*IJCNN*), pp. 1–10.

Velickovic, P., G. Cucurull, A. Casanova, et al. 2017. "Graph Attention Networks." *Stat* 1050, no. 20: 10–48550.

Vig, J., S. Gehrmann, Y. Belinkov, et al. 2020. "Investigating Gender Bias in Language Models Using Causal Mediation Analysis." *Advances in Neural Information Processing Systems* 33: 12388–12401.

Wale, N., I. A. Watson, and G. Karypis. 2008. "Comparison of Descriptor Spaces for Chemical Compound Retrieval and Classification." *Knowledge and Information Systems* 14: 347–375.

Wang, D., Z. Chen, J. Ni, et al. 2023c. "Hierarchical Graph Neural Networks for Causal Discovery and Root Cause Localization." *arXiv Preprint* arXiv:2302.01987.

Wang, H., R. Liu, S. X. Ding, Q. Hu, Z. Li, and H. Zhou. 2023d. "Causal-Trivial Attention Graph Neural Network for Fault Diagnosis of Complex Industrial Processes." *IEEE Transactions on Industrial Informatics* 20, no. 2: 1987–1996. https://doi.org/10.1109/TII.2023.3282979.

Wang, H., Y. Pan, M. Ma, and P. Wang. 2022c. "When Dynamic Causality Comes to Graph-Temporal Neural Network." In *2022 International Joint Conference on Neural Networks (IJCNN)*, 1–9. IEEE. https://doi.org/10.1109/IJCNN55064.2022.9892477.

Wang, L., A. Adiga, J. Chen, A. Sadilek, S. Venkatramanan, and M. Marathe. 2022a. "CausalGNN: Causal-Based Graph Neural Networks for Spatio-Temporal Epidemic Forecasting," in *Proceedings of the AAAI Conference on Artificial Intelligence*, 36.11.12191–12199.

Wang, L., M. Xu, Q. Zhang, Y. Shi, and Q. Wu. 2025a. "Causal Disentanglement for Regulating Social Influence Bias in Social Recommendation." *Neurocomputing* 618: 129133.

Wang, S., Q. Zhao, Y. Han, and J. Wang. 2023a. "Root Cause Diagnosis for Process Faults Based on Multisensor Time-Series Causality Discovery." *Journal of Process Control* 122: 27–40.

Wang, X., Y. Wu, A. Zhang, F. Feng, X. He, and T.-S. Chua. 2022b. "Reinforced Causal Explainer for Graph Neural Networks." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 45, no. 2: 2297–2309.

Wang, Y., Z. Chu, X. Ouyang, et al. 2024. "Llmrg: Improving Recommendations Through Large Language Model Reasoning Graphs." *Proceedings of the AAAI Conference on Artificial Intelligence* 38, no. 17: 19189–19196.

Wang, Y., S. He, Q. Luo, et al. 2025b. "Causal Invariant Geographic Network Representations With Feature and Structural Distribution Shifts." *Future Generation Computer Systems* 169: 107814. https://doi.org/10.1016/j.future.2025.107814.

Wang, Y., L. Rui, J. Ma, and Q. Jin. 2023b. "A Short-Term Residential Load Forecasting Scheme Based on the Multiple Correlation-Temporal Graph Neural Networks." *Applied Soft Computing* 146: 110629.

Wang, Z., Z. Yin, Y. Zhang, et al. 2025c. "Graph Fairness via Authentic Counterfactuals: Tackling Structural and Causal Challenges." *ACM SIGKDD Explorations Newsletter* 26, no. 2: 89–98.

Wei, B., W. Zeng, Y. Shi, and H. Zhang. 2025. "Causal Invariance Guides Interpretable Graph Contrastive Learning in Fmri Analysis." *Alexandria Engineering Journal* 117: 635–647.

Wein, S., W. M. Malloni, A. M. Tomé, et al. 2021. "A Graph Neural Network Framework for Causal Inference in Brain Networks." *Scientific Reports* 11, no. 1: 8061.

Wen, H., T. Chen, L. K. Chai, S. Sadiq, K. Zheng, and H. Yin. 2023. "To Predict or to Reject: Causal Effect Estimation With Uncertainty on Networked Data." In *2023 IEEE International Conference on Data Mining (ICDM)*, 1415–1420. IEEE.

