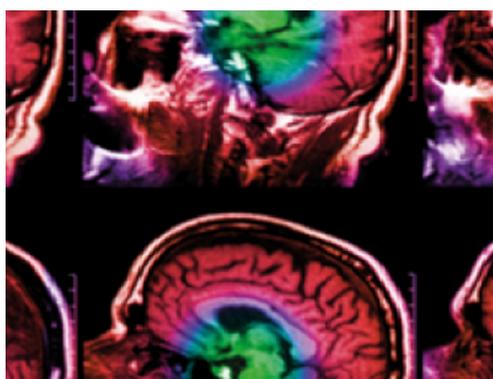


TOPICAL REVIEW • **OPEN ACCESS**

Anatomy-aided deep learning for medical image segmentation: a review

To cite this article: Lu Liu *et al* 2021 *Phys. Med. Biol.* **66** 11TR01

View the [article online](#) for updates and enhancements.



IPEM | IOP

Series in Physics and Engineering in Medicine and Biology

Your publishing choice in medical physics,
biomedical engineering and related subjects.

Start exploring the collection—download the
first chapter of every title for free.



TOPICAL REVIEW

OPEN ACCESS

RECEIVED
19 January 2021REVISED
19 April 2021ACCEPTED FOR PUBLICATION
27 April 2021PUBLISHED
26 May 2021

Original content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](#).

Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.



Anatomy-aided deep learning for medical image segmentation: a review

Lu Liu^{1,2} , Jelmer M Wolterink¹ , Christoph Brune¹  and Raymond N J Veldhuis² ¹ Applied Analysis, Department of Applied Mathematics, Faculty of Electrical Engineering, Mathematics and Computer Science, University of Twente, Drienerlolaan 5, 7522 NB, Enschede, The Netherlands² Data Management and Biometrics, Department of Computer Science, Faculty of Electrical Engineering, Mathematics and Computer Science, University of Twente, Drienerlolaan 5, 7522 NB, Enschede, The NetherlandsE-mail: l.liu-2@utwente.nl**Keywords:** medical image segmentation, anatomical information, deep learning

Abstract

Deep learning (DL) has become widely used for medical image segmentation in recent years. However, despite these advances, there are still problems for which DL-based segmentation fails. Recently, some DL approaches had a breakthrough by using anatomical information which is the crucial cue for manual segmentation. In this paper, we provide a review of anatomy-aided DL for medical image segmentation which covers systematically summarized anatomical information categories and corresponding representation methods. We address known and potentially solvable challenges in anatomy-aided DL and present a categorized methodology overview on using anatomical information with DL from over 70 papers. Finally, we discuss the strengths and limitations of the current anatomy-aided DL approaches and suggest potential future work.

1. Introduction

Generally, segmentation is defined as an operation that separates images into several parts with different meanings. In the medical field, segmentation as a method of medical imaging analysis is crucial for diagnosis and treatment. As soon as computers could load medical images as digital files, multiple pieces of research have explored systems for automated medical image segmentation. For medical images, segmentation is usually a binary problem since only the parts of interest are important. It can also be multi-target segmentation if the number of targets of interest is more than one.

Initially, segmentation was done with mathematical models (e.g. Fuzzy C-mean clustering, K-means clustering) and low-level pixel processing (e.g. region growing, edge detection methods, watershed) (Lee *et al* 2015). At the end of the 20th century, machine learning (ML) and pattern recognition techniques were applied to segmentation using training data to develop a model. These supervised techniques are still very popular and many commercial medical image analysis applications are based on them. The extracted features are designed or selected by researchers in these approaches and referred to as handcrafted features. Although they are usually human-understandable, they may not be ideal features for segmentation. Recently, with the development of neural networks and deep learning (DL), computers can extract representative features from images. Many deep neural networks for image processing are designed based on the concept that networks of many layers transform input images to output labels by learning high-level features. The most representative type of model for image analysis to date is the convolutional neural network (CNN). The CNN as an important model in has solved many key commercial applications and showed its ability in many contests. In medical image analysis, the CNN and other DL methods started to show their ability at many workshops, challenges, and conferences.

As the number of papers increases rapidly, a few review papers are trying to summarize these applications. There are general review papers on DL in medical image analysis published by Shen *et al* (2017) and Litjens *et al* (2017). Review papers are focusing on medical image segmentation using DL published recently by Taghanaki *et al* (2020), Haque and Neubert (2020), and Hesamian *et al* (2019). Some review papers have more specific

focuses. For example, the work by Zhuang *et al* (2019) and Chen *et al* (2020a) summarized cardiac image segmentation networks; Yi *et al* (2019) discussed the use of generative adversarial networks (GANs) in medical imaging; Zhang *et al* (2019) focused on the small sample problem in biomedical image analysis; Karimi *et al* (2020) explored the problem of noisy labels in medical image analysis; Tajbakhsh *et al* (2020) investigated DL solutions for imperfect datasets in medical image segmentation; Cheplygina *et al* (2019) surveyed not-so-supervised networks using semi-supervised, multi-instance and transfer learning in medical image analysis. Jurdia *et al* (2020) categorized high-level prior-based loss functions for medical image segmentation according to the nature of the prior: shape, size, topology, and the inter-region constraints. Bohlender *et al* (2021) reviewed shape-constrained DL for medical image segmentation. Some review papers focus on specific segmentation tasks in certain modalities. Yedavalli *et al* (2020) reviewed artificial intelligence in stroke imaging. Zhang *et al* (2020a) surveyed DL methods for isointense infant brain segmentation in magnetic resonance imaging (MRI), while Ma *et al* (2020) focused on liver tumor segmentation in computed tomography (CT) images. Vrtovec *et al* (2020) explored the segmentation of organs at risk for head and neck radiotherapy planning. Ebrahimkhani *et al* (2020) summarized segmentation methods of knee articular cartilage. As brain tumor segmentation is an active topic, several review papers on this topic are published by Sun *et al* (2019), Hameurlaine and Moussaoui (2019), Saman and Narayanan (2019), Jiang *et al* (2020), and Li *et al* (2020a). Feo and Giove (2019) reviewed segmentation methods of small rodents brains, and Lin and Li (2019) reviewed brain segmentation methods from multi-atlas to DL. Chen *et al* (2020b) summarized thyroid gland segmentation and thyroid nodule segmentation methods for ultrasound images. Two recent reviews of retinal blood vessel segmentation are published by Soomro *et al* (2019) and Samuel and Veeramalai (2020). From these review papers, we see that there are many possibilities in this field and various review focuses indicates various research directions. However, none of them focus on the use of anatomical information which is the main cue of segmentation for experts. In some segmentation tasks, anatomical information is critical. For example, when experts segmenting epicardial fat in CT images, as epicardial fat and paracardial fat have a very similar appearance, the pericardium is the only thing separating the two fat tissues. But pericardium is rarely visible in CT images. Experts need anatomy knowledge of the heart to estimate the pericardium and epicardial fat as well.

