Survey Paper

A survey on incorporating domain knowledge into deep learning for medical image analysis

Xiaozheng Xie[a], Jianwei Niu[a,b], Xuefeng Liu[a,*], Zhengsu Chen[a], Shaojie Tang[c], Shui Yu[d]

[a] *State Key Laboratory of Virtual Reality Technology and Systems, School of Computer Science and Engineering, Beihang University, 37 Xueyuan Road, Haidian District, Beijing 100191, China*
[b] *Beijing Advanced Innovation Center for Big Data and Brain Computing (BDBC) and Hangzhou Innovation Institute of Beihang University, 18 Chuanghui Street, Binjiang District, Hangzhou 310000, China*
[c] *Jindal School of Management, The University of Texas at Dallas, 800 W Campbell Rd, Richardson, TX 75080-3021, USA*
[d] *School of Computer Science, University of Technology Sydney, 15 Broadway, Ultimo NSW 2007, Australia*

ARTICLE INFO

ABSTRACT

Although deep learning models like CNNs have achieved great success in medical image analysis, the small size of medical datasets remains a major bottleneck in this area. To address this problem, researchers have started looking for external information beyond current available medical datasets. Traditional approaches generally leverage the information from natural images via transfer learning. More recent works utilize the domain knowledge from medical doctors, to create networks that resemble how medical doctors are trained, mimic their diagnostic patterns, or focus on the features or areas they pay particular attention to. In this survey, we summarize the current progress on integrating medical domain knowledge into deep learning models for various tasks, such as disease diagnosis, lesion, organ and abnormality detection, lesion and organ segmentation. For each task, we systematically categorize different kinds of medical domain knowledge that have been utilized and their corresponding integrating methods. We also provide current challenges and directions for future research.

© 2021 Elsevier B.V. All rights reserved.

Contents

---

* Corresponding author.
  E-mail address: liu_xuefeng@buaa.edu.cn (X. Liu).

1. Introduction

Recent years have witnessed a tremendous progress in computer-aided detection/diagnosis (CAD) in medical imaging and diagnostic radiology, primarily thanks to the advancement of deep learning techniques. Having achieved great success in computer vision tasks, various deep learning models, mainly convolutional neural networks (CNNs), soon be applied to CAD. Among the applications are the early detection and diagnosis of breast cancer, lung cancer, glaucoma, and skin cancer (Esteva et al., 2017; Shin et al., 2019; Zhou et al., 2017; Zhu et al., 2018).

However, the small size of medical datasets continues to be an issue in obtaining satisfactory deep learning model for CAD; in general, bigger datasets result in better deep learning models (Halevy et al., 2009). In traditional computer vision tasks, there are many large-scale and well-annotated datasets, such as ImageNet[1] (over 14M labeled images from 20k categories) and COCO[2] (with more than 200k annotated images across 80 categories). In contrast, some popular publicly available medical datasets are much smaller (see Table 1). For example, among the datasets for different tasks, only ChestX-ray14 and DeepLesion, contain more than 100k labeled medical images, while most datasets only have a few thousands or even hundreds of medical images.

The lack of medical datasets is represented in three aspects. First, the number of medical images in datasets is usually small. This problem is mainly due to the high cost associated with the data collection. Medical images are collected from computerized tomography (CT), Ultrasonic imaging (US), magnetic resonance imaging (MRI) scans, positron emission tomography (PET), all of which are expensive and labor-intensive. Second, only a small portion of medical images are annotated. These annotations including classification labels (e.g., benign or malignant), the segmentation annotations of lesion areas, etc., require efforts from experienced doctors. Third, it is difficult to collect enough positive cases for some rare diseases to obtain the balanced datasets.

One direct consequence of the lack of well annotated medical data is that the trained deep learning models can easily suffer from the overfitting problem (Chen et al., 2016a). As a result, the models perform very well on training datasets, but fail when dealing with new data from the problem domain. Correspondingly, many existing studies on medical image analysis adopt techniques from computer vision to address overfitting, such as reducing the complexity of the network (Ciompi et al., 2015; Shin et al., 2016), adopting some regularization techniques (Erickson et al., 2017), or using data augmentation strategies (Zhu et al., 2017).

However, in essence, both decreasing model complexity and leveraging data augmentation techniques only focus on the target task on the given datasets, but *do not introduce any new information into deep learning models*. Nowadays, introducing more information, beyond the given medical datasets has become a more promising approach to address the problem of small-sized medical datasets.

The idea of introducing external information to improve the performance of deep learning models for CAD is not new. For example, it is common practice to first train a deep learning model on some natural image datasets like ImageNet, and then fine tune them on target medical datasets (Huynh et al., 2016a). This process, called transfer learning (Pan and Yang, 2009), implicitly introduces information from natural images. Besides natural images, multi-modal medical datasets or medical images from different but related diseases can also be used to improve the performance of deep learning models (Samala et al., 2018; Yu et al., 2019).

Moreover, as experienced medical doctors (e.g., radiologists, ophthalmologists, and dermatologists) can generally give fairly accurate results, it is not surprising that their knowledge may help deep learning models to better accomplish the designated tasks. The domain knowledge of medical doctors includes the way they browse images, the particular areas they usually focus on, the features they give special attention to, and the anatomical prior knowledge they used. These types of knowledge are accumulated, summarized, and validated by a large number of practitioners over many years based on a huge amount of cases. Note that in this survey any network that incorporate one of these types of knowledge in their training or designing pro-

---

[1] http://www.image-net.org/
[2] http://mscoco.org/

Table 1
Examples of popular datasets in the medical domain.

Name	Purpose	Type	Imaging	Number of Images
ADNI (Weiner et al., 2017)	Classification	Brain	Multiple	1921 patients
ABIDE (Di Martino et al., 2014)	Classification	Brain	MRI	539 patients and 573 controls
ACDC (MICCAI, 2017)	Classification	Cardiac	MRI	150 patients
ChestX-ray14 (Wang et al., 2017a)	Detection	Chest	X-ray	112,120 images from 30,805 patients
LIDC-IDRI (Armato III et al., 2011)	Detection	Lung	CT, X-ray	1018 patients
LUNA16 (Setio et al., 2017)	Detection	Lung	CT	888 images
MURA (Rajpurkar et al., 2017)	Detection	Musculo-skeletal	X-ray	40,895 images from 14,982 patients
BraTS2018 (Bakas et al., 2018)	Segmentation	Brain	MRI	542 images
STARE (Hoover et al., 2000)	Segmentation	Eye	SLO	400 images
DDSM (Heath et al., 2000)	Classification Detection	Breast	Mammography	2500 patients
DeepLesion (Yan et al., 2018)	Classification Detection	Multiple	CT	32,735 images from 4427 patients
Cardiac MRI (Alexander A., 2008)	Classification Segmentation	Cardiac	MRI	7980 images from 33 cases
ISIC 2018 (Codella et al., 2019)	Classification Detection Segmentation	Skin	Dermoscopic	13,000 images

cess should be regarded as the one incorporated medical domain knowledge.

In this survey, we focus on the three main tasks of medical image analysis: (1) disease diagnosis, (2) lesion, organ and abnormality detection, and (3) lesion and organ segmentation. We also include other related tasks such as the image reconstruction, image retrieval and report generation. This survey demonstrates that, for almost all tasks, identifying and carefully integrating one or more types of domain knowledge related to the designated task will improve the performance of deep learning models. We organize existing works according to the following three aspects: the types of tasks, the types of domain knowledge that are introduced, and the ways of introducing the domain knowledge.

More specifically, in terms of the types of domain knowledge, some of them are of high-level such as training pattern (Maicas et al., 2018; Tang et al., 2018; Wang et al., 2018a) and diagnostic pattern. Some domain knowledge are low-level, such as particular features and special areas where medical doctors pay more attention to (Hsu et al., 2019). In particular, in disease diagnosis tasks, high-level domain knowledge is widely utilized. For an object detection task, the low-level domain knowledge, such as detection patterns and specific features where medical doctors give special attention is more commonly adopted. For lesion or organ segmentation tasks, anatomical priors and the knowledge from different modalities seem to be more useful (Chen et al., 2019a; Larrazabal et al., 2019; Valverde et al., 2019).

In terms of the integrating methods, various approaches have been designed to incorporate different types of domain knowledge into networks (Ji et al., 2020). For example, a simple approach is to concatenate hand-crafted features with the ones extracted from deep learning models (Xie et al., 2016). In some works, network architectures are revised to simulate the pattern of radiologists when they read images (Guan et al., 2018). Attention mechanism, which allows a network to pay more attention to a certain region of an image, is a powerful technique to incorporate domain knowledge of radiologists (Li et al., 2019b). In addition, multi-task learning and meta learning are also widely used to introduce medical domain knowledge into deep learning models (Chen et al., 2016c; Yue et al., 2019).

Although there are a number of reviews on deep learning for medical image analysis, including (Debelee et al., 2019; Litjens et al., 2017; Shen et al., 2017; Suzuki, 2017), they all describe the existing works from the application point of view, i.e., how deep learning techniques are applied to various medical applications. To the best of our knowledge, there is no review that gives systematic introduction on *how medical domain knowledge can help deep learning models*. This aspect, we believe, is the unique feature that distinguishes deep learning models for CAD from those for general computer vision tasks.

Fig. 1 gives the overview on how we organize the related researches. At the top level, existing studies are classified into three main categories according to their purposes: (1) disease diagnosis, (2) lesion, organ and abnormality detection, and (3) lesion and organ segmentation. In each category, we organize them into several groups based on the types of extra knowledge have been incorporated. At the bottom level, they are further categorized according to the different integrating approaches of the domain knowledge.
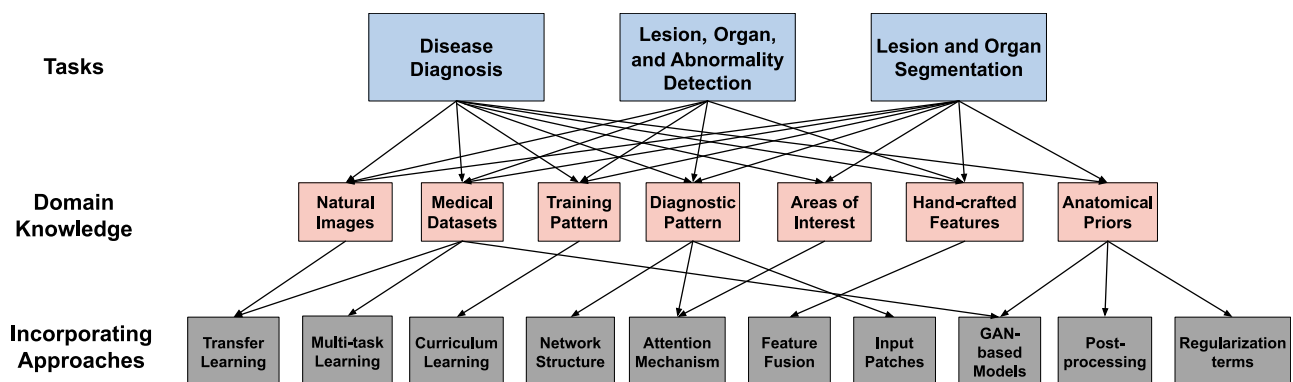

Fig. 1. Methods of information categorization and incorporating methods in disease diagnosis; lesion, organ, and abnormality detection; lesion and organ segmentation.
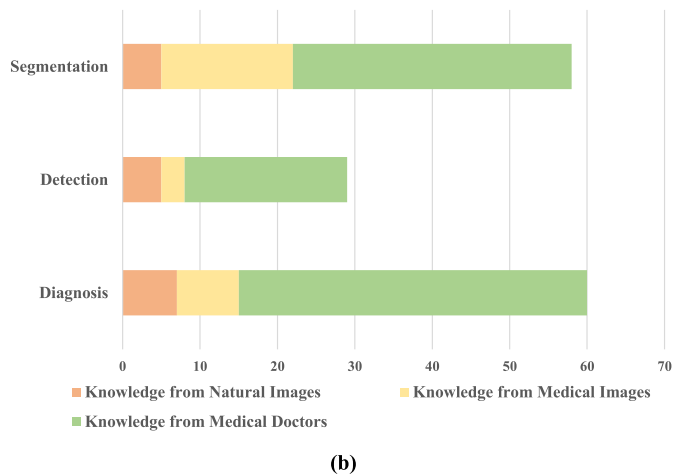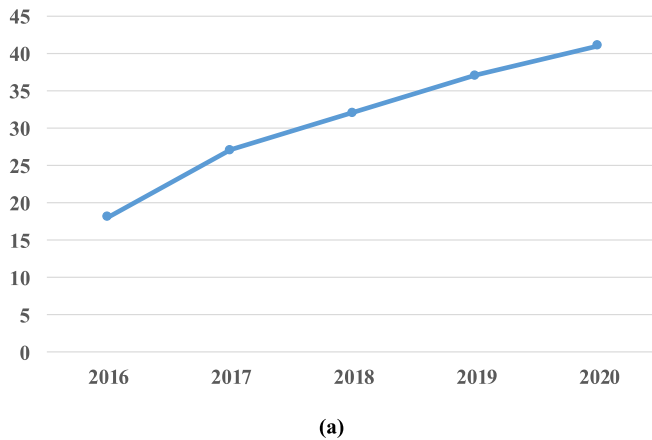
Fig. 2. (a) Number of papers arranged chronically (2016–2020). (b) The distribution of selected papers in different applications of medical image analysis.

This survey contains more than 200 papers (163 are with domain knowledge), most of which are published recently (2016–2020), on a wide variety of applications of deep learning techniques for medical image analysis. In addition, most of the corresponding works are from the conference proceedings for MICCAI, EMBC, ISBI and some journals such as TMI, Medical Image Analysis, JBHI and so on.

The distribution of these papers are shown in Fig. 2(a). It can be seen that the number of papers increases rapidly from 2016 to 2020. With respect to the applications, most of them are related to disease diagnosis and lesion/organ segmentation (shown in Fig. 2(b)). To sum up, with this survey we aim to:

- summarize and classify different types of domain knowledge in medical areas that are utilized to improve the performance of deep learning models in various applications;
- summarize and classify different ways of introducing medical domain knowledge into deep learning models;
- give the outlook of challenges and future directions in integrating medical domain knowledge into deep learning models.

The rest of the survey is organized as follows. Sections 2, 3 and 4 introduce the existing works for the major three tasks in medical image analysis. Besides these three major tasks, other tasks in medical image analysis are described in Section 5. In each section, we first introduce the general architectures of deep learning models for a task, and then categorize related works according to the types of the domain knowledge to be integrated. Various incorporating methods for each type of domain knowledge are then described. Section 6 discusses research challenges, and gives the outlook of future directions. Lastly, Section 7 concludes this survey. The structure of this survey is shown in Fig. 3.

2. Disease diagnosis

Disease diagnosis refers to the task of determining the type and condition of possible diseases based on the medical images provided. In this section, we give an overview of the deep learning models that generally used for disease diagnosis. Concretely, subsection 2.1 outlines the general structures of deep learning models used for disease diagnosis. Subsection 2.2 introduces the works that utilize knowledge from natural images or other medical datasets. Deep learning models that leverage knowledge from medical doctors are introduced in Subsection 2.3 in detail. Lastly, Subsection 2.4 summarizes the research of disease diagnosis.

2.1. General structures of deep learning models used for disease diagnosis

In the last decades, deep learning techniques, especially CNNs, have achieved a great success in disease diagnosis.

Fig. 4 shows the structure of a typical CNN that used for disease diagnosis in chest X-ray image. The CNN employs alternating convolutional and pooling layers, and contains trainable filter banks per layer. Each individual filter in a filter bank is able to generate a feature map. This process is alternated and the CNN can learn increasingly more and more abstract features that will later be used by the fully connected layers to accomplish the classification task.

Different types of CNN architectures, from AlexNet (Krizhevsky et al., 2012), GoogLeNet (Szegedy et al., 2015), VGGNet (Simonyan and Zisserman, 2014), ResNet (He et al., 2016) to DenseNet (Huang et al., 2017), have achieved a great success in the diagnosis of various diseases. For example, GoogLeNet, ResNet, and VGGNet models are used in the diagnosis of canine ulcerative keratitis (Kim et al., 2019), and most of them achieve accuracies of over 90% when classifying superficial and deep corneal ulcers. DenseNet is adopted to diagnose lung nodules on chest X-ray radiograph (Li et al., 2019d), and experimental results show that more than 99% of lung nodules can be detected. In addition, it is found that VGGNet and ResNet are more effective than other network structures for many medical diagnostic tasks (Liu et al., 2018; Xie et al., 2019; Guan et al., 2018; Mitsuhara et al., 2019).

However, the above works generally directly apply CNNs to medical image analysis or slightly modified CNNs (e.g., by modifying the number of kernals, the number of channels or the size of filters), and no medical knowledge is incorporated. In addition, these methods generally require large medical datasets to achieve a satisfactory performance.

In the following subsections, we systematically review on the research that utilizes medical domain knowledge for the disease diagnosis. The types of knowledge and the incorporating methods are summarized in Table 2.

2.2. Incorporating knowledge from natural datasets or other medical datasets

Despite the disparity between natural and medical images, it has been demonstrated that CNNs comprehensively trained on the large-scale well-annotated natural image datasets can still be helpful for disease diagnosis tasks (Wimmer et al., 2016). Intrinsically speaking, this transfer learning process introduces knowledge from natural images into the network for medical image diagnosis.
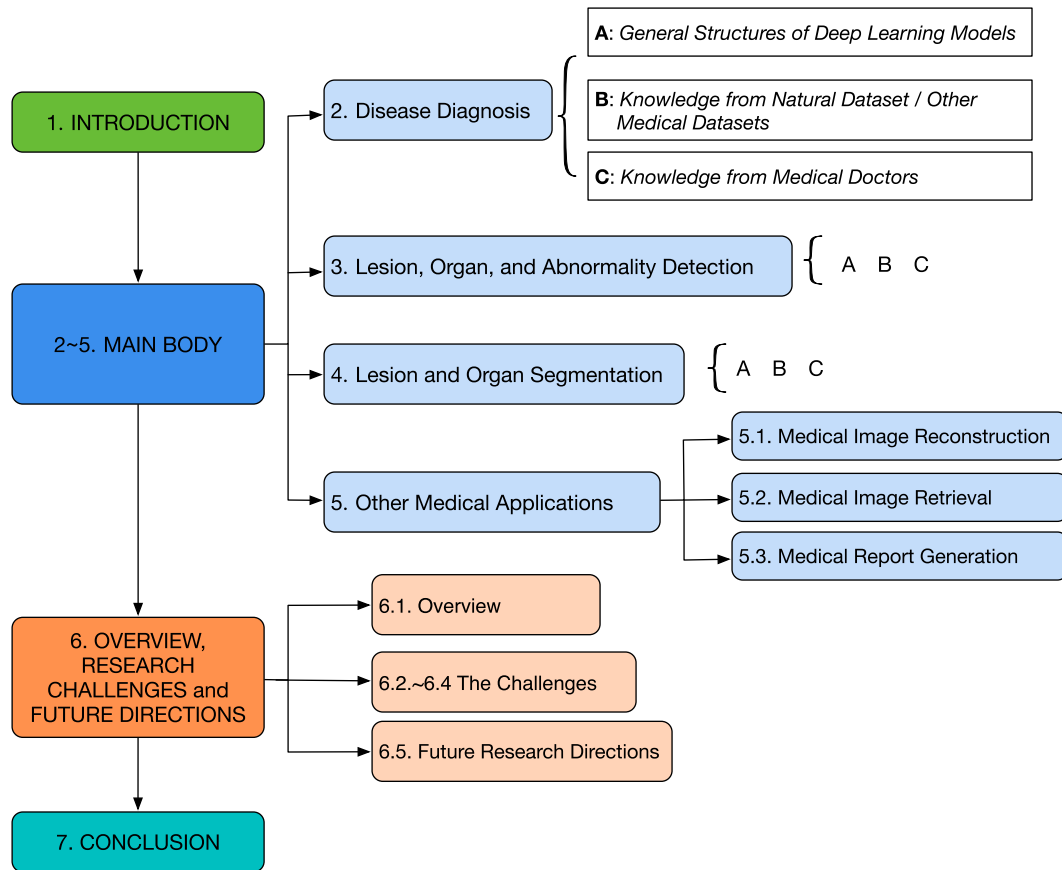
Fig. 3. The organizational structure of this survey.

According to (Litjens et al., 2017), the networks pre-trained on natural images can be leveraged via two different ways: by utilizing them as fixed feature extractors, and as an initialization which will then be fine-tuned on target medical datasets. These two strategies are illustrated in Fig. 5(a) and Fig. 5(b), respectively.

The first strategy takes a pre-trained network, removes its last fully-connected layer, and treats the rest of the network as a fixed feature extractor. Extracted features are then fed into a linear classifier (e.g., support vector machine (SVM)), which is trained on the target medical datasets. Applications in this category include mammography mass lesion classification (Huynh et al., 2016b) and chest pathology identification (Bar et al., 2015).

