A review on the construction and application of medical knowledge graph

Zhi Qi

University of Electronic Science and Technology of China, UESTC, China

qizhiuestc@163.com

Abstract. With the development of Artificial Intelligence, Knowledge Graph has become a popular direction in recent years, with which many practical problems have been solved. As a special type of Knowledge Graph, Medical Knowledge Graph is an important research field. While most of related research describes the construction and application of Medical Knowledge Graph (MKG) for a certain disease, there lacks reviews on previous research on Medical Knowledge Graph in recent years. In order to investigate recent research of Medical Knowledge Graph, this paper reviews studies of the construction and application of Medical Knowledge Graph in recent years, dividing them into six parts according to the content. Subsequently, some suggestions are put forward for the development of Knowledge Graph.

Keywords: Artificial Intelligence, Knowledge Graph, Medicine, Construction, Application.

1. Introduction

In recent years, the application of Knowledge Graph in finance, medicine and other fields has been maturer and maturer with the development of Knowledge Graph. Compared with popular artificial intelligence methods such as reinforcement learning and deep learning, Knowledge Graph is more explicable and thus more suitable for the application in medical field where high accuracy is required. For instance, Knowledge Graph can be used to analyze new diseases through connecting it with diseases which have had a mature method of treatment, such as the research on the medicine for COVID-19. In addition, Knowledge Graph can be used to analyze health conditions, construct question-and-answer platforms, and automatically generate reports. The application of Knowledge Graph in medical field frees doctors from trivia.

However, due to the lack of comprehensive summary of recent research on the combination of medicine and Knowledge Graph, some important aspects like automatic report generation based on Knowledge Graph retain to be further explored.

The present study first collects previous research on the application of Knowledge Graph in medical field, then analyzes the development of the usage of Knowledge Graph in medical field in recent years, and finally proposes some suggestions for future study of Medical Knowledge Graph (MKG).
The contribution of the present study is as follows.

It comprehensively reviews previous research on KG in the medical field.

In addition to traditional construction of KG and Q&A system, it also covers recent studies of KG combined with medical imaging and automatic report generation, providing some perspectives for future research.

2. Overview

2.1. Research motivation

As one of the most popular directions in the field of Artificial Intelligence, Knowledge Graph has been explored by many researchers. Shengguang Peng et al.'s study summarizes the development of Knowledge Graph in the early stage after it was proposed in 2012 and its application in various subdivided fields of medicine, and introduces the concept of Knowledge Graph as well as the classification of Knowledge Graph [1]. However, Shengguang Peng's study didn't cover some excellent research related to Medical Knowledge Graph such as automatic report generation. Stephen Bonner's research mainly focuses on the application of Knowledge Graph in new drug research [2]. Nevertheless, they didn't mention the application of Medical Knowledge Graph, such as the question-and-answer system and automatic report generation. The aforementioned fields will be detailed later in this article. It is worth mentioning that Medical Knowledge Graph based on Traditional Chinese Medicine (TCM) has emerged and attracted much attention in recent years. Lirong Jia's study examines the use of Knowledge Graph in TCM. As the empirically-based medicine, TCM can be better used than modern Western medicine in research on Knowledge Graph [3]. However, Lirong Jia's study only focuses on the TCM field and doesn't cover the western medicine. Actually, western medicine plays a crucial role in modern medicine and has a deep connection with Knowledge Graph.
In conclusion, existing research on Medical Knowledge Graph hasn't covered many important advances in medicine field in recent years.

