The Report Committee for Song Wang
Certifies that this is the approved version of the following Report:

Knowledge Graph Applications in Medical Imaging Analysis: A Scoping Review

APPROVED BY
SUPERVISING COMMITTEE:

Joydeep Ghosh, Supervisor

Ying Ding

Knowledge Graph Applications in Medical Imaging Analysis: A Scoping Review

by

Song Wang

Report

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Master of Science in Engineering

The University of Texas at Austin

December 2021

Abstract

Knowledge Graph Applications in Medical Imaging Analysis: A Scoping Review

Song Wang, MSE
The University of Texas at Austin, 2021

Supervisor: Joydeep Ghosh

There is an increasing trend to represent domain knowledge in structured graphs, which provide efficient knowledge representations for many downstream tasks. Knowledge graphs are widely used to model prior knowledge in the form of nodes and edges to represent semantically connected knowledge entities, which several works have adopted into different medical imaging applications. We systematically search over five databases to find relevant articles that apply knowledge graphs to medical imaging analysis. After screening, evaluating, and reviewing the selected articles, we performed a systematic analysis. We look at four applications in medical imaging analysis, including disease classification, disease localization and segmentation, report generation, and image retrieval. We also identify limitations of current work, such as the limited amount of available annotated data and weak generalizability to other tasks. We further identify the potential future directions according to the identified limitations, including employing semi-supervised frameworks to alleviate the need for annotated data and exploring task-agnostic models to provide better generalizability. We hope that our article will provide

the readers with aggregated documentation of the state-of-the-art knowledge graph applications for medical imaging.

Table of Contents

List of Tables

List of Figures

Introduction

In recent years, incorporating structured domain knowledge into downstream tasks has attracted great research attention from both industry and academia (Ji, et al. 2021). This is because domain knowledge provides a proper understanding of a field which can be represented as a knowledge graph that can facilitate efficient inference to empower downstream tasks.

A knowledge graph is a structured graph representing facts, consisting of entities (e.g., abstract concepts and real-world objects) and the relationships between the presented entities (Ji, et al. 2021). It provides semantically structured information that computers can interpret and promises to build more intelligent systems to solve numerous real-world problems.

Knowledge graphs (viewed as the graph structure) differ from knowledge bases in terms of the involvement of formal semantics for interpretation and inference over facts (Figure 1). Knowledge graphs (KGs) such as Wikidata (Vrandečić and Krötzsch 2015), NELL (Carlson, et al. 2010), and DBPedia (Auer, et al. 2007) have recently played impactful roles in many machine learning-based applications, including search and information retrieval (Sumithra and Sridhar 2020), information extraction (Fei, et al. 2021, Jaradeh, et al. 2021, Bastos, et al. 2021), question answering (Ma, et al. 2021, Banerjee and Baral 2020), and recommendation (Wang, et al. 2019, Wang, et al. 2020, Xiang, et al. 2021).

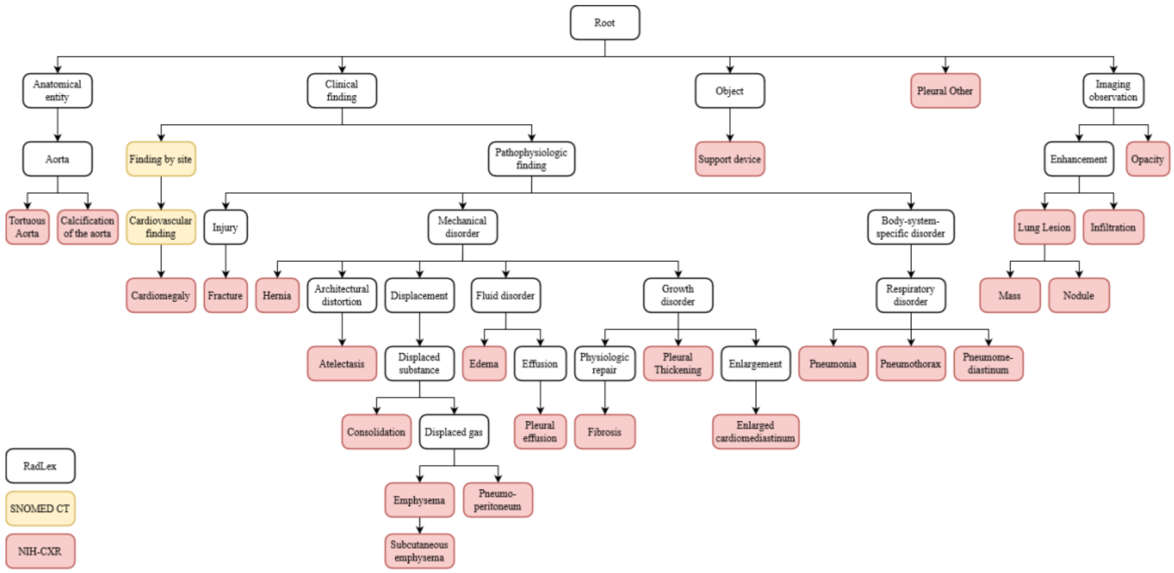

Figure 1. Example of the radiology knowledge graph of NIH Chest X-ray labels based on RadLex and SNOMED_CT

Within a biomedical setting, knowledge graphs can greatly help researchers deal with many real-world clinical problems, such as exploring new treatments for existing drugs (Dai, et al. 2021, Yu, et al. 2021), aiding efforts to diagnose patients (Xi, et al. 2021), and identifying underlying associations between biomolecules and diseases (Chilińskiab, Senguptab and Plewczynski 2021). In many scenarios, the solutions will learn to map the knowledge graphs and represent in a low-dimensional space, the process of which is named representational learning (Ji, et al. 2021). This aims to encode and retain the local and/or global graph structure that has close relevance to the problem, while mapping the graph into a representation space that machine learning methods can utilize to build predictors. Among various knowledge graph applications in biomedicine, medical imaging is one of the most significant diagnostic aids available to physicians (Xie, et al. 2021). Medical imaging includes different technologies such as computed tomography (CT), conventional radiography, ultrasound, mammography, and magnetic resonance imaging (MRI) (Xie, et al. 2021).

Artificial intelligence in medical imaging is a way of training a computer to identify abnormal regions and tissue variations, which is similar to how humans are trained (Goodfellow, Bengio and Courville 2016). The model is trained based on the past medical records and diagnoses made by pathologists or radiologists. The algorithm learns this with huge amounts of data, and after analyzing thousands of iterations of different images and diagnoses, it eventually will learn to make some diagnoses. In the medical imaging analysis domain, knowledge graphs have drawn a lot of research attention. According to our review, most studies applied knowledge graphs to specific topics such as disease detection, localization, and report generation. In this review, we describe various applications applying knowledge graphs in medical imaging analysis. We then point out future directions that have yet to be explored.

Background

KNOWLEDGE GRAPHS IN GENERAL

Recently, knowledge graphs have become a predominant part in many information systems that require prior knowledge. We can trace the concept of graphical knowledge representation back to 1956 when Richens proposed the idea of Semantic Net (Richens 1956); however, the community realized the importance of his work only belatedly. One of the most famous rule-based expert systems for medical diagnosis, MYCIN (Shortliffe 1976) has a knowledge base containing about 600 rules (Ji, et al. 2021). Many researchers promoted the idea of graph-based knowledge representation aiming to assemble human knowledge. Resource Description Framework (RDF) (E. Miller 2005) and Web Ontology Language (OWL) (McGuinness 2004) were later released and became the mainstay of Semantic Web.

In 2009, the concept of Linked Data was proposed with the aim of linking various datasets in the Semantic Web with each other and treat it as a single large, global knowledge graph (Berners-Lee, Bizer and Heath 2009). Subsequently, many ontologies or open knowledge bases were published, such as WordNet (G. Miller 1995), YAGO (Suchanek, Kasneci and Weikum 2007), DBpedia (Auer, et al. 2007), and Freebase (Bollacker, et al. 2008), to realize the idea of structured knowledge representation in the form of a graph. In 2012, Google proposed Knowledge Graph (Knowledge Vault) to utilize semantic knowledge in the web search setting, and the concept gained great popularity (Dong, et al. 2014). Entities in the texts are identified and disambiguated through the usage of Google's knowledge graph, where the search results are enriched with semantically structured text summaries, and the links to related entities are further provided in the exploratory search (Ji, et al. 2021). Recently, many companies such as Microsoft, Amazon, and Pinterest have started investing massive resources to build knowledge graphs for their commercial applications (Färber 2019, Cui and Shrouty 2020, Dong, et al. 2020).