The success of leveraging information from natural images for disease diagnosis can be attributed to the fact that a network pre-trained on natural images, especially in the earlier layers, contain more generic features (e.g., edge detectors and color blob detectors) (Yosinski et al., 2014).

In the second strategy, the weights of the pre-trained network are fine-tuned based on the medical datasets. It is possible to fine-tune the weights of all layers in the network, or to keep some of the earlier layers fixed and only fine-tune some higher-level portion of the network. This can be applied to the classification of skin cancer (Esteva et al., 2017), breast cancer (Cao et al., 2018), thorax diseases (Wang et al., 2017), prostate cancer (Li et al., 2018) and interstitial lung diseases (Hussein et al., 2017).

Besides the information from natural images, using images from other medical datasets is also quite popular.

Medical datasets containing images of the same or similar modality as target images have similar distribution and therefore can be helpful. For example, to classify malignant and benign breast masses in digitized screen-film mammograms (SFMs), a multi-task transfer learning DCNN is proposed to incorporate the information from digital mammograms (DMs) (Samala et al., 2017). It is found to have significantly higher performance compared to the single-task transfer learning DCNN which only utilizes SFMs.

In addition, even medical images with different modalities can provide complementary information. For example, (Samala et al., 2018) uses a model pre-trained on a mammography dataset to show that it could obtain better results than models trained solely on the target dataset comprising digital breast tomosynthesis (DBT) images. Another example is in prostate cancer classification, where the radiofrequency ultrasound images are first used to train the DCNN, then the model is fine-tuned on B-mode ultrasound images (Azizi et al., 2017). Other examples of using the images from different modalities can be found in (Hadad et al., 2017; Han et al., 2020; Li et al., 2020).

Furthermore, as datasets of different classes can help each other in classification tasks (Xiao et al., 2020), medical datasets featuring images of a variety of diseases can also have similar morphological structures or distribution, which may be beneficial for other
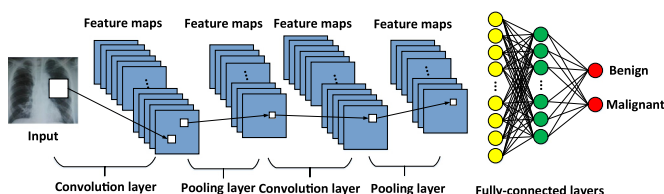

Fig. 4. A typical CNN architecture for medical disease diagnosis.

Table 2

A compilation of the knowledge and corresponding incorporating methods used in disease diagnosis.

Knowledge Source	Knowledge Type	Incorporating Method	References
Natural datasets	Natural images	Transfer learning	(Bar et al., 2015)(Esteva et al., 2017) (Wimmer et al., 2016)(Cao et al., 2018) (Wang et al., 2017) (Hussein et al., 2017) (Li et al., 2018) (Huynh et al., 2016b)
Medical datasets	Multi-modal images	Transfer learning	(Samala et al., 2018)(Hadad et al., 2017) (Samala et al., 2017)(Azizi et al., 2017) (Li et al., 2020e)(Han et al., 2020)
Medical doctors	Datasets from other diseases	Multi-task learning	(Li et al., 2019) (Liao et al., 2019)
	Training pattern	Curriculum learning	(Maicas et al., 2018)(Tang et al., 2018) (Jiménez-Sánchez et al., 2019) (Haarburger et al., 2019)(Zhao et al., 2020b) (Jiménez-Sánchez et al., 2020)(Wei et al., 2020) (Qi et al., 2020)
	Diagnostic patterns	Network design	(Guan et al., 2018)(González-Díaz, 2018) (Wang et al., 2020)(Huang et al., 2020) (Yang et al., 2020)(Liu et al., 2020b)
	Areas doctors focus on	Attention mechanism	(Li et al., 2019)(Mitsuhara et al., 2019) (Fang et al., 2019)(Cui et al., 2020) (Xie et al., 2020)(Zhang et al., 2020a)
	Features doctors docus on	Decision level fusion	(Huynh et al., 2016b) (Moradi et al., 2016) (Majtner et al., 2016)(Xie et al., 2018) (Antropova et al., 2017)(Xia et al., 2020)
		Feature level fusion	(Xie et al., 2016)(Hagerty et al., 2019) (Cao et al., 2018) (Chai et al., 2018) (Buty et al., 2016)(Saba et al., 2020)
		Input level fusion	(Yang et al., 2019) (Xie et al., 2019) (Tan et al., 2019)(Liu et al., 2019b) (Feng et al., 2020)
		As labels of CNNs	(Chen et al., 2016c)(Hussein et al., 2017) (Murthy et al., 2017)
	Other related information	Multi-task learning /network design	(Liu et al., 2018) (Wang et al., 2018) (Zhang et al., 2017)(Wu et al., 2019) (Yu et al., 2020)

tasks. For example, a multi-task deep learning (MTDL) method is proposed in (Liao et al., 2019). MTDL can simultaneously utilize multiple cancer datasets so that hidden representations among these datasets can provide more information to small-scale cancer datasets, and enhance the classification performance. Another example is a cross-disease attention network (CANet) proposed in (Li et al., 2019). CANet characterizes and leverages the relationship between diabetic retinopathy (DR) and diabetic macular edema (DME) in fundus images using a special designed disease-dependent attention module. Experimental results on two public datasets show that CANet outperforms other methods on diagnosing both of the two diseases.

Fig. 5. Two strategies to utilize the pre-trained network on natural images: (a) as a feature extractor and (b) as an initialization which will be fine-tuned on the target dataset.

2.3. Incorporating knowledge from medical doctors

Experienced medical doctors can give fairly accurate conclusion on the given medical images, mainly thanks to the training they have received and the expertise they have accumulated over many years. In general, they often follow some certain patterns or take some procedures when reading medical images. Incorporating these types of knowledge can improve the diagnostic performance of deep learning models.

The types of medical domain knowledge utilized in deep learning models for disease diagnosis can be summarized into the following five categories:

1. the training pattern,
2. the general diagnostic patterns they view images,
3. the areas on which they usually focus,
4. the features (e.g., characteristics, structures, shapes) they give special attention to, and
5. other related information for diagnosis.

The research works for each category will be described in the following sections.

2.3.1. Training pattern of medical doctors

The training process of medical students has a character: they are trained by tasks with increasing difficulty. For example, students begin with some easier tasks, such as deciding whether an image contains lesions, later are required to accomplish more challenging tasks, such as determining whether the lesions are benign or malignant. Over time, they will advance to more difficult tasks, such as determining the subtypes of lesions in images.

This training pattern can be introduced in the training process of deep neural networks via curriculum learning (Bengio et al., 2009). Curriculum determines a sequence of training samples
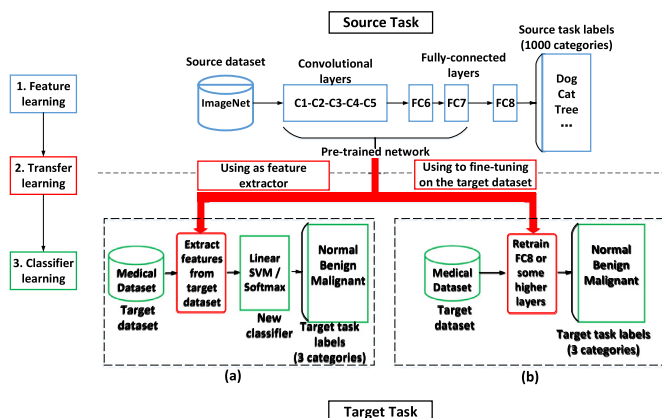
ranked in ascending order of learning difficulty. Curriculum learning has been an active research topic in computer vision and has been recently utilized for medical image diagnosis.

For example, a teacher-student curriculum learning strategy is proposed for breast screening classification from DCE-MRI (Maicas et al., 2018). The deep learning model is trained on simpler tasks before introducing the hard problem of malignancy detection. This strategy shows the better performance when compared with the other methods.

Similarly, (Tang et al., 2018) presents a CNN based attention-guided curriculum learning framework by leveraging the severity-level attributes mined from radiology reports. Images in order of difficulty (grouped by different severity-levels) are fed into the CNN to boost the learning process gradually.

In (Jiménez-Sánchez et al., 2019), the curriculum learning is adopted to support the classification of proximal femur fracture from X-ray images. The approach assigns a degree of difficulty to each training sample. By first learning 'easy' examples and then 'hard' ones, the model can reach a better performance even with fewer data. Other examples of using curriculum learning for disease diagnosis can be found in (Haarburger et al., 2019; Qi et al., 2020; Zhao et al., 2020; Jiménez-Sánchez et al., 2020; Wei et al., 2020).

2.3.2. General diagnostic patterns of medical doctors

Experienced medical doctors generally follow some patterns when they read medical images. These patterns can be integrated into deep learning models with appropriately modified architectures.

For example, radiologists generally follow a three-staged approach when they read chest X-ray images: first browsing the whole image, then concentrating on the local lesion areas, and finally combining the global and local information to make decisions. This pattern is incorporated in the architecture design of the network for thorax disease classification (Guan et al., 2018) (see Fig. 6). The proposed network has three branches, one is used to view the whole image, the second for viewing the local areas, and the third one for combining the global and local information together. The network yields state-of-the-art accuracy on the ChestX-ray14 dataset. In addition, besides the information from the whole image and local lesion area, the information from lung area is also leveraged in (Wang et al., 2020). In particular, a segmentation subnetwork is first used to locate the lung area from the whole image, and then lesion areas are generated by using an attention heatmap. Finally, the most discriminative features are fused for final disease prediction. Another example is a Dual-Ray Net proposed to deal with the front and lateral chest radiography simultaneously (Huang et al., 2020), which also mimics the reading pattern of radiologists.
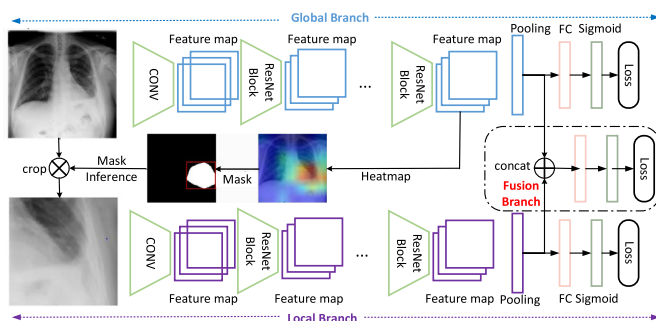

Fig. 6. The example of leveraging the diagnostic pattern of radiologists for thorax disease diagnosis (Guan et al., 2018), where three branches are used to extract and combine the global and local features.
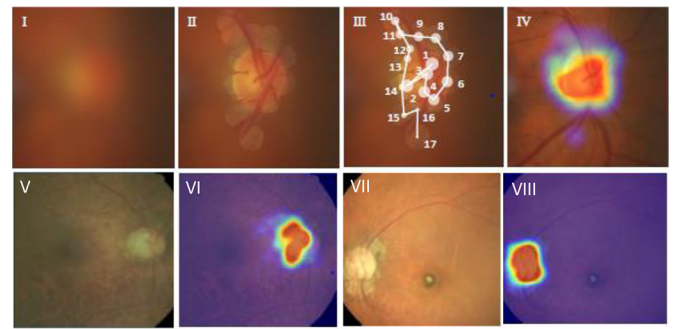

Fig. 7. Example of capturing the attention maps of an ophthalmologist in glaucoma diagnosis (Li et al., 2019). I, II, III and IV are the original blurred fundus image, the fixations of ophthalmologists with cleared regions, the order of clearing the blurred regions, and the generated attention map based on the captured fixations, respectively. V and VII represent the original fundus images. VI and VIII are the corresponding attention maps of V and VII generated by using the method in I-IV.

In the diagnosis of skin lesions, experienced dermatologists generally first locate lesions, then identify dermoscopic features from the lesion areas, and finally make diagnosis based on the features. This pattern is mimicked in the design of the network for the diagnosis of skin lesions (González-Díaz, 2018). The proposed network, DermaKNet, comprised several subnetworks with dedicated tasks: lesion-skin segmentation, detection of dermoscopic features, and global lesion diagnosis. The DermaKNet achieves higher performance compared to the traditional CNN models.

In addition, in mass identification in mammogram, radiologists generally analyze the bilateral and ipsilateral views simultaneously. To emulate this reading practice, (Yang et al., 2020) proposes MommiNet to perform end-to-end bilateral and ipsilateral analysis of mammogram images. In addition, symmetry and geometry constraints are also aggregated from these views. Experiments show the state-of-the-art mass identification accuracy on DDSM. Another example of leveraging this diagnostic pattern of medical doctors can be found in skin lesion diagnosis and thorax disease classification (Liu et al., 2020b).

2.3.3. The areas medical doctors usually focus on

When experienced medical doctors read images, they generally focus on a few specific areas, as these areas are more informative than other places for the purpose of disease diagnosis. Therefore, the information about where medical doctors focus may help deep learning models yield better results.

The knowledge above is generally represented as 'attention maps', which are annotations given by medical doctors indicating the areas they focus on when reading images. For example, a CNN named AG-CNN explicitly incorporates the 'attention maps' for glaucoma diagnosis (Li et al., 2019). The attention maps of ophthalmologists are collected through a simulated eye-tracking experiment, which are used to indicate where they focus when reading images. An example of capturing the attention maps of an ophthalmologist in glaucoma diagnosis is shown in Fig. 7. To incorporate the attention maps, an attention prediction subnet in AG-CNN is designed, and the attention prediction loss measuring the difference between the generated and ground truth attention maps (provided by ophthalmologists) is utilized to supervise the training process. Experimental results show that AG-CNN significantly outperforms the state-of-the-art glaucoma detection methods.

Another example in this category is the lesion-aware CNN (LACNN) for the classification of retinal optical coherence tomography (OCT) images (Fang et al., 2019). The LACNN simulates the pattern of ophthalmologists' diagnosis by focusing on local lesion-related regions. Concretely, the 'attention maps' are firstly repre-

sented as the annotated OCT images delineating the lesion regions using bounding polygons. To incorporate the information, the LACNN proposes a lesion-attention module to enhance the features from local lesion-related regions while still preserving the meaningful structures in global OCT images. Experimental results on two clinically acquired OCT datasets demonstrate the effectiveness of introducing attention maps for retinal OCT image classification, with 8.3% performance gain when compared with the baseline method.

Furthermore, (Mitsuhara et al., 2019) proposes an Attention Branch Network (ABN) to incorporate the knowledge given by the radiologists in diabetic retinopathy. ABN introduces a branch structure which generates attention maps that highlight the attention regions of the network. During the training process, ABN allows the attention maps to be modified with semantic segmentation labels of disease regions. The semantic labels are also annotated by radiologists as the ground truth attention maps. Experimental results on the disease grade recognition of retina images show that ABN achieves 93.73% classification accuracy and its interpretability is clearer than conventional approaches.

Other examples of incorporating attention maps of medical doctors can be found in the diagnosis of radiotherapy-related esophageal fistula (Cui et al., 2020), breast cancer diagnosis (Xie et al., 2020), and short-term lesion change detection in melanoma screening (Zhang et al., 2020a).

2.3.4. Features that medical doctors give special attention to

In the last decades, many guidelines and rules have gradually developed in various medical fields to point out some important features for diagnosis. These features are called *'hand-crafted features'* as they are designated by medical doctors. For example, the popular ABCD rule (Nachbar et al., 1994) is widely used by dermatologists to classify melanocytic tumors. The ABCD rule points out four distinguishing features, namely asymmetry, border, color and differential structures, to determine whether a melanocytic skin lesion under the investigation is benign or malignant.

Another example is in the field of breast cancer diagnosis. Radiologists use the BI-RADS (Breast Imaging Reporting and Data System) score (Mendelson et al., 2013) to place abnormal findings into different categories, with a score of 1 indicating negative findings and a score of 6 indicating breast cancer. More importantly, for each imaging modality, BI-RADS indicates some features closely related to its scores, including margin, shape, micro-calcification, and echo pattern. For example, for breast ultrasound images, tumors with smooth, thin and regular margins are more likely to be benign, while tumors with irregular and thick margins are highly suspected to be malignant. Other features that can help to classify benign and malignant breast tumors are shown in Table 3. Similarly, for the benign-malignant risk assessment of lung nodules in (Hussein et al., 2017), six high-level nodule features, including calcification, sphericity, margin, lobulation, spiculation and texture, have shown a tightly connection with malignancy scores (see Fig. 8).
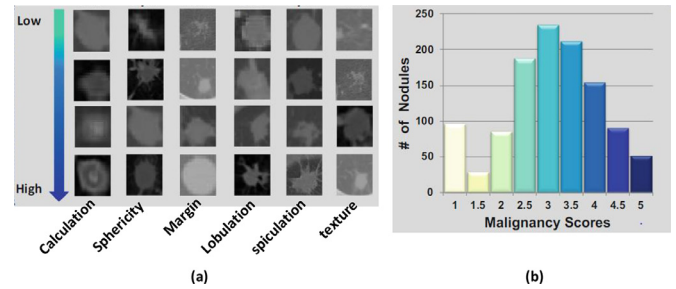

Fig. 8. Lung nodule attributes with different malignancy scores (Hussein et al., 2017).(a) From top to the bottom, six different nodule features attribute from missing to highest prominence. (b) The number of nodules with different malignancy scores.

These different kinds of hand-crafted features have been widely used in many traditional CAD systems. These systems generally first extract these features from medical images, and then feed them into some classifiers like SVM or Random Forest (Breiman, 2001; Cortes and Vapnik, 1995). For example, for the lung nodule classification on CT images, many CAD systems utilize features including the size, shape, morphology, and texture from the suspected lesion areas (Alilou et al., 2017; Xie et al., 2019). Similarly, in the CAD systems for the diagnosis of breast ultrasound images, features such as intensity, texture and shape are used (Hsu et al., 2019).

When using deep learning models like CNNs, which have the ability to automatically extract representative features, there are four approaches to combining 'hand-crafted features' with features extracted from CNNs.

- Decision-level fusion: the two types of features are fed into two classifiers separately, then the decisions from the two classifiers are combined.
- Feature-level fusion: the two types of features are directly combined via techniques such as concatenation.
- Input-level fusion: the hand-crafted features are represented as image patches, which are then taken as inputs to the CNNs.
- Usage of features as labels of CNNs: the hand-crafted features are first annotated and then utilized as labels for CNNs during the training process.

Decision-level fusion: The structure of this approach is illustrated in Fig. 9. In this approach, the hand-crafted features and the features extracted from CNNs are separately fed into two classifiers. Then, the classification results from both classifiers are combined using some decision fusion techniques to produce final classification results.

For example, a skin lesion classification model proposed in (Majtner et al., 2016) combines the results from two SVM classifiers. The first one uses hand-crafted features (i.e., RSurf features and local binary patterns (LBP)) as input and the second employs features derived from a CNN. Both of the classifiers predict the category for each tested image with a classification score. These scores are subsequently used to determine the final classification result.

Table 3
Features in the BI-RADS guideline to classify benign and malignant breast tumors in ultrasound images (Mendelson et al., 2013).

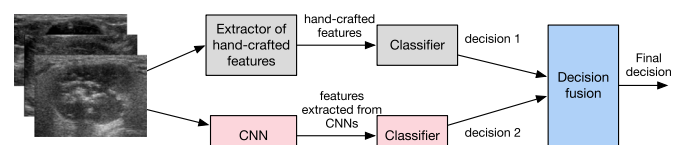|	Benign	Malignant
Margin	smooth, thin, regular	irregular, thick
Shape	round or oval	irregular
Microcalcification	no	yes
Echo Pattern	clear	unclear
Acoustic Attenuation	not obvious	obvious
Side Acoustic Shadow	obvious	not obvious

Fig. 9. Decision-level fusion: the decisions from two classifiers, one based on hand-crafted features, and the other on the CNNs, are combined.

Similarly, a mammographic tumor classification method also combines features in decision-level fusion (Huynh et al., 2016b). After individually performing classification with CNN features and analytically extracted features (e.g., contrast, texture, and margin spiculation), the method adopts the soft voting to combine the outputs from both individual classifiers. Experimental results show that the performance of the ensemble classifier was significantly better than the individual ones. Other examples that utilize this approach include lung nodule diagnosis (Xie et al., 2018), breast cancer diagnosis (Antropova et al., 2017), the classification of cardiac slices (Moradi et al., 2016) and the prediction of the invasiveness risk of stage-I lung adenocarcinomas (Xia et al., 2020).

Feature-level fusion: In this approach, hand-crafted features and features extracted from CNNs are concatenated, and the combined features are fed into a classifier for diagnosis. The structure of this approach is illustrated in Fig. 10.