Wu, B., K.-M. Chao, and Y. Li. 2024a. "Heterogeneous Graph Neural Networks for Fraud Detection and Explanation in Supply Chain Finance." *Information Systems* 121: 102335.

Wu, H., C. Geng, and H. Fang. 2023. "Causality and Correlation Graph Modeling for Effective and Explainable Session-Based Recommendation." *ACM Transactions on the Web* 18, no. 1: 1–25.

Wu, L., P. Cui, J. Pei, L. Zhao, and X. Guo. 2022. "Graph Neural Networks: Foundation, Frontiers and Applications," in *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pp. 4840–4841.

Wu, P., and J. Liu. 2021. "Learning Causal Temporal Relation and Feature Discrimination for Anomaly Detection." *IEEE Transactions on Image Processing* 30: 3513–3527.

Wu, Y., Y. Liu, Z. Zhao, et al. 2024b. "De-Biased Attention Supervision for Text Classification With Causality." *Proceedings of the AAAI Conference on Artificial Intelligence* 38, no. 17: 19279–19287.

Wu, Y., and J. Zhou. 2023. "CogTrans: A Cognitive Transfer Learning-Based Self-Attention Mechanism Architecture for Knowledge Graph Reasoning," Annual Meeting of the Cognitive Science Society, 45.

Wu, Z., S. Pan, F. Chen, G. Long, C. Zhang, and S. Y. Philip. 2020. "A Comprehensive Survey on Graph Neural Networks." *IEEE Transactions on Neural Networks and Learning Systems* 32, no. 1: 4–24.

Wu, Z., B. Ramsundar, E. N. Feinberg, et al. 2018. "Moleculenet: A Benchmark for Molecular Machine Learning." *Chemical Science* 9, no. 2: 513–530.

Xia, R., and Z. Ding. 2019. "Emotion-Cause Pair Extraction: A New Task to Emotion Analysis in Texts." In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, 1003–1012. Association for Computational Linguistics. https://doi.org/10.18653/v1/P19-1096.

Xia, Y., Y. Liang, H. Wen, et al. 2023. "Deciphering Spatio-Temporal Graph Forecasting: A Causal Lens and Treatment." *Advances in Neural Information Processing Systems* 36: 37068–37088.

Xiang, W., W. Yingxin, Z. An, H. Xiangnan, and C. Tat-seng. 2021. "Causal Screening to Interpret Graph Neural Networks," in *International Conference on Learning Representations*. https://www.openreview.net/forum?id=nzKv5vxZfge.

Xiao, F., S. Chen, J. Yang, et al. 2025. "Grain: Graph Neural Network and Reinforcement Learning Aided Causality Discovery for Multi-Step Attack Scenario Reconstruction." *Computers & Security* 148: 104180.

Xiong, Y., and H. Wang. 2024. "Spatio-Temporal Contextual Conditions Causality and Spread Delay-Aware Modeling for Traffic Flow Prediction." *IEEE Access* 12: 21250–21261.

Xu, H., Y. Huang, Z. Duan, J. Feng, and P. Song. 2020. "Multivariate Time Series Forecasting Based on Causal Inference With Transfer Entropy and Graph Neural Network." *arXiv Preprint* arXiv:2005.01185.

Xu, K., W. Hu, J. Leskovec, and S. Jegelka. 2019. "How Powerful Are Graph Neural Networks?," in *International Conference on Learning Representations*. openreview.net/forum?id=ryGs6iA5Km.

Xu, M., Y. Ma, R. Li, G. Qi, X. Meng, and H. Jin. 2023. "Traffnet: Learning Causality of Traffic Generation for Road Network Digital Twins." *arXiv Preprint* arXiv:2303.15954.

Xu, Z., X. Wei, J. Hao, et al. 2024. "Dgformer: A Physics-Guided Station Level Weather Forecasting Model With Dynamic Spatial-Temporal Graph Neural Network." *GeoInformatica* 28, no. 3: 499–533.