Knowledge Graphs for Medical Imaging Analysis

Machine learning techniques have recently been used in nearly every radiotherapy stage, including diagnostic imaging, at-risk organ delineation which potentially uses image registration, followed by the automated planning and outcome assessment (Chan, Witztum and Valdes 2020, Shan, et al. 2020). Valuable input for treatment planning refinements towards personalized medicine can be provided through the process. It is promising to combine deep learning-based medical imaging together with artificial intelligence-driven radiotherapy. From the perspective of precision radiotherapy, it is vital to conduct quantitative analysis of the comprehensive features obtained from image data (e.g., MRI, CT images) and other forms of relevant data (Shan, et al. 2020). One important insight is that there tends to be significantly more information contained in images and other forms of data, than what can be visually identified by human oncologists and radiologists. We can effectively identify and harvest the underlying information using sophisticated algorithms to improve diagnosis and treatments (Shan, et al. 2020). There are two reasons that AI-based radiotherapy is believed to outperform conventional workflows. First, many latent features that human readers can not perceive somehow can be well utilized by the radiomic analysis, enabling the AI-based radiotherapy to model and learn from the features in a more sophisticated manner. Second, structured prior knowledge and structured constraints can be used in a data-driven and end-to-end manner, making AI-based radiotherapy more powerful (Shan, et al. 2020). Big data is generated during the radiotherapeutic process on functional, cellular, molecular, anatomical, metabolic, and pathological features, especially in genetic profiles, tomographic images, and medical reports (Skripcak, et al. 2014). We can structure these data into patterns and primitives, that can be understood as a general-sense ‘biological languages’. One major clinical challenge is that it can be complicated to extract and present those ‘biological languages’ in a meaningful manner that oncologists and/or radiologists can interpret. Also, pathological, radiological, and oncological reports are manually written in natural languages, hence the standards of sensitivity and the degrees of specificity of medical decisions, medical reports, and treatment plans require further improvements for better treatment (Shan, et al. 2020). Synergizing expertise in imaging analysis and natural language processing will be helpful to improve the prognosis of patients (Shan, et al. 2020).

An active investigation is demanded by many interesting medical topics. Treatment-related image features can be extracted and further organized as graphs to capture and encode the temporal and spatial dependencies among image features. To be more specific, a graph-based learning system can be developed to quantify the states of patients, and to predict the outcomes (Shan, et al. 2020). For example, a ResNet model pre-trained on ImageNet can be employed to extract the imaging biomarkers from medical images (Shan, et al. 2020). The pre-trained model can be fine-tuned with other real data to enhance the transferability, and the soft activation maps can be visualized to provide better interpretability. The learned features can also be visualized in a low-dimensional manifold with respect to the classification labels using t-SNE. The edge weights in the graphs can be learned and adjusted through graph learning, reflecting the strength between disease growth, response patterns, and other specific features (Shan, et al. 2020).

A medical knowledge graph constructed from a patient’s electronic medical records and reports is invaluable for reasoning and planning (Shan, et al. 2020). Given these text inputs, we can identify and extract the domain knowledge to develop a library of knowledge graphs through NLP techniques. For example, to extract knowledge from medical reports, we can use online services such as the Watson Natural Language Understanding platform or the Amazon Comprehend Medical instead of starting from scratch. Based on these cloud services and other available systems, we can then distill and query not only the high-quality domain-specific rules, but also unstructured or even semi-structured contents from patient data, such as medical images, medical conditions, medication details, clinical reports, etc. (Shan, et al. 2020).

With the efforts mentioned above, treatment-related feature graphs and knowledge graphs can be built from both medical images and medical text data. The feature graphs and knowledge graphs fall in two different domains: feature graphs are from images and data that are clinically informative; while the knowledge graphs are from the professional languages that are directly interpreted (Shan, et al. 2020). Hence, an across-domain graph transformation is needed for bridging these two domains. To this end, a graph-based encoder-decoder network can be employed, including graph convolutions and graph pooling operations. The information from a radiomic graph will be extracted by the graph encoder, while the graph decoder will reconstruct the corresponding graph (Shan, et al. 2020).

Methods

DATA SOURCES AND SEARCH STRATEGIES

Following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines, we conducted a comprehensive search of English-language articles published between 2006 and 2021 from five databases. The databases include IEEE Xplore[1], PubMed[2], Arxiv[3], Google Scholar[4], the ACM Digital Library[5]. We excluded several article types, including the review, editorial, erratum, letter, note, and comment. The search strategy is to iteratively search keywords for relevant articles and related citations. The keywords include knowledge graph(s) and medical imaging, knowledge graph(s) and medical image(s), graph(s) and medical imaging, graph(s) and medical image(s).

ARTICLE SELECTION

After acquiring the list of potential articles, we screened the abstracts of articles to do article filtering and selection. Articles were excluded based on several criteria, including articles not related to applying knowledge graphs in medical imaging, duplicates, unavailable full-text articles, and conference abstracts. Article titles and abstracts screening was conducted first, followed by full text and relevance screening. The selected articles were then reviewed. During the screening, all conflicting opinions among reviewers were discussed until we reached a consensus.

---

1 https://ieeexplore.ieee.org/Xplore/home.jsp
2 https://pubmed.ncbi.nlm.nih.gov/
3 https://arxiv.org/
4 https://scholar.google.com/
5 https://dl.acm.org/

IDENTIFICATION OF INCLUDED ARTICLES

We filtered the articles manually based on the titles and abstracts to check whether the articles were related to the knowledge graphs applied in medical imaging, and 21 articles remained for subsequent full-text reviews.

Results

Statistical Characteristics of the Included Articles

All the articles included in this work are published from 2006 to 2021, with a noticeable increment in the number of papers published per year (Figure 2). The included publications are across nine countries, with most contributions coming from China (48%) (Figure 3). Among all the included articles, the most common application of knowledge graphs in medical imaging is disease classification (56.5%), followed by disease localization and segmentation (17.4%), report generation (17.4%), image retrieval (8.7%) (Figure 4).