2.2. Research structure

The present study collects articles on the construction and application of Medical Knowledge Graph in recent years, and analyzes the development of Medical Knowledge Graph from the following six research directions, namely researching a disease based on Knowledge Graph, Optimization of construction scheme of Knowledge Graph, Optimization of Knowledge Graph efficiency, Question-and-answer system based on Knowledge Graph, Combination of Knowledge Graph and medical image, and Automatic report generation based on Knowledge Graph. The application and construction of Knowledge Graph in these fields are shown in Figure 1.
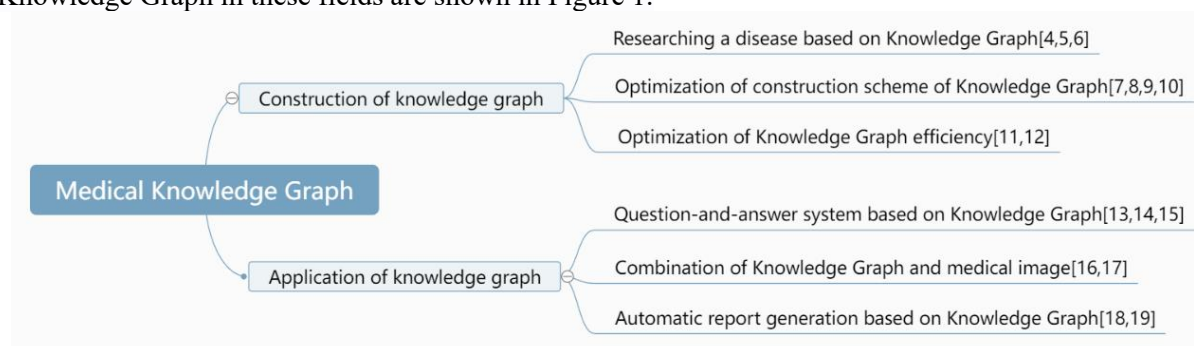

Figure 1. Structure of the present study.

The following presents the six directions shown in Fig.1. The research on a disease based on Knowledge Graph uses KG to study the symptoms and treatment of a specific disease. Optimization of construction scheme of Knowledge Graph refers to building a better data storage and structure. Optimization of Knowledge Graph efficiency is to make it more efficient to perform data calculation and resolve other operational problems on KG. Question-and-answer system based on Knowledge Graph is using KG to build a platform which can automatically answer a user's question. Combination of Knowledge Graph

and medical image is using KG to analyse medical images. Automatic report generation based on Knowledge Graph is generating diagnosis results after reading the KG constructed from diagnosis information.

3. Overview of related studies

3.1. Construction of Medical Knowledge Graph
This part reviews studies related to how to construct a MKG from both a medical view and a technological view.

3.1.1. Researching a certain disease based on Knowledge Graph. In order to better study a disease, KG can be used to analyze it to get a result by taking more situations into consideration. Binjie Cheng constructed a framework based on Knowledge Graph to find the relation between medical entities, which could help build question-and-answer systems and decision-making systems [4]. This study fills the gap of research on stoke disease. Meanwhile, the model has a high precision on TransE model and TransD model. Yunrong Yang proposed a model SS-DYGIE++ and used it to construct a CPHE-KG (COVID-19 Public Health Evidence-Knowledge Graph) [5]. The knowledge graph can show dynamic spread of the COVID-19 pandemic across different regions and countries and present effects of various interventions. Taejin Kim constructed a model Co-BERT (COVID-19-Bidirectional Encoder Representation from Transformer), which can be used to connect entities related to COVID-19 [6]. This framework is better than original BERT when it is used in COVID-19.

3.1.2. Optimization of construction scheme of Knowledge Graph. Some scholars have tried to get a better structure to build the KG. Irene Y. Chen proposed a method from non-linear models using causal inference principles, which can be used for analyzing disease and symptoms from electronic medical records based on knowledge graph [7]. This method has helped researchers construct Knowledge Graph and robustly extract knowledge, which is better than previous models such as LR (logistic regression) model and NB (naive Bayes) model in AUPRC (area under the precision-recall curve). Aynur Guluzade constructed a framework demographic aware probabilistic medical knowledge embedding (DARLING) [8]. Compared with traditional frameworks, this work considers patient demographics, which is superior to traditional ways in medicines and treatments. Fan Gong proposed a framework SMR (safe medicine recommendation) [9]. The robustness of traditional methods is questionable because of the graphs may be incomplete. In contrast, this work can reduce the influence of such incompleteness and improve the stability of recommendation systems. Wanheng Liu proposed a framework based on deep neural networks, which can be used to generate Chinese Knowledge Graph [10]. This model fills the gap in Chinese Medical Knowledge Graph research, which is better than previous models such as conditional random field (CRF) and BiLSTM-CNNs in terms of accuracy.

3.1.3. Optimization of Knowledge Graph efficiency. Solving data calculation and other operational problems can make KG more efficient. Fenglin Liu constructed an unsupervised model KGAE (knowledge graph auto-encoder) based on Knowledge Graph, providing an approach that doesn't rely on paired dataset. This model can not only work on more training sets than traditional approaches, but also work on supervised settings and semi-supervised settings [11]. Linfeng Li proposed a algorithm PrTransX used in Medical Knowledge Graph [12]. Traditional methods have difficulty constructing Knowledge Graph when connecting head entities and tail entities. Nevertheless, this model can reduce the influence of this problem, which is better than TransX algorithm in all of evaluation metrics [12].