Yan, C.-G., X. Chen, L. Li, et al. 2019. "Reduced Default Mode Network Functional Connectivity in Patients With Recurrent Major Depressive Disorder." *Proceedings of the National Academy of Sciences of the United States of America* 116, no. 18: 9078–9083.

Yanardag, P., and S. V. N. Vishwanathan. 2015. "Deep Graph Kernels," in *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 1365–1374.

Yang, D., G. Yu, J. Wang, Z. Yan, and M. Guo. 2023. "Causal Discovery by Graph Attention Reinforcement Learning," in *Proceedings of the 2023 SIAM International Conference on Data Mining* (*SDM*), pp. 28–36.

Yang, J., H. Xiong, H. Zhang, M. Hu, and N. An. 2022. "Causal Pattern Representation Learning for Extracting Causality From Literature," in *Proceedings of the 2022 5th International Conference on Machine Learning and Natural Language Processing*, pp. 229–233.

Yao, L., Z. Chu, S. Li, Y. Li, J. Gao, and A. Zhang. 2021. "A Survey on Causal Inference." *ACM Transactions on Knowledge Discovery From Data* 15, no. 5: 1–46.

Yelp. 2022. Yelp Database [Database]. https://www.yelp.com/dataset [Accessed February 4, 2025].

Ying, Z., D. Bourgeois, J. You, M. Zitnik, and J. Leskovec. 2019. "Gnnexplainer: Generating Explanations for Graph Neural Networks." *Advances in Neural Information Processing Systems* 32: 9240–9251.

Yu, D., Q. Li, H. Yin, and G. Xu. 2023. "Causality-Guided Graph Learning for Session-Based Recommendation," in *Proceedings of the 32nd ACM International Conference on Information and Knowledge Management*, pp. 3083–3093.

Yu, Y., J. Chen, T. Gao, and M. Yu. 2019. "DAG-GNN: Dag Structure Learning With Graph Neural Networks," in *International Conference on Machine Learning*, pp. 7154–7163.

Yuan, H., H. Yu, S. Gui, and S. Ji. 2022. "Explainability in Graph Neural Networks: A Taxonomic Survey." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 45, no. 5: 5782–5799.

Yuan, X., K. Chen, W. Zuo, and Y. Zhang. 2023. "Tc-Gat: Graph Attention Network for Temporal Causality Discovery." In *2023 International Joint Conference on Neural Networks (IJCNN)*, 1–8. IEEE. https://doi.org/10.1109/IJCNN54540.2023.10191712.

Zečević, M., D. S. Dhami, P. Veličković, and K. Kersting. 2021. "Relating Graph Neural Networks to Structural Causal Models." *arXiv Preprint* arXiv:2109.04173.

Zeyu, H., L. Yan, F. Wendi, Z. Wei, F. Alenezi, and P. Tiwari. 2023. "Causal Embedding of User Interest and Conformity for Long-Tail Session-Based Recommendations." *Information Sciences* 644: 119167.

Zhai, P., Y. Yang, and C. Zhang. 2023. "Causality-Based Ctr Prediction Using Graph Neural Networks." *Information Processing & Management* 60, no. 1: 103137.

Zhang, J., X. Zhang, G. Chen, and Q. Zhao. 2022e. "Granger-Causality-Based Multi-Frequency Band Eeg Graph Feature Extraction and Fusion for Emotion Recognition." *Brain Sciences* 12, no. 12: 1649.

Zhang, J., X. Zhang, and Q. Zhao. 2022a. "Improved Graph Convolutional Neural Networks Based on Granger Causality Analysis for EEG Emotion Recognition," in *International Conference on Computer Engineering and Artificial Intelligence* (*ICCEAI*), pp. 684–688.

Zhang, K., G. Chuai, J. Zhang, X. Chen, Z. Si, and S. Maimaiti. 2022b. "Dic-St: A Hybrid Prediction Framework Based on Causal Structure Learning for Cellular Traffic and Its Application in Urban Computing." *Remote Sensing* 14, no. 6: 1439.

Zhang, L., K. Fu, T. Ji, and C.-T. Lu. 2022c. "Granger Causal Inference for Interpretable Traffic Prediction," in *IEEE 25th International Conference on Intelligent Transportation Systems* (*ITSC*), pp. 1645–1651.