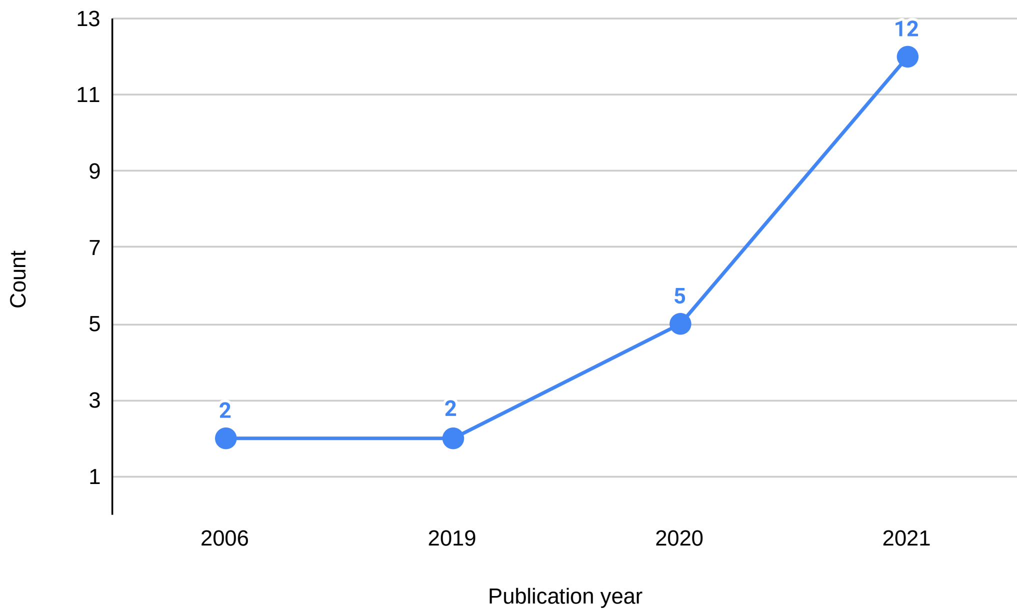

Figure 2. Year trend of reviewed articles

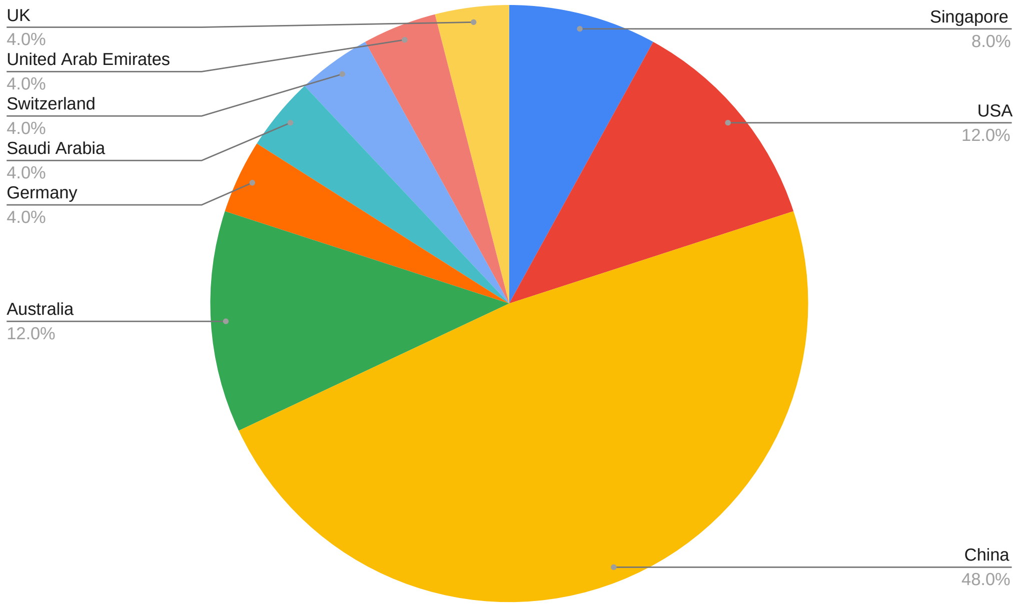

Figure 3. Publication country distributions

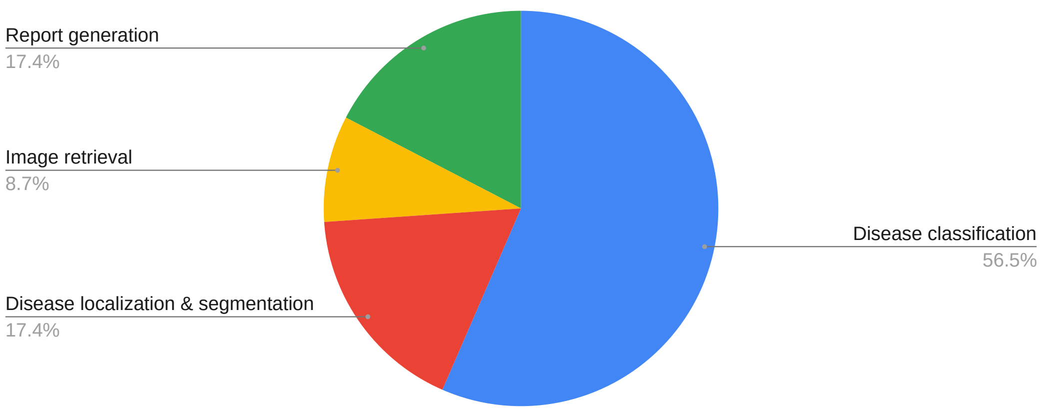

Figure 4. Application topic distributions

DISEASE CLASSIFICATION

Classifying images into appropriate categories is one of the most common tasks in computer vision (Obaid, Zeebaree and Ahmed 2020). Disease classification is especially of vital importance in medical imaging to assist diagnosis (Xie, et al. 2021). The types and sizes of image datasets are increasing dramatically. Hence often, we need to classify images from unseen classes into the correct categories based on the relationships between the seen and unseen classes. Our world contains millions of visual concepts. Due to its complex and dynamic characteristics, it is impossible to build a large dataset for every concept to ameliorate various computer vision tasks. Prior knowledge is the key to building semantic relationships between classes, which can be of great help, especially when we have limited training data. Knowledge graphs contain rich knowledge, modeling the relationships among classes or concepts. Incorporating disease classification in medical imaging with knowledge graphs has been explored by researchers and has shown promising results (Xie, et al. 2021). Table 1 lists the overview of datasets used by the included articles related to disease classification in this work. The most commonly used datasets are IU X-Ray (Demner-Fushman, et al. 2015), CheXpert (Irvin, et al. 2019) and NIH Chest X-Ray/ChestX-Ray 14 (Wang, et al. 2017), and they all include medical images, associated medical reports, and disease labels.

There are five articles included in this review that explored binary disease classification incorporated with knowledge graphs. Xie et al. constructed a knowledge-based collaborative sub-model for nodule classification. Three types of image patches were designed to fine-tune three pre-trained ResNet-50 networks to respectively characterize the nodules' overall appearance, voxel, and shape heterogeneity (Xie, et al. 2019). In this way, they proposed a multi-view collaborative deep model to separate the malignant nodules from the benign ones using very limited chest CT data. Yu et al. aimed to facilitate the process of diagnosing pneumonia (Yu, Wang and Zhang 2021). A graph-based feature reconstruction module was employed that takes the produced image features from a trained convolutional neural network (CNN) as input. The resulting combined features will be fed to a one-layer graph neural network (GNN) to classify chest X-ray images into normal and pneumonia. According to Chen et al., most existing work manually built a population graph to aggregate structural information where the relationship between nodes was represented by the graph adjacency matrix (Chen, et al. 2021). Chen et al. automatically constructed the population graph and further utilized the fusion of multimodal information, which improved the diagnostic accuracy for Autism Spectrum Disorder and breast cancer. Specifically, an encoder was proposed to select the appropriate phenotypic measures according to their spatial distributions and calculate the edge weights between nodes using the text-similarity awareness mechanism. Liu et al. claimed to outperform previous work on the Mammogram mass classification task (Liu, et al. 2021). They introduced Bipartite Graph Convolutional Network (BGN) to model the intrinsic geometric and semantic relations of ipsilateral views. Given the fact that clinical practice widely adopted visual asymmetry of bilateral views to assist the breast lesions diagnosis, they proposed an Inception Graph Convolutional Network (IGN) to model the structural similarities of bilateral views to construct the graph structure. Further based on the constructed graphs, the multi-view information can be systematically propagated through nodes, which allows the features learned from the examined view the ability of multi-view reasoning. Fu et al. pointed out that most existing methods focus on one single modality (i.e., image) and other modalities are ignored, or the complementary information from both modalities are not fully leveraged (Fu, et al. 2021). They proposed to exploit the inter-category relationships in the 7-point visual category checklist (7PC) for Melanoma diagnosis. Specifically, they proposed to use a graph-based relational module to leverage inter-categorical and inter-modal relations, and further the visual structure details from dermoscopy were prioritized by encoding the category representations in a graph network. Another category embedding learning module was also employed to capture the specialized representations for each category and support the graph-based relational module. In this way, they claimed to outperform the state-of-the-art performance at 7PC categories classifications and diagnosis.

Six articles explored multi-label classifications in medical imaging. Zhang et al. utilized a pre-constructed graph embedding module on multiple disease findings to assist the disease classification task (Zhang, et al. 2020). The incorporation of a knowledge graph allowed for the dedicated feature learning for each disease finding. Similarly, Hou et al. employed the graph convolutional network (GCN) to model the correlations among different disease labels. The disease label embeddings were pre-trained on the radiology reports. The graph features were initialized by fusing semantic features together with the encoded image features in a transformer encoder (Hou, Zhao and Hu 2021). To enhance the representation ability of the graph, they mined additional medical terms from radiology reports and added the mined terms to the graph model as auxiliary nodes without changing the actual output space size.