For example, a combined-feature based classification approach called CFBC is proposed for lung nodule classification by (Xie et al., 2016). In CFBC, the hand-crafted features (including texture and shape descriptors) and the features learned by a nine-layer CNN are combined and fed into a back-propagation neural network. Experimental results derived from a publicly available dataset show that compared with a purely CNN model, incorporating hand-crafted features improves the accuracy, sensitivity, and specificity by 3.87%, 6.41%, and 3.21%, respectively.

Another example in this category is the breast cancer classification in histology images (Cao et al., 2018). More specifically, two hand-crafted features, namely the parameter-free threshold adjacency statistics and the gray-level co-occurrence matrix, are fused with the five groups of deep learning features extracted from five different deep models. The results show that after incorporating hand-crafted features, the accuracy of the deep learning model can be significantly improved.

Other examples of employing feature-level fusion can be found in glaucoma diagnosis (Chai et al., 2018), skin lesion classification (Hagerty et al., 2019), lung nodule classification (Buty et al., 2016) and brain tumor diagnosis (Saba et al., 2020).

Input-level fusion: In this approach, hand-crafted features are first represented as patches highlighting the corresponding features. Then, these patches are taken as input to CNNs to make the final conclusion. Using this approach, the CNNs are expected to rely more on the hand-crafted features. It should be noted that generally speaking, one CNN is required for each type of hand-crafted feature, and information obtained from these CNNs need to be combined in some manner to make a final decision. The structure of this approach is illustrated in Fig. 11.

For example, in (Tan et al., 2019), three types of hand-crafted features, namely the contrast information of the initial nodule candidates (INCs) and the outer environments, the histogram of oriented gradients (HOG) feature, and the LBP feature, are transformed into the corresponding patches and are taken as inputs of multiple CNNs. The results show that this approach outperforms
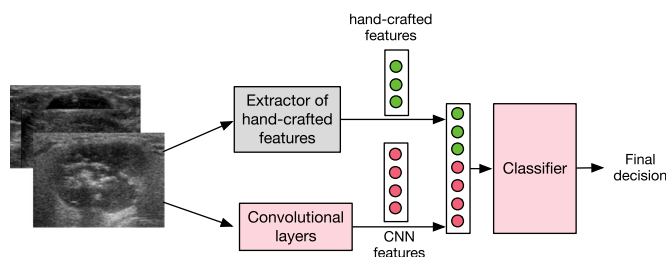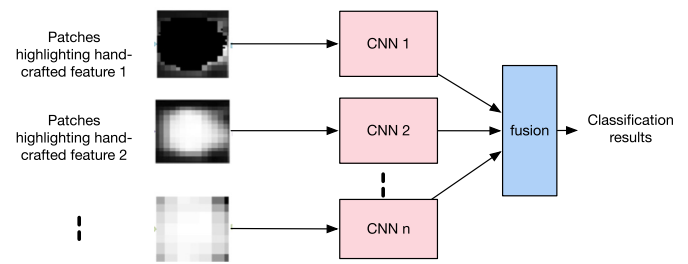

Fig. 11. Input-level fusion: the hand-crafted features are represented as image patches that are taken as inputs to the CNNs.

both conventional CNN-based approaches and traditional machine-learning approaches based on hand-crafted features.

Another example using input-level fusion approach is the MV-KBC (multi-view knowledge-based collaborative) algorithm proposed for lung nodule classification (Xie et al., 2019). The MV-KBC first extracts patches corresponding to three features: the overall appearance (OA), nodule's heterogeneity in voxel values (HVV) and heterogeneity in shapes (HS). Each type of patch is fed into a ResNet. Then, the outputs of these ResNets are combined to generate the final classification results.

Moreover, (Yang et al., 2019) proposes the dual-path semi-supervised conditional generative adversarial networks (DScGAN) for thyroid classification. More specifically, the features from the ultrasound B-mode images and the ultrasound elastography images are first extracted as the OB patches (indicating the region of interest (ROI) in B-mode images), OS patches (characterizing the shape of the nodules), and OE patches (indicating the elasticity ROI according to the B-mode ROI position). Then, these patches are used as the input of the DScGAN. Using the information from these patches is demonstrated to be effective to improve the classification performance. Other examples employ input-level fusion can be found in thyroid nodule diagnosis (Liu et al., 2019b) and breast cancer diagnosis on multi-sequence MRI (Feng et al., 2020).

Usage of features as labels of CNNs: In this approach, besides the original classification labels of images, medical doctors also provide labels for some hand-crafted features. This extra information is generally incorporated into deep learning models via the multi-task learning structure.

For example, in (Chen et al., 2016c), the nodules in lung images were quantitatively rated by radiologists with regard to nine hand-crafted features (e.g., spiculation, texture, and margin). The multi-task learning framework is used to incorporate the above information into the main task of lung nodule classification.

In addition, for the benign-malignant risk assessment of lung nodules in low-dose CT scans (Hussein et al., 2017), the binary labels about the presence of six high-level nodule attributes, namely the calcification, sphericity, margin, lobulation, spiculation and texture, are utilized. Six CNNs are designed and each one aims at learning one attribute. The discriminative features automatically learned by CNNs for these attributes are fused in a multi-task learning framework to obtain the final risk assessment scores.

Similarly, in (Murthy et al., 2017), each glioma nuclear image is exclusively labeled in terms of the shapes and attributes for the centermost nuclei of the image. These features are then learned by a multi-task CNN. Experiments demonstrate that the proposed method outperforms the baseline CNN.

2.3.5. Other types of information related to diagnosis

In this section, we summarize other types of information that can help deep learning models to improve their diagnostic performance. These types of information include extra category labels and clinical diagnostic reports.

Fig. 10. Feature-level fusion: the hand-crafted features are combined with the features extracted from CNNs as the new feature vectors.

Table 4

The comparison of the quantitative metrics for some disease diagnosis methods after integrating domain knowledge.

Reference	Baseline Model/With Domain Knowledge	Accuracy	AUC
(Guan et al., 2018)	AG-CNN only with global branch/AG-CNN	–/–	0.840/0.871
(González-Díaz, 2018)	ResNet-50/DermaKNet	–/–	0.874/0.908
(Hadad et al., 2017)	Fine-tuned VGG-Net/Fine-tuned MG-Net	0.900/0.930	0.950/0.970
(Hagerty et al., 2019)	ResNet-50/ResNet-50 with handcrafted features	–/–	0.830/0.940
(Li et al., 2019)	CNN without using attention map/AG-CNN	0.908/0.953	0.966/0.975
(Li et al., 2019)	ResNet50/CANet	0.828/0.851	–/–
(Liu et al., 2018)	VGG16/Multi-task VGG16	0.829 /0.833	–/–
(Maicas et al., 2018)	DenseNet/BMSL	–/–	0.850/0.900
(Samala et al., 2018)	CNN with single-stage/multi-stage transfer learning	–/–	0.850/0.910
(Tang et al., 2018)	CNN/AGCL	–/–	0.771/0.803
(Yang et al., 2019)	DScGAN without/with domain knowledge	0.816/0.905	0.812/0.914

Extra category labels

For medical images, besides a classification label (i.e., normal, malignant or benign), radiologists may give some extra category labels. For example, in the ultrasonic diagnosis of breast cancer, an image usually has a BI-RADS label that classifies the image into 0~6 (Mendelson et al., 2013)—category 0 suggests re-examination, categories 1 and 2 indicate that it is prone to be a benign lesion, category 3 suggests probably benign findings, categories 4 and 5 are suspected to be malignant, category 6 definitely suggests malignant). These labels also contain information about the condition of lesions. In (Liu et al., 2018), the BI-RADS label for each medical image is incorporated in a multi-task learning structure as the label of an auxiliary task. Experimental results show that incorporating these BI-RADS labels can improve the diagnostic performance of major classification task. Another example of using the information from BI-RADS labels can be found in (Wu et al., 2019).

In addition, during the process of image annotation, consensus or disagreement among experts regarding images reflects the gradeability and difficulty levels of the image, which is also a representation of medical domain knowledge. To incorporate this information,(Yu et al., 2020) proposes a multi-branch model structure to predict the most sensitive, most specifical and a balanced fused result for glaucoma images. Meanwhile, the consensus loss is also used to encourage the sensitivity and specificity branch to generate consistent and contradictory predictions for images with consensus and disagreement labels, respectively.

Extra clinical diagnostic reports

A clinical report is a summary of all the clinical findings and impressions determined during examination of a radiography study. It usually contains rich information and reflects the findings and descriptions of medical doctors. Incorporating clinical reports into CNNs designed for disease diagnosis is typically beneficial. As medical reports are generally handled by recurrent neural networks (RNNs), to incorporate information from medical reports, generally hybrid networks containing both CNNs and RNNs are needed.

For example, the Text-Image embedding network (TieNet) is designed to classify the common thorax disease in chest X-rays (Wang et al., 2018). TieNet has an end-to-end CNN-RNN architecture enabling it to integrate information of radiological reports. The classification results are significantly improved (with about a 6% increase on average in AUCs) compared with the baseline CNN purely based on medical images.

In addition, using semantic information from diagnostic reports is explored in (Zhang et al., 2017) for understanding pathological bladder cancer images. A dual-attention model is designed to facilitate the high-level interaction of semantic information and visual information. Experiments demonstrate that incorporating information from diagnostic reports significantly improves the cancer diagnostic performance over the baseline method.

2.4. Summary

In the previous sections, we introduced different kinds of domain knowledge and the corresponding integrating methods into the deep learning models for disease diagnosis. Generally, almost all types of domain knowledge are proven to be effective in boosting the diagnostic performance, especially using the metrics of accuracy and AUC, some examples and their quantitative improvements are listed in Table 4.

With respect to type of domain knowledge for disease diagnosis, the knowledge from natural images is widely incorporated in deep learning models. However, considering the gap between natural images and medical ones, using information from external medical datasets of the same diseases with similar modalities (e.g., SFM and DM) (Samala et al., 2017), with different modalities (DBT and MM, Ultrasound) (Samala et al., 2018), and even with different diseases (Liao et al., 2019) may be more effective. To incorporate the above information is relatively easy, and both transfer learning and multi-task learning have been widely adopted. However, there are still no comparative studies on how different extra datasets can improve the performance of deep learning models.

For the domain knowledge of medical doctors, the high-level domain knowledge (e.g., the training pattern and the diagnostic pattern) is generally common for different diseases. On the other hand, the low-level domain knowledge, such as the areas in images and features of diseases that medical doctors usually pay attention to, is generally different for different diseases. There is generally a trade-off between the versatility and the utility of domain knowledge. To diagnose a certain disease, the benefit of incorporating a versatile type of domain knowledge suitable for many diseases may not be as significant as using the one that is specific for the disease. However, identifying such specific domain knowledge may not be easy, and generally requires more efforts from medical doctors (e.g., to identify hand-crafted features or attention maps).

We believe that more kinds of medical domain knowledge can be explored and utilized for disease diagnosis. In addition, comparative studies on some benchmark datasets should be carried out with respect to different types of domain knowledge and different incorporating methods for disease diagnosis. The results can provide further insights about the utility of medical domain knowledge for deep learning models.

3. Lesion, organ, and abnormality detection

Detecting lesions in medical images is an important task and a key part for disease diagnosis in many conditions. Similarly, organ detection is an essential preprocessing step for image registration, organ segmentation, and lesion detection. Detecting abnormalities

Table 5
List of studies on lesion, organ, and abnormality detection and the knowledge they incorporated.

Knowledge Source	Knowledge Type	Incorporating Method	References
Natural datasets	Natural images	Transfer learning	(Shin et al., 2016)(Yap et al., 2018) (Näppi et al., 2016)(Zhang et al., 2016) (Tajbakhsh et al., 2016)
Medical datasets	Multi-modal images	Transfer learning	(Zhang et al., 2018a) (Ben-Cohen et al., 2019) (Zhao et al., 2020a)
Medical doctors	Training pattern	Curriculum learning	(Tang et al., 2018)(Jesson et al., 2017) (Astudillo et al., 2020)
		Using images collected under different settings	(Li et al., 2019)(Ni et al., 2020)
	Detection patterns	Comparing bilateral or cross-view images	(Liu et al., 2019)(Liu et al., 2020) (Lisowska et al., 2017a,b)(Li et al., 2020d)
		Analyzing adjacent slices	(Li et al., 2017)
	Features doctors focus on	Feature level fusion	(Fu et al., 2017)(Kooi et al., 2017) (Ghatwary et al., 2019) (Liu et al., 2019b) (Chao et al., 2020) (Sonora-Mengan Sr et al., 2020)
	Other related information	Multi-task learning /training pattern design	(Tang et al., 2018) (Hwang and Kim, 2016) (Bakalo et al., 2019)(Liang et al., 2020)
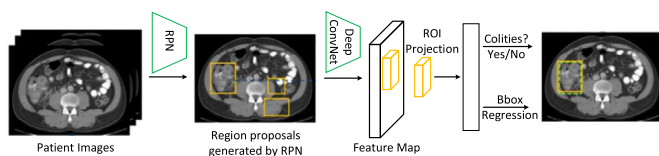
Fig. 12. The workflow of colitis detection method by using the Faster R-CNN structure (Liu et al., 2017).

in medical images, such as cerebral microbleeds in brain MRI images and hard exudates in retinal images, is also required in many applications.

In this section, the deep learning models widely used for object detection in medical images are first described in Subsection 3.1. Then, the existing works on utilizing domain knowledge from natural and medical datasets, and from medical doctors are introduced in detail in Subsection 3.2 and Subsection 3.3, respectively. Lastly, we summarize and discuss these different types of domain knowledge and the associated incorporating methods in Subsection 3.4.

3.1. General structures of deep learning models for object detection in medical images

Deep learning models designed for object detection in natural images have been directly applied to detect objects in medical images. These applications include pulmonary nodule detection in CT images (Setio et al., 2016), retinal diseases detection in retinal fundus (Gulshan et al., 2016) and so on.

One popular type of model is the two-stage detectors like the Faster R-CNN (Ren et al., 2015) and the Mask R-CNN (He et al., 2017). They generally consist of a region proposal network (RPN) that hypothesizes candidate object locations and a detection network that refines region proposals. Applications in this category include colitis detection in CT images (Liu et al., 2017), intervertebral disc detection in X-ray images (Sa et al., 2017) and the detection of architectural distortions in mammograms (Ben-Ari et al., 2017). Fig. 12 shows an example of using Faster R-CNN structure for colitis detection (Liu et al., 2017).

Another category is one-stage models like YOLO (You Only Look Once) (Redmon et al., 2016), SSD (Single Shot MultiBox Detector) (Liu et al., 2016) and RetinaNet (Lin et al., 2017). These networks skip the region proposal stage and run detection directly by

considering the probability that the object appears at each point in image. Compared with the two-stage models, models in this approach are generally faster and simpler. Examples in this category can be found in breast tumor detection in mammograms (Platania et al., 2017), pulmonary lung nodule detection in CT (Li et al., 2017), and the detection of different lesions (e.g., liver lesion, lung lesion, bone lesion, abdomen lesion) in CT images (Cai et al., 2019). An example of using one-stage structure for lesion detection is shown in Fig. 13.

In the following subsections, we will introduce related works that incorporate external knowledge into deep learning models for object detection in medical images. A summary of these works is shown in Table 5.

3.2. Incorporating knowledge from natural datasets or other medical datasets

Similar to disease diagnosis, it is quite popular to pre-train a large natural image dataset (generally ImageNet) to introduce information for object detection in medical domain. Examples can be found in lymph node detection (Shin et al., 2016), polyp and pulmonary embolism detection (Tajbakhsh et al., 2016), breast tumor detection (Yap et al., 2018), colorectal polyps detection (Näppi et al., 2016; Zhang et al., 2016) and so on.

In addition, utilizing other medical datasets (i.e., multi-modal images) is also quite common. For example, PET images are used to help the lesion detection in CT scans of livers (Ben-Cohen et al., 2019). Specifically, PET images are first generated from CT scans using a combined structure of FCN and GAN, then the synthesized
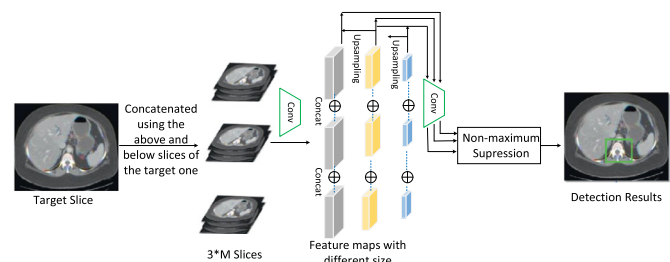
Fig. 13. An example of using one-stage structure for lesion detection in CT images (Cai et al., 2019).

PET images are used in a false-positive reduction layer for detecting malignant lesions. Quantitative results show a 28% reduction in the average false positive per case. Besides, (Zhang et al., 2018a) develops a strategy to detect breast masses from digital tomosynthesis by fine-tuning the model pre-trained on mammography datasets. Another example of using multi-modal medical images can be found in liver tumor detection (Zhao et al., 2020a).

3.3. Incorporating knowledge from medical doctors

In this subsection, we summarize the existing works on incorporating the knowledge of medical doctors to more effectively detect objects in medical images. The types of domain knowledge from medical doctors can be grouped into the following four categories:

1. the training pattern,
2. the general detection patterns they view images,
3. the features (e.g., locations, structures, shapes) they give special attention to, and
4. other related information regarding detection.

3.3.1. Training pattern of medical doctors

The training pattern of medical doctors, which is generally characterized as being given tasks with increasing difficulty, is also used for object detection in medical images. Similar to the disease diagnosis, most works utilize the curriculum learning to introduce this pattern. For example, an attention-guided curriculum learning (AGCL) framework is presented to locate the lesion in chest X-rays (Tang et al., 2018). During this process, images in order of difficulty (grouped by different severity-levels) are fed into the CNN gradually, and the disease heatmaps generated from the CNN are used to locate the lesion areas.

Another work is called as CASED proposed for lung nodule detection in chest CT (Jesson et al., 2017). CASED adopts a curriculum adaptive sampling technique to address the problem of extreme data imbalance. In particular, CASED lets the network to first learn how to distinguish nodules from their immediate surroundings, and then it adds a greater proportion of difficult-to-classify global context, until uniformly samples from the empirical data distribution. In this way, CASED tops the LUNA16 challenge leaderboard with an average sensitivity score of 88.35%. Another example of using curriculum learning can be found in cardiac landmark detection (Astudillo et al., 2020).

3.3.2. General detection patterns of medical doctors

When experienced medical doctors are locating possible lesions in medical images, they also display particular patterns, and these patterns can be incorporated into deep learning models for object detection. Experienced medical doctors generally have the following patterns: (1) they usually take into account images collected under different settings (e.g., brightness and contrast), (2) they often compare bilateral images or use cross-view images, and (3) they generally read adjacent slices.

For example, to locate possible lesions during visual inspection of the CT images, radiologists would combine images collected under different settings (e.g., brightness and contrast). To imitate the above process, a multi-view feature pyramid network (FPN) is proposed in (Li et al., 2019), where multi-view features are extracted from images rendered with varied brightness and contrast. The multi-view information is then combined using a position-aware attention module. Experiments show that the proposed model achieves an absolute gain of 5.65% over the previous state-of-the-art method on the NIH DeepLesion dataset. Another example of using images under different settings can be found in the COVID-19 pneumonia-based lung lesion detection (Ni et al., 2020).
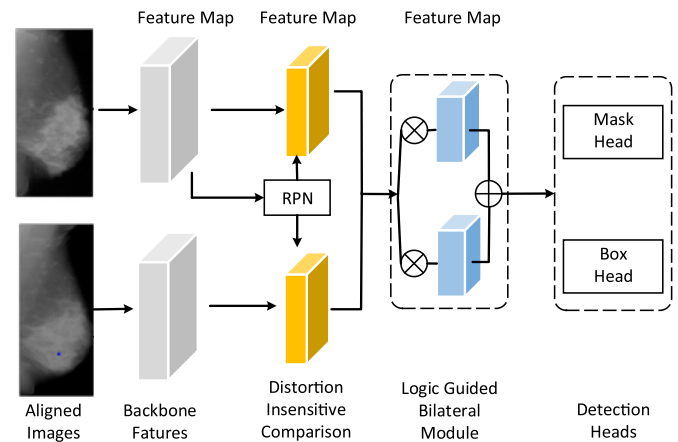

Fig. 14. The workflow of mammogram mass detection by integrating the bilateral information (Liu et al., 2019), where the aligned images are fed into two networks seperately to extract features for further detection.

In addition, the bilateral information is widely used by radiologists. For example, in standard mammographic screening, images are captured from both two breasts, and experienced radiologists generally compare bilateral mammogram images to find masses. To incorporate this pattern, a contrasted bilateral network (CBN) is proposed in (Liu et al., 2019), in which the bilateral images are first coarsely aligned and then fed into a pair of networks to extract features for the following detection steps (shown in Fig. 14). Experimental results demonstrate the effectiveness of incorporating the bilateral information.