In the very beginning, only experts who are experienced and have anatomical knowledge can do the segmentation manually for medical images. Thus, ideally, the segmentation networks should mimic the experts when predicting the labels for medical images. Considering that most of the segmentation networks are typically trained with pixel-wise or voxel-wise loss functions (e.g. cross-entropy, dice losses), it may limit the ability to learn features that is representative anatomically or structurally. Some of the above review papers mentioned the use of anatomical information with examples and pointed out that applying anatomical constraints to networks shows improved performance and robustness. However, they rarely analyze or summarize the use of anatomical information in DL for medical image segmentation in depth. This review paper is made to fill this gap.

Figure 1 illustrates the overview of this paper. In this paper, we provide an overview of state-of-the-art anatomy-aided DL techniques for medical image segmentation by introducing various anatomical information types (section 2), summarizing challenges in this field (section 3), summarizing and analyzing the methodology of using anatomical information in DL (section 4) with its weaknesses, strengths, and uncertainty. At the end (section 5), we discuss the strengths and weaknesses of anatomy-aided DL for medical image segmentation and potential directions for future work.

2. Anatomical information

In this section, we introduce anatomical information of four categories: shape, appearance, motion, and context. In general, human anatomy studies the morphology of the human body. Anatomy knowledge is concluded after long-time observation of the human body or part of the human body. As the object of study and observation way varies, the two main branches of human anatomy are gross anatomy and microscopic anatomy. Gross anatomy studies anatomical structures that are obvious to the naked human eye such as body parts and organs. Nowadays, many noninvasive techniques like MRI, CT, ultrasound (US), or x-ray are used to image inside the living body. After the microscope was invented, people study minute anatomical structures such as tissues and cells with the assistance of it. This is the so-called microscopic anatomy.

For segmentation tasks in medical images, plenty of anatomical information is available. However, when applying, we need to describe or model the information properly. For different objects, different information is considered informative or useful to distinguish the target from other structures. Thus, not all anatomical information can contribute. In the following subsections, we list the anatomical information that has been used in medical image segmentation networks and introduce their ways of description.

As shown in figure 2, anatomical information can be divided into four categories and the sub-categories are listed as well.

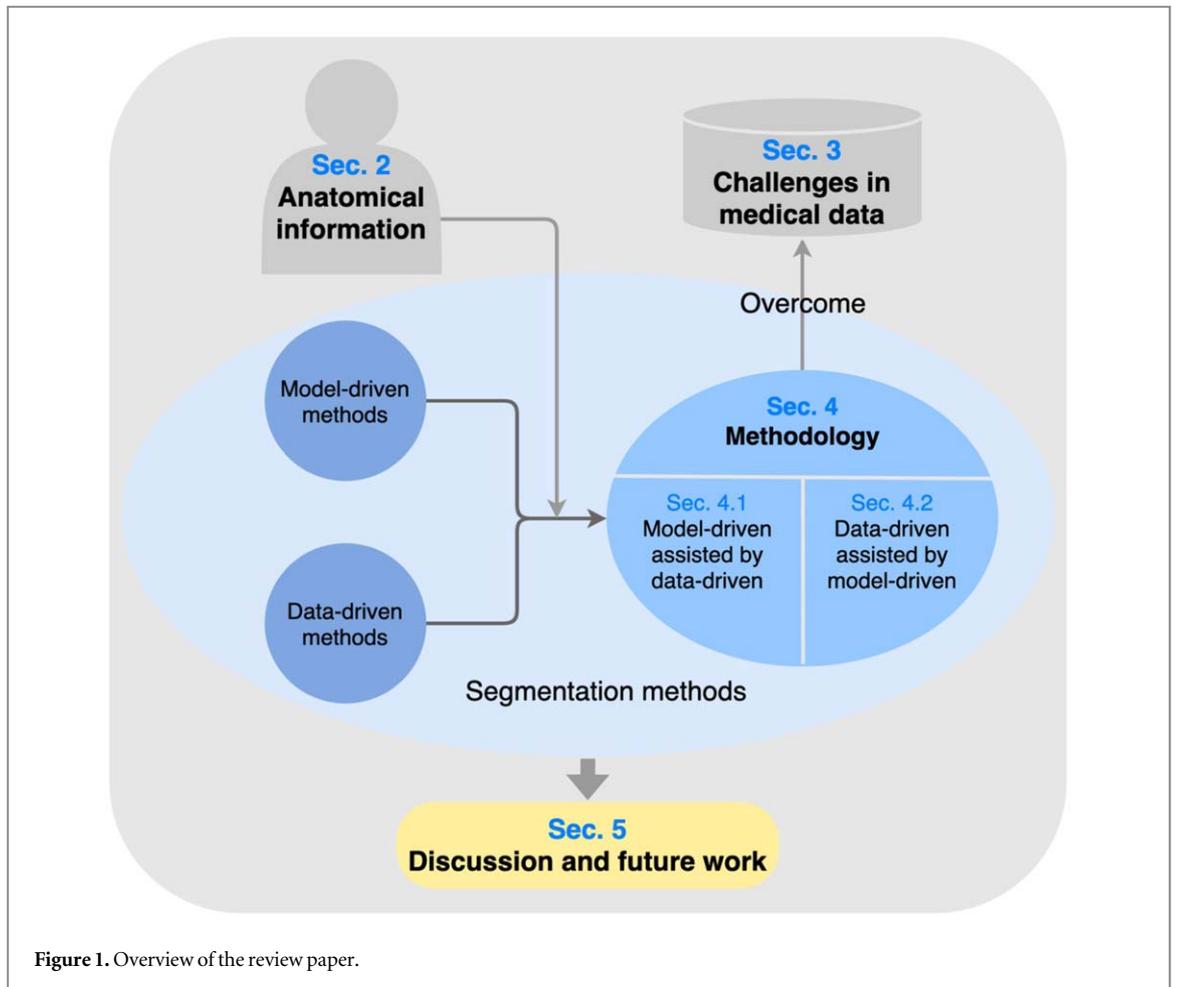


Figure 1. Overview of the review paper.