Zhang, S., H. Tong, J. Xu, and R. Maciejewski. 2019. "Graph Convolutional Networks: A Comprehensive Review." *Computational Social Networks* 6, no. 1: 1–23.

Zhang, T., H.-R. Shan, and M. A. Little. 2022f. "Causal Graphsage: A Robust Graph Method for Classification Based on Causal Sampling." *Pattern Recognition* 128: 108696.

Zhang, X., Q. Liu, and R. Han. 2025a. "Mmgcf: Generating Counterfactual Explanations for Molecular Property Prediction via Motif Rebuild." *Journal of Computer and Communications* 13, no. 1: 152–168.

Zhang, X., M. Wang, X. Zhuang, X. Zeng, and Q. Li. 2025b. "Cdea: Causality-Driven Dialogue Emotion Analysis via Llm." *Symmetry* 17, no. 4: 489.

Zhang, X., and M. Zitnik. 2020. "Gnnguard: Defending Graph Neural Networks Against Adversarial Attacks." *Advances in Neural Information Processing Systems* 33: 9263–9275.

Zhang, Y., H. Tong, Y. Xia, Y. Zhu, Y. Chi, and L. Ying. 2022g. "Batch Active Learning With Graph Neural Networks via Multi-Agent Deep Reinforcement Learning." *Proceedings of the AAAI Conference on Artificial Intelligence* 36, no. 8: 9118–9126.

Zhang, Z., X. Wang, Z. Zhang, H. Li, Z. Qin, and W. Zhu. 2022d. "Dynamic Graph Neural Networks Under Spatio-Temporal Distribution Shift." *Advances in Neural Information Processing Systems* 35: 6074–6089.

Zhao, T., G. Liu, D. Wang, W. Yu, and M. Jiang. 2022a. "Learning From Counterfactual Links for Link Prediction," in *International Conference on Machine Learning*, pp. 26911–26926.

Zhao, T., D. Luo, X. Zhang, and S. Wang. 2023a. "Towards Faithful and Consistent Explanations for Graph Neural Networks," in *Proceedings of the Sixteenth ACM International Conference on Web Search and Data Mining*, pp. 634–642.

Zhao, W., S. Zhang, B. Zhou, and B. Wang. 2023b. "Spatio-Temporal Causal Graph Attention Network for Traffic Flow Prediction in Intelligent Transportation Systems." *PeerJ Computer Science* 9: e1484.

Zhao, Y., Y. Yu, H. Wang, et al. 2022b. "Machine Learning in Causal Inference: Application in Pharmacovigilance." *Drug Safety* 45, no. 5: 459–476.

Zhao, Z., P. Wang, H. Wen, Y. Zhang, Z. Zhou, and Y. Wang. 2024. "A Twist for Graph Classification: Optimizing Causal Information Flow in Graph Neural Networks." *Proceedings of the AAAI Conference on Artificial Intelligence* 38, no. 15: 17042–17050.

Zheng, K., S. Yu, and B. Chen. 2024. "Ci-Gnn: A Granger Causality-Inspired Graph Neural Network for Interpretable Brain Network-Based Psychiatric Diagnosis." *Neural Networks* 172: 106147.

Zheng, W.-L., W. Liu, Y. Lu, B.-L. Lu, and A. Cichocki. 2018. "Emotionmeter: A Multimodal Framework for Recognizing Human Emotions." *IEEE Transactions on Cybernetics* 49, no. 3: 1110–1122.

Zhou, J., G. Cui, S. Hu, et al. 2020. "Graph Neural Networks: A Review of Methods and Applications." *AI Open* 1: 57–81.

Zhou, Z., Q. Huang, K. Yang, et al. 2023. "Maintaining the Status Quo: Capturing Invariant Relations for Ood Spatiotemporal Learning," in *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pp. 3603–3614.

Zhu, Q., S. Chen, T. Guo, Y. Lv, and W. Du. 2024. "A Spatio-Temporal Approach With Self-Corrective Causal Inference for Flight Delay Prediction." *IEEE Transactions on Intelligent Transportation Systems* 25, no. 12: 20820–20831. https://doi.org/10.1109/TITS.2024.3443261.