

1. ***Title: Improving Graph Neural Network Expressivity via Subgraph Isomorphism Counting***

2023 IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE

Index Terms: Graph classification, graph isomorphism, graph neural networks, graph substructures, network analysis, neural network expressivity

a. Research background:

GNNs have limitations in capturing graph structures.

b. Past methods, problems and motivations:

Traditional GNNs cannot capture graph structures well and cannot compute subgraph isomorphism.

c. Proposed research method:

Improving the expressive power of GNNs by introducing a subgraph isomorphism counting module. The SCGNN model is proposed for graph classification and molecular property prediction.

d. Performance of the method in the task:

The SCGNN model exhibits excellent performance on multiple benchmark datasets and can be used as a feature extractor for downstream tasks.

2. 2022 ICLR

Keywords: graph neural networks, subgraph isomorphism counting, edge-centric message passing, graph modulation

3. ***Title: A Comprehensive Survey on Graph Neural Networks***

2019

Keywords: Deep Learning, Graph Neural Networks, Graph Convolutional Networks, Graph Representation Learning, Graph Autoencoder, Network Embedding

This paper discusses graph neural networks, including recursive graph neural networks (RecGNN), convolutional graph neural networks (ConvGNN), graph autoencoders (GAE), and spatiotemporal graph neural networks (STGNN).

Past methods such as graph embedding and graph kernel methods have limitations. Graph neural networks are a type of deep learning model that combines graph and node features and can be used for tasks such as node classification.

This paper mainly introduces different types of graph neural networks and their roles in processing various types of graph data, as well as their performance on different datasets.

4. ***Title: A Survey on Subgraph Counting: Concepts, Algorithms, and Applications to Network Motifs and Graphlets***

2021

Keywords: Subgraph census, network motifs, graphlets

a. Research background:

Subgraph counting is a fundamental task in network analysis, which involves determining the frequency of small connected subgraphs in larger networks.

b. Past methods, problems and motivations:

Subgraph counting is a computationally complex task, and some efficient algorithms and strategies have been proposed to make it applicable to larger subgraphs and networks.

c. *Proposed research method:*

This paper provides a comprehensive overview of existing subgraph counting methods, including exact sequential algorithms, approximate methods sacrificing accuracy and execution time, and parallel strategies for addressing imbalanced search spaces. The authors identify and describe the main conceptual approaches, provide insights into their advantages and limitations, and offer pointers to existing implementations.

5. **Title:** *A Learned Sketch for Subgraph Counting*

SIGMOD 2021

Keywords: Subgraph counting, Deep learning

a. *Theoretical basis of the study:*

The theoretical basis of this study is the use of machine learning and deep learning techniques for subgraph counting.

b. *Method proposed in the paper:*

The paper proposes an Active Learning Subgraph Counting Schematic (ALSS) method, which uses a neural network regression model and an active learner. ALSS consists of two main components: a learning schematic for subgraph counting (LSS) and an active learner (AL). LSS learns the schematic using a neural network regression model, extracts features for each substructure, predicts the count of the query graph, and aggregates representations for all substructures. AL uses uncertainty sampling to select a subset of informative query graphs for enhanced training data.

6. **Title:** *Learned sketch for subgraph counting: a holistic approach*

VLDB Journal 2023

Keywords: Subgraph counting, Graph neural network, Active learning

a. *Importance of the work:*

The significance of this paper is to propose a machine learning method LSS for subgraph counting, which supports various types of node and/or edge labels, directed/undirected graphs, and homomorphic/subgraph isomorphism, while achieving efficient and accurate counting on large-scale graphs and better query plan recommendations.

b. *Innovation, performance and workload:*

LSS is faster and more accurate than traditional relational and graph-based algorithms, providing query optimization suggestions for multi-dimensional self-join queries and reducing the cost by three orders of magnitude compared to traditional methods.

c. *Research conclusions (key points):*

LSS is a fast and accurate subgraph counting method.

LSS supports various types of node and/or edge labels, directed/undirected graphs, and homomorphic/subgraph isomorphism.

LSS can effectively assist the query optimizer in finding better query plans on MDL queries and is more efficient than traditional methods.

7. **Title:** *Can Graph Neural Networks Count Substructures*

NeurIPS 2020

Keywords: graph neural networks, substructure counting, attributed graphs, induced subgraphs, GNN architectures

a. *Importance of research:*

This research provides a theoretical framework for defining the expressive power of graph neural networks (GNNs) and proves that some GNN architectures are not suitable for certain scientific problems, which has important theoretical guidance significance.

b. *Research conclusions:*

Through the study of the expressive power of GNN architectures, a theoretical framework is proposed to measure the ability of GNNs to handle substructure counting problems, providing guidance for designing better GNNs. However, further research is needed on the optimization and learning process of GNNs. This paper explores the theoretical framework for the expressive power of GNN architectures on substructure counting problems, focusing on induced subgraph counting and subgraph counting. The paper proves that some well-known GNN architectures cannot handle

specific substructure counting problems and proposes a new mathematical model - Local Relational Pooling (LRP), which has certain innovation and application prospects in the field. The theoretical framework for evaluating the expressive power of GNNs provided by the paper provides inspiration for solving other scientific problems, but does not address issues such as GNN optimization and learning processes, so there are still many issues to be studied.

d. Performance of the method in the task: The SCGNN model exhibits excellent performance on multiple benchmark datasets and can be used as a feature extractor for downstream tasks.