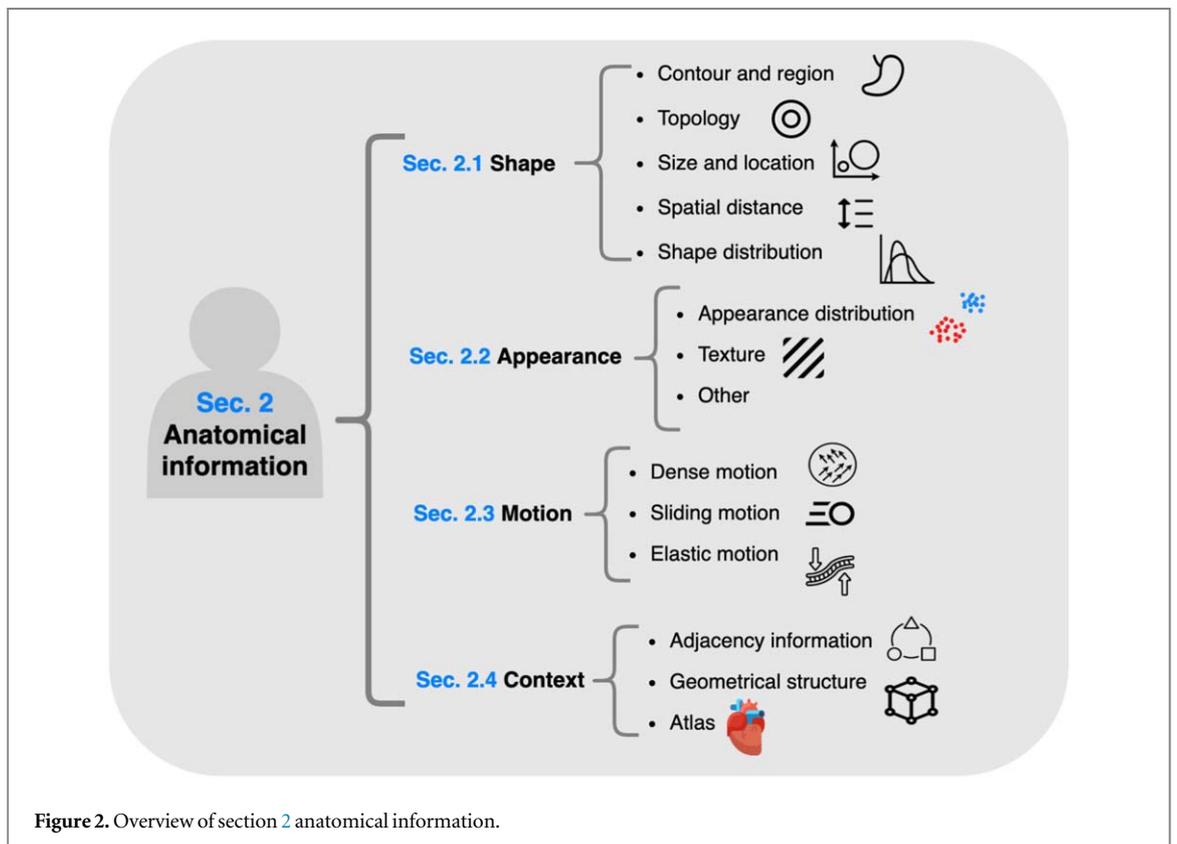


Figure 2. Overview of section 2 anatomical information.

2.1. Shape information

Shape information is a crucial descriptor for target objects in a medical image. For many segmentation targets, the shape is a basic element for experts to distinguish the target from other structures. In this section, we discuss shape information from five aspects: contour and region (section 2.1.1), topology (section 2.1.2), size and location (section 2.1.3), spatial distance (section 2.1.4), and shape distribution (section 2.1.5).

2.1.1. Contour and region

There are mainly two ways to model the shape geometrically: parametric way and non-parametric way. Most relatively regular shapes can be modeled parametrically. For example, 2D shapes like ellipse, circle, rectangular, etc and 3D shapes like sphere, cylinder, cube, etc can be described with parameters like center coordinate, radius, height, width, orientation, etc.

For not-so-regular shapes, a non-parametric way like a level set (LS) representation may be a better description. The idea of an LS representation is to represent contour C by a function φ . The boundary C of an object Ω is defined as the zero set of φ . i.e.

$$C = \{x \in \Omega: \varphi(x) = 0\}, \quad (1)$$

where Ω denotes the entire image plane. The sign of $\varphi(x)$ determines whether x is in $\tilde{\Omega}$ or outside $\tilde{\Omega}$,

$$\text{sign}(\varphi(x)) = \begin{cases} +1 & \text{if } x \text{ is in } \tilde{\Omega} \\ 0 & \text{if } x \text{ is on the boundary.} \\ -1 & \text{if } x \text{ is in } \tilde{\Omega}^c \end{cases} \quad (2)$$

The LS representation is widely used in optimization-based segmentation methods, and it is particularly effective for convex objects like optic cum and disc segmentation (Nosrati and Hamarneh 2016). With LS representation, the active contour model (ACM) is widely used for image segmentation.

ACM (or deformable model) is one of the conventional segmentation methods (McInerney and Terzopoulos 1996, Jayadevappa *et al* 2011, Chen *et al* 2020a). Compared to basic models such as simple thresholding, region-growing, or edge detection, ACMs have shown better performance and robustness. Representative models are snakes (Kass *et al* 1988), the Mumford–Shah model (Mumford and Shah 1989), and the active contour without edges (ACWE) (Chan and Vese 2001). Subsequently, many extensions and variations of ACM tried to solve the problem efficiently. Examples of well-known solvers are dual projection and graph cut (Morar *et al* 2012). These models formulate segmentation as an energy minimization problem. Variational methods and partial differential equations (PDEs) are used to solve them.

Most ACMs start with an initial guess boundary which is one or multiple closed contours. During the process of energy minimization, the contour is modified to more accurate and closer to the desired boundary under the constraints and a penalty given in the energy function. The common penalty terms control the smoothness of the contour, and the curvature of the contour, etc. Based on the information used in the models, there are two categories of ACMs—edge-based models and region-based models (Le *et al* 2020). Edge-based models use image gradient as edge information to constrain the contour to the target boundary. A well-known model in this category is the geodesic ACM (Caselles *et al* 1997). The ACWE model (Chan and Vese 2001) is an example of region-based models. This model utilizes an LS representation for the contour of target objects. Similar to most region-based models, its energy function consists of two parts: regularization and energy minimization. It uses the statistical information inside and outside the boundary to guide the modification. Since an image gradient is not involved, comparing to edge-based models, region-based models are more robust against noise and better at detecting weak boundaries.

Nowadays, the group of ACM methods is still growing. Many recent models have better performance than years before and can handle many image segmentation problems. However, they have many limitations. Generally, the segmentation results are obtained by minimizing a certain energy function using gradient descent (Zhou *et al* 2013). Since most of the ACMs are not convex, the segmentation results may get stuck in local minima. They are unstable when dealing with occluded images. There are many parameters chosen empirically. The segmentation results rely on the parameters, number of iterations, and image quality. They may generate unpredictable or wrong results when handling complex images. Especially, the accuracy decreases dramatically with images in the wild.

2.1.2. Topology

Many anatomical objects in medical images have fixed topological characteristics that are supposed to be maintained in the segmentation results. For example, when segmenting the airway wall in transverse CT slices, the airway wall has a doughnut shape which needs to be preserved in the segmentation results. Topology studies the properties of geometric objects that are invariant during continuous deformations in topological spaces. The two main topological properties are connectivity and compactness. Connectivity describes whether an object is

connected (e.g. one circle is connected, while two non-intersecting circles are not connected). Compactness describes whether an object is closed and bounded (e.g. a circle is compact, while a line is not). There are many tools from topological data analysis (TDA) to describe the topological characteristics. TDA relates topology and geometry to extract information from datasets of high-dimension. The main tool of TDA is persistent homology (PH). The review papers Wasserman (2018) and Chazal and Michel (2017) include fundamental and practical aspects about TDA.