However, Zhou et al. pointed out that developing a robust and reliable intelligent diagnosis system was hindered by the inconsistent appearances and high complexities of different lesions in chest X-rays (Zhou, et al. 2021). It is a promising direction to attend to the abnormal regions of high probabilities and exploit the prior of a related knowledge graph. Hence, they proposed one contrastive network to learn the left-right lung intra-attentive abnormal features for better identifying the most common thoracic diseases, whose lesions rarely appear on both sides symmetrically. To obtain the abnormal attention map, they also employed an inter-contrastive abnormal attention model to compare the query scan with multiple anchor scans with no lesions present. Once the intra-contrastive and inter-contrastive attentions were weighted over the features, a chest radiology graph was constructed for dual-weighting graph reasoning, in addition to the basic visual-spatial convolution. Following the same direction, Agu et al. noted that most existing models only looked at the entire chest X-ray images for classification, failing to utilize the vital anatomical information (Agu, et al. 2021). They utilized a GCN which enables the learning of the label dependencies and the correlations between the anatomical regions in the chest X-ray images. They efficiently created an adjacency matrix for the anatomical regions based on the correlation of the labels across different regions. Combining this with a detection module, a novel multi-label chest X-ray classification model was presented that can classify image findings and localize them to their anatomical regions.

According to Sekuboyina et al., learning to map images to binary labels made it challenging to utilize auxiliary information such as annotation uncertainty or label dependencies (Sekuboyina, et al. 2021). A multimodal knowledge graph was constructed using chest X-ray images and their labels. They posed the task of multi-label disease classification as a link prediction problem. They claimed that incorporating auxiliary information can then be achieved by adding additional nodes to the graph and at the same time, adding the relations among these nodes. Similarly, Chen et al. noted that most state-of-the-art works only focused on regression from the input image to the binary output labels, but such valuable graph-structured information was not fully utilized due to the complexity of graph data (Chen, et al. 2020). As a result, they explicitly explored the pathology dependencies for the multi-label chest X-ray image classification task. They introduced the word embeddings of pathologies and multi-layer graph information propagation, in which way the relationships between pathologies can be generalized into a set of classifier scores. During end-to-end training, it can be flexibly integrated into the image feature embedding module and then the multi-label outputs can be adaptively recalibrated with these scores (Chen, et al. 2020).

Since 2020, knowledge graphs have also been explored in COVID-19 related research and shown noticeable performance improvements. Zheng et al. pointed out that current deep learning-based work suffered from the multimodal data adequacy issues and that multimodal information should be considered together to make accurate inferences (Zheng, et al. 2021). To solve this, they proposed a multimodal knowledge graph attention embedding specifically for diagnosing COVID-19. Their method not only learned the relational embedding from nodes in a constituted knowledge graph but also had access to the medical knowledge, aiming to improve the classification performance through the medical knowledge attention mechanism. According to Mudiyanselage et al., limited correlations of the transferred features from the pre-trained model to a specific medical imaging domain, and the overfitting on fewer data can be the reasons for the poor performance for unseen data in COVID-19 classification (Mudiyanselage, et al. 2021). They proposed to exploit the relational information between data instances and their features as the knowledge, and to apply graph convolutions to learn the graph data representations, which is not feasible with conventional convolutions on the Euclidean domain.

The aforementioned work incorporated knowledge graphs with disease classification using three approaches: embed visual features to pre-constructed prior knowledge abnormality graph (Zhang, et al. 2020, Hou, Zhao and Hu 2021, Zhou, et al. 2021, Agu, et al. 2021, Sekuboyina, et al. 2021, Fu, et al. 2021, Chen, et al. 2020), region graph (Liu, et al. 2021, Xie, et al. 2019), pathology graph (Chen, et al. 2020), population graph (Chen, et al. 2021); extract and use visual features as graph nodes (Yu, Wang and Zhang 2021, Mudiyanselage, et al. 2021); use images and/or text descriptions of diagnose as graph nodes (Zheng, et al. 2021, Sekuboyina, et al. 2021). Though these work applied knowledge graphs in various ways, the results showed that applying knowledge graphs to disease classification helped boost the classification performance; for example, Zhang et al. achieved 1.4% improvement on average AUC, 4.7% AUC improvement on cardiomegaly, and 4.5% AUC improvement on atelectasis after adding knowledge graphs to the baseline DenseNet (Irvin, et al. 2019) model (Zhang, et al. 2020). Zhou et al. achieved a 3.77% improvement on average AUC when incorporating disease identification with prior knowledge on the NIH Chest X-Ray dataset and a 3% average AUC improvement on the CheXpert dataset (Zhou, et al. 2021).

Ref	Year	Task	Dataset	Dataset Info
(Xie, et al. 2019)	2019	Binary Classification	LIDC-IDRI (Armato III, et al. 2011)	1,018 clinical chest CT scans with lung nodules.
(Zhang, et al. 2020)	2020	Multi-label Classification	IU X-Ray (Demner-Fushman, et al. 2015)	3,955 radiology reports, 7470 chest X-ray images.
(Yu, Wang and Zhang 2021)	2020	Binary Classification	COVID-19 CT Report (Yang, et al. 2020)	728 images (349 COVID-19 and 379 Non-COVID) and their corresponding Chinese reports.
			Chest X-Ray Images (Pneumonia) (Kermany, Zhang and Goldbaum 2018)	5,863 chest X-ray images with two categories (Pneumonia and Normal).
(Chen, et al. 2020)	2020	Multi-label Classification	CheXpert (Irvin, et al. 2019)	224,316 chest radiographs of 65,240 patients, with 14 common disease labels.
			ChestX-Ray14 (Wang, et al. 2017)	112,120 frontal-view X-ray images, with the text-mined 14 common disease labels.

(Hou, Zhao and Hu 2021)	2021	Multi-label Classification	IU X-Ray (Demner-Fushman, et al. 2015)	3,955 radiology reports, 7470 chest X-ray images.
			MIMIC-CXR (Johnson, et al. 2019)	377,110 chest X-ray images associated with 227,835 reports.
(Zhou, et al. 2021)	2021	Multi-label Classification	CheXpert (Irvin, et al. 2019)	224,316 chest radiographs of 65,240 patients, with 14 common disease labels.
			NIH Chest X-Ray (Wang, et al. 2017)	112,120 frontal-view X-ray images with the text-mined 14 common disease labels.
(Agu, et al. 2021)	2021	Multi-label Classification	Chest ImaGenome (Wu, et al. 2021)	A scene graph data structure to describe 242,072 images.
(Chen, et al. 2021)	2021	Binary Classification	Autism Brain Imaging Data Exchange (ABIDE) (Martino, et al. 2014)	This dataset shares fMRI and the corresponding phenotypic data (e.g., age and gender) of 1,112 subjects, and notes whether these subjects have Autism Spectrum Disorder (ASD).
(Sekuboyina, et al. 2021)	2021	Multi-label Classification	CheXpert (Irvin, et al. 2019)	224,316 chest radiographs of 65,240 patients, with 14 common disease labels.
(Liu, et al. 2021)	2021	Binary Classification	DDSM (Lee, et al. 2017)	2,620 scanned film mammography studies.
(Mudiyanselage, et al. 2021)	2021	Binary Classification and Multi-label Classification	COVID-19 (Cohen, Morrison and Dao 2020), COVID-19 Radiography (Chowdhury,	150 CXR of Covid-19, 150 other pneumonia and another 150 instances for normal CXR images.

			et al. 2020) (Rahman, et al. 2020)	
(Zheng, et al. 2021)	2021	Binary Classification	COVID-19 multi-modal dataset	1,393 doctor–patient dialogues and 3706 images (347 X-ray + 2,598 CT + 761 ultrasound) about COVID-19 patients and 607 non-COVID-19 patient dialogues and 10,754 images (9658 X-ray + 494 CT + 761 ultrasound), and the fine-grained labels of all.
(Fu, et al. 2021)	2021	Multi-label Classification	7PC (Kawahara, et al. 2019)	1,011 lesion cases, and report comprehensive results over all the 7-point criteria and diagnosis

Table 1. Overview of datasets used in the disease classification articles.