8. **Title:** *Neural Subgraph Isomorphism Counting*

SIGKDD 2020

Keywords: Subgraph Isomorphism, Dynamic Memory, Neural Network

a. Theoretical foundation of the research:

Different representation learning architectures are used to handle small and large-scale graphs, and an attention-based neural network architecture is used to learn distributed representations of graphs and patterns and to handle pairwise interactions between them.

b. Technical approach proposed in the paper (step-by-step explanation):

The paper introduces two encoding methods for sequences and graphs, and discusses the use of CNNs, LSTMs, Transformers, RGCNs, and GINs to extract high-level features. However, due to the quadratic computational cost of attention mechanisms, the paper proposes a new neural network model called Dynamic Intermediate Attention Memory Network (DIAMNet), which introduces additional memory mechanisms to reduce the complexity to approximately linear while maintaining sufficient performance.

9. **Title:** *Neural Subgraph Matching*

2020

Keywords: subgraph matching, graph neural network, embeddings, voting mechanism

a. Theoretical basis of the study:

The theoretical basis of this study is graph neural networks (GNNs), which learn embedding representations of nodes and subgraphs to predict topological isomorphism.

b. Technical route of the article (step by step):

NeuroMatch consists of two stages: first, learning embedding representations of nodes and subgraphs through GNNs, and then using a subgraph prediction function to predict the existence of topological isomorphism in the embedding space.

In the prediction results, NeuroMatch also uses a voting mechanism for matching nodes to improve accuracy.

The model training is conducted using a curriculum learning approach, and training data is generated by randomly sampling subgraphs of large target graphs.

Performance evaluation is carried out using benchmark datasets, and the results show that NeuroMatch exhibits excellent performance, far better than traditional methods.

10. **Title:** *Neural Subgraph Counting with Wasserstein Estimator*

SIGMOD 2022

Keywords: Subgraph Counting, Graph Neural Network

a. Significance of the work:

This paper proposes a new approximate subgraph counting algorithm NeurSC, which improves the estimation accuracy and scalability of the subgraph counting problem.

b. Innovation, performance, and workload:

The innovation of this paper lies in proposing a subgraph counting algorithm applicable to query and data graphs, and comparing NeurSC with other techniques, demonstrating that the NeurSC model performs well in terms of prediction accuracy and robustness.

11. **Title:** *On the Substructure Countability of Graph Neural Networks*

2015 IEEE Transactions on Knowledge and Data Engineering

Keywords: Graph neural networks; Substructure counting; Weisfeiler-Leman algorithm

a. Research background of this article:

This article explores the theoretical capabilities of graph neural networks (GNNs) in substructure counting tasks.

b. Past methods, their problems, and motivation:

Previous methods were unable to count specific substructure quantities, hence the need to investigate the theoretical capabilities of GNNs.

c. Research methodology proposed in this paper:

This paper studies the conditions under which k -dimensional Weisfeiler-Leman (k -WL) equivalent GNNs can count substructures, with a focus on 3-WL equivalent GNNs, and proposes an algorithm to determine the countability of substructures.

d. Task and performance achieved by the methods in this paper:

This paper proposes the Layer-wise Permutation Pooling (LPP) model to improve substructure counting performance, and compares it with several state-of-the-art GNNs on various datasets, finding that LPP is on average superior to the baseline by 84%.

12. **Title:** *REVOCABLE DEEP REINFORCEMENT LEARNING WITH AFFINITY REGULARIZATION FOR OUTLIER-ROBUST GRAPH MATCHING*

ICLR 2023

a. Work significance:

This paper proposes a new method for solving node matching problems, which is of significant importance for solving combinatorial optimization problems.

b. Innovation, performance, and workload:

The innovation of this paper lies in the use of reinforcement learning methods, which are rarely used in graph matching problems, and the use of Double Dueling DQN algorithm for optimization and training. It outperforms other current methods and reduces workload.

c. Research conclusions (list points):

This paper proposes a new reinforcement graph matching (RGM) algorithm based on reinforcement learning.

The method uses association graphs to transform the node matching problem into a node selection problem.

This paper provides detailed steps for constructing the state, action, and reward functions.

The proposed Double Dueling DQN (D3QN) algorithm outperforms other methods and has good scalability.

This paper provides a new direction for the application of reinforcement learning in combinatorial optimization problems.

13. **Title:** *LMKG: Learned Models for Cardinality Estimation in Knowledge Graphs*

2021

Keywords: Cardinality Estimation, Knowledge Graphs, Deep Learning

a. Topic and features:

The topic of this paper is the application of cardinality estimation techniques in knowledge graph predicate queries.

b. Historical development:

There have been some cardinality estimation techniques already developed. However, the problem of cardinality estimation in knowledge graph predicate queries remains to be solved.

c. Past methods used:

In the past, cardinality estimation has been tackled using some relatively mature techniques.

d. Deficiencies of previous research:

There is a lack of cardinality estimation techniques suitable for knowledge graph predicate queries.

e. Current problems to be solved:

There is a need to improve the accuracy of cardinality estimation for knowledge graph predicate queries.

14. ***Title: SimGNN: A Neural Network Approach to Fast Graph Similarity Computation***

WSDM 2019

Keywords: network embedding, neural networks, graph similarity computation, graph edit distance

Taking the first step towards bridging the gap, by tackling the core operation of graph similarity computation , via a novel neural network based approach. The central idea is to learn a neural network based function that is representation-invariant, inductive, and adaptive to the specific similarity metric, which takes any two graphs as input and outputs their similarity score. Our model runs very fast compared to existing classic algorithms on approximate Graph Edit Distance computation, and achieves very competitive accuracy