3.2. Application of medical knowledge graph
This part is mainly about how KG is applied to solve the different problems, and the the advantages of using KG.

3.2.1. Question-and-answer system based on Knowledge Graph. One of the most important applications of KG is the building of QA systems. Qiming Bao constructed a HHH (healthcare helper system with a hybrid question-and-answer model) framework based on Knowledge Graph, which can answer complex questions about medicine [13]. This framework is better than BERT (bidirectional encoder representation from transformers) and MaLSTM (Manhattan LSTM Model). Wenbo Zheng proposed a Knowledge Graph based on attention embedding, which can be used in COVID-19 diagnosis [14]. This model can make a quick diagnosis at early time of disease, which is better than other methods such as DGLM and GMMF. Zhixue Jiang constructed a system based on Knowledge Graph and crawler technology, connecting these techniques with professional knowledge of medicine to answer medical questions [15]. The system using Aho-Corasick(AC) algorithm to collect entities types, which performs well in many datasets such as 39 health network and Clove Garden datasets.

3.2.2. Combination of Knowledge Graph and medical image. As KG can be used in medical image processing, there are some excellent related studies in this field recently. Daibing Hu constructed a VSE-GCN (visual-semantic embedded based on graph convolutional network) model to diagnose CXR (Chest X-Ray), which is better than competitive approaches tested on public datasets [16]. Yixiao Zhang constructed a framework based on Knowledge Graph and proposed a evaluation metric MIRQI (Medical Image Report Quality Index) [17], which can be used in generating radiology report automatically. The framework is superior to other models such as ChestXray8 and TieNet; and the metrics is better than others like CARG and CoAtt.

3.2.3. Automatic report generation based on Knowledge Graph. KG can also be used in automatic report generation. Yinyu lan proposed two reasoning algorithms about Knowledge Graph, which can help build effective semantic representation for sparse Medical Knowledge Graphs of MedKGC (medical knowledge graph completion) [18]. Haifeng Zhao constructed a framework using MIM (multi-modal interaction module) to connect image and semantic feature based on Knowledge Graph [19], which can be used in medical report generation, and performs better than previous models such as HRGR and ADAATT.

3.3. Discussion

Most of the existing research focuses on the improvement of previous methods, while there lacks research combining Knowledge Graph with new medical directions.

In order to better apply KG, it is necessary to combine it with other fields to propose new applications, like using KG to analyze health condition of people and giving some advises. KG can also be applied in some other research directions, such as combining it with medical image and report generation. The emergence of these new directions can bring vitality to the research on Medical Knowledge Graph.

4. Conclusion

The present study collects and classifies the recent articles on Medical Knowledge Graph, and summarizes the main development of Medical Knowledge Graph in recent years. It also provides some perspectives for researchers to understand the status of Medical Knowledge Graph and carry out new research in this field.

References

[1]   Peng Shengguang et al., "Overview of Medical Knowledge Mapping Research", 2020 International Conference on Intelligent Computing and Human-Computer Interaction (ICHCI), 2020, pp.231-234.
[2]   Tephen Bonner, Ian P Barrett, Cheng Ye, Rowan Swiers, Ola Engkvist, Andreas Bender4, Charles Tapley Hoyt, and William Hamilton, "A Review of Biomedical Datasets Relating to Drug Discovery: A Knowledge Graph Perspective", arXiv:2102.10062v4 [cs.AI] 2021, pp.1-34.

[3]   Lirong Jia, Jing Liu, Tong Yu, Yan Dong, Ling Zhu, Bo Gao, and Lihong Liu, "Construction of traditional Chinese knowledge graph", Journal of Medical Informatics 2015, pp.51-59.

[4]   Binjie Cheng, Jin Zhang, Hong Liu, Meiling Cai, Ying Wang,2020 "Research on Medical Knowledge Graph for Stroke", Journal of Healthcare Engineering, vol. 2021, pp.1-10.

[5]   Yunrong Yang, Zhidong Cao, Pengfei Zhao, Dajun Daniel Zeng, Qingpeng Zhang, Yin Luo, 2021 "Constructing public health evidence knowledge graph for decision-making support from COVID-19 literature of modelling study", Journal of Safety Science and Resilience, pp.146–156.

[6]   Taejin Kim, Yeoil Yun, and Namgyu Kim, 2021 "Deep Learning-Based Knowledge Graph Generation for COVID-19", 13, 2276, pp.1-19.

[7]   Irene Y. Chen, Monica Agrawal, Steven Horng, David Sontag, 2020 "Robustly Extracting Medical Knowledge from EHRs: A Case Study of Learning a Health Knowledge Graph", Pacific Symposium on Biocomputing, pp.19-30.

[8]   Aynur Guluzade, Endri Kacupaj, Maria Maleshkova, 2021 "Demographic Aware Probabilistic Medical Knowledge Graph Embeddings of Electronic Medical Records", arXiv:2103.11951v1 [cs.IR] 22 Mar 2021, pp.1-14.

[9]   Fan Gong, Meng Wang, Haofen Wang, Sen Wang, Mengyue Liu, 2021 "SMR: Medical Knowledge Graph Embedding for Safe Medicine Recommendation", Big Data Research 23 (2021) 100174, pp.1-8.

[10]   Wanheng Liu, Ling Yin, Cong Wang, Fulin Liu, Zhiyu Ni, 2021 "Medical Knowledge Graph in Chinese Using Deep Semantic Mobile Computation Based on IoT and WoT", Wireless Communications and Mobile Computing Volume 2021, Article ID 5590754, pp.1-13.

[11]   Fenglin Liu, Chenyu You, Xian Wu, Shen Ge, Sheng Wang, Xu Sun, 2021"Auto-Encoding Knowledge Graph for Unsupervised Medical Report Generation", In proceeding of 35th Conference on Neural Information Processing Systems (NeurIPS 2021), pp.1-14.

[12]   Linfeng Li, Peng Wang, Yao Wang, Shenghui Wang, Jun Yan, Jinpeng Jiang, 2020 "A Method to Learn Embedding of a Probabilistic Medical Knowledge Graph_ Algorithm Development", JMIR Med Inform 2020;8(5): e17645, pp.1-12.

[13]   Qiming Bao, Lin Ni, Jiamou Liu, 2020 "HHH: An Online Medical Chatbot System based on Knowledge Graph and Hierarchical Bi-Directional Attention", ACSW 2020, February 4–6, 2020, Melbourne, VIC, Australia, pp.1-10.

[14]   Wenbo Zheng, Lan Yan, Chao Gou, Zhi-Cheng Zhang, Jun Jason Zhang, Ming Hu,Fei-Yue Wang, 2021 "Pay attention to doctor–patient dialogues Multi-modal knowledge graph attention image-text embedding for COVID-19 diagnosis", Information Fusion 75 (2021) pp.168–185.

[15]   Zhixue Jiang, Chengying Chi, Yunyun Zhan, 2021 "Research on Medical Question Answering System Based on Knowledge Graph", pp.21094-21101.

[16]   Daibing Hou, Zijian Zhao, Sanyuan Hu, 2020 "Multi-Label Learning with Visual-Semantic Embedded Knowledge Graph for Diagnosis of Radiology Imaging", pp.15720-15730.

[17]   Yixiao Zhang, Xiaosong Wang, Ziyue Xu, Qihang Yu, Alan Yuille, Daguang Xu, "When Radiology Report Generation Meets Knowledge Graph", The Thirty-Fourth AAAI Conference on Artificial Intelligence (AAAI-20), pp.12910-12917.

[18]   Yinyu Lan, Shizhu He, Kang Liu, Xiangrong Zeng, Shengping Liu, Jun Zhao, 2020"Path-based knowledge reasoning with textual semantic information for medical knowledge graph completion", The China Conference on Health Information Processing (CHIP) 2020 Shenzhen, Guangdong, China. 30-31 November 2020, pp.1-10.

[19]   Haifeng Zhao, Jie Chen, Lili Huang, Tingting Yang, Wanhai Ding, Chuanfu Li, 2021 "Automatic Generation of Medical Report with Knowledge Graph", In 2021 10th International Conference on Comp uting and Pattern Recognition (ICCPR '21), October 15–17, 2021, Shanghai, China, pp.1-10.