DISEASE LOCALIZATION AND SEGMENTATION

In medical imaging, disease localization and segmentation are useful for clinical diagnosis, disease assessment, and treatment planning (Sharma and Aggarwal 2010). Previous supervised methods suffered from the lack of finely annotated data, and weakly-supervised methods often generated inaccurate or incomplete regions (Qi, et al. 2021). One can obtain complementary information by considering the relationship between anatomical regions and between different images, leading to more accurate disease localizations. This also aligns with the prior knowledge in the medical domain: radiologists are trained to read many X-ray images and analyze them by recognizing and comparing the differences (e.g., shapes, textures, and contrast), including comparing multiple images and comparing different regions of one single image. Infusing knowledge graphs into the system offers the potential for more accurate localizations and segmentations. Table 2 lists the overview of included articles and the datasets that are used. The most commonly used datasets for this task are CheXpert (Irvin, et al. 2019) and NIH Chest X-Ray/ChestX-Ray 14 (Wang, et al. 2017). Both datasets include medical images and associated medical reports, and disease labels.

Peng et al. used lung anatomy prior knowledge to help identify the fissure region of interest, then an oriented derivative of stick filter was applied to isolate the plate-like structures from clutters for lobar fissure verification. Finally, the lung lobe segmentation was implemented using a surface fitting model to complete the incomplete fissure surface (Peng, et al. 2021). Qi et al. noted that one reason for incomplete localization regions was the neglect of the pathological implications hidden in the relationship across anatomical regions within each image and the relationship across images (Qi, et al. 2021). Hence, they proposed to model the inter-image relationships using an inter-image graph to compare multiple images and to model the intra-image relationships using an intra-image graph to compare different regions. These cross-image and cross-region relationships were used as the contextual and compensating knowledge and were incorporated for disease localizations. Through ablation study, they showed that the model employing the intra-image and inter-image prior knowledge outperformed the localization accuracy of the baseline model by 0.08, 0.11, and 0.1 when the intersection over union (IoU) threshold was 0.3, 0.5, and 0.7. Zhao et al. also noted that general weakly-supervised disease localization methods failed to consider the characteristics of chest X-ray images, such as the highly structural attributes (Zhao, Qi and Li 2021). They used a very similar method to Qi et al. (Qi, et al. 2021), which integrated the intra-image anatomical structural knowledge and inter-image knowledge information into a unified end-to-end framework.

The aforementioned articles incorporated knowledge graphs with disease localization and segmentation using two approaches: embed visual features to pre-constructed prior knowledge region graph (Qi, et al. 2021, Peng, et al. 2021, Zhao, Qi and Li 2021); use images as graph nodes (Qi, et al. 2021, Zhao, Qi and Li 2021).

Ref	Year	Method	Dataset	Dataset Info
(Zhou, et al. 2021)	2021	Visual spatial convolution, dual-weighting graph convolution	CheXpert (Irvin, et al. 2019)	224,316 chest radiographs of 65,240 patients, with 14 common disease labels.
			NIH Chest X-Ray (Wang, et al. 2017)	112,120 frontal-view X-ray images with the text-mined 14 common disease labels.
(Peng, et al. 2021)	2021	Fissure verification, Surface fitting	LObe and Lung Analysis 2011 (LOLA11) (Lassen, et al. 2011)	A dataset of chest CT scans with varying abnormalities for which reference standards of lung and lobe segmentations have been established.
(Qi, et al. 2021), (Zhao, Qi and Li 2021)	2021	U-Net (Ronneberger, Fischer and Brox 2015)	NIH Chest X-Ray (Wang, et al. 2017)	112,120 frontal-view X-ray images with the text-mined 14 common disease labels.

Table 2. Overview of datasets used in the disease localization and segmentation articles.

REPORT GENERATION

Natural language captioning aims to summarize visual information in one sentence or generate one topic-related paragraph (Li, et al. 2019). Medical report generation translates the medical images to human-readable medical reports, which requires an increased capability to cover accurate abnormal terminologies, understand the medical domain knowledge, and describe the findings in a semantic-coherent and fine-grained manner that should satisfy both medical common sense and human logic (Li, et al. 2020). Outstanding challenges associated with automatic medical report generation are how to successfully detect visual groundings and incorporate medical domain knowledge. To write a medical image report, radiologists will first check a patient's images, carefully inspect the abnormal regions to identify the findings, then describe the abnormal findings in detail based on prior medical experiences and medical knowledge. Only employing the global images as input and training the language model with the dataset's corpora alone cannot provide the underlying prior knowledge vital for accurate reporting. Several works infused knowledge graphs into report generation and showed the performance gain on the quality of generated reports. The datasets used by the included articles related to report generation are listed in Table 3, and the most used dataset is IU X-Ray (Demner-Fushman, et al. 2015).

Zhang et al. utilized a pre-constructed graph embedding module on multiple disease findings as the prior knowledge to assist report generation. The incorporation of a knowledge graph allowed for dedicated feature learning for each disease finding and the relationship modeling between them (Zhang, et al. 2020). The knowledge graph module improved the baseline SentSAT model (Yuan, et al. 2019) on nearly all report generation evaluation metrics, especially 0.036 improvements on the CIDEr metric. Li et al. noted the significant challenges towards bridging visual and linguistic modalities; hence they proposed to encode visual features as an abnormality knowledge graph, which incorporated the visual features with prior medical knowledge, and was then used to guide the report template retrieval-paraphrase process or used for disease classification (Li, et al. 2019). Similarly, Liu et al. noted that visual and textual data biases remained a challenge for data-driven report generation systems, so in addition to the disease-tag attended visual features and the disease-attended textual features, they proposed to explore the prior knowledge from a pre-defined medical knowledge graph guided by attended-image features, and further adaptively distill the knowledge for report generation (Liu, et al. 2021). Through ablation study, they show that removing the prior knowledge graph module from the proposed model will cause a significant drop in all evaluation metrics, especially 0.66 drop on CIDEr, 0.34 drop on BLEU-1. Li et al. pointed out that previous methods suffered from the deviation that taught models to generate inessential sentences regularly. Therefore, inspired by Generative Pre-Training, they proposed integrating internal visual feature fusion and external medical linguistic information to guide medical knowledge transfer and learning through a medical graph encoder (Li, et al. 2020). The included articles incorporated knowledge graphs with report generation using one common approach: embed visual features to pre-constructed prior knowledge abnormality graph (Li, et al. 2019, Zhang, et al. 2020, Liu, et al. 2021, Li, et al. 2020).

Ref	Year	Method	Dataset	Dataset Info
(Li, et al. 2019)	2019	Graph Transformer	CX-CHR dataset	Private dataset. Consists of 35,609 patients and 45,598 images with corresponding reports.
			IU X-Ray (Demner-Fushman, et al. 2015)	3,955 radiology reports, 7,470 chest X-ray images.
(Zhang, et al. 2020)	2020	Two-level LSTM	IU X-Ray (Demner-Fushman, et al. 2015)	3,955 radiology reports, 7,470 chest X-ray images.
(Li, et al. 2020)	2020	Generative Pre-Training (Radford, et al. 2018)	CX-CHR dataset	Private dataset. Consists of 35,609 patients and 45,598 images with corresponding reports.
			COVID-19 CT Report (Yang, et al. 2020)	728 images (349 COVID-19 and 379 Non-COVID) and their corresponding Chinese reports.

(Liu, et al. 2021)	2021	Multi-Head Attention, Feed-Forward Network	IU X-Ray (Demner-Fushman, et al. 2015)	3,955 radiology reports, 7,470 chest X-ray images.
			MIMIC-CXR (Johnson, et al. 2019)	377,110 chest X-ray images associated with 227,835 reports.

Table 3. Overview of datasets used in the report generation articles.