Similarly, to detect acute stroke signs in non-contrast CT images, experienced neuroradiologists routinely compare the appearance and Hounsfield Unit intensities of the left and right hemispheres, and then find the regions most commonly affected in stroke episodes. This pattern is mimicked by (Lisowska et al., 2017) for the detection of dense vessels and ischaemia. The experimental results show that introducing the pattern greatly improves the performance for detecting ischaemia. Other examples of integrating the bilateral feature comparison or the symmetry constrains can be found in thrombus detection (Lisowska et al., 2017a) and hemorrhagic lesion detection (Li et al., 2020d) in brain CT images.

Besides the bilateral images, the information from cross views (i.e., mediolateral-oblique and cranio-caudal) is highly related and complementary, and hence is also used for mammogram mass detection. In (Liu et al., 2020), a bipartite graph convolutional network is introduced to endow the existing methods with cross-view reasoning ability of radiologists. Concretely, the bipartite node sets are constructed to represent the relatively consistent regions, and the bipartite edge are used to model both inherent cross-view geometric constraints and appearance similarities between correspondences. This process can enables spatial visual features equipped with cross-view reasoning ability. Experimental results on DDSM dataset achieve the state-of-the-art performance (with a recall of 79.5 at 0.5 false positives per image).

When looking for small nodules in CT images, radiologists often observe each slice together with adjacent slices, similar to detecting an object in a video. This workflow is imitated in (Li et al., 2017) to detect pulmonary lung nodule in CT images, where the state-of-the-art object detector SSD is applied in this process. This method obtains state-of-the-art result with the FROC score of 0.892 in LUNA16 dataset.

3.3.3. Features that medical doctors give special attention to

Similar to disease diagnosis, medical doctors also use many 'hand-crafted features' to help them to find target objects (e.g., nodules or lesions) in medical images.

Table 6

The comparison of the quantitative metrics for some medical object detection methods after incorporating domain knowledge.

Reference	Baseline model/With domain knowledge	Sensitivity	Recall
(Fu et al., 2017)	CNN/CNN with hand-crafted features	0.890/0.909	–/–
(Ghatwary et al., 2019)	ResNet/CNN with handcrafted features	–/–	0.940/0.950
(Li et al., 2017)	SSD/MSSD	0.927/0.976	–/–
(Liu et al., 2019)	Mask R-CNN/CBD	–/–	0.869/0.890
(Liu et al., 2020)	Mask R-CNN/BG-RCNN	–/–	0.918/0.945
(Tang et al., 2018)	CNN/AGCL	–/–	0.660/0.730

The performance is evaluated at 4 false positives per image in (Fu et al., 2017; Liu et al., 2019, 2020; Samala et al., 2018).

For example, in (Kooi et al., 2017), to detect mammographic lesions, different types of hand-crafted features (e.g., contrast features, geometrical features, and location features) are first extracted, and then concatenated with those learned by a CNN. The results show that these hand-crafted features, particularly the location and context features, can complement the network generating a higher specificity over the CNN alone.

Similarly, (Ghatwary et al., 2019) presents a deep learning model based on Faster R-CNN to detect abnormalities in the esophagus from endoscopic images. In particular, to enhance texture details, the proposed detection system incorporates the Gabor handcrafted features with the CNN features through concatenation in the detection stage. The experimental results on two datasets (Kvasir and MICCAI 2015) show that the model is able to surpass the state-of-the-art performance.

Another example can be found in (Fu et al., 2017) for the detection of lung nodules, where 88 hand-crafted features, including intensity, shape, texture are extracted and combined with features extracted by a CNN and then fed into a classifier. Experimental results demonstrate the effectiveness of the combination of hand-crafted features and CNN features.

In the automated detection of thyroid nodules, the size and shape attribute of nodules are considered in (Liu et al., 2019b). To incorporate the information above, the generating process of region proposals is constrained, and the detection results on two different datasets show high accuracy.

Furthermore, in lymph node gross tumor volume detection (GTV_{LN}) in oncology imaging, some attributes of lymph nodes (LNs) are also utilized (Chao et al., 2020). Motivated by the prior clinical knowledge that LNs from a connected lymphatic system, and the spread of cancer cells among LNs often follows certain pathways, a LN appearance and inter-LN relationship learning framework is proposed for GTV_{LN} detection. More specifically, the instance-wise appearance features are first extracted by a 3D CNN, then a graph neural network (GNN) is used to model the inter-LN relationships, and the global LN-tumor spatial priors are included in this process. This method significantly improves over state-of-the-art method. Another example of combining handcrafted features and deep features can be found in lung lesion detection (Sonora-Mengan Sr et al., 2020).

3.3.4. Other types of information related to detection

Similar with that in disease diagnosis, there are other information (e.g., radiological reports, extra labels) can also be integrated into the lesion detection process.

For example in (Tang et al., 2018), to locate thoracic diseases on chest radiographs, the difficulty of each sample, represented as the severity level of the thoracic disease, is first extracted from radiology reports. Then, the curriculum learning technique is adopted, in which the training samples are presented to the network in order of increasing difficulty. Experiments on the ChestXray14 database validate the effectiveness on significant performance improvement over baseline methods.

Example of using extra labels can be found in (Hwang and Kim, 2016). In this method, the information of the classification labels is incorporated to help the lesion localization in chest X-rays and mammograms. In particular, a framework named as self-transfer learning (STL) is proposed, which jointly optimizes both classification and localization networks to help the localization network focus on correct lesions. Experimental results show that STL can achieve significantly better localization performance compared to previous weakly supervised localization approaches. More examples of using extra labels can be found in detection in mammograms (Bakalo et al., 2019; Liang et al., 2020).

3.4. Summary

In the previous sections, we introduced different kinds of domain knowledge and the corresponding integrating methods into the deep learning models for object detection in medical images. Table 6 illustrates the quantitative improvements, in terms of sensitivity and recall, of some typical work over the baseline methods for object detection in medical images. From the results we can see that in general, integrating domain knowledge can be beneficial for detection tasks.

Similar to disease diagnosis, the high-level training pattern of medical doctors is generic and can be utilized for detecting different diseases or organs. In contrast, the low-level domain knowledge, like the detection patterns that medical doctors adopt and some hand-crafted features they give more attention when searching lesions, are generally different for different diseases. For example, the pattern of comparing bilateral images can only be utilized for detecting organs with symmetrical structures (Lisowska et al., 2017; Liu et al., 2019). In addition, we can see from Table 6 that leveraging pattern of medical doctors on average shows better performance when compared with integrating hand-crafted features. This may indicate that there is still large room to explore more effective features for object detection in medical images.

4. Lesion and organ segmentation

Medical image segmentation devotes to identifying pixels of lesions or organs from the background, and is generally regarded as a prerequisite step for the lesion assessment and disease diagnosis. Segmentation methods based on deep learning models have become the dominant technique in recent years and have been widely used for the segmentation of lesions such as brain tumors (Havaei et al., 2017), breast tumors (Zhang et al., 2018), and organs such as livers (Christ et al., 2017) and pancreas (Roth et al., 2015).

In Subsection 4.1, we describe the models that are generally used for object segmentation in the medical domain. Then in Subsection 4.2, the works of utilizing domain knowledge from natural and other medical datasets are introduced. Then, the models utilizing domain knowledge from medical doctors are introduced in Subsection 4.3. A summary of this section is provided in Subsection 4.4.
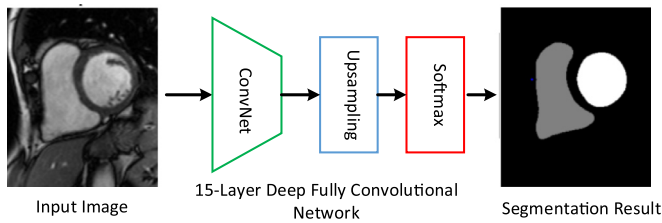
Fig. 15. The schematic diagram of using FCN structure for cardiac segmentation (Tran, 2016).

Fig. 16. The network structure of U-Net (Ronneberger et al., 2015).

4.1. General structures of deep learning models for object segmentation in medical images

The deep learning models utilized for medical image segmentation are generally divided into three categories: the FCN (fully convolutional network) (Long et al., 2015) based models, the U-Net (Ronneberger et al., 2015) based models, and the GAN (Goodfellow et al., 2014) based models.

In particular, the FCN has been proven to perform well in various medical image segmentation tasks (Chen et al., 2016b; Gibson et al., 2017). Some variants of FCN, such as cascaded FCN (Christ et al., 2016), parallel FCN (Kamnitsas et al., 2017b) and recurrent FCN (Yang et al., 2017c) are also widely used for segmentation tasks in medical images. Fig. 15 illustrates an example of using FCN based model for cardiac segmentation.

In addition, the U-Net (Ronneberger et al., 2015) (shown in Fig. 16) and its variants are also widely utilized for medical image segmentation. U-Net builds upon FCN structure, mainly consists of a series of convolutional and deconvolutional layers, and with the short connections between the layers of equal resolution.
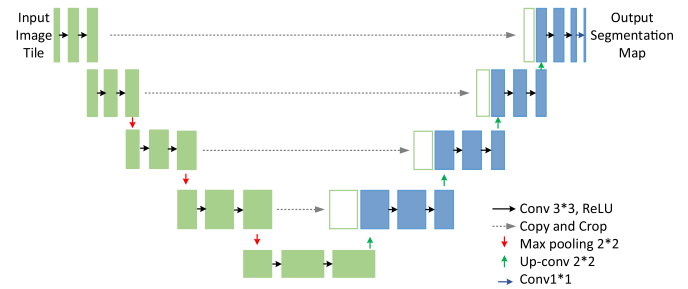
U-Net and its variants like UNet++(Zhou et al., 2018) and recurrent U-Net (Alom et al., 2018) perform well in many medical image segmentation tasks (Gordienko et al., 2018).

In the GAN-based models (Yang et al., 2017; Zhao et al., 2018), the generator is used to predict the mask of the target based on some encoder-decoder structures (such as FCN or U-Net). The discriminator serves as a shape regulator that helps the generator to obtain satisfactory segmentation results. Applications of using GAN-based models in medical image segmentation include brain segmentation (Kamnitsas et al., 2017a), skin lesion segmentation (Izadi et al., 2018), vessel segmentation (Lahiri et al., 2017) and anomaly segmentation in retinal fundus images (Schlegl et al., 2017). Fig. 17 is an example of using GAN-based model for vessel segmentation in miscroscopy images.

In the following sections, we will introduce research studies that incorporate domain knowledge into deep segmentation models. The summary of these works is shown in Table 7.

Table 7
The list of researches of lesion, organ segmentation and the knowledge they incorporated.

| Knowledge Source	Knowledge Type	Incorporating Method	References
Natural datasets	Natural images	Transfer learning	(Chen et al., 2016b)(Chen et al., 2017) (Wu et al., 2017)(Zeng et al., 2017) (Tajbakhsh et al., 2016)
Medical datasets	Multi-modal images	Transfer learning Multi-task/-modal learning Modality transformation /synthesis	(Ghafoorian et al., 2017)(Valverde et al., 2019) (Moeskops et al., 2016)(Valindria et al., 2018) (Chen et al., 2018)(Chen et al., 2020) (Jiang et al., 2018) (Kamnitsas et al., 2017a) (Yan et al., 2019)(Yang et al., 2019) (Li et al., 2020a,b)(Hu et al., 2020) (Zhang et al., 2018)(Chartsias et al., 2020)
	Datasaets of other diseases	Transfer learning Disease domain transformation	(Chen et al., 2019b) (Yu et al., 2019)
Medical doctors	Training pattern	Curriculum learning	(Berger et al., 2018) (Wang et al., 2018) (Li et al., 2020a)(Kervadec et al., 2019) (Zhao et al., 2019)(Zhang et al., 2020b)
	Diagnostic patterns	Using different views as input Attention mechanism Network design	(Wu et al., 2018)(Chen et al., 2019) (Hatamizadeh et al., 2019) (Zhu et al., 2020)(Jin et al., 2020)
	Anatomical priors	Incorporated in post-processing	(Huang et al., 2018) (Painchaud et al., 2019, 2020) (Larrazabal et al., 2019)
		Incorporated in loss function	(Oktay and Ferrante, 2017)(Ravishankar et al., 2017) (BenTaieb and Hamarneh, 2016) (Mirikharaji and Hamarneh, 2018) (Yue et al., 2019)(Zheng et al., 2019) (Oktay and Ferrante, 2017) (Zotti et al., 2019)
		Generative models	(Chen et al., 2019) (Dalca et al., 2018) (He et al., 2019)(Luo et al., 2020) (Song et al., 2020)(Boutillon et al., 2020) (Pham et al., 2020)(Engin et al., 2020) (Gao et al., 2020)
	Hand-crafted features	Feature level fusion Input level fusion	(Kushibar et al., 2018)(Rezaei et al., 2019) (Khan et al., 2020)(Narotamo et al., 2020)
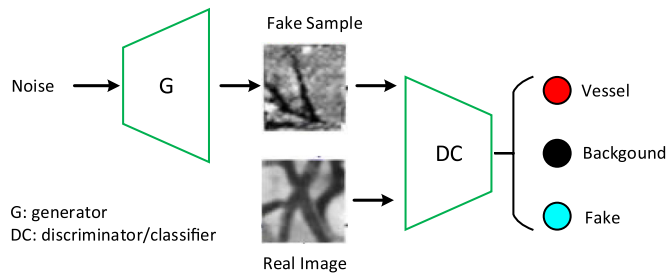
Fig. 17. An example of using a GAN-based model for vessel segmentation (Lahiri et al., 2017).

4.2. Incorporating knowledge from natural datasets or other medical datasets

It is also quite common that deep learning segmentation models are firstly trained on a large-scale natural image dataset (e.g., ImageNet) and then fine-tuned on the target datasets. Using the above transfer learning strategy to introduce knowledge from natural images has demonstrated to achieve a better performance in medical image segmentation. Examples can be found in intima-media boundary segmentation (Tajbakhsh et al., 2016) and prenatal ultrasound image segmentation(Wu et al., 2017). Besides ImageNet, (Chen et al., 2016b) adopts the off-the-shelf DeepLab model trained on the PASCAL VOC dataset for anatomical structure segmentation in ultrasound images. This pre-trained model is also used in the deep contour-aware network (DCAN), which is designed for the gland segmentation in histopathological images (Chen et al., 2017).

Besides using models pre-trained on 'static' datasets like ImageNet and PASCAL VOC, many deep neural networks, especially those designed for the segmentation of 3D medial images, leverage models pre-trained on large-scale video datasets. For example, in the automatic segmentation of proximal femur in 3D MRI, the C3D pre-trained model is adopted as the encoder of the proposed 3D U-Net (Zeng et al., 2017). Notably, the C3D model is trained on the Sports-1M dataset, which is the largest video classification benchmark with 1.1 million sports videos in 487 categories (Tran et al., 2015).

In addition to natural images, using knowledge from external medical datasets with different modalities and with different diseases is also quite popular.

For example, (Ghafoorian et al., 2017) investigates the transferability of the acquired knowledge of a CNN model initially trained for WM hyper-intensity segmentation on legacy low-resolution data to new data from the same scanner but with higher image resolution. Likewise, the images with different MRI scanners and protocols are used in (Valverde et al., 2019) to help the multi sclerosis segmentation process via transfer learning.

In (Moeskops et al., 2016), the multi-task learning is adopted, where the data of brain MRI, breast MRI and cardiac CT angiography (CTA) are used simultaneously as multiple tasks. On the other hand, (Valindria et al., 2018) adopts a multi-modal learning structure for organ segmentation. A dual-stream encoder-decoder architecture is proposed to learn modality-independent, and thus, generalisable and robust features shared among medical datasets with different modalities (MRI and CT images). Experimental results prove the effectiveness of this multi-modal learning structure.

Moreover, many works adopt GAN-based models to achieve the domain transformation among datasets with different modalities. For example, a model named SeUDA (unsupervised domain adaptation) is proposed for the left/right lung segmentation process (Chen et al., 2018). It leverages the semantic-aware GAN to transfer the knowledge from one chest dataset to another. In particular, target images are first mapped towards the source data space via

the constraint of a semantic-aware GAN loss. Then the segmentation results are obtained from the segmentation DNN learned on the source domain. Experimental results show that the segmentation performance of SeUDA is highly competitive.

More examples of using the knowledge from images with other modalities can be found in brain MRI segmentation (Hu et al., 2020; Kamnitsas et al., 2017a), cardiac segmentation (Chartsias et al., 2020; Li et al., 2020a,b; Yan et al., 2019; Zhang et al., 2018), liver segmentation (Yang et al., 2019), lung tumor segmentation (Jiang et al., 2018), cardiac substructure and abdominal multi-organ segmentation (Chen et al., 2020).

There are also a few works utilize the datasets of other diseases. For instance, (Chen et al., 2019b) first builds a union dataset (3DSeg-8) by aggregating eight different 3D medical segmentation datasets, and designs the Med3D network to co-train based on 3DSeg-8. Then the pre-trained models obtained from Med3D are transferred into lung and liver segmentation tasks. Experiments show that this method not only improves the accuracy, but also accelerates the training convergence speed. In addition, the annotated retinal images are used to help the cardiac vessel segmentation without annotations (Yu et al., 2019). In particular, a shape-consistent generative adversarial network (SC-GAN) is used to generate the synthetic images and the corresponding labels. Then the synthetic images are used to train the segmentor. Experiments demonstrate the supreme accuracy of coronary artery segmentation.

4.3. Incorporating knowledge from medical doctors

The domain knowledge of medical doctors is also widely adopted when designing deep learning models for segmentation tasks in medical images. The types of domain knowledge from medical doctors utilized in deep segmentation models can be divided into four categories:

1. the training pattern,
2. the general diagnostic patterns they view images,
3. the anatomical priors (e.g., shape, location, topology) of lesions or organs, and
4. other hand-crafted features they give special attention to.

4.3.1. Training pattern of medical doctors

Similar to disease diagnosis and lesion/organ detection, many research works for the object segmentation in medical images also mimic the training pattern of medical doctors, which involves assigning tasks that increase in difficulty over time. In this process, the curriculum learning technique or its derivative methods like self-paced learning (SPL) are also utilized (Kumar et al., 2010).

For example, for the segmentation of multi-organ CT images (Berger et al., 2018), each annotated image is divided into small patches. During the training process, patches producing large error by the network are selected with a high probability. In this manner, the network can focus sampling on difficult regions, resulting in improved performance. In addition, (Wang et al., 2018) combines the SPL with the active learning for the pulmonary segmentation in 3D images. This system achieves the state-of-the-art segmentation results.

Moreover, a three-stage curriculum learning approach is proposed for liver tumor segmentation (Li et al., 2020a). The first stage is performed on the whole input volume to initialize the network, then the second stage of learning focuses on tumor-specific features by training the network on the tumor patches, and finally the network is retrained on the whole input in the third stage. This approach exhibits significant improvement when compared with the commonly used cascade counterpart in MICCAI 2017 liver tumor segmentation (LiTS) challenge dataset. More examples can also be

found in left ventricle segmentation (Kervadec et al., 2019), finger bones segmentation (Zhao et al., 2019) and vessel segmentation (Zhang et al., 2020b).

4.3.2. General diagnostic patterns of medical doctors

In the lesion or organ segmentation tasks, some specific patterns that medical doctors adopted are also incorporated into the network.

For example, during the visual inspection of CT images, radiologists often change window widths and window centers to help to make decision on uncertain nodules. This pattern is mimicked in (Wu et al., 2018). In particular, image patches of different window widths and window centers are stacked together as the input of the deep learning model to gain rich information. The evaluation implemented on the public LIDC-IDRI dataset indicates that the proposed method achieves promising performance on lung nodule segmentation compared with the state-of-the-art methods.

In addition, experienced clinicians generally assess the cardiac morphology and function from multiple standard views, using both long-axis (LA) and short-axis (SA) images to form an understanding of the cardiac anatomy. Inspired by the above observation, a cardiac MR segmentation method is proposed which takes three LA and one SA views as the input (Chen et al., 2019). In particular, the features are firstly extracted using a multi-view autoencoder (MAE) structure, and then are fed in a segmentation network. Experimental results show that this method has a superior segmentation accuracy over state-of-the-art methods.

Furthermore, expert manual segmentation usually relies on the boundaries of anatomical structures of interest. For instance, radiologists segmenting a liver from CT images would usually trace liver edges first, and then deduce the internal segmentation mask. Correspondingly, boundary-aware CNNs are proposed in (Hatamizadeh et al., 2019) for medical image segmentation. The networks are designed to account for organ boundary information, both by providing a special network edge branch and edge-aware loss terms. The effectiveness of these boundary aware segmentation networks are tested on BraTS 2018 dataset for the task of brain tumor segmentation.

Recently, the diagnostic pattern named as 'divide-and-conquer manner' is mimicked in the GTV_{LN} detection and segmentation method (Zhu et al., 2020). Concretely, the GTV_{LN} is first divided into two subgroups of 'tumor-proximal' and 'tumor-distal', by means of binary or soft distance gating. Then a multi-branch detection-by-segmentation network is trained with each branch specializing on learning one GTV_{LN} category features. After fusing the outs from multi-branch, the method shows significant improvements on the mean recall from 72.5% to 78.2%. Another example of using the diagnostic pattern of medical doctors can be found in gross tumor and clinical target volume segmentation (Jin et al., 2020).

4.3.3. Anatomical priors of lesions or organs

In comparison to non-medical images, medical images have many anatomical priors such as the shape, position and topology of organs or lesions. Experienced medical doctors greatly rely on these anatomical priors when they are doing segmentation tasks on these images. Incorporating the knowledge of anatomical priors into deep learning models has been demonstrated to be an effective way for accurate medical image segmentation. Generally speaking, there are three different approaches to incorporate these anatomical priors into deep learning models: (1) incorporating anatomical priors in the post-processing stage, (2) incorporating anatomical priors as regularization terms in the loss function and (3) learning anatomical priors via generative models.

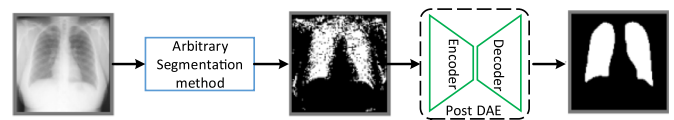**Incorporating anatomical priors in post-processing stage**

Fig. 18. The example of integrating the shape prior in the post-process stage (Larrazabal et al., 2019).

The first approach is to incorporate the anatomical priors in the post processing stage. The result of a segmentation network is often blurry and post-processing is generally needed to refine the segmentation result.

For example, according to the pathology, most of breast tumors begin in glandular tissues and are located inside the mammary layer (Sharma et al., 2010). This position feature is utilized by (Huang et al., 2018) in its post-processing stage where a fully connected conditional random field (CRF) model is employed. In particular, the position of tumors and their relative locations with mammary layer are added as a new term in CRF energy function to obtain better segmentation results.

Besides, some research first learn the anatomical priors, and then incorporate them into the post-processing stage to help produce anatomically plausible segmentation results (Larrazabal et al., 2019; Painchaud et al., 2019, 2020). For instance, the latent representation of anatomically correct cardiac shape is first learned by using adversarial variational autoencoder (aVAE), then be used to convert erroneous segmentation maps into anatomically plausible ones (Painchaud et al., 2019). Experiments manifest that aVAE is able to accommodate any segmentation method, and convert its anatomically implausible results to plausible ones without affecting its overall geometric and clinical metrics.

Another example in (Larrazabal et al., 2019) introduces the post-processing step based on denoising autoencoders (DAE) for lung segmentation. In particular, the DAE is trained using only segmentation masks, then the learned representations of anatomical shape and topological constraints are imposed on the original segmentation results (as shown in Fig. 18). By applying the Post-DAE on the resulting masks from arbitrary segmentation methods, the lung anatomical segmentation of X-ray images shows plausible results.

Incorporating anatomical priors as regularization terms in the loss function

The second approach is incorporating anatomical priors as regularization terms in the objective function of segmentation networks. For example, for the segmentation of cardiac MR images, a network called as SRSCN is proposed (Yue et al., 2019). SRSCN comprises a shape reconstruction neural network (SRNN) and a spatial constraint network (SCN). SRNN aims to maintain a realistic shape of the resulting segmentation and the SCN is adopted to incorporate the spatial information of the 2D slices. The loss of the SRSCN comes from three parts: the segmentation loss, the shape reconstruction (SR) loss for shape regularization, and the spatial constraint (SC) loss to assist segmentation. The results using images from 45 patients demonstrate the effectiveness of the SR and SC regularization terms, and show the superiority of segmentation performance of the SRSCN over the conventional schemes.

Another example in this category is the one designed for skin lesion segmentation (Mirikharaji and Hamarneh, 2018). In this work, the star shape prior is encoded as a new loss term in a FCN to improve its segmentation of skin lesions from their surrounding healthy skin. In this manner, the non-star shape segments in FCN prediction maps are penalized to guarantee a global structure in segmentation results. The experimental results on the ISBI 2017 skin segmentation challenge dataset demonstrate the advantage of regularizing FCN parameters by the star shape prior.

Table 8

The comparison of the quantitative metrics for some medical object segmentation methods after incorporating domain knowledge.

Reference	Baseline Model/With Domain Knowledge	Dice score
(Chen et al., 2019)	3D U-Net/MV U-Net	0.923/0.926
(He et al., 2019)	V-Net/DPA-DenseBiasNet	0.787/0.861
(Jiang et al., 2018)	Masked cycle-GAN/Tumore aware semi-unsupervised	0.630/0.800
(Valindria et al., 2018)	modality specific method/dual-stream encoder-decoder multi-model method	0.838/0.860
(Yue et al., 2019)	U-Net/SRSCN	0.737/0.830
(Yu et al., 2019)	U-Net/SC-GAN	0.742/0.824
(Zhang et al., 2018)	cycle- and shape-consistency GAN trained without/with synthetic data	0.678/0.744

More examples in this category can be found in gland segmentation (BenTaieb and Hamarneh, 2016), kidney segmentation (Ravishankar et al., 2017), liver segmentation (Zheng et al., 2019) and cardiac segmentation (Oktay and Ferrante, 2017; Zotti et al., 2019).

Learning anatomical priors via generative models

In the third approach, the anatomical priors (especially the shape prior) are learned by some generative models first and then incorporated into segmentation networks.

For example, in the cardiac MR segmentation process, a shape multi-view autoencoder (MAE) is proposed to learn shape priors from MR images of multiple standard views (Chen et al., 2019). The information encoded in the latent space of the trained shape MAE is incorporated into multi-view U-Net (MV U-Net) in the fuse block to guide the segmentation process.

Another example is shown in (Luo et al., 2020), where the shape constrained network (SCN) is proposed to incorporate the shape prior into the eye segmentation network. More specifically, the prior information is first learned by a VAE-GAN, and then the pre-trained encoder and discriminator are leveraged to regularize the training process.

More examples can be found in brain geometry segmentation in MRI (Dalca et al., 2018), 3D fine renal artery segmentation (He et al., 2019), overlapping cervical cytoplasms segmentation (Song et al., 2020), scapula segmentation (Boutillon et al., 2020), liver segmentation (Pham et al., 2020), carotid segmentation (Engin et al., 2020), and head and neck segmentation (Gao et al., 2020).

4.3.4. Other hand-crafted features

Besides anatomical priors, some hand-crafted features are also utilized for segmentation tasks. Generally speaking, there are two ways to incorporate the hand-crafted features into deep learning models: the feature-level fusion and the input-level fusion.

In the feature-level fusion, the hand-crafted features and the features learned by the deep models are concatenated. For example, for the gland segmentation in histopathology images (Rezaei et al., 2019), two handcrafted features, namely invariant LBP features as well as *H*&*E* components, are firstly calculated from images. Then these features are concatenated with the features generated from the last convolutional layer of the network for predicting the segmentation results. Similarly, in the brain structure segmentation (Kushibar et al., 2018), the spatial atlas prior is first represented as a vector and then concatenated with the deep features.

For the input-level fusion, the hand-crafted features are transformed into the input patches. Then the original image patches and the feature-transformed patches are fed into a deep segmentation network. For example in (Khan et al., 2020), for automatic brain tumor segmentation in MRI images, three handcrafted features (i.e., mean intensity, LBP and HOG) are firstly extracted. Based on these features, a SVM is employed to generate confidence surface modality (CSM) patches. Then the CSM patches and the original patches from MRI images are fed into a segmentation network. This method achieves good performance on BRATS2015 dataset.

In addition, using handcrafted features by input-level fusion is also adopted in cell nuclei segmentation (Narotamo et al., 2020). In particular, as nuclei are expected to have an approximately round shape, a map of gradient convergence is computed and be used by CNN as an extra channel besides the fluorescence microscopy image. Experimental results show higher F1-score when compared with other methods. Another example in this category can be found in brain tumor segmentation (Khan et al., 2020).

4.4. Summary

The aforementioned sections described researches of incorporating domain knowledge for object (lesion or organ) segmentation in medical images. The segmentation performance of some methods is shown in Table 8, where the Dice score is used with a higher score indicating a better performance.

We can see that similar to disease diagnosis, using information from natural images like ImageNet is quite popular for lesion and organ segmentation tasks. The reason behind it may be that segmentation can be seen as a specific classification problem. Meanwhile, besides the ImageNet, some video datasets can also be utilized for segmenting 3D medical images (e.g., (Zeng et al., 2017)). Using extra medical datasets with different modalities has also been proven to be helpful, although most applications are limited in using MRI to help segmentation tasks in CT images (Valindria et al., 2018). Leveraging domain knowledge from medical doctors is also widely used in segmentation tasks. In particular, the anatomical priors of organs are widely adopted. However, anatomical priors are only suitable for the segmentation of organs with fixed shapes like hearts or lungs (Chen et al., 2019).

5. Other medical applications

In this section, we briefly introduce the works on incorporating domain knowledge in other medical images analysis applications, like medical image reconstruction, medical image retrieval and medical report generation.

5.1. Medical image reconstruction

The objective of medical image reconstruction is reconstructing a diagnostic image from a number of measurements (e.g., X-ray projections in CT or the spatial frequency information in MRI). Deep learning based methods have been widely applied in this field (Qin et al., 2018; Schlemper et al., 2017). It is also quite common that external information is incorporated into deep learning models for medical image reconstruction.

Some methods incorporate hand-crafted features in the medical image reconstruction process. For example, a network model called as DAGAN is proposed for the reconstruction of compressed sensing magnetic resonance imaging (CS-MRI) (Yang et al., 2017). In the DAGAN, to better preserve texture and edges in the reconstruction

process, the adversarial loss is coupled with a content loss. In addition, the frequency-domain information is incorporated to enforce similarity in both the image and frequency domains. Experimental results show that the DAGAN method provides superior reconstruction with preserved perceptual image details.

In (Yedder et al., 2019), a new image reconstruction method is proposed to solve the limited-angle and limited sources breast cancer diffuse optical tomography (DOT) image reconstruction problem in a strong scattering medium. By adaptively focusing on important features and filtering irrelevant and noisy ones using the Fuzzy Jaccard loss, the network is able to reduce false positive reconstructed pixels and reconstruct more accurate images.

Similarly, a GAN-based method is proposed to recover MRI images of the target contrast (Dar et al., 2018). The method simultaneously leverages the relatively low-spatial-frequency information available in the collected evidence for the target contrast and the relatively high-spatial frequency information available in the source contrast. Demonstrations on brain MRI datasets indicate the proposed method outperforms state-of-the-art reconstruction methods, with enhanced recovery of high-frequency tissue structure, and improved reliability against feature leakage or loss.

5.2. Medical image retrieval

The hospitals are producing large amount of imaging data and the development of medical image retrieval, especially the content based image retrieval (CBIR) systems can be of great help to aid clinicians in browsing these large datasets. Deep learning methods have been applied to CBIR and have achieved high performance due to their superior capability for extracting features automatically.

It is also quite common that these deep learning models for CBIR utilize external information beyond the given medical datasets. Some methods adopt the transfer learning to utilize the knowledge from natural images or external medical datasets (Ahmad et al., 2017; Khatami et al., 2018; Swati et al., 2019). For example, the VGG model pre-trained based on ImageNet is used in brain tumor retrieval process (Swati et al., 2019), where a block-wise fine-tuning strategy is proposed to enhance the retrieval performance on the T1-weighted CE-MRI dataset. Another example can be found in x-ray image retrieval process (Khatami et al., 2018), where a model pre-trained on the large augmented dataset is fine-tuned on the target dataset to extract general features.

Besides, as features play an important role in the similarly analysis in CBIR, some methods fuse prior features with deep features. In particular, in the chest radiograph image retrieval process, the decision values of binary features and texture features are combined with the deep features in the form of decision-level fusion (Anavi et al., 2015). Similarly, the metadata such as patients' age and gender is combined with the image-based features extracted from deep CNN for X-ray chest pathology image retrieval (Anavi et al., 2016). Furthermore, the features extracted from saliency areas can also be injected into the features extracted from the whole image for the high retrieval accuracy (Ahmad et al., 2017).

5.3. Medical report generation

Recently, deep learning models for image captioning have been successfully applied for automatic generation of medical reports (Jing et al., 2017; Liu et al., 2019a). It is also found that incorporating external knowledge can help deep learning models to generate better medical reports.

For example, some methods try to incorporate specific or general patterns that doctors adopt when writing reports. For example, radiologists generally write reports using certain templates.

Therefore, some templates are used during the sentence generation process (Li et al., 2018, 2019). Furthermore, as the explanation given by doctors is fairly simple and phrase changing does not change their meaning, a model-agnostic method is presented to learn the short text description to explain this decision process (Gale et al., 2018).

In addition, radiologists follow some procedures when writing reports: they generally first check a patient's images for abnormal findings, then write reports by following certain templates, and adjust statements in the templates for each individual case when necessary (Hong and Kahn, 2013). This process is mimicked in (Li et al., 2019a), which first transfers the visual features of medical images into an abnormality graph, then retrieves text templates based on the abnormalities and their attributes for chest X-ray images.

In (Zhang et al., 2020c), a pre-constructed graph embedding module (modeled with a graph CNN) on multiple disease findings is utilized to assist the generation of reports. The incorporation of knowledge graph allows for dedicated feature learning for each disease finding and the relationship modeling between them. Experiments on the publicly accessible dataset (IU-RR) demonstrate the superior performance of the method integrated with the proposed graph module.

6. Research challenges and future directions

The aforementioned sections reviewed research studies on deep learning models that incorporate medical domain knowledge for various tasks. Although using medical domain knowledge in deep learning models is quite popular, there are still many difficulties about the selection, representation and incorporating method of medical domain knowledge. In the following sections, we summarize challenges and future directions in this area.

6.1. The challenges related to the identification and selection of medical domain knowledge

Identifying medical domain knowledge is not an easy task. Firstly, the experiences of medical doctors are generally subjective and fuzzy. Not all medical doctors can give accurate and objective descriptions on what kinds of experiences they have leveraged to finish a given task. In addition, experiences of medical doctors can vary significantly or even contradictory to each other. Furthermore, medical doctors generally utilize many types of domain knowledge simultaneously. Finally, currently the medical domain knowledge is identified manually, and there is no existing work on the automatically and comprehensively identifying medical domain knowledge for a given area.

One solution to the automatic identifying medical knowledge is through text mining techniques on the guidelines, books, and medical reports related to different medical areas. Guidelines or books are more robust than individual experiences. Medical reports generally contain specific terms (usually adjectives) that describe the characteristics of tumors. These terms, containing important information to help doctors to make diagnosis, can potentially be beneficial for deep learning models.

Besides the identification of medical domain knowledge, how to select appropriate knowledge to help image analysis tasks is also challenging. It should be noted that a common practice of medical doctors may not be able to help deep learning models because *the domain knowledge might be learned by the deep learning model from training data*. We believe that the knowledge that is not easily learned by a deep learning model can greatly help the model to improve its performance.

6.2. The challenges related to the representation of medical domain knowledge

The original domain knowledge of medical doctors is generally in the form of descriptive sentences like 'we will focus more on the margin areas of a tumor to determine whether it is benign or malignant', or 'we often compare bilateral images to make decision'. How to transform the knowledge into appropriate representations and incorporate it into deep learning models need a careful design.

There are generally four ways to represent a certain type of medical domain knowledge. One is to represent knowledge as patches or highlighted images (as in (Tan et al., 2019)). This is generally used when doctors pay more attention to specific areas. The second approach is to represent knowledge as feature vectors (Majtner et al., 2016). This way is suitable when the selected domain knowledge can be described as certain features. The third approach is to represent domain knowledge as extra labels (Hussein et al., 2017; Liu et al., 2018), which is suitable for the knowledge in clinical reports or extra feature attributes of diseases. The last approach is to embed medical domain knowledge in network structure design, which is suitable to represent high-level domain knowledge like the training pattern and diagnostic pattern of medical doctors(Jiménez-Sánchez et al., 2019; Liu et al., 2019, 2020).

6.3. The challenges related to the incorporating methods of medical domain knowledge

Currently, there are four ways to incorporate medical domain knowledge. The first is to transform the knowledge into certain patches or highlighted images and put them as extra inputs (Xie et al., 2019). The second approach is via concatenation (Rezaei et al., 2019). The domain knowledge are generally transformed into feature vectors and concatenated with those extracted by deep learning models. The third way is the attention mechanism (Li et al., 2019). The approach is applicable when doctors focus on certain areas of medical images or focus on certain time slots on medical videos. The last one is to learn the domain knowledge by using some specific network structures like generative models (Chen et al., 2019; Luo et al., 2020).

However, most of the existing works only incorporate a single type of medical domain knowledge, or a few types of medical domain knowledge of the same modality (e.g., a number of hand-crafted features). In practice, experienced medical doctors usually combine different experience in different stages.

There are some researches that simultaneously introduce high-level domain knowledge (e.g., diagnostic pattern, training pattern) and the low-level one (e.g., hand-crafted features, anatomical priors). In particular, the high-level domain knowledge is incorporated as input images, and low-level one is learned by using specific network structures (Chen et al., 2019). In addition, besides incorporating into network directly, the information from low-level domain knowledge can also be used to design the training orders when combined with the easy-to-hard training pattern (Tang et al., 2018). We believe that simultaneously incorporating multiple kinds of medical domain knowledge can better help deep learning models in various medical applications.

6.4. Future research directions

Besides the above challenges, there are several directions that we feel need further investigation in the future.

Domain adaptation

Domain adaptation is developed to transfer the information from a source domain to a target one. Via techniques like adversarial learning (Goodfellow et al., 2014), domain adaptation is able to narrow the domain shift between the source domain and the target one in input space (Hoffman et al., 2018), feature space (Long et al., 2016; Tzeng et al., 2017) and output space (Luo et al., 2019; Tsai et al., 2018). It can be naturally adopted to transfer knowledge of one medical dataset to another (Li et al., 2020), even when they have different imaging modes or belong to different diseases (Ghafoorian et al., 2017; Jiang et al., 2018).

In addition, unsupervised domain adaptation (UDA) is a promising avenue to enhance the performance of deep neural networks on the target domain, using labels only from the source domain. This is especially useful for medical field, as annotating the medical images is quite labor-intensive and the lack of annotations is quite common in medical datasets. Some examples have demonstrated the effectiveness of UDA in disease diagnosis and organ segmentation (Chen et al., 2018; Liu et al., 2020; Yang et al., 2019; Zhang et al., 2020), but further depth study needs to be implemented in the future.

The knowledge graph

We believe that the knowledge graph (Wang et al., 2014)(), with the character of embedding different types of knowledge, is a generic and flexible approach to incorporate multi-modal medical domain knowledge. Although rarely used at present, it also shows advantage in medical image analysis tasks, especially in medical report generation (Li et al., 2019a). We believe that more types of knowledge graph can be used to represent and learn domain knowledge in medical image analysis tasks.

According to different relationships in graphs, there are three possible types of knowledge graphs can be established. The first knowledge graph reflects the relationship among different kinds of medical domain knowledge with respect to a certain disease. This knowledge graph can help us identify a few key types of knowledge that may help to improve the performance deep learning models. The second type of knowledge graph may reflects the relationship among different diseases. This knowledge graph can help us to find out the potential domain knowledge from other related diseases. The third type one can describe the relationship among medical datasets. These datasets can belong to different diseases and in different imaging modes (e.g., CT, MRI, ultrasound). This type of knowledge graph will help to identify the external datasets that may help to improve the performance of the current deep learning model.

The generative models

The generative models, like GAN and AE, have shown great promise to be applied to incorporate medical domain knowledge into deep learning models, especially for segmentation tasks. GAN has shown its capability to leverage information from extra datasets with different imaging modes (e.g., using a MRI dataset to help segmenting CT images (Chen et al., 2018; Jiang et al., 2018)). In addition, GAN is able to learn important features contained in medical images in a weakly or fully unsupervised manner and therefore is quite suitable for medical image analysis.

AE-based models have already achieved a great success in extracting features, especially the shape priors in objects like organs or lesions in a fully unsupervised manner (Chen et al., 2019; Luo et al., 2020). The features learning by AE can also be easily integrated into the training process of networks.

Network architecture search (NAS)

At last, we mentioned in the previous section that one challenge is to find appropriate network architectures to incorporate medical domain knowledge. We believe one approach to address this problem is the technique of network architecture search (NAS). NAS has demonstrated its capability to automatically find a good network architecture in many computer vision tasks (Wistuba et al., 2019) and has a great promise in the medical do-

main (Guo et al., 2020). For instance, when some hand-crafted features are used as the domain knowledge, with the help of NAS, a network structure can be identified with the special connections between domain knowledge features and deep features. In addition, instead of designing the feature fusion method (feature-level fusion, decision-level fusion or input-level fusion) for these two kinds of features, the integrating phase and integrating intensity of these two kinds of features can also be determined during the searching process.

7. Conclusion

In this paper, we give a comprehensive survey on incorporating medical domain knowledge into deep learning models for various medical image analysis tasks ranging from disease diagnosis, lesion, organ and abnormality detection to lesion and organ segmentation. In addition, some other tasks such as medical image reconstruction, medical image retrieval and medical report generation are also included. For each task, we first introduce different types of medical domain knowledge, and then review some works of introducing domain knowledge into target tasks by using different incorporating methods. From this survey, we can see that with appropriate integrating methods, different kinds of domain knowledge can help deep learning models to better accomplish corresponding tasks.

Besides reviewing current works on incorporating domain knowledge into deep learning models, we also summarize challenges of using medical domain knowledge, and introduce the identification, selection, representation and incorporating method of medical domain knowledge. Finally, we give some future directions of incorporating domain knowledge for medical image analysis tasks.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Xiaozheng Xie: Conceptualization, Investigation, Methodology, Resources, Writing - original draft. **Jianwei Niu:** Writing - review & editing, Supervision, Funding acquisition. **Xuefeng Liu:** Writing - review & editing, Supervision, Formal analysis, Funding acquisition. **Zhengsu Chen:** Writing - review & editing, Resources. **Shaojie Tang:** Writing - review & editing, Visualization. **Shui Yu:** Writing - review & editing.

Acknowledgments

References

Ahmad, J., Sajjad, M., Mehmood, I., Baik, S.W., 2017. Sinc: saliency-injected neural codes for representation and efficient retrieval of medical radiographs. PLoS ONE 12 (8).

Alexander A., J.K.T., 2008. Efficient and generalizable statistical models of shape and appearance for analysis of cardiac mri. Med. Image Anal. 12 (3), 335–357. doi:10.1016/j.media.2007.12.003.

Alilou, M., Orooji, M., Madabhushi, A., 2017. Intra-perinodular textural transition (ipris): a 3d descriptor for nodule diagnosis on lung ct. International Conference on Medical Image Computing and Computer Assisted Intervention. Springer, pp. 647–655.

Anavi, Y., Kogan, I., et al., 2016. Visualizing and enhancing a deep learning framework using patients age and gender for chest x-ray image retrieval. In: Medical Imaging 2016: Computer-Aided Diagnosis, 9785. International Society for Optics and Photonics, p. 978510.

Anavi, Y., Kogan, I., Gelbart, E., Geva, O., Greenspan, H., 2015. A comparative study for chest radiograph image retrieval using binary texture and deep learning classification. International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE, pp. 2940–2943.

Antropova, N., Huynh, B.Q., Giger, M.L., 2017. A deep feature fusion methodology for breast cancer diagnosis demonstrated on three imaging modality datasets. Med. Phys. 44 (10), 5162–5171.

Armato III, S.G., McLennan, G., Bidaut, L., McNitt-Gray, M.F., Meyer, C.R., Reeves, A.P., Zhao, B., Aberle, D.R., Henschke, C.I., Hoffman, E.A., et al., 2011. The lung image database consortium (lidc) and image database resource initiative (idri): a completed reference database of lung nodules on ct scans. Med. Phys. 38 (2), 915–931.

Astudillo, P., Mortier, P., De Beule, M., et al., 2020. Curriculum deep reinforcement learning with different exploration strategies: A feasibility study on cardiac landmark detection. In: BIOIMAGING 2020: 7th International Conference on Bioimaging.

Azizi, S., Mousavi, P., Yan, P., Tahmasebi, A., Kwak, J.T., Xu, S., Turkbey, B., Choyke, P., Pinto, P., Wood, B., et al., 2017. Transfer learning from rf to b-mode temporal enhanced ultrasound features for prostate cancer detection. Int. J. Comput. Assist. Radiol. Surg. 12 (7), 1111–1121.

Bakas, S., Reyes, M., Jakab, A., Bauer, S., Rempfler, M., Crimi, A., Shinohara, R.T., Berger, C., Ha, S. M., Rozycki, M., et al., 2018. Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the brats challenge. arXiv:1811.02629

Bakalo, R., Ben-Ari, R., Goldberger, J., 2019. Classification and detection in mammograms with weak supervision via dual branch deep neural net. International Symposium on Biomedical Imaging (ISBI). IEEE, pp. 1905–1909.

Bar, Y., Diamant, I., Wolf, L., Greenspan, H., 2015. Deep learning with non-medical training used for chest pathology identification. Medical Imaging 2015: Computer-Aided Diagnosis, 9414. International Society for Optics and Photonics. 94140V

Ben-Ari, R., Akselrod-Ballin, A., Karlinsky, L., Hashoul, S., 2017. Domain specific convolutional neural nets for detection of architectural distortion in mammograms. In: IEEE International Symposium on Biomedical Imaging (ISBI). IEEE, pp. 552–556.

Ben-Cohen, A., Klang, E., Raskin, S.P., Soffer, S., Ben-Haim, S., Konen, E., Amitai, M.M., Greenspan, H., 2019. Cross-modality synthesis from ct to pet using fcn and gan networks for improved automated lesion detection. Eng Appl Artif Intell 78, 186–194.

Bengio, Y., Louradour, J., Collobert, R., Weston, J., 2009. Curriculum learning. Proceedings of the 26th annual international conference on machine learning, ACM. 41–48.

BenTaieb, A., Hamarneh, G., 2016. Topology aware fully convolutional networks for histology gland segmentation. International Conference on Medical Image Computing and Computer Assisted Intervention. Springer, pp. 460–468.

Berger, L., Eoin, H., Cardoso, M.J., Ourselin, S., 2018. An adaptive sampling scheme to efficiently train fully convolutional networks for semantic segmentation. In: Annual Conference on Medical Image Understanding and Analysis. Springer, pp. 277–286.

Boutillon, A., Borotikar, B., Burdin, V., Conze, P.-H., 2020. Combining shape priors with conditional adversarial networks for improved scapula segmentation in mr images. International Symposium on Biomedical Imaging (ISBI). IEEE, pp. 1164–1167.

Breiman, L., 2001. Random forests. Mach. Learn. 45 (1), 5–32.

Buty, M., Xu, Z., Gao, M., Bagci, U., Wu, A., Mollura, D.J., 2016. Characterization of lung nodule malignancy using hybrid shape and appearance features. Medical Image Computing and Computer Assisted Intervention. Springer, pp. 662–670.

Cai, G., Chen, J., Wu, Z., Tang, H., Liu, Y., Wang, S., Su, S., 2019. One stage lesion detection based on 3d context convolutional neural networks. Comput. Electr. Eng. 79, 106449.

Cao, H., Bernard, S., Heutte, L., Sabourin, R., 2018. Improve the performance of transfer learning without fine-tuning using dissimilarity-based multi-view learning for breast cancer histology images. In: International conference image analysis and recognition, pp. 779–787.

Chai, Y., Liu, H., Xu, J., 2018. Glaucoma diagnosis based on both hidden features and domain knowledge through deep learning models. Knowl. Based Syst. 161, 147–156.

Chao, C.-H., Zhu, Z., Guo, D., Yan, K., Ho, T.-Y., Cai, J., Harrison, A.P., Ye, X., Xiao, J., Yuille, A., et al., 2020. Lymph node gross tumor volume detection in oncology imaging via relationship learning using graph neural network. International Conference on Medical Image Computing and Computer Assisted Intervention. Springer, pp. 772–782.

Chartsias, A., Papanastasiou, G., Wang, C., Semple, S., Newby, D., Dharmakumar, R., Tsaftaris, S., 2020. Disentangle, align and fuse for multimodal and semi-supervised image segmentation. IEEE Trans. Med. Imaging.

Chen, C., Biffi, C., Tarroni, G., Petersen, S.E., Bai, W., Rueckert, D., et al., 2019. Learning shape priors for robust cardiac mr segmentation from multi-view images. International Conference on Medical Image Computing and Computer Assisted Intervention. Springer, pp. 523–531.

Chen, C., Dou, Q., Chen, H., Heng, P.-A., 2018. Semantic-aware generative adversarial nets for unsupervised domain adaptation in chest x-ray segmentation. In: International Workshop on Machine Learning in Medical Imaging. Springer, pp. 143–151.

Chen, C., Dou, Q., Chen, H., Qin, J., Heng, P.A., 2020. Unsupervised bidirectional cross-modality adaptation via deeply synergistic image and feature alignment for medical image segmentation. IEEE Trans. Med. Imaging.

Chen, H., Qi, X.J., Cheng, J.Z., Heng, P.A., 2016a. Deep contextual networks for neuronal structure segmentation. In: Thirtieth AAAI conference on artificial intelligence. AAAI, pp. 1167–1173.

Chen, H., Qi, X., Yu, L., Dou, Q., Qin, J., Heng, P.-A., 2017. Dcan: deep contour-aware networks for object instance segmentation from histology images. Med. Image Anal. 36, 135–146.

Chen, H., Zheng, Y., Park, J.-H., Heng, P.-A., Zhou, S.K., 2016b. Iterative multi-domain regularized deep learning for anatomical structure detection and segmentation from ultrasound images. International Conference on Medical Image Computing and Computer Assisted Intervention. Springer, pp. 487–495.

Chen, S., Ma, K., Zheng, Y., 2019b. Med3d: Transfer learning for 3d medical image analysis. arXiv:1904.00625

Chen, S., Qin, J., Ji, X., Lei, B., Wang, T., Ni, D., Cheng, J.-Z., 2016c. Automatic scoring of multiple semantic attributes with multi-task feature leverage: a study on pulmonary nodules in ct images. IEEE Trans. Med. Imaging 36 (3), 802–814.

Christ, P.F., Ettlinger, F., Grün, F., et al., 2017. Automatic liver and tumor segmentation of ct and mri volumes using cascaded fully convolutional neural networks. arXiv:1702.05970

Christ, P. F., Elshaer, M.E.A., Ettlinger, F., Tatavarty, S., Bickel, M., Bilic, P., Rempfler, M., Hofmann, F., Danastasi, M., Sommer, W.H., Ahamadi, S., Menze, B.H., 2016. Automatic liver and lesion segmentation in ct using cascaded fully convolutional neural networks and 3d conditional random fields. International Conference on Medical Image Computing and Computer Assisted Intervention. Springer, pp. 415–423.

Ciompi, F., de Hoop, B., van Riel, S.J., Chung, K., Scholten, E.T., Oudkerk, M., de Jong, P.A., Prokop, M., van Ginneken, B., 2015. Automatic classification of pulmonary peri-fissural nodules in computed tomography using an ensemble of 2d views and a convolutional neural network out-of-the-box. Med. Image Anal. 26 (1), 195–202.

Codella, N., Rotemberg, V., Tschandl, P., Celebi, M.E., Dusza, S., Gutman, D., Helba, B., Kalloo, A., Liopyris, K., Marchetti, M., et al., 2019. Skin lesion analysis toward melanoma detection 2018: A challenge hosted by the international skin imaging collaboration (isic). arXiv:1902.03368

Cortes, C., Vapnik, V., 1995. Support-vector networks. Mach. Learn. 20 (3), 273–297.

Dar, S.U.H., Yurt, M., Shahdloo, M., Ildız, M.E., Çukur, T., 2018. Synergistic reconstruction and synthesis via generative adversarial networks for accelerated multi-contrast mri. arXiv:1805.10704

Cui, H., Xu, Y., Li, W., Wang, L., Duh, H., 2020. Collaborative learning of cross-channel clinical attention for radiotherapy-related esophageal fistula prediction from ct. International Conference on Medical Image Computing and Computer Assisted Intervention. Springer, pp. 212–220.

Dalca, A.V., Guttag, J., Sabuncu, M.R., 2018. Anatomical priors in convolutional networks for unsupervised biomedical segmentation. IEEE Conference on Computer Vision and Pattern Recognition. IEEE, pp. 9290–9299.

Debelee, T.G., Schwenker, F., Ibenthal, A., Yohannes, D., 2019. Survey of deep learning in breast cancer image analysis. Evolving Systems 1–21.

Di Martino, A., Yan, C.-G., Li, Q., Denio, E., Castellanos, F.X., Alaerts, K., Anderson, J.S., Assaf, M., Bookheimer, S.Y., Dapretto, M., et al., 2014. The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism. Mol. Psychiatry 19 (6), 659–667.

Engin, M., Lange, R., Nemes, A., Monajemi, S., Mohammadzadeh, M., Goh, C.K., Tu, T.M., Tan, B.Y., Paliwal, P., Yeo, L.L., et al., 2020. Agan: an anatomy corrector conditional generative adversarial network. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 708–717.

Erickson, B.J., Korfiatis, P., Akkus, Z., Kline, T.L., 2017. Machine learning for medical imaging. Radiographics 37 (2), 505–515.

Esteva, A., Kuprel, B., Novoa, R.A., Ko, J.M., Swetter, S.M., Blau, H.M., Thrun, S., 2017. Dermatologist-level classification of skin cancer with deep neural networks. Nature 542 (7639), 115–118.

Fang, L., Wang, C., Li, S., Rabbani, H., Chen, X., Liu, Z., 2019. Attention to lesion: lesion-aware convolutional neural network for retinal optical coherence tomography image classification. IEEE Trans. Med. Imaging 38 (8), 1959–1970. doi:10.1109/TMI.2019.2898414.

Feng, H., Cao, J., Wang, H., Xie, Y., Yang, D., Feng, J., Chen, B., 2020. A knowledge–driven feature learning and integration method for breast cancer diagnosis on multi-sequence mri. Magn. Reson. Imaging 69, 40–48.

Fu, L., Ma, J., Ren, Y., Han, Y.S., Zhao, J., 2017. Automatic detection of lung nodules: false positive reduction using convolution neural networks and handcrafted features. In: Medical Imaging 2017: Computer-Aided Diagnosis, 10134. International Society for Optics and Photonics, p. 101340A.

Gale, W., Oakden-Rayner, L., Carneiro, G., Bradley, A.P., Palmer, L.J., 2018. Producing radiologist-quality reports for interpretable artificial intelligence. arXiv:1806.00340

Gao, Y., Huang, R., Yang, Y., Zhang, J., Shao, K., Tao, C., Chen, Y., Metaxas, D.N., Li, H., Chen, M., 2020. Focusnetv2: imbalanced large and small organ segmentation with adversarial shape constraint for head and neck ct images. Med. Image Anal. 67, 101831.

Ghafoorian, M., Mehrtash, A., Kapur, T., Karssemeijer, N., Marchiori, E., Pesteie, M., Guttmann, C.R., de Leeuw, F.-E., Tempany, C.M., van Ginneken, B., et al., 2017. Transfer learning for domain adaptation in mri: Application in brain lesion segmentation. International Conference on Medical Image Computing and Computer Assisted Intervention. Springer, pp. 516–524.

Ghatwary, N., Ye, X., Zolgharni, M., 2019. Esophageal abnormality detection using densenet based faster r-cnn with gabor features. IEEE Access 7, 84374–84385.

Gibson, E., Giganti, F., Hu, Y., Bonmati, E., Bandula, S., Gurusamy, K., Davidson, B.R., Pereira, S.P., Clarkson, M.J., Barratt, D.C., 2017. Towards image-guided pancreas and biliary endoscopy: automatic multi-organ segmentation on abdominal ct with dense dilated networks. International Conference on Medical Image Computing and Computer Assisted Intervention. Springer, pp. 728–736.

González-Díaz, I., 2018. Dermaknet: incorporating the knowledge of dermatologists to convolutional neural networks for skin lesion diagnosis. IEEE J. Biomed. Health Inform. 23 (2), 547–559.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y., 2014. Generative adversarial nets. In: Advances in neural information processing systems, pp. 2672–2680.

Gordienko, Y., Gang, P., Hui, J., Zeng, W., Kochura, Y., Alienin, O., Rokovyi, O., Stirenko, S., 2018. Deep learning with lung segmentation and bone shadow exclusion techniques for chest x-ray analysis of lung cancer. In: International Conference on Computer Science, Engineering and Education Applications. Springer, pp. 638–647.

Gulshan, V., Peng, L., Coram, M., Stumpe, M.C., Wu, D., Narayanaswamy, A., Venugopalan, S., Widner, K., Madams, T., Cuadros, J., et al., 2016. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA 316 (22), 2402–2410.

Guo, D., Jin, D., Zhu, Z., Ho, T.-Y., Harrison, A.P., Chao, C.-H., Xiao, J., Lu, L., 2020. Organ at risk segmentation for head and neck cancer using stratified learning and neural architecture search. IEEE Conference on Computer Vision and Pattern Recognition. IEEE, pp. 4223–4232.

Haarburger, C., Baumgartner, M., Truhn, D., Broeckmann, M., Schneider, H., Schrading, S., Kuhl, C., Merhof, D., 2019. Multi scale curriculum cnn for context-aware breast mri malignancy classification. Medical Image Computing and Computer Assisted Intervention. Springer, pp. 495–503.

Hadad, O., Bakalo, R., Ben-Ari, R., Hashoul, S., Amit, G., 2017. Classification of breast lesions using cross-modal deep learning. International Symposium on Biomedical Imaging. IEEE, pp. 109–112.

Hagerty, J., Stanley, J., et al., 2019. Deep learning and handcrafted method fusion: higher diagnostic accuracy for melanoma dermoscopy images. IEEE J. Biomed. Health Inform..

Halevy, A., Norvig, P., Pereira, F., 2009. The unreasonable effectiveness of data. IEEE Intell. Syst. 24 (2), 8–12.

Han, X., Wang, J., Zhou, W., Chang, C., Ying, S., Shi, J., 2020. Deep doubly supervised transfer network for diagnosis of breast cancer with imbalanced ultrasound imaging modalities. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 141–149.

Hatamizadeh, A., Terzopoulos, D., Myronenko, A., 2019. End-to-end boundary aware networks for medical image segmentation. In: International Workshop on Machine Learning in Medical Imaging. Springer, pp. 187–194.

Havaei, M., Davy, A., Warde-Farley, D., Biard, A., Courville, A., Bengio, Y., Pal, C., Jodoin, P.-M., Larochelle, H., 2017. Brain tumor segmentation with deep neural networks. Med. Image Anal. 35, 18–31.

He, K., Gkioxari, G., Dollar, P., Girshick, R.B., 2017. Mask r-cnn. International Conference on Computer Vision 2980–2988.

He, Y., Yang, G., Chen, Y., Kong, Y., Wu, J., Tang, L., Zhu, X., Dillenseger, J., Shao, P., Zhang, S., et al., 2019. Dpa-densebiasnet: semi-supervised 3d fine renal artery segmentation with dense biased network and deep priori anatomy. In: Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 139–147.

He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition. In: IEEE Conference on Computer Vision and Pattern Recognition 770–778.

Heath, M., Bowyer, K., Kopans, D., Moore, R., Kegelmeyer, P., 2000. The digital database for screening mammography. Proceedings of the Fourth International Workshop on Digital Mammography doi:10.1007/978-94-011-5318-8_75.

Hoffman, J., Tzeng, E., Park, T., Zhu, J.-Y., Isola, P., Saenko, K., Efros, A., Darrell, T., 2018. Cycada: Cycle-consistent adversarial domain adaptation. In: International conference on machine learning. PMLR, pp. 1989–1998.

Hong, Y., Kahn, C.E., 2013. Content analysis of reporting templates and free-text radiology reports. J. Digit. Imaging 26 (5), 843–849.

Hoover, A., Kouznetsova, V., Goldbaum, M., 2000. Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response. IEEE Trans. Med. Imaging 19 (3), 203–210.

Hsu, S.-M., Kuo, W.-H., Kuo, F.-C., Liao, Y.-Y., 2019. Breast tumor classification using different features of quantitative ultrasound parametric images. Int. J. Comput. Assist. Radiol. Surg. 14 (4), 623–633.

Hu, M., Maillard, M., Zhang, Y., Ciceri, T., La Barbera, G., Bloch, I., Gori, P., 2020. Knowledge distillation from multi-modal to mono-modal segmentation networks. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 772–781.

Huang, K., Cheng, H.-D., Zhang, Y., Zhang, B., Xing, P., Ning, C., 2018. Medical knowledge constrained semantic breast ultrasound image segmentation. In: 2018 24th International Conference on Pattern Recognition (ICPR). IEEE, pp. 1193–1198.

Huang, X., Fang, Y., Lu, M., Yan, F., Yang, J., Xu, Y., 2020. Dual-ray net: automatic diagnosis of thoracic diseases using frontal and lateral chest x-rays. J. Med. Imaging Health Inform. 10 (2), 348–355.

Huang, G., Liu, Z., Der Maaten, L.V., Weinberger, K.Q., 2017. Densely connected convolutional networks. IEEE Conference on Computer Vision and Pattern Recognition 2261–2269.

Hussein, S., Cao, K., Song, Q., Bagci, U., 2017. Risk stratification of lung nodules us-

ing 3d cnn-based multi-task learning. In: International Conference Information Processing 249–260.

Huynh, B., Drukker, K., Giger, M., 2016. Mo-de-207b-06: computer-aided diagnosis of breast ultrasound images using transfer learning from deep convolutional neural networks. Med. Phys. 43. 3705–3705.

Huynh, B.Q., Li, H., Giger, M.L., 2016. Digital mammographic tumor classification using transfer learning from deep convolutional neural networks. J. Med. Imaging 3 (3), 034501.

Hwang, S., Kim, H.-E., 2016. Self-transfer learning for weakly supervised lesion localization. International Conference on Medical Image Computing and Computer Assisted Intervention. Springer, pp. 239–246.

Izadi, S., Mirikharaji, Z., Kawahara, J., Hamarneh, G., 2018. Generative adversarial networks to segment skin lesions. The IEEE International Symposium on Biomedical Imaging (ISBI). IEEE, pp. 881–884.

Jesson, A., Guizard, N., Ghalehjegh, S.H., Goblot, D., Soudan, F., Chapados, N., 2017. Cased: curriculum adaptive sampling for extreme data imbalance. Medical Image Computing and Computer Assisted Intervention. Springer, pp. 639–646.

Ji, C., Basodi, S., Xiao, X., Pan, Y., 2020. Infant sound classification on multi-stage cnns with hybrid features and prior knowledge. In: International Conference on AI and Mobile Services. Springer, pp. 3–16.

Jiang, J., Hu, Y.-C., Tyagi, N., Zhang, P., Rimner, A., Mageras, G.S., Deasy, J.O., Veeraraghavan, H., 2018. Tumor-aware, adversarial domain adaptation from ct to mri for lung cancer segmentation. International Conference on Medical Image Computing and Computer Assisted Intervention. Springer, pp. 777–785.

Jiménez-Sánchez, A., Mateus, D., Kirchhoff, S., Kirchhoff, C., Biberthaler, P., Navab, N., Ballester, M.A.G., Piella, G., 2019. Medical-based deep curriculum learning for improved fracture classification. In: Med Image Comput Comput Assist Interv. Springer, pp. 694–702.

Jin, D., Guo, D., Ho, T.-Y., Harrison, A.P., Xiao, J., Tseng, C.-k., Lu, L., 2020. Deeptarget: gross tumor and clinical target volume segmentation in esophageal cancer radiotherapy. Med. Image Anal. 101909.

Jiménez-Sánchez, A., Mateus, D., Kirchhoff, S., Kirchhoff, C., Biberthaler, P., Navab, N., Ballester, M.A.G., Piella, G., 2020. Curriculum learning for annotation-efficient medical image analysis: scheduling data with prior knowledge and uncertainty. arXiv:2007.16102

Jing, B., Xie, P., Xing, E., 2017. On the automatic generation of medical imaging reports. arXiv:1711.08195

Kamnitsas, K., Baumgartner, C., Ledig, C., et al., 2017. Unsupervised domain adaptation in brain lesion segmentation with adversarial networks. In: International conference on information processing in medical imaging. Springer, pp. 597–609.

Kamnitsas, K., Ledig, C., Newcombe, V.F., Simpson, J.P., Kane, A.D., Menon, D.K., Rueckert, D., Glocker, B., 2017. Efficient multi-scale 3d cnn with fully connected crf for accurate brain lesion segmentation. Med. Image Anal. 36, 61–78.

Kervadec, H., Dolz, J., Granger, E., Ayed, I.B., 2019. Curriculum semi-supervised segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 568–576.

Khan, H., Shah, P.M., Shah, M.A., ul Islam, S., Rodrigues, J.J., 2020. Cascading handcrafted features and convolutional neural network for iot-enabled brain tumor segmentation. Comput. Commun..

Khatami, A., Babaie, M., Tizhoosh, H.R., Khosravi, A., Nguyen, T., Nahavandi, S., 2018. A sequential search-space shrinking using cnn transfer learning and a radon projection pool for medical image retrieval. Expert Syst. Appl. 100, 224–233.

Kim, J.Y., Lee, H.E., Choi, Y.H., Lee, S.J., Jeon, J.S., 2019. Cnn-based diagnosis models for canine ulcerative keratitis. Sci. Rep. 9 (1), 1–7.

Kooi, T., Litjens, G., Van Ginneken, B., Gubern-Mérida, A., Sánchez, C.I., Mann, R., den Heeten, A., Karssemeijer, N., 2017. Large scale deep learning for computer aided detection of mammographic lesions. Med. Image Anal. 35, 303–312.

Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. Imagenet classification with deep convolutional neural networks. Neural Information Processing Systems 141 (5), 1097–1105.

Kumar, M.P., Packer, B., Koller, D., 2010. Self-paced learning for latent variable models. In: Advances in Neural Information Processing Systems, pp. 1189–1197.

Kushibar, K., Valverde, S., González-Villà, S., Bernal, J., Cabezas, M., Oliver, A., Lladó, X., 2018. Automated sub-cortical brain structure segmentation combining spatial and deep convolutional features. Med. Image Anal. 48, 177–186.

Lahiri, A., Ayush, K., Kumar Biswas, P., Mitra, P., 2017. Generative adversarial learning for reducing manual annotation in semantic segmentation on large scale miscroscopy images: Automated vessel segmentation in retinal fundus image as test case. IEEE Conference on Computer Vision and Pattern Recognition Workshops. IEEE, pp. 42–48.

Li, Y., Liang, X., Hu, Z., Xing, E.P., 2018. Hybrid retrieval-generation reinforced agent for medical image report generation. Advances in Neural Information Processing Systems. NIPS, pp. 1530–1540.

Li, C.Y., Liang, X., Hu, Z., Xing, E.P., 2019. Knowledge-driven encode, retrieve, paraphrase for medical image report generation. In: Proceedings of the AAAI Conference on Artificial Intelligence, 33, pp. 6666–6673.

Li, H., Liu, X., Boumaraf, S., Liu, W., Gong, X., Ma, X., 2020. A new three-stage curriculum learning approach for deep network based liver tumor segmentation. In: 2020 International Joint Conference on Neural Networks (IJCNN). IEEE, pp. 1–6.

Li, L., Xu, M., Wang, X., Jiang, L., Liu, H., 2019. Attention based glaucoma detection: A large-scale database and cnn model. IEEE Conference on Computer Vision and Pattern Recognition. IEEE, pp. 10571–10580.

Li, K., Yu, L., Wang, S., Heng, P.-A., 2020. Towards cross-modality medical image segmentation with online mutual knowledge distillation. In: Proceedings of the AAAI Conference on Artificial Intelligence, 34, pp. 775–783.

Li, K., Wang, S., Yu, L., Heng, P.-A., 2020. Dual-teacher: integrating intra-domain and inter-domain teachers for annotation-efficient cardiac segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 418–427.

Li, L., Wei, M., Liu, B., Atchaneeyasakul, K., Zhou, F., Pan, Z., Kumar, S., Zhang, J., Pu, Y., Liebeskind, D.S., et al., 2020. Deep learning for hemorrhagic lesion detection and segmentation on brain ct images. IEEE J. Biomed. Health Inform..

Li, X., Qin, G., He, Q., Sun, L., Zeng, H., He, Z., Chen, W., Zhen, X., Zhou, L., 2020. Digital breast tomosynthesis versus digital mammography: integration of image modalities enhances deep learning-based breast mass classification. Eur. Radiol. 30 (2), 778–788.

Li, X., Shen, L., Xie, X., Huang, S., Xie, Z., Hong, X., Yu, J., 2019. Multi-resolution convolutional networks for chest x-ray radiograph based lung nodule detection. Artif. Intell. Med. 101744.

Larrazabal, A.J., Martinez, C.E., Ferrante, E., 2019. Anatomical priors for image segmentation via post-processing with denoising autoencoders. International Conference on Medical Image Computing and Computer Assisted Intervention. Springer, pp. 585–593.

Li, X., Hu, X., Yu, L., Zhu, L., Fu, C.-W., Heng, P.-A., 2019. Canet: cross-disease attention network for joint diabetic retinopathy and diabetic macular edema grading. IEEE Trans. Med. Imaging 39 (5), 1483–1493. doi:10.1109/TMI.2019.2951844.

Li, W., Li, J., Sarma, K.V., Ho, K.C., Shen, S., Knudsen, B.S., Gertych, A., Arnold, C.W., 2018. Path r-cnn for prostate cancer diagnosis and gleason grading of histological images. IEEE Trans. Med Imaging 38 (4), 945–954.

Li, Z., Zhang, S., Zhang, J., Huang, K., Wang, Y., Yu, Y., 2019. Mvp-net: Multi-view fpn with position-aware attention for deep universal lesion detection. In: Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 13–21.

Liang, G., Wang, X., Zhang, Y., Jacobs, N., 2020. Weakly-supervised self-training for breast cancer localization. International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC). IEEE, pp. 1124–1127.

Liao, Q., Ding, Y., Jiang, Z.L., Wang, X., Zhang, C., Zhang, Q., 2019. Multi-task deep convolutional neural network for cancer diagnosis. Neurocomputing 348, 66–73.

Lin, T.-Y., Goyal, P., Girshick, R., He, K., Dollár, P., 2017. Focal loss for dense object detection. In: Proceedings of the IEEE international conference on computer vision, pp. 2980–2988.

Lisowska, A., Beveridge, E., Muir, K., Poole, I., 2017. Thrombus detection in ct brain scans using a convolutional neural network.. In: BIOIMAGING, pp. 24–33.

Lisowska, A., Oneil, A., Dilys, V., Daykin, M., Beveridge, E., Muir, K., Mclaughlin, S., Poole, I., 2017. Context-aware convolutional neural networks for stroke sign detection in non-contrast ct scans. In: Annual Conference on Medical Image Understanding and Analysis. Springer, pp. 494–505.

Litjens, G., Kooi, T., Bejnordi, B.E., Setio, A.A.A., Ciompi, F., Ghafoorian, M., Van Der Laak, J.A., Van Ginneken, B., Sánchez, C.I., 2017. A survey on deep learning in medical image analysis. Med. Image Anal. 42, 60–88.

Liu, D., Zhang, D., Song, Y., Zhang, F., ODonnell, L., Huang, H., Chen, M., Cai, W., 2020. Pdam: a panoptic-level feature alignment framework for unsupervised domain adaptive instance segmentation in microscopy images. IEEE Trans. Med Imaging.

Li, N., Liu, H., Qiu, B., Guo, W., Zhao, S., Li, K., He, J., 2017. Detection and attention: Diagnosing pulmonary lung cancer from ct by imitating physicians. arXiv:1712.05114

Liu, G., Hsu, T. H., Mcdermott, M.B.A., Boag, W., Weng, W., Szolovits, P., Ghassemi, M., 2019a. Clinically accurate chest x-ray report generation. arXiv:1904.02633

Liu, J., Li, W., Zhao, N., Cao, K., Yin, Y., Song, Q., Chen, H., Gong, X., 2018. Integrate domain knowledge in training cnn for ultrasonography breast cancer diagnosis. International Conference on Medical Image Computing and Computer Assisted Intervention. Springer, pp. 868–875.

Liu, J., Wang, D., Lu, L., Wei, Z., Kim, L., Turkbey, E.B., Sahiner, B., Petrick, N.A., Summers, R.M., 2017. Detection and diagnosis of colitis on computed tomography using deep convolutional neural networks. Med. Phys. 44 (9), 4630–4642.

Liu, Q., Yu, L., Luo, L., Dou, Q., Heng, P.A., 2020. Semi-supervised medical image classification with relation-driven self-ensembling model. IEEE Trans. Med. Imaging.

Liu, T., Guo, Q., Lian, C., Ren, X., Liang, S., Yu, J., Niu, L., Sun, W., Shen, D., 2019. Automated detection and classification of thyroid nodules in ultrasound images using clinical-knowledge-guided convolutional neural networks. Med. Image Anal. 58, 101555.

Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y., Berg, A.C., 2016. Ssd: Single shot multibox detector. In: European conference on computer vision. Springer, pp. 21–37.

Liu, Y., Zhang, F., Zhang, Q., Wang, S., Yu, Y., 2020. Cross-view correspondence reasoning based on bipartite graph convolutional network for mammogram mass detection. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, pp. 3812–3822.

Liu, Y., Zhou, Z., Zhang, S., Luo, L., Zhang, Q., Zhang, F., Li, X., Wang, Y., Yu, Y., 2019. From unilateral to bilateral learning: Detecting mammogram masses with contrasted bilateral network. International Conference on Medical Image Computing and Computer Assisted Intervention. Springer, pp. 477–485.

Long, J., Shelhamer, E., Darrell, T., 2015. Fully convolutional networks for semantic segmentation. IEEE Conference on Computer Vision and Pattern Recognition 3431–3440.

Long, M., Zhu, H., Wang, J., Jordan, M.I., 2016. Unsupervised domain adaptation with residual transfer networks. In: Advances in Neural Information Processing Systems, pp. 136–144.

Luo, B., Shen, J., Cheng, S., Wang, Y., Pantic, M., 2020. Shape constrained network for

eye segmentation in the wild. In: The IEEE Winter Conference on Applications of Computer Vision, pp. 1952–1960.

Luo, Y., Zheng, L., Guan, T., Yu, J., Yang, Y., 2019. Taking a closer look at domain shift: Category-level adversaries for semantics consistent domain adaptation. IEEE Conference on Computer Vision and Pattern Recognition. IEEE, pp. 2507–2516.

Maicas, G., Bradley, A.P., Nascimento, J.C., Reid, I., Carneiro, G., 2018. Training medical image analysis systems like radiologists. Medical Image Computing and Computer Assisted Intervention Society. Springer, pp. 546–554.

Majtner, T., Yildirim-Yayilgan, S., Hardeberg, J.Y., 2016. Combining deep learning and hand-crafted features for skin lesion classification. In: 2016 Sixth International Conference on Image Processing Theory, Tools and Applications (IPTA). IEEE, pp. 1–6.

MICCAI, C., 2017. Automated cardiac diagnosis challenge (ACDC). https://www.creatis.insa-lyon.fr/Challenge/acdc/databases.html

Mendelson, E.B., Bohm-Vlez, M., B.W., E.A., 2013. Acr bi-rads ultrasound. in: acr bi-rads atlas, breast imaging reporting and data system.. Reston, VA, American College of Radiology.

Mirikharaji, Z., Hamarneh, G., 2018. Star shape prior in fully convolutional networks for skin lesion segmentation. International Conference on Medical Image Computing and Computer Assisted Intervention. Springer, pp. 737–745.

Moeskops, P., Wolterink, J.M., van der Velden, B.H., Gilhuijs, K.G., Leiner, T., Viergever, M.A., Isgum, I., 2016. Deep learning for multi-task medical image segmentation in multiple modalities. International Conference on Medical Image Computing and Computer Assisted Intervention. Springer, pp. 478–486.

Moradi, M., Gur, Y., Wang, H., Prasanna, P., Syeda-Mahmood, T., 2016. A hybrid learning approach for semantic labeling of cardiac ct slices and recognition of body position. In: IEEE Int Symp Biomedical Imaging. IEEE, pp. 1418–1421.

Murthy, V., Hou, L., Samaras, D., Kurc, T.M., Saltz, J.H., 2017. Center-focusing multi-task cnn with injected features for classification of glioma nuclear images. In: 2017 IEEE Winter Conference on Applications of Computer Vision (WACV). IEEE, pp. 834–841.

Nachbar, F., Stolz, W., Merkle, T., Cognetta, A.B., Vogt, T., Landthaler, M., Bilek, P., Braun-Falco, O., Plewig, G., 1994. The abcd rule of dermatoscopy: high prospective value in the diagnosis of doubtful melanocytic skin lesions. J. Am. Acad. Dermatol. 30 (4), 551–559.

Näppi, J.J., Hironaka, T., Regge, D., Yoshida, H., 2016. Deep transfer learning of virtual endoluminal views for the detection of polyps in ct colonography. In: Medical Imaging 2016: Computer-Aided Diagnosis, 9785. International Society for Optics and Photonics, p. 97852B.

Narotamo, H., Sanches, J.M., Silveira, M., 2020. Combining deep learning with hand-crafted features for cell nuclei segmentation. International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE, pp. 1428–1431.

Ni, Q., Sun, Z.Y., Qi, L., Chen, W., Yang, Y., Wang, L., Zhang, X., Yang, L., Fang, Y., Xing, Z., et al., 2020. A deep learning approach to characterize 2019 coronavirus disease (covid-19) pneumonia in chest ct images. Eur Radiol 30 (12), 6517–6527.

Oktay, O., Ferrante, E., et al., 2017. Anatomically constrained neural networks (acnns): application to cardiac image enhancement and segmentation. IEEE Trans. Med. Imaging 37 (2), 384–395.

Painchaud, N., Skandarani, Y., Judge, T., Bernard, O., Lalande, A., Jodoin, P.-M., 2019. Cardiac mri segmentation with strong anatomical guarantees. International Conference on Medical Image Computing and Computer Assisted Intervention. Springer, pp. 632–640.

Painchaud, N., Skandarani, Y., Judge, T., Bernard, O., Lalande, A., Jodoin, P.-M., 2020. Cardiac segmentation with strong anatomical guarantees. IEEE Trans. Med. Imaging 39 (11), 3703–3713.

Pan, S.J., Yang, Q., 2009. A survey on transfer learning. IEEE Trans. Knowl. Data Eng. 22 (10), 1345–1359.

Pham, D., Dovletov, G., Pauli, J., 2020. Liver segmentation in ct with mri data: zero-shot domain adaptation by contour extraction and shape priors. International Symposium on Biomedical Imaging (ISBI). IEEE, pp. 1538–1542.

Platania, R., Shams, S., Yang, S., Zhang, J., Lee, K., Park, S.-J., 2017. Automated breast cancer diagnosis using deep learning and region of interest detection (bc-droid). In: Proceedings of the 8th ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics. ACM, pp. 536–543.

Qi, Q., Lin, X., Chen, C., Xie, W., Huang, Y., Ding, X., Liu, X., Yu, Y., 2020. Curriculum feature alignment domain adaptation for epithelium-stroma classification in histopathological images. IEEE J. Biomed. Health Inform..

Rajpurkar, P., Irvin, J., Bagul, A., Ding, D., Duan, T., Mehta, H., Yang, B., Zhu, K., Laird, D., Ball, R.L., et al., 2017. Mura: Large dataset for abnormality detection in musculoskeletal radiographs. arXiv:1712.06957

Qin, C., Schlemper, J., Caballero, J., Price, A., Hajnal, J.V., Rueckert, D., 2018. Convolutional recurrent neural networks for dynamic mr image reconstruction. IEEE Trans. Med. Imaging 38 (1), 280–290.

Ravishankar, H., Venkataramani, R., Thiruvenkadam, S., Sudhakar, P., Vaidya, V., 2017. Learning and incorporating shape models for semantic segmentation. International Conference on Medical Image Computing and Computer Assisted Intervention. Springer, pp. 203–211.

Redmon, J., Divvala, S., Girshick, R., Farhadi, A., 2016. You only look once: Unified, real-time object detection. IEEE Conference on Computer Vision and Pattern Recognition. IEEE, pp. 779–788.

Ren, S., He, K., Girshick, R., Sun, J., 2015. Faster r-cnn: towards real-time object detection with region proposal networks. arXiv preprint arXiv:1506.01497.

Rezaei, S., Emami, A., Zarrabi, H., Rafiei, S., Najarian, K., Karimi, N., Samavi, S., Soroushmehr, S., 2019. Gland segmentation in histopathology images using deep networks and handcrafted features. International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE, pp. 1031–1034.

Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: Convolutional networks for biomedical image segmentation. International Conference on Medical Image Computing and Computer Assisted Intervention. Springer, pp. 234–241.

Roth, H.R., Farag, A., Lu, L., Turkbey, E.B., Summers, R.M., 2015. Deep convolutional networks for pancreas segmentation in ct imaging. In: Medical Imaging 2015: Image Processing, 9413. International Society for Optics and Photonics, p. 94131G.

Sa, R., Owens, W., Wiegand, R., Studin, M., Capoferri, D., Barooha, K., Greaux, A., Rattray, R., Hutton, A., Cintineo, J., et al., 2017. Intervertebral disc detection in x-ray images using faster r-cnn. In: International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE, pp. 564–567.

Saba, T., Mohamed, A.S., El-Affendi, M., Amin, J., Sharif, M., 2020. Brain tumor detection using fusion of hand crafted and deep learning features. Cogn. Syst. Res. 59, 221–230.

Samala, R.K., Chan, H.-P., Hadjiiski, L.M., Helvie, M.A., Cha, K.H., Richter, C.D., 2017. Multi-task transfer learning deep convolutional neural network: application to computer-aided diagnosis of breast cancer on mammograms. Physics in Medicine & Biology 62 (23), 8894.

Samala, R.K., Chan, H.-P., Hadjiiski, L., Helvie, M.A., Richter, C.D., Cha, K.H., 2018. Breast cancer diagnosis in digital breast tomosynthesis: effects of training sample size on multi-stage transfer learning using deep neural nets. IEEE Trans. Med. Imaging 38 (3), 686–696.

Schlegl, T., Seeböck, P., Waldstein, S.M., Schmidt-Erfurth, U., Langs, G., 2017. Unsupervised anomaly detection with generative adversarial networks to guide marker discovery. In: International Conference on Information Processing in Medical Imaging. Springer, pp. 146–157.

Schlemper, J., Caballero, J., Hajnal, J.V., Price, A.N., Rueckert, D., 2017. A deep cascade of convolutional neural networks for dynamic mr image reconstruction. IEEE Trans. Med. Imaging 37 (2), 491–503.

Setio, A.A.A., Ciompi, F., Litjens, G., Gerke, P., Jacobs, C., Van Riel, S.J., Wille, M.M.W., Naqibullah, M., Sánchez, C.I., van Ginneken, B., 2016. Pulmonary nodule detection in ct images: false positive reduction using multi-view convolutional networks. IEEE Trans. Med. Imaging 35 (5), 1160–1169.

Setio, A., Traverso, A., Bel, T., Berens, M., Bogaard, C., et al., 2017. Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: The LUNA16 challenge. Med. Image Anal. 1–13.

Sharma, G.N., Dave, R., Sanadya, J., Sharma, P., Sharma, K., 2010. Various types and management of breast cancer: an overview. Journal of advanced pharmaceutical technology & research 1 (2), 109.

Shen, D., Wu, G., Suk, H.-I., 2017. Deep learning in medical image analysis. Annu. Rev. Biomed. Eng. 19, 221–248.

Shin, S.Y., Lee, S., Yun, I.D., Kim, S.M., Lee, K.M., 2019. Joint weakly and semi-supervised deep learning for localization and classification of masses in breast ultrasound images. IEEE Trans. Med. Imaging 38 (3), 762–774. doi:10.1109/TMI.2018.2872031.

Shin, H.-C., Roth, H.R., Gao, M., Lu, L., Xu, Z., Nogues, I., Yao, J., Mollura, D., Summers, R.M., 2016. Deep convolutional neural networks for computer-aided detection: cnn architectures, dataset characteristics and transfer learning. IEEE Trans. Med. Imaging 35 (5), 1285–1298.

Simonyan, K., Zisserman, A., 2014. Very deep convolutional networks for large-scale image recognition. arXiv:1409.1556.

Song, Y., Zhu, L., Lei, B., Sheng, B., Dou, Q., Qin, J., Choi, K.-S., 2020. Shape mask generator: learning to refine shape priors for segmenting overlapping cervical cytoplasms. International Conference on Medical Image Computing and Computer Assisted Intervention. Springer, pp. 639–649.

Sonora-Mengan Sr, A., Gonidakis, P., Jansen, B., Garcia-Naranjo, J., Vandemeulebroucke, J., 2020. Evaluating several ways to combine handcrafted features-based system with a deep learning system using the luna16 challenge framework. In: Medical Imaging 2020: Computer-Aided Diagnosis, 11314. International Society for Optics and Photonics, p. 113143T.

Suzuki, K., 2017. Survey of deep learning applications to medical image analysis. Medical Imaging Technology 35 (4), 212–226.

Swati, Z.N.K., Zhao, Q., Kabir, M., Ali, F., Ali, Z., Ahmed, S., Lu, J., 2019. Content-based brain tumor retrieval for mr images using transfer learning. IEEE Access 7, 17809–17822.

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S.E., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A., 2015. Going deeper with convolutions. IEEE Conference on Computer Vision and Pattern Recognition 1–9.

Tajbakhsh, N., Shin, J.Y., Gurudu, S.R., Hurst, R.T., Kendall, C.B., Gotway, M.B., Liang, J., 2016. Convolutional neural networks for medical image analysis: full training or fine tuning? IEEE Trans. Med. Imaging 35 (5), 1299–1312.

Tan, J., Huo, Y., Liang, Z., Li, L., 2019. Expert knowledge-infused deep learning for automatic lung nodule detection. J. Xray Sci. Technol. (Preprint) 1–20.

Tang, Y., Wang, X., Harrison, A.P., Lu, L., Xiao, J., Summers, R.M., 2018. Attention-guided curriculum learning for weakly supervised classification and localization of thoracic diseases on chest radiographs. In: International Workshop on Machine Learning in Medical Imaging. Springer, pp. 249–258.

Tran, D., Bourdev, L., Fergus, R., Torresani, L., Paluri, M., 2015. Learning spatiotemporal features with 3d convolutional networks. In: Proceedings of the IEEE international conference on computer vision, pp. 4489–4497.

Tran, P.V., 2016. A fully convolutional neural network for cardiac segmentation in short-axis mri. arXiv:1604.00494

Tsai, Y.-H., Hung, W.-C., Schulter, S., Sohn, K., Yang, M.-H., Chandraker, M., 2018. Learning to adapt structured output space for semantic segmentation.

IEEE Conference on Computer Vision and Pattern Recognition. IEEE, pp. 7472–7481.

Tzeng, E., Hoffman, J., Saenko, K., Darrell, T., 2017. Adversarial discriminative domain adaptation. IEEE Conference on Computer Vision and Pattern Recognition. IEEE, pp. 7167–7176.

Valindria, V.V., Pawlowski, N., et al., 2018. Multi-modal learning from unpaired images: Application to multi-organ segmentation in ct and mri. In: 2018 IEEE Winter Conference on Applications of Computer Vision (WACV). IEEE, pp. 547–556.

Valverde, S., Salem, M., Cabezas, M., Pareto, D., Vilanova, J.C., Ramiotorrenta, L., Rovira, A., Salvi, J., Oliver, A., Llado, X., 2019. One-shot domain adaptation in multiple sclerosis lesion segmentation using convolutional neural networks. NeuroImage: Clinical 21, 101638.

Wang, W., Lu, Y., Wu, B., Chen, T., Chen, D.Z., Wu, J., 2018. Deep active self-paced learning for accurate pulmonary nodule segmentation. International Conference on Medical Image Computing and Computer Assisted Intervention. Springer, pp. 723–731.

Wang, X., Peng, Y., Lu, L., Lu, Z., Bagheri, M., Summers, R.M., 2017. Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. IEEE Conference on Computer Vision and Pattern Recognition. IEEE, pp. 2097–2106.

Wang, X., Peng, Y., Lu, L., Lu, Z., Summers, R.M., 2018. Tienet: text-image embedding network for common thorax disease classification and reporting in chest x-rays. IEEE Conference on Computer Vision and Pattern Recognition. IEEE, pp. 9049–9058.

Wang, Z., Zhang, J., Feng, J., Chen, Z., 2014. Knowledge graph and text jointly embedding. Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP) 1591–1601.

Wang, K., Zhang, X., Huang, S., Chen, F., Zhang, X., Huangfu, L., 2020. Learning to recognize thoracic disease in chest x-rays with knowledge-guided deep zoom neural networks. IEEE Access 8, 159790–159805.

Mitsuhara, M., Fukui, H., Sakashita, Y., Ogata, T., Hirakawa, T., Yamashita, T., Fujiyoshi, H., 2019. Embedding human knowledge in deep neural network via attention map. arXiv:1905.03540

Wang, X., Peng, Y., Lu, L., Lu, Z., Bagheri, M., Summers, R., 2017a. Chestx-ray14: hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 2097-2106.

Wei, J., Suriawinata, A., Ren, B., Liu, X., Lisovsky, M., Vaickus, L., Brown, C., Baker, M., Nasir-Moin, M., Tomita, N., et al., 2020. Learn like a pathologist: curriculum learning by annotator agreement for histopathology image classification. arXiv:2009.13698

Alom, M.Z., Hasan, M., Yakopcic, C., Taha, T.M., Asari, V.K., 2018. Recurrent residual convolutional neural network based on u-net (r2u-net) for medical image segmentation. arXiv:1802.06955

Guan, Q., Huang, Y., Zhong, Z., Zheng, Z., Zheng, L., Yang, Y., 2018. Diagnose like a radiologist: Attention guided convolutional neural network for thorax disease classification. arXiv:1801.09927

Weiner, M.W., Veitch, D.P., Aisen, P.S., Beckett, L.A., Cairns, N.J., Green, R.C., Harvey, D., Jack, C.R., Jagust, W., Morris, J.C., Petersen, R.C., Salazar, J., Saykin, A.J., Shaw, L.M., Toga, A.W., Trojanowski, J.Q., 2017. The alzheimer's disease neuroimaging initiative 3: continued innovation for clinical trial improvement. Alzheimer's and Dementia 13 (5), 561–571. doi:10.1016/j.jalz.2016.10.006.

Wimmer, G., Hegenbart, S., Vécsei, A., Uhl, A., 2016. Convolutional neural network architectures for the automated diagnosis of celiac disease. In: International Workshop on Computer-Assisted and Robotic Endoscopy. Springer, pp. 104–113.

Wistuba, M., Rawat, A., Pedapati, T., 2019. A survey on neural architecture search. arXiv:1905.01392

Wu, N., Phang, J., Park, J., Shen, Y., Huang, Z., Zorin, M., Jastrzebski, S., Févry, T., Katsnelson, J., Kim, E., et al., 2019. Deep neural networks improve radiologists performance in breast cancer screening. IEEE Trans. Med. Imaging 39 (4), 1184–1194.

Wu, L., Xin, Y., Li, S., Wang, T., Heng, P.-A., Ni, D., 2017. Cascaded fully convolutional networks for automatic prenatal ultrasound image segmentation. The IEEE International Symposium on Biomedical Imaging (ISBI). IEEE, pp. 663–666.

Xiao, X., Ji, C., Mudiyanselage, T.B., Pan, Y., 2020. Pk-gcn: Prior knowledge assisted image classification using graph convolution networks. arXiv:2009.11892

Wu, B., Zhou, Z., Wang, J., Wang, Y., 2018. Joint learning for pulmonary nodule segmentation, attributes and malignancy prediction. The IEEE International Symposium on Biomedical Imaging (ISBI). IEEE, pp. 1109–1113.

Xia, X., Gong, J., Hao, W., Yang, T., Lin, Y., Wang, S., Peng, W., 2020. Comparison and fusion of deep learning and radiomics features of ground-glass nodules to predict the invasiveness risk of stage-i lung adenocarcinomas in ct scan. Front. Oncol. 10.

Xie, X., Niu, J., Liu, X., Li, Q., Wang, Y., Han, J., Tang, S., 2020. Dg-cnn: introducing margin information into cnn for breast cancer diagnosis in ultrasound images. Journal of Computer Science and Engineering.

Xie, Y., Xia, Y., Zhang, J., Song, Y., Feng, D., Fulham, M.J., Cai, W., 2019. Knowledge-based collaborative deep learning for benign-malignant lung nodule classification on chest ct. IEEE Trans. Med. Imaging 38 (4), 991–1004.

Xie, Y., Zhang, J., Liu, S., Cai, W., Xia, Y., 2016. Lung Nodule Classification by Jointly Using Visual Descriptors and Deep Features. In: Medical Computer Vision and Bayesian and Graphical Models for Biomedical Imaging. Springer, pp. 116–125.

Xie, Y., Zhang, J., Xia, Y., Fulham, M., Zhang, Y., 2018. Fusing texture, shape and deep model-learned information at decision level for automated classification of lung nodules on chest ct. Information Fusion 42, 102–110.

Yan, W., Wang, Y., Gu, S., Huang, L., Yan, F., Xia, L., Tao, Q., 2019. The domain shift problem of medical image segmentation and vendor-adaptation by unet-gan. International Conference on Medical Image Computing and Computer Assisted Intervention. Springer, pp. 623–631.

Yan, K., Wang, X., Lu, L., Summers, R.M., 2018. Deeplesion: automated mining of large-scale lesion annotations and universal lesion detection with deep learning. J. Med. Imaging 5 (3), 036501.

Yang, Z., Cao, Z., Zhang, Y., Han, M., Xiao, J., Huang, L., Wu, S., Ma, J., Chang, P., 2020. Momminet: mammographic multi-view mass identification networks. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 200–210.

Yang, J., Dvornek, N.C., Zhang, F., Chapiro, J., Lin, M., Duncan, J.S., 2019. Unsupervised domain adaptation via disentangled representations: Application to cross-modality liver segmentation. International Conference on Medical Image Computing and Computer Assisted Intervention. Springer, pp. 255–263.

Yang, D., Xu, D., Zhou, S.K., Georgescu, B., Chen, M., Grbic, S., Metaxas, D., Comaniciu, D., 2017. Automatic liver segmentation using an adversarial image-to-image network. International Conference on Medical Image Computing and Computer Assisted Intervention. Springer, pp. 507–515.

Yang, G., Yu, S., Dong, H., Slabaugh, G., Dragotti, P.L., Ye, X., Liu, F., Arridge, S., Keegan, J., Guo, Y., et al., 2017. Dagan: deep de-aliasing generative adversarial networks for fast compressed sensing mri reconstruction. IEEE Trans. Med. Imaging 37 (6), 1310–1321.

Yang, X., Yu, L., Wu, L., Wang, Y., Ni, D., Qin, J., Heng, P.-A., 2017. Fine-grained recurrent neural networks for automatic prostate segmentation in ultrasound images. In: Thirty-First AAAI Conference on Artificial Intelligence.

Yang, W., Zhao, J., Qiang, Y., Yang, X., Dong, Y., Du, Q., Shi, G., Zia, M.B., 2019. Dscgans: integrate domain knowledge in training dual-path semi-supervised conditional generative adversarial networks and s3vm for ultrasonography thyroid nodules classification. International Conference on Medical Image Computing and Computer Assisted Intervention. Springer, pp. 558–566.

Yap, M.H., Pons, G., Marti, J., Ganau, S., Sentis, M., Zwiggelaar, R., Davison, A.K., Marti, R., 2018. Automated breast ultrasound lesions detection using convolutional neural networks. IEEE J. Biomed. Health Inform. 22 (4), 1218–1226.

Yedder, H.B., Shokoufi, M., Cardoen, B., Golnaraghi, F., Hamarneh, G., 2019. Limited-angle diffuse optical tomography image reconstruction using deep learning. Medical Image Computing and Computer Assisted Intervention Society. Springer, pp. 66–74.

Yosinski, J., Clune, J., Bengio, Y., Lipson, H., 2014. How transferable are features in deep neural networks? In: Advances in neural information processing systems, pp. 3320–3328.

Yu, F., Zhao, J., Gong, Y., Wang, Z., Li, Y., Yang, F., Dong, B., Li, Q., Zhang, L., 2019. Annotation-free cardiac vessel segmentation via knowledge transfer from retinal images. International Conference on Medical Image Computing and Computer Assisted Intervention. Springer, pp. 714–722.

Yu, S., Zhou, H.Y., Ma, K., Bian, C., Chu, C., Liu, H., Zheng, Y., 2020. Difficulty-aware glaucoma classification with multi-rater consensus modeling. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 741–750.

Yue, Q., Luo, X., Ye, Q., Xu, L., Zhuang, X., 2019. Cardiac segmentation from lge mri using deep neural network incorporating shape and spatial priors. International Conference on Medical Image Computing and Computer Assisted Intervention. Springer, pp. 559–567.

Zeng, G., Yang, X., Li, J., Yu, L., Heng, P.-A., Zheng, G., 2017. 3d u-net with multi-level deep supervision: fully automatic segmentation of proximal femur in 3d mr images. In: International workshop on machine learning in medical imaging. Springer, pp. 274–282.

Zhang, Z., Chen, P., Sapkota, M., Yang, L., 2017. Tandemnet: distilling knowledge from medical images using diagnostic reports as optional semantic references. Medical Image Computing and Computer Assisted Intervention. Springer, pp. 320–328.

Zhang, Z., Saha, A., Zhu, Z., Mazurowski, M.A., 2018. Hierarchical convolutional neural networks for segmentation of breast tumors in mri with application to radiogenomics. IEEE Trans. Med. Imaging 38 (2), 435–447.

Zhang, B., Wang, Z., Gao, J., Rutjes, C., Nufer, K., Tao, D., Feng, D.D., Menzies, S.W., 2020. Short-term lesion change detection for melanoma screening with novel siamese neural network.. IEEE Trans. Med. Imaging.

Zhang, J., Cain, E.H., Saha, A., Zhu, Z., Mazurowski, M.A., 2018. Breast mass detection in mammography and tomosynthesis via fully convolutional network-based heatmap regression. In: Medical Imaging 2018: Computer-Aided Diagnosis, 10575. International Society for Optics and Photonics, p. 1057525.

Zhang, J., Wang, G., Xie, H., Zhang, S., Huang, N., Zhang, S., Gu, L., 2020. Weakly supervised vessel segmentation in x-ray angiograms by self-paced learning from noisy labels with suggestive annotation. Neurocomputing 417, 114–127.

Zhang, Y., Wei, Y., Wu, Q., Zhao, P., Niu, S., Huang, J., Tan, M., 2020. Collaborative unsupervised domain adaptation for medical image diagnosis. IEEE Trans. Image Process. 29, 7834–7844.

Zhang, Z., Yang, L., Zheng, Y., 2018. Translating and segmenting multimodal medical volumes with cycle-and shape-consistency generative adversarial network. IEEE Conference on Computer Vision and Pattern Recognition. IEEE, pp. 9242–9251.

Zhang, R., Zheng, Y., Mak, T.W.C., Yu, R., Wong, S.H., Lau, J.Y., Poon, C.C., 2016. Automatic detection and classification of colorectal polyps by transferring low-level cnn features from nonmedical domain. IEEE J. Biomed. Health Inform. 21 (1), 41–47.

Zhang, Y., Wang, X., Xu, Z., Yu, Q., Yuille, A., Xu, D., 2020c. When radiology report generation meets knowledge graph. arXiv:2002.08277

Zhao, J., Li, D., Kassam, Z., Howey, J., Chong, J., Chen, B., Li, S., 2020. Tripartite-gan: synthesizing liver contrast-enhanced mri to improve tumor detection. Med. Image Anal. 101667.

Zhao, R., Chen, X., Chen, Z., Li, S., 2020. Egdcl: An adaptive curriculum learning framework for unbiased glaucoma diagnosis. In: European Conference on Computer Vision. Springer, pp. 190–205.

Zhao, M., Wang, L., Chen, J., Nie, D., Cong, Y., Ahmad, S., Ho, A., Yuan, P., Fung, S.H., Deng, H.H., et al., 2018. Craniomaxillofacial bony structures segmentation from mri with deep-supervision adversarial learning. International Conference on Medical Image Computing and Computer Assisted Intervention. Springer, pp. 720–727.

Zhao, Z., Zhang, X., Chen, C., Li, W., Peng, S., Wang, J., Yang, X., Zhang, L., Zeng, Z., 2019. Semi-supervised self-taught deep learning for finger bones segmentation. In: 2019 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI). IEEE, pp. 1–4.

Zhu, J.-Y., Park, T., Isola, P., Efros, A.A., 2017. Unpaired image-to-image translation using cycle-consistent adversarial networks. In: ICCV 2017, pp. 2223–2232.

Zheng, H., Lin, L., Hu, H., Zhang, Q., Chen, Q., Iwamoto, Y., Han, X., Chen, Y., Tong, R., Wu, J., 2019. Semi-supervised segmentation of liver using adversarial learning with deep atlas prior. Medical Image Computing and Computer Assisted Intervention. Springer, pp. 148–156.

Zhou, Z., Shin, J.Y., Zhang, L., Gurudu, S.R., Gotway, M.B., Liang, J., 2017. Fine-tuning convolutional neural networks for biomedical image analysis: actively and incrementally. IEEE Conference on Computer Vision and Pattern Recognition 4761–4772.

Zhou, Z., Siddiquee, M.M.R., Tajbakhsh, N., Liang, J., 2018. Unet++: a nested u-net architecture for medical image segmentation. In: Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. Springer, pp. 3–11.

Zhu, Z., Jin, D., Yan, K., Ho, T.-Y., Ye, X., Guo, D., Chao, C.-H., Xiao, J., Yuille, A., Lu, L., 2020. Lymph node gross tumor volume detection and segmentation via distance-based gating using 3d ct/pet imaging in radiotherapy. International Conference on Medical Image Computing and Computer Assisted Intervention. Springer, pp. 753–762.

Zhu, W., Liu, C., Fan, W., Xie, X., 2018. Deeplung: deep 3d dual path nets for automated pulmonary nodule detection and classification. workshop on applications of computer vision 673–681.

Zotti, C., Luo, Z., Lalande, A., Jodoin, P., 2019. Convolutional neural network with shape prior applied to cardiac mri segmentation. IEEE J. Biomed. Health Inform. 23 (3), 1119–1128.