2.1.3. Size and location

Size and location are the most basic information of an object. In many cases, size and location can be used as a constraint to filter or remove redundant and unrelated objects. Moreover, some segmentation methods using shape priors may over-correct the segmentation results to make them fit the input shape prior. As an example, when there are pathological cases in the input image, an abnormal part that deviates from the shape before, may lead to healthy cases not being segmented correctly. Using size and location constraints may be an alternative to reduce undesirable results (Nosrati and Hamarneh 2016).

The parameter for describing size varies as image modality or target changes. It could be the length, width, height, area, and volume, etc. Similarly, there are many parameters for describing location such as coordinates, and the centroid. In the case of having rough information about the size and location, soft constraints such as a size range or a location range can be applied.

2.1.4. Spatial distance

There mainly two types of spatial distances that are relatively widely incorporated in segmentation methods: minimum distance and maximum distance (Nosrati and Hamarneh 2016). The minimum distance between two objects can be used as a constraint to enforce the separation of regions or objects. The maximum distance between regions or boundaries is known in many cases. For example, in cardiac CT, the maximum distance between the left ventricle and its myocardium can be estimated. Other types of spatial distance are derived from minimum and maximum distance. In some cases, the distance between objects is supposed to be controlled in a specific range. The idea of attractive force and repulsive force in physics could be used to model the spatial relationships (Zeng *et al* 1998). More models can be used to control the distance between objects such as a deformable model, etc. Since recently there is little work based on them for medical image segmentation, it is not described in this article.

2.1.5. Shape distribution

In practice, target objects in medical image segmentation hardly have regular shapes. Most objects from different sample images are not identical or rigid. Even relatively regular objects (in medical image segmentation) like organs have various shapes from one to another. Thus, a fixed geometrical model may not be appropriate for such objects. One way to handle such intra-class variation is to form a shape probability model by adding a probability distribution to the model.

There are two parts in most of the shape probability models: shape representations and probability distributions. For shape representations, there are many choices like the LS, point cloud, surface mesh, etc. For probability distributions, common models are Gaussian distribution, Gaussian mixture model, etc.

2.2. Appearance information

Appearance is one of the most important and obvious visual information to distinguish various objects and structures in medical images. Appearance is influenced by many factors such as intensity, color, brightness, texture, and saturation. There are many ways to formulate appearance models for image segmentation. Here we introduce appearance distributions, texture models, and other common ideas of extracting appearance information.

2.2.1. Appearance distribution

The appearance distribution is usually learned or estimated by observing the distribution of appearance features in small samples. Assuming that $F_i(x)$ represents a set of appearance features of the object i and the probability $P(x|F_i(x))$ of every pixel or voxel to each class is known, this is the appearance distribution. Examples of the most direct appearance features are a gray-scale value, an RGB value, or other values of every pixel or voxel. To use the distribution in segmentation networks, for example, we can force the segmentation distribution to fit the prior distribution by minimizing the distance between them.

2.2.2. Texture

Texture in medical images is one of the direct visual cues to distinguish many objects such as tissues and lesions. Many models are used to represent the texture of objects. Most of them are used in ML methods to represent texture features, while some inspired research using DL for image segmentation.

The recent review on texture feature extraction methods (Humeau-Heurtier 2019) classified methods into seven classes: statistical approaches, structure approaches, transform-based approaches, model-based approaches, graph-based approaches, learning-based approaches, and entropy-based approaches. Many of the texture features mentioned in the review have been used in medical image segmentation. Here we give several examples. There are simple texture models. For example, the model proposed by Bigün *et al* (1991) utilizes the Jacobian matrix and a Gaussian kernel to generate a three-channel texture feature. Some advanced texture features are widely used in image segmentation before DL became the main trend. The group of texture features based on Haar and Gabor filters has shown effectiveness in medical image segmentation (Santner *et al* 2009, Yang *et al* 2014, Ibragimov *et al* 2017).

2.2.3. Other

Many other appearance features are extracted for image segmentation. For instance, the Fourier transformation, a bag of visual words, the local binary pattern (LBP), the histogram of oriented gradient, the scale-invariant feature transform, etc are used to extract appearance features (Nosrati and Hamarneh 2016). Appearance features are mainly extracted from three domains: the spatial domain, the time domain, and the scale domain (Nosrati and Hamarneh 2016). For different targets, different features can be selected or designed to reach better segmentation performances. These appearance features that are designed manually are considered hand-crafted features. Nowadays, with DL, segmentation networks can learn appearance features automatically.

2.3. Motion information

Life is in motion. There are three types of motion in our body for image analysis: the dense motion, the sliding motion, and the elastic motion. A typical example of dense motion is the particle movement in the fluid. Applied to medical images, it could be for example cells moving in blood or other fluid. The standard representation of dense motion in computer vision is the optical flow (Szeliski 2010). Another basic motion type is the sliding motion. Usually, physical models with velocity and locations are used to describe such motion. Elastic motion is the deformation of objects caused by force. There are many physical models available for describing various motion types. However, as the human body is complex, it is not easy to capture and utilize proper motion information for assisting medical image segmentation.

Some targets in medical image segmentation move regularly (e.g. heart, lung) or irregularly (e.g. fetus). For many cardiac, chest, and thoracic images, the imaging technique electrocardiogram (ECG)-gating (Desjardins and Kazerooni 2004) is applied to CT and MRI to solve the problem of heart motion throughout the cardiac cycle. With ECG-gating, the stop motion image is taken during the time slot of the cardiac cycle when the heart is not moving. Apart from the ECG-gating technique in cardiac imaging, there are other examples of using the motion before representing motions with physical models. Some target objects (mainly tissues) in medical images have special physical characteristics so that they can be modeled as additional prior information. Some research tried to use vibrational spatial deformations, elastic shape models, etc with other models like statistic models for image segmentation (Nosrati and Hamarneh 2016).

The acquisition of motion information is difficult in many cases. For regular motions like the heart motion, we have ECG to get the motion prior, while for irregular motion like the fetus motion, the motion information needs to be captured using other tools. Thus, in this kind of application, there is no general approach to capture or utilize the motion information. However, it is possible to apply irregular motion before segmentation methods. For example, Nosrati *et al* (2014) introduced a multi-organ segmentation method in multi-view endoscopic videos with priors captured pre-operatively.

2.4. Context information

In many cases, not only the information of the target object is valuable for segmentation but also the relationships between the target objects and the context. Below, we discuss the simple adjacency relationships and a more complex geometrical structure or atlas for segmentation.

2.4.1. Adjacency information

As anatomy is to study the structure of the human body, in many cases the relationship between one object and its adjacent structures is known. For example, the location of organs in normal cases is fixed. There are three ways to represent the adjacent information: labels, distances, and models. In section 2.4.2, we introduce models in detail. Thus, in this section, we only discuss labels and distances.