IMAGE RETRIEVAL

Medical image retrieval systems have enormous potential in medical domain applications (Qayyum, et al. 2017). It can be beneficial for the clinical decision-making process to find other related images that belong to the same modality, fall in the same anatomic region, and belong to the same disease. Medical image retrieval systems can greatly assist doctors' diagnosis process by retrieving images with known pathologies that are similar to a patient's image(s) (Qayyum, et al. 2017). Furthermore, visual retrieval methods could help researchers find relevant images from large repositories in teaching and research. Visual features not only make possible the retrieval of cases with patients having similar diagnoses, but also enable retrieval of cases with visual similarity but different diagnoses. Current image retrieval systems generally use primitive features such as color or texture, or logical features such as object and their relationships, to represent images (Hwang, Lee and Choi 2012). No medical knowledge was used in this process; hence such systems provide poor results in the medical domain. This loss of information can be called the semantic gap, which can be reduced by exploiting all sources of information (Putzu, Piras and Giacinto 2020). Table 4 lists the overview of datasets used in the image retrieval articles included in this review.

Lacoste et al. presented their work on medical image retrieval mainly based on incorporating medical knowledge in the system within a fusion framework (Lacoste, et al. 2006). The text knowledge infused was from the Unified Medical Language System (UMLS) sources. The visual knowledge was from semantic features learned from examples and did not rely on robust region segmentation. UMLS concepts allowed the system to work at a higher semantic level and standardize the semantic index of medical data, facilitating the communication between visual end textual indexing and retrieval. Racoceanu et al. employed similar methods by using global indexing to access image modality and local indexing to access local semantic features to fuse the textual and visual knowledge into image retrieval (Racoceanu, et al. 2006). Two included articles incorporated knowledge graphs with image retrieval using one common approach: represent images and texts in UMLS graphs (Lacoste, et al. 2006, Racoceanu, et al. 2006).

Ref	Year	Method	Dataset	Dataset Info
(Lacoste, et al. 2006), (Racoceanu, et al. 2006)	2006	Support Vector Machine (SVM)	Clef Medical Image Database	50,000 medical images with the associated medical report in English, German, French.

Table 4. Overview of datasets used in the image retrieval articles.

Limitations and Future Directions

DISEASE CLASSIFICATION

Most articles share some common limitations in this review. The small sizes of datasets (on average, 129,788 images, ranging from 450-384,580) still hinder the model ability to produce more convincing results. The graph construction and the feature reconstruction are vital parts of most of the work. However, the graphs were constructed on a given dataset, making it inconvenient to extend to other domains. For example, some graphs were designed as components of the proposed model for diagnosing chest diseases, which would not work for a brain tumor diagnosis task. If other researchers want to deal with a problem in another area, building a new graph using a similar approach would be necessary. Also, an encoding component pre-trained towards a specific task (multi-label classification) could result in representations that do not generalize well across tasks. Furthermore, global classifications can be unreliable even when the label is correct, as the classifier might find the correct label for the wrong reason at an irrelevant spot (Agu, et al. 2021).

For the future direction, one can consider a semi-supervised learning framework to reduce the need for data annotation. Also, we can think of considering more sophisticated graph structures (with more detailed disease relationship modeling) in the future. Other approaches can be explored to better incorporate visual and semantic features. It is worth exploring a task-agnostic representation learning framework for better generalizability. Combining the encoding and embedding modules resulting in a fully end-to-end formulation is also a future research direction.

DISEASE LOCALIZATION AND SEGMENTATION

The included articles regarding disease localization and segmentation in this review did not consider the label uncertainty, which is worth exploring to improve the performance. The reported results showed that small targets (e.g., Atelectasis, Effusion, Nodule) were more challenging to localize due to their relatively smaller size. Algorithms applicable to the localization of small targets (e.g., Atelectasis, Effusion, Nodule) are worth exploring.

Existing work like (Peng, et al. 2021) heavily depends on airway segmentation, which is an arduous task and sensitive to the image quality. They employ lung anatomy knowledge to segment pulmonary fissures, which is time-consuming. Poorly segmented airways and pulmonary arteries may cause parts of fissures to be undetected in some cases. Verified fissures largely determined the fitting accuracy of the estimated lobe boundaries. To conclude, a better-segmented method of fissure detection and lung lobe segmentation is needed to pursue.

REPORT GENERATION

Currently, most work applying knowledge graphs into report generation uses visual features for graph feature initializations. It is worth exploring different fusion methods to combine knowledge graphs with multimodal features. We can also explore a general captioning framework guided by auxiliary signals to encode and decode general corpora knowledge for report generation tasks.

IMAGE RETRIEVAL

In this review, two included articles regarding image retrieval used global and local indexing to infuse additional visual, textual, and knowledge graph features into image retrieval. One can further explore the potential of an early fusion scheme using appropriate clustering methods. It is also worth exploring the visual filtering based on the local information from the semantic local indexing module to distill visual features for better performance.

Conclusion

This review discussed the current work on knowledge graph applications in medical imaging analysis and identified the limitations and future directions. We looked at the proven success of applying knowledge graphs into four medical imaging tasks: disease classification, disease localization and segmentation, report generation, and image retrieval. We identified the limitations due to limited annotated data for some supervised tasks and weak generalizability. We also identified potential future directions, for example, employing semi-supervised framework, exploring different fusion methods, exploring task-agnostic models that may improve the opportunities for better performance.

Bibliography

Agu, Nkechinyere N., Joy T. Wu, Hanqing Chao, Ismini Lourentzou, Arjun Sharma, Mehdi Moradi, Pingkun Yan, and James Hendler. 2021. "AnaXNet: Anatomy Aware Multi-label Finding Classification in Chest X-ray." *Proceedings of the 24th International Conference on Medical Image Computing and Computer Assisted Intervention.* Strasbourg, France.

Armato III, Samuel G, Geoffrey McLennan, Luc Bidaut, Michael F McNitt-Gray, Charles R Meyer, Anthony P Reeves, Binsheng Zhao, Denise R Aberle, Claudia I Henschke, and Eric A Hoffman. 2011. "The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI): A completed reference database of lung nodules on CT scans." *Medical physics.*

Auer, Sören, Christian Bizer, Georgi Kobilarov, Jens Lehmann, Richard Cyganiak, and Zachary Ives. 2007. "DBpedia: A Nucleus for a Web of Open Data." *Asian Semantic Web Conference.*

Banerjee, Pratyay, and Chitta Baral. 2020. "Self-supervised Knowledge Triplet Learning for Zero-shot Question Answering." *The 2020 Conference on Empirical Methods in Natural Language Processing.*

Bastos, Anson, Abhishek Nadgeri, Kuldeep Singh, Isaiah Onando Mulang, Saeedeh Shekarpour, Johannes Hoffart, and Manohar Kaul. 2021. "RECON: Relation Extraction using Knowledge Graph Context in a Graph Neural Network." *The World Wide Web Conference 2021*.

Berners-Lee, Tim, Christian. Bizer, and Tom Heath. 2009. "Linked Data - The Story So Far." *Int. J. Semantic Web Inf. Syst.*

Bollacker, Kurt, Colin Evans, Praveen Paritosh, Tim Sturge, and Jamie Taylor. 2008. "Freebase: a collaboratively created graph database for structuring human knowledge." *SIGMOD Conference.*

Carlson, Andrew, Justin Betteridge, Bryan Kisiel, Burr Settles, Estevam R. Hruschka Jr., and Tom M. Mitchell. 2010. "Toward an Architecture for Never-Ending Language Learning." *Proceedings of the Twenty-Fourth AAAI Conference on Artificial Intelligence.*

Chan, Maria F., Alon Witztum, and Gilmer Valdes. 2020. "Integration of AI and Machine Learning in Radiotherapy QA." *Frontiers in Artificial Intelligence.*

Chen, Bingzhi, Jinxing Li, Guangming Lu, Hongbing Yu, and David Zhang. 2020. "Label Co-Occurrence Learning With Graph Convolutional Networks for Multi-Label Chest X-Ray Image Classification." *IEEE Journal of Biomedical and Health Informatics* 2292-2302.

Chen, Hao, Fu-Zhen Zhuang, Li Xiao, Ling Ma, Haiyan Liu, Ruifang Zhang, Huiqin Jiang, and Qing He. 2021. "AMA-GCN: Adaptive Multi-layer Aggregation Graph Convolutional Network for Disease Prediction." *Proceedings of 2021 International Joint Conference on Artificial Intelligence.*

Chilińskiab, Mateusz, Kaustav Senguptab, and Dariusz Plewczynski. 2021. "From DNA human sequence to the chromatin higher order organisation and its biological meaning: Using biomolecular interaction networks to understand the influence of structural variation on spatial genome organisation and its functional effect." *Seminars in Cell & Developmental Biology.*

Chowdhury, Muhammad EH, Tawsifur Rahman, Amith Khandakar, Rashid Mazhar, Muhammad Abdul Kadir, Zaid Bin Mahbub, Khandakar Reajul Islam, Muhammad Salman Khan, Atif Al Iqbal, and Nasser Emadi. 2020. "Can AI help in screening Viral and COVID-19 pneumonia?"

Cohen, Joseph Paul, Paul Morrison, and Lan Dao. 2020. "COVID-19 Image Data Collection."

Cui, Song, and Dhananjay Shrouty. 2020. "Interest Taxonomy: A knowledge graph management system for content understanding at Pinterest."

Dai, Yuanfei, Chenhao Guo, Wenzhong Guo, and Carsten Eickhoff. 2021. "Drug–drug interaction prediction with Wasserstein Adversarial Autoencoder-based knowledge graph embeddings." *Briefings in Bioinformatics.*

Demner-Fushman, Dina, Marc D. Kohli, Marc B. Rosenman, Sonya E. Shooshan, Laritza M. Rodriguez, AntaniS., George R. Thom, and Clement J. McDonald. 2015. "Preparing a collection of radiology examinations for distribution and retrieval." *Journal of the American Medical Informatics Association.*

Dong, Xin Luna, Xiang He, Andrey Kan, Xian Li, Yan Liang, Jun Ma, Yifan Ethan Xu, et al. 2020. "AutoKnow: self-driving knowledge collection for products of thousands of types."

Dong, Xin, Evgeniy Gabrilovich, Geremy Heitz, Wilko Horn, Ni Lao, Kevin Murphy, Thomas Strohmann, Shaohua Sun, and Wei Zhang. 2014. "Knowledge vault: a web-scale approach to probabilistic knowledge fusion." *The 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.* New York, NY, USA.

Färber, Michael. 2019. "The Microsoft Academic Knowledge Graph: {A} Linked Data Source with 8 Billion Triples of Scholarly Data." *Proceedings of the 18th International Semantic Web Conference.*

Fei, Hao, Yafeng Ren, Yue Zhang, Donghong Ji, and Xiaohui Liang. 2021. "Enriching contextualized language model from knowledge graph for biomedical information extraction." *Briefings in Bioinformatics.*

Fu, Xiaohang, Lei Bi, Ashnil Kumar, Michael Fulham, and Jinman Kim. 2021. "Graph-Based Intercategory and Intermodality Network for Multilabel Classification and Melanoma Diagnosis of Skin Lesions in Dermoscopy and Clinical Images."

Goodfellow, Ian, Yoshua Bengio, and Aaron Courville. 2016. *Deep Learning.* The MIT Press.

Hou, Daibing, Zijian Zhao, and Sanyuan Hu. 2021. "Multi-Label Learning With Visual-Semantic Embedded Knowledge Graph for Diagnosis of Radiology Imaging." *IEEE Access* (IEEE) 9.

Hwang, Kyung Hoon, Haejun Lee, and Duckjoo Choi. 2012. "Medical Image Retrieval: Past and Present." *Health Informatics Research.*

Irvin, Jeremy, Pranav Rajpurkar, Michael Ko, Yifan Yu, Silviana Ciurea-Ilcus, Chris Chute, Henrik Marklund, Behzad Haghgoo, Robyn Ball, and Katie Shpanskaya. 2019. "CheXpert: A Large Chest Radiograph Dataset with Uncertainty Labels and Expert Comparison."

Jaradeh, Mohamad Yaser, Kuldeep Singh, Markus Stocker, Andreas Both, and Sören Auer. 2021. "Better Call the Plumber: Orchestrating Dynamic Information Extraction Pipelines." *International Conference on Web Engineering.*

Ji, Shaoxiong, Shirui Pan, Erik Cambria, Pekka Marttinen, and Philip S. Yu. 2021. "A Survey on Knowledge Graphs: Representation, Acquisition, and Applications."

Johnson, Alistair EW, Tom J Pollard, Nathaniel R Greenbaum, Matthew P Lungren, Chih-ying Deng, Yifan Peng, Zhiyong Lu, Roger G Mark, Seth J Berkowitz, and Steven Horng. 2019. "MIMIC-CXR-JPG, a large publicly available database of labeled chest radiographs." *arXiv preprint arXiv:1901.07042.*

Kawahara, Jeremy, Sara Daneshvar, Giuseppe Argenziano, and Ghassan Hamarneh. 2019. "Seven-Point Checklist and Skin Lesion Classification Using Multitask Multimodal Neural Nets." *IEEE Journal of Biomedical and Health Informatics.*

Kermany, Daniel, Kang Zhang, and Michael Goldbaum. 2018. "Labeled Optical Coherence Tomography (OCT) and Chest X-Ray Images for Classification."

Lacoste, Caroline, Jean-Pierre Chevallet, Joo-Hwee Lim, Xiong Wei, Daniel Racoceanu, Diem Thi Hoang Le, Roxana Teodorescu, and Nicolas Vuillemenot. 2006. "IPAL Knowledge-based Medical Image Retrieval in ImageCLEFmed 2006." *7th Workshop of the Cross-Language Evaluation Forum.* CLEF 2006.

Lassen, Bianca, Eva M van Rikxoort, Michael Schmidt, Sjoerd Kerkstra, Bram van Ginneken, and Jan-Martin Kuhnigk. 2011. "Automatic Segmentation of the Lungs and Lobes from Thoracic CT Scans."

Lee, Rebecca Sawyer, Francisco Gimenez, Assaf Hoogi, Kanae Kawai Miyake, Mia Gorovoy, and Daniel L. Rubin. 2017. "A curated mammography data set for use in computer-aided detection and diagnosis research."

Li, Christy, Xiaodan Liang, Zhiting Hu, and Eric Xing. 2019. "Knowledge-Driven Encode, Retrieve, Paraphrase for Medical Image Report Generation." Proceedings of the AAAI Conference on Artificial Intelligence. 6666-6673.

Li, Mingjie, Fuyu Wang, Xiaojun Chang, and Xiaodan Liang. 2020. "Auxiliary Signal-Guided Knowledge Encoder-Decoder for Medical Report Generation."

Li, Sheng, Zhiqiang Tao, Kang Li, and Yun Fu. 2019. "Visual to Text: Survey of Image and Video Captioning." *IEEE Transactions on Emerging Topics in Computational Intelligence.*

Liu, Fenglin, Xian Wu, Shen Ge, Wei Fan, and Yuexian Zou. 2021. "Exploring and Distilling Posterior and Prior Knowledge for Radiology Report Generation." *Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR2021).* CVPR 2021.

Liu, Yuhang, Fandong Zhang, Chaoqi Chen, Siwen Wang, Yizhou Wang, and Yizhou Yu. 2021. "Act Like a Radiologist: Towards Reliable Multi-view Correspondence

Reasoning for Mammogram Mass Detection." *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI).*

Ma, Kaixin, Filip Ilievski, Jonathan Francis, Yonatan Bisk, Eric Nyberg, and Alessandro Oltramari. 2021. "Knowledge-driven Data Construction for Zero-shot Evaluation in Commonsense Question Answering." *35th AAAI Conference on Artificial Intelligence.*

Martino, Adriana Di, Chao-Gan Yan, Qingyang Li, Erin Denio, Francisco Xavier Castellanos, Kaat Alaerts, Jason Anderson, Michal Assaf, Susan Bookheimer, and Mirella Dapretto. 2014. "The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism."

McGuinness, D. 2004. *OWL Web ontology language overview.*

Miller, Eric. 2005. *An Introduction to the Resource Description Framework.*

Miller, G. 1995. "WordNet: a lexical database for English." *Commun. ACM.*

Mudiyanselage, Thosini Bamunu, Nipuna Senanayake, Chunyan Ji, Yi Pan, and Yanqing Zhang. 2021. "Covid-19 Detection from Chest X-ray and Patient Metadata using Graph Convolutional Neural Networks."

Obaid, Kavi B., Subhi R. M. Zeebaree, and Omar M. Ahmed. 2020. "Deep Learning Models Based on Image Classification: A Review." *International Journal of Science and Business.*

Peng, Yuanyuan, Hualan Zhong, Zheng Xu, Hongbin Tu, Xiong Li, and Lan Peng. 2021. "Pulmonary Lobe Segmentation in CT Images Based on Lung Anatomy Knowledge." *Mathematical Problems in Engineering.*

Putzu, Lorenzo, Luca Piras, and Giorgio Giacinto. 2020. "Convolutional neural networks for relevance feedback in content based image retrieval." *Multimedia Tools and Applications.*

Qayyum, Adnan, Syed Muhammad Anwar, Muhammad Awais, and Muhammad Majid. 2017. "Medical Image Retrieval using Deep Convolutional Neural Network." *Nerucomputing.*

Qi, Baolian, Gangming Zhao, Xin Wei, Chaowei Fang, Chengwei Pan, Jinpeng Li, Huiguang He, and Licheng Jiao. 2021. "GREN: Graph-Regularized Embedding Network for Weakly-Supervised Disease Localization in X-ray images."

Racoceanu, Daniel, Caroline Lacoste, Roxana Teodorescu, and Nicolas Vuillemenot. 2006. "A Semantic Fusion Approach Between Medical Images and Reports Using UMLS." *Information Retrieval Technology.*

Radford, Alec, Karthik Narasimhan, Tim Salimans, and Ilya Sutskever. 2018. "Improving Language Understanding by Generative Pre-Training."

Rahman, Tawsifur, Amith Khandakar, Yazan Qiblawey, Anas Tahir, Serkan Kiranyaz, Saad Bin Abul Kashem, Mohammad Tariqul Al Islam, Somaya Maadeed, Susu M Zughaier, and Muhammad Salman Khan. 2020. "Exploring the Effect of Image Enhancement Techniques on COVID-19 Detection using Chest X-ray Images."

Richens, R. H. 1956. "Preprogramming for mechanical translation." *Mechanical Translation and Computational Linguistics.*

Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. 2015. "U-Net: Convolutional Networks for Biomedical Image Segmentation." *MICCAI.*

Sekuboyina, Anjany, Daniel Oñoro-Rubio, Jens Kleesiek, and Brandon Malone. 2021. "A Relational-learning Perspective to Multi-label Chest X-ray Classification." *IEEE ISBI.*

Shan, Hongming, Xun Jia, Pingkun Yan, Yunyao Li, Harald Paganetti, and Ge Wang. 2020. "Synergizing medical imaging and radiotherapy with deep learning." *Machine Learning: Science and Technology.*

Sharma, Neeraj, and Lalit M. Aggarwal. 2010. "Automated medical image segmentation techniques." *Journal of Medical Physics.*

Shortliffe, Edward. 1976. *Computer-based medical consultations, MYCIN.*

Skripcak, Tomas, Claus Belka, Walter Bosch, Carsten Brink, Thomas Brunner, Volker Budach, Daniel Büttner, Jürgen Debus, Andre Dekker, and Cai Grau. 2014. "Creating a data exchange strategy for radiotherapy research: towards federated databases and anonymised public datasets." *Radiother Oncol.*

Suchanek, Fabian M., Gjergji Kasneci, and Gerhard Weikum. 2007. "Yago: a core of semantic knowledge."

Sumithra, M. Kani, and Rajeswari Sridhar. 2020. "Information Retrieval in Financial Documents." *Evolving Technologies for Computing, Communication and Smart World.*

Vrandečić, Denny, and Markus Krötzsch. 2015. "Wikidata: a free collaborative knowledgebase." *Communications of the ACM.*

Wang, Pengfei, Yu Fan, Long Xia, Wayne Xin Zhao, ShaoZhang Niu, and Jimmy Huang. 2020. "KERL: A Knowledge-Guided Reinforcement Learning Model for Sequential Recommendation." *ACM SIGIR Conference on Research and Development in Information Retrieval.*

Wang, Xiang, Xiangnan He, Yixin Cao, Meng Liu, and Tat-Seng Chua. 2019. "KGAT: Knowledge Graph Attention Network for Recommendation." *KDD '19: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining.*

Wang, Xiaosong, Yifan Peng, Le Lu, Zhiyong Lu, Mohammadhadi Bagheri, and Ronald M. Summer. 2017. "ChestX-ray8: Hospital-scale Chest X-ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases."

Wu, Joy T, Nkechinyere Nneka Agu, Ismini Lourentzou, Arjun Sharma, Joseph Alexander Paguio, Jasper Seth Yao, Edward Christopher Dee, William G Mitchell, Satyananda Kashyap, and Andrea Giovannini. 2021. "Chest ImaGenome Dataset for Clinical Reasoning."

Xi, Jianing, Liping Ye, Qinghua Huang, and Xuelong Li. 2021. "Tolerating Data Missing in Breast Cancer Diagnosis from Clinical Ultrasound Reports via Knowledge Graph Inference." *KDD '21: Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining.*

Xiang, Wang, Tinglin Huang, Dingxian Wang, Yancheng Yuan, Zhenguang Liu, Xiangnan He, and Tat-Seng Chua. 2021. "Learning Intents behind Interactions with Knowledge Graph for Recommendation." *WWW '21: Proceedings of the Web Conference 2021*.

Xie, Xiaozheng, Jianwei Niu, Xuefeng Liu, Zhengsu Chen, Shaojie Tang, and Shui Yu. 2021. "A survey on incorporating domain knowledge into deep learning for medical image analysis." *Medical Image Analysis.*

Xie, Yutong, Yong Xia, Jianpeng Zhang, Yang Song, Dagan Fang, Michael Fulham, and Weidong Cai. 2019. "Knowledge-based Collaborative Deep Learning for Benign-Malignant Lung Nodule Classification on Chest CT." *IEEE Transactions on Medical Imaging* 991-1004.

Yang, Xingyi, Xuehai He, Jinyu Zhao, Yichen Zhang, Shanghang Zhang, and Pengtao Xie. 2020. "COVID-CT-Dataset: A CT Scan Dataset about COVID-19."

Yu, Xiang, Shui-Hua Wang, and Yu-Dong Zhang. 2021. "CGNet: A graph-knowledge embedded convolutional neural network for detection of pneumonia." *Information Processing and Management* 58 (1).

Yu, Yue, Kexin Huang, Chao Zhang, Lucas M Glass, Jimeng Sun, and Cao Xiao. 2021. "SumGNN: multi-typed drug interaction prediction via efficient knowledge graph summarization." *Bioinformatics.*

Yuan, Jianbo, Haofu Liao, Rui Luo, and Jiebo Luo. 2019. "Automatic Radiology Report Generation based on Multi-view Image Fusion and Medical Concept Enrichment." *MICCAI 2019.*

Zhang, Yixiao, Xiaosong Wang, Ziyue Xu, Qihang Yu, Alan Yuille, and Daguang Xu. 2020. "When Radiology Report Generation Meets Knowledge Graph." Proceedings of the AAAI Conference on Artificial Intelligence. 12910-12917.

Zhao, Gangming, Baolian Qi, and Jinpeng Li. 2021. "Cross Chest Graph for Disease Diagnosis with Structural Relational Reasoning."

Zheng, Wenbo, Lan Yan, Chao Gou, Zhi-Cheng Zhang, Jun Jason Zhang, Ming Hu, and Fei-Yue Wang. 2021. "Pay attention to doctor-patient dialogues: Multi-modal knowledge graph attention image-text embedding for COVID-19 diagnosis." *Information Fusion.*

Zhou, Yi, Tianfei Zhou, Tao Zhou, Huazhu Fu, Jiacheng Liu, and Ling Shao. 2021. "Contrast-Attentive Thoracic Disease Recognition With Dual-Weighting Graph Reasoning." *IEEE Transactions on Medical Imaging* (IEEE) 40: 1196-1206.