Table 1. Challenges list for table 2.

Extrinsic variability	1	Spatial consistency reduces
	2	Diverse equipment protocols and field inhomogeneity
	3	Partial volume effect
	4	Contrast around the border varies
	5	Missing boundaries
	6	Low SNR, speckles, shadows
	7	Low soft tissue contrast
Intrinsic variability	8	No shape prior
	9	Heterogeneous appearance
	10	Multiple positions
Spatial complexity	11	High complexity
	12	Locally obscured
	13	Artifacts
Motion	14	Non-rigid movements
	15	Floating spatial relationships
	16	Motion blurring
Other	17	Relatively small in terms of volumes or area
	18	Adjacent tissue with similar intensity

The idea of using labels is to describe ordering constraints and adjacency relations for semantic segmentation. For example, ‘cat’ and ‘rat’ are less likely to be close to each other. Thus, the conversion between ‘cat’ and ‘rat’ is supposed to be constrained in some way (Nosrati and Hamarneh 2016). This can be applied to the multi-object segmentation context. Distances in this context are the 2D or 3D distance between two objects. As the adjacency relationships are known, the distance between two objects can be controlled or constrained according to the prior.

2.4.2. Geometrical structure and atlas

Geometrical structure and the atlas consist of anatomical information such as shape information, adjacency information, size, location, spatial relationships, etc. It has shown success in many medical image analysis applications. To describe and use a geometrical structure, one way is to formulate it in geometrical models, which is mentioned in section 2.1.1. Another way is the graph neural networks (GNNs) (Scarselli *et al* 2008). More information about GNN can be found in section 4.2.6. The segmentation approaches using multiple atlases are named multi-atlas segmentation (MAS). Before DL started being popular for medical image segmentation, atlas-based approaches are widely used in biomedical image segmentation, especially for heart segmentation (Iglesias and Sabuncu 2015, Chen *et al* 2020a). MAS considers the entire labeled training dataset as a set of the atlas which is different from some average models. In this way, the anatomical variation is preserved. When applying MAS, there are many challenges such as the selection of the proper atlas, image registration, label fusion, and high computational cost (Yang *et al* 2016, Ding *et al* 2020).

3. Challenges in medical data

Medical image segmentation is a challenging task due to many challenges in the data. In table 1, we listed and categorized the common challenges for medical image segmentation. In table 2, we summarized some common challenges in data for various targets in US, CT, and MRI. The numbers in this table indicate the index of challenges listed in table 1. The common challenges for all DL applications, like high computation cost and lack of interpretability, are not discussed here. In this section, the challenges are summarized in eight categories: extrinsic variability, intrinsic variability, spatial complexity, moving or deforming targets, extremely small targets, and similar adjacent structures. Apart from the challenges mentioned above, the data limitation of medical images is an important challenge. Though the data and label collection is difficult in many other segmentation tasks, medical image and label collection is more time-consuming and labor-intensive. As this challenge is common for all targets and modalities, it is not listed in the tables below.

3.1. Extrinsic variability

Extrinsic variability indicates the challenges caused by outside uncertainty or the physics of the imaging modality. Challenges in this category are modality-related. Outside uncertainty includes the diversity of

Table 2. Common challenges for various targets on US, CT, and MRI.

Targets	US	CT	MRI
All targets	1, 6, 7		2, 3, 7
Organs	4, 5, 11, 13, 14, 16, 18		11, 13, 16, 18
Epithelial tissue	—		8, 9, 11, 18
Muscle tissue	—		18
Nervous tissue	—	—	8, 9
Connective tissue	—		8, 9, 11, 18
Tumor		8, 9, 10, 11, 18	
Nodule	18		8, 9, 10, 11, 17
Vessels	—		5, 9, 17
Bone	18		8, 11, 13, 17
Joints	—		11, 13
Fetus	8, 9, 11, 14, 15	—	8, 9, 11, 15, 16

Note. ‘—’ indicates that there is no or not enough research or not applied.

equipment protocols, motion and field inhomogeneity, noises, etc (Milletari *et al* 2016, Yang *et al* 2017, Šprem *et al* 2018, Chen *et al* 2020a) Images like CT and MRI are complained about for being acquired from very different imaging equipment. People rarely have the same position or body shape when taking images. In histopathology images like hematoxylin and eosin (H&E) stained images, the color, brightness, and saturation are hard to be unified. This kind of diversity may present non-negligible differences among images. In the meantime, noise is inevitable, and it may lead to a low signal-noise ratio, speckles, and shadows. No imaging modality is perfect. The physics of imaging modalities determines their defects and sometimes there are unexpected artifacts. For example, due to the physics of ultrasound, the spatial consistency reduces along with the directions which are orthogonal to the acoustic beam in ultrasound images (Szabo 2004, Yang *et al* 2017). This may cause difficulties to all segmentation tasks in ultrasound. Some segmentation targets like organs are supposed to have clear boundaries. But the variability of contrast on boundaries or even missing boundaries could be challenges for segmentation. In CT and MRI, the partial volume effect may lead to too simplistic borders of objects (Šprem *et al* 2018). And the low soft tissue contrast problem is complained about in almost all imaging modalities (Chen *et al* 2020a).

3.2. Intrinsic variability

Intrinsic variability indicates the challenges caused by the diversity of the targets. Three challenges in this category are listed in table 1: no shape prior, heterogeneous appearance, and multiple positions. Segmentation of tumors in CT or MRI is an example. There are targets like organs that have a relatively certain closed shape and size, while there are targets like tumors which could be of many shapes, sizes and positions. Thus, for these targets, no shape prior can be used to assist segmentation. As the tumor is a region that suffered from damages, it could have various appearances, fuzzy boundaries, and heterogeneous densities (Li *et al* 2015). It may be known inside a specific organ, but the precise position is unknown.

3.3. Spatial complexity

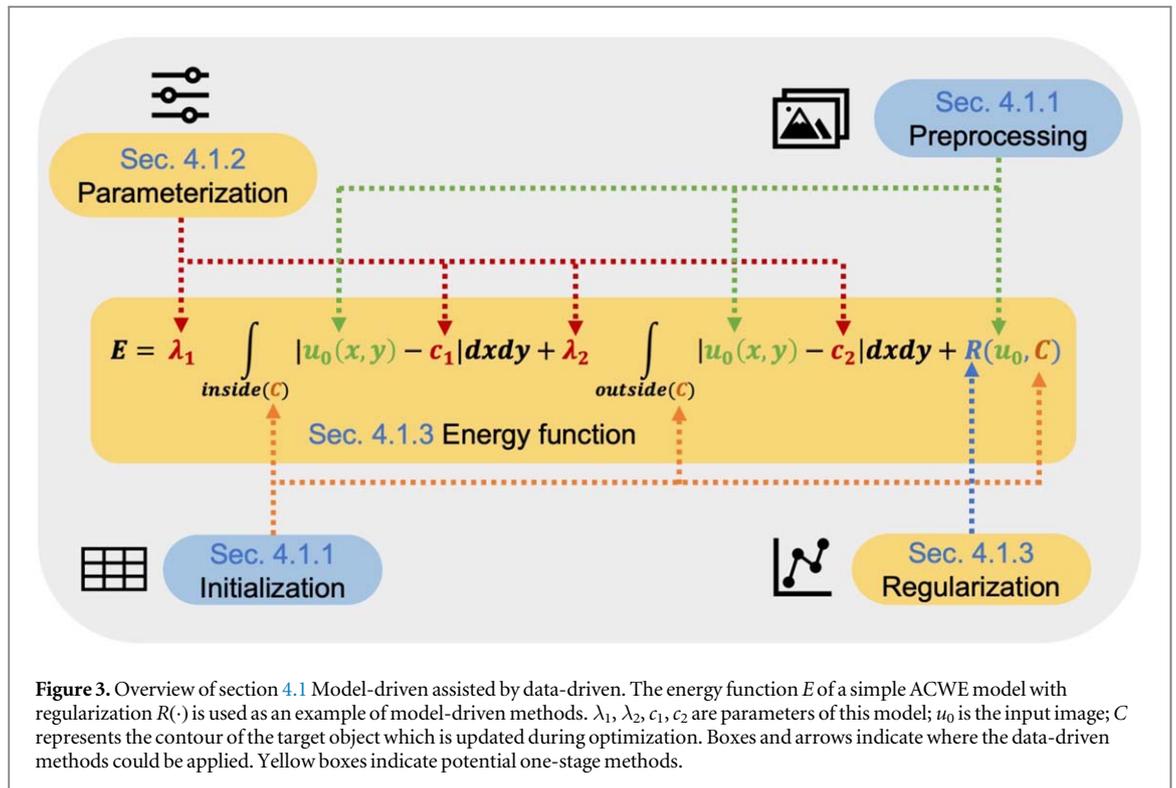
In some cases, there is expected or unexpected spatial complexity in medical image segmentation. Examples of targets with expected high complexity are fetuses, joints, and skulls. These targets contain many parts of various sizes, shapes, and appearances, which are challenging to distinguish. Some targets may be locally obscured. An example is that in thrombus segmentation, sometimes the thrombotic surface is locally obscured (López-Linares *et al* 2018). Another example of spatial complexity is artifacts inside the human body. Some patients may have artifacts like vasculature stents, metal restorations, osteosynthesis materials, or even free pieces of bones (Ibragimov *et al* 2017, Egger *et al* 2018). Thus, images of these people look different from images of others, and it is more challenging for automatic segmentation methods to work on such images.

3.4. Moving or deforming target

Some targets like the heart and fetus may move or deform during the image acquisition. This may lead to difficulties in multiple modalities. Sometimes ultrasound videos are used to analyze moving targets (Yang *et al* 2017). But for static images like CT and MRI, it may cause motion blurring (Chen *et al* 2020a).

3.5. Extremely small target

Some segmentation targets are relatively small in terms of their volumes or area. Coronary calcium is an example of an extremely small target. In every 2D CT slice, there are only several pixels or no pixels labeled as coronary



calcium, which means most pixels in the image are negative samples. Considering that pixel-wise losses are commonly used for training deep neural networks, extremely small targets can be ignored by the network (Chen *et al* 2020a).

3.6. Similar adjacent structure

Commonly, there are adjacent tissue or structures with similar intensity or texture around the segmentation targets. As mentioned above, there are multiple reasons that the boundaries of the target are not visible clearly. A similar adjacent structure is one of them. An example is epicardial fat segmentation (Commandeur *et al* 2018). In non-contrast CT images, epicardial fat and thoracic fat are very similar. They are separated by the pericardium which is very thin and not always visible. Thus, the adjacent thoracic fat makes epicardial fat segmentation difficult.

4. Methodology

In this section, we discuss the methods of using anatomical information with DL. By the backbone of these methods, we separately discuss model-driven assisted by data-driven methods and data-driven assisted by model-driven methods.

4.1. Model-driven assisted by data-driven

Before DL became popular for image segmentation, traditional model-driven segmentation techniques like active contour based segmentation have been widely used. These model-driven techniques make use of region information, edge information, shape constraints, appearance information, etc in a straightforward and explainable manner. They have shown promising performance in many segmentation tasks, but they are unsupervised approaches that strongly depend on manually selected parameters and initialization. In contrast, supervised DL requires large datasets with ground truth and learns features automatically from the data. Recently, researchers tried to assist model-driven approaches by the data-driven DL to boost performance and robustness. In this section, we discuss existing segmentation methods that use model-driven approaches as their frameworks or backbones and DL as assistance. Figure 3 shows the overview of this section. Data-driven methods are used as assistance in four ways: preprocessing, initialization, parameterization, energy function, and regularization.

4.1.1. Preprocessing and initialization

Many model-driven techniques are strongly influenced by the quality of input images. This problem is critical, especially in medical image segmentation. Some segmentation targets in medical images are relatively small, or the background is too large in the original image. Thus, the extraction of the region of interest (ROI) is a common preprocessing step in medical image segmentation. Using cardiac cine magnetic resonance, Ngo *et al* (2017) proposed a left ventricle segmentation method that uses a deep belief network for locating ROI and distance regularized LSs for segmentation. The combination takes advantage of both approaches that require small labeled data and generate accurate results. In the optic disc (OD) segmentation method by Zhang *et al* (2018), a faster R-CNN is trained to locate the OD with a bounding box and a shape-constrained LS algorithm is applied to segment the boundary of the OD.

Initialization and reinitialization are important for many model-driven algorithms. For instance, ACMs require an initial contour or contours as the start of evolution, and region-grow models require an initial seed as the starting point for growth. Usually, the initialization is determined by experts or empirical values, which means these models are empirical and even some are not entirely automatic. Since neural networks are designed to mimic the human brain, many researchers tried to replace the manual or empirical initialization with results learned by neural networks. Early work by Cha *et al* (2016a, 2016b) proposed a segmentation approach using a trained CNN to generate likelihood maps. After thresholding and hole-filling, the likelihood maps are fed as initial contours for 3D and 2D LS models. This approach has been applied to both bladder segmentation and bladder cancer segmentation. Hu *et al* (2017) introduced a method that applies a trained 3D CNN to automatically locate the organs of interest via a probability map and the map is fed into time-implicit LSs to obtain a fine segmentation of multiple organs. Later, in the deep nested LSs for segmentation of cardiac MRI in patients with pulmonary hypertension published by Duan *et al* (2018), CNNs are used for predicting three region probability maps and an edge probability map. Then the probability maps are incorporated into a single nested LS optimization framework for multi-region segmentation. More recent work by Gordon *et al* (2019) used a likelihood map generated by CNNs as the base to get proper initialization for LSs to segment inner and outer bladder walls in CT. Cai *et al* (2019) proposed a saliency-guided LS model for object segmentation. For this model, the initialization is automatically generated by the deep hierarchical saliency network proposed by Liu and Han (2016) followed by graph cut. Having a similar idea, Han *et al* (2019) initialized the LS function in their model with the probability maps generated from fully convolutional networks (FCNs). A more straightforward initialization for ACMs is to use the detection results from neural networks. Xu *et al* (2019) presented a segmentation method on breast histopathological images where a region-based active contour is initialized by nuclear patches detected by CNNs. Recent work by Xie and Chen (2020) used CNNs to detect a myocardial central-line and used the detection as initialization for their central-line guided LS approach to segment left ventricles in MRI.

Weakness, strength, and uncertainty: Using DL for preprocessing and initialization seems like a simple way to apply data-driven methods to classic model-driven methods. The data consistency is maintained as the physics model is the core for generating segmentation. Methods in this category keep most of the advantages of the traditional model-driven methods such as data consistency, interpretability, robustness to noise, etc. By adding data information, many works show improved accuracy. However, it divides the segmentation algorithm into multiple stages. None of these methods is a unified method that incorporates DL and model-driven methods, and they are hardly fully automatic.

4.1.2. Parameterization

Classical segmentation methods like variational methods are often dependent on good initializations and an adequate manual setting of hyperparameters. Although those methods are mathematically elegant they often cannot be used in a purely automatic manner. Therefore, some researchers proposed to make use of the power of DL to learn optimized parameters for such segmentation models. Hoogi *et al* (2016) proposed to generalize an LS segmentation approach by adaptively estimating active contour parameters using a CNN. In this method, the CNN is trained to predict the probability for each of the three classes: inside the object and far from its boundaries (p_1), close to the boundaries (p_2), or outside the object and far from its boundaries (p_3). Then p_1 , p_2 , and p_3 are used to set the weighting parameters of the energy function. This method was demonstrated for liver lesion segmentation in MRI and CT. More recent work (Ramírez *et al* 2018) has a similar idea for brain tumor segmentation. In this work, a U-net Ronneberger *et al* (2015) is used to output a spatially adaptive $\delta(x)$ function which is for the saliency term in the energy function of a variational model. A CNN followed by a multi-layer perceptron is trained to estimate the remaining parameters directly. With a similar idea, Hatamizadeh *et al* (2019) developed a framework called deep active lesion segmentation. In this framework, a U-net-like CNN is used to produce segmentation probability maps. The probability maps are transformed into a signed distance map to initialize an ACM. Two weighting parameters are estimated by extending the approach from Hoogi *et al* (2016). Another similar work is done by Deng *et al* (2019) for liver tumor segmentation. Research by Xie *et al* (2020)

employed DL to learn parameters of cost functions in the graph model for multiple surface segmentation. The model demonstrated promising results on spectral domain optical coherence tomography retinal layer segmentation and intravascular ultrasound vessel wall segmentation. In the LevelSet R-CNN segmentation for instance proposed by Homayounfar *et al* (2020), a neural network is trained to predict a truncated signed distance function initialization, a deep feature tensor, and a set of instance aware adaptive hyperparameters for each detection. These outputs are fed into an unrolled ACWE model for the final segmentation.

Weakness, strength, and uncertainty: Compared to using DL for initialization and preprocessing, using DL for parameterization is more complex. Similarly, these methods have many advantages over traditional model-driven methods. Compared to methods in section 4.1.1, these methods incorporated data information more deeply and took one step further on automation as some or all parameters are learned from the data instead of setting by a human. However, most of them are not fully automatic.

4.1.3. Energy function and regularization

Apart from the parameters, there are more complex terms or representations that can be learned in energy functions or the optimization procedure. In the work by Rupprecht *et al* (2016), a simple seven-layer CNN is trained to learn the mapping from input images to a flow field. The predictions from the CNN form a vector field that is used for the contour evolution in the Sobolev ACM for segmentation. This method was evaluated on both medical (STACOM dataset) and non-medical datasets. Another similar work uses an FCN to guide contour evolution for liver segmentation presented by Guo *et al* (2019). Another similar work called recurrent active contour evolution network by Chakravarty and Sivaswamy (2018) generalized the level-set-based deformable models evolving as a recurrent neural network (RNN). Apart from using DL for guiding contour evolution in ACMs, the regularizer of energy functions is a choice to apply DL. Boink *et al* (2019) proposed a joint approach for photo-acoustic reconstruction and segmentation in which a neural network is trained to learn the primal-dual optimization. In their work, DL is involved in both optimization and regularization for a variational model.

Some works use DL to learn part of the energy function. In the methods published by Cai *et al* (2019), a deep hierarchical saliency network is trained for initialization and a global saliency-guided energy term. The global saliency-guided energy term can guide the contour evolution of objects in color images and it improves the efficiency and robustness of the model. The approach by Han *et al* (2019) generates a shape prior mask by fitting the probability map from FCNs in a specific image of the global affine transformation. A shape energy term uses the shape prior mask to guarantee that the final segmentation is close to the shape prior.

Weakness, strength, and uncertainty: Some of the methods (Rupprecht *et al* 2016, Boink *et al* 2019, Homayounfar *et al* 2020) mentioned in sections 4.1.2 and 4.1.3 are also called physics-informed neural networks (PINNs). PINN can be trained to solve nonlinear PDEs (Raissi *et al* 2019). In our context, it is to minimize the energy functional of variational models. PINNs are relatively easy to implement and train, and as they are based on traditional segmentation models, researchers who know traditional segmentation methods could understand them without effort. Data consistency is easily guaranteed using this type of method. The amount of data required for training such neural networks is much less than data-driven methods in section 4.2. And clearly, we have more control as the physics model is known. However, there are still many problems and questions here. The physics model is crucial for such methods. As Rupprecht *et al* (2016) reported, their method has problems on the object boundary with some details. It is not sure whether this is the deficiency of their physics model or the trained neural network is not good enough. Another important problem is how to choose or design neural networks for such methods. Though some research shows neural networks could solve PDEs, they are not replacements of classical numerical methods. There are many unsolved questions behind these methods.

4.2. Data-driven assisted by model-driven

As DL has shown its success for image segmentation, many researchers attempted to boost segmentation networks by adding anatomical constraints or by using anatomical information in other ways. In this section, we discuss the segmentation methods whose main framework is deep neural networks. Anatomical information and related model-driven approaches are used as assistance for the main framework. In figure 4, a flowchart shows the process of image segmentation using neural networks and the overview of this section. Referring to the purposes of the assistance, the following content is separated into six parts: data augmentation and preprocessing, postprocessing, loss function, and regularization, model as a module, multi-task network, and GNN.

4.2.1. Data augmentation and preprocessing

Data augmentation is crucial to many medical image analysis applications using DL, as in many cases, the acquisition of labeled data is labor-expensive and time-consuming. Thus, it is necessary to generate more image samples to increase both the amount and diversity of the training samples (Zhang *et al* 2019). Anatomical

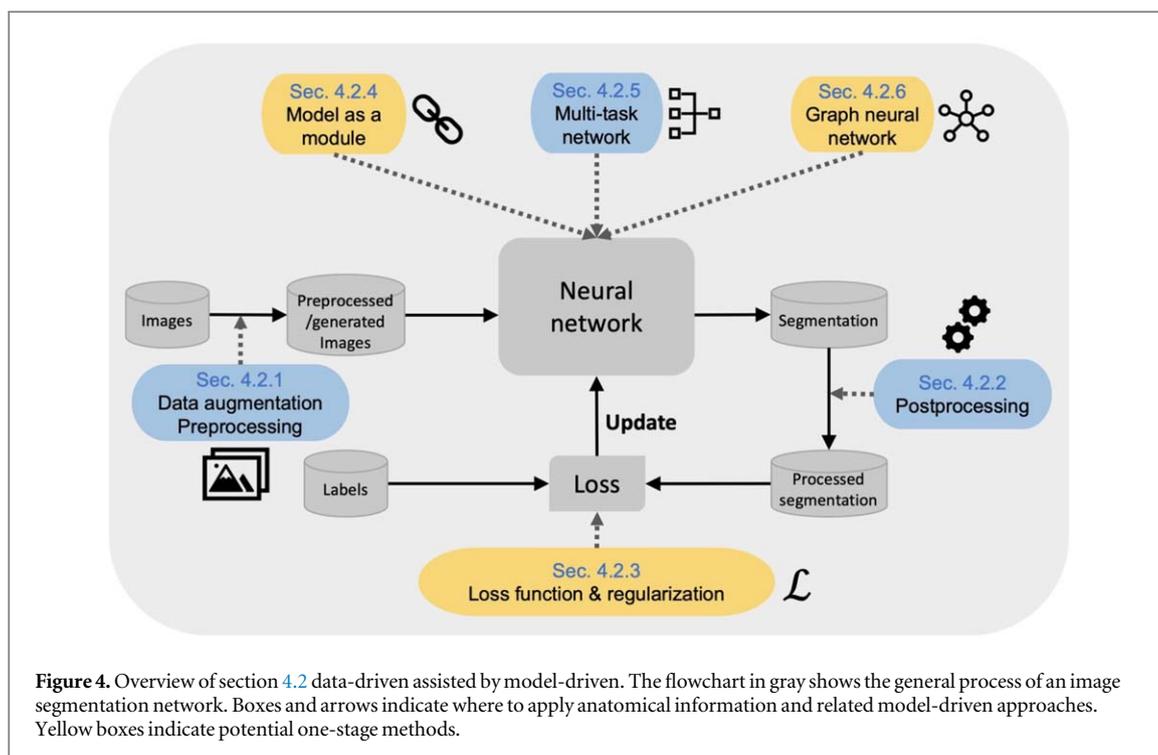


Figure 4. Overview of section 4.2 data-driven assisted by model-driven. The flowchart in gray shows the general process of an image segmentation network. Boxes and arrows indicate where to apply anatomical information and related model-driven approaches. Yellow boxes indicate potential one-stage methods.

information is the base of data augmentation as the visual variations of the objects of interest guide the augmentation direction. The variations include many aspects such as scale variation, deformation, illumination variation, rotation, translation, etc. Classical transformations for data augmentation involve (horizontal or vertical) flipping, random rotation, cropping, scaling, translations, shearing, and elastic deformations (Zhang *et al* 2019). To select the proper transformations for the object of interest, anatomical information mentioned in section 2 especially shape information and contour information is the main clue. It is also possible to add prior information to the data. Clinical prior represented by probability maps are used as additional training data in Saha *et al* (2020) for prostate cancer detection.

Preprocessing of medical images is common in DL segmentation. Anatomical information of the target objects is usually helpful to choose the proper preprocessing techniques. Common preprocessing steps for medical image segmentation include ROI extraction, thresholding, denoising, enhancement, intensity standardization, bias correction, etc. But not all preprocessing techniques use anatomical information. The related widely used anatomical information includes location, size, adjacency information, motion, shape, etc. Considering the use of anatomical information and related model-driven approaches, not all techniques are covered in this work. Here we give examples of preprocessing techniques assisted by anatomical information and related model-driven algorithms.

Extraction of ROI is one of the most powerful preprocessing steps. It is more meaningful than many other preprocessing steps as it can remove non-related regions and reduce the computational cost significantly. Generally, the extraction of the ROI is also a segmentation task and there are many works on ROI extraction with or without DL. As a preprocessing step, usually, the accuracy is not required to be very high. An example of using ROI extraction is to segment lung lobes before segmenting smaller structures like lung nodules. Thresholding is another widely used preprocessing step. As many structures like bones and tissues have a specific range of intensity in medical images, thresholding can be applied to filter out all the other non-related parts. Apart from the techniques mentioned above, other preprocessing steps are available. An example is that in the colon glands segmentation approach by Kainz *et al* (2015), the RGB H&E stained images are deconvolved to extract a robust representation of the tissue structures. Overall, the preprocessing steps should be selected considering both the modality and the target objects.

Weakness, strength, and uncertainty: Using anatomical information and related model-driven approaches for data augmentation and preprocessing does not change the core of the neural network. Usually, it is neither complicated nor time-consuming to do data augmentation or preprocessing, while they could lead to a huge improvement in accuracy or reduce computation time significantly. Not all anatomical information applies to data augmentation and preprocessing for neural networks. Commonly-used anatomical information usually consists of simple representative features like location, size, contrast, etc.

4.2.2. Postprocessing

Anatomical information and related model-driven approaches are widely used as postprocessing steps to obtain fine segmentation results in DL methods. Commonly useful anatomical information are contour and region. Many works (Li *et al* 2017, Xu *et al* 2018, Zhao *et al* 2018) use the conditional random field (CRF) to produce delicate delineation of boundaries. The famous DeepLab (Chen *et al* 2017) for semantic image segmentation also used fully connected CRFs to improve localization performance both qualitatively and quantitatively. Graph cut methods are popular for optimizing the location of a contour. For medical image segmentation, the graph cut is a common postprocessing step with examples (Ma *et al* 2018, Močnik *et al* 2018, Zabihollahy *et al* 2019). Location information can be useful for postprocessing too. Song *et al* (2016) were aware that a tumor rarely happens completely systematically in the brain. They used symmetric difference and thresholding to generate a rough segmentation. Then four types of voxel-wise features—appearance, texture, location, and context—are extracted to perform further voxel classification into five subcategories (necrosis, edema, non-enhancing tumor, enhancing tumor, and other tissues). Finally, they applied a pathology-guided refinement scheme (edema is usually not inside the active cores, and non-enhancing cores often surround active cores) to correct mislabeling. LS models are another choice for postprocessing. In the pulmonary nodule segmentation method published by Roy *et al* (2019), the shape-driven evolution of LSs was designed to produce an accurate segmentation with the coarse segmentation from the FCN as initialization. Another work by Hu *et al* (2019) for tumor segmentation in breast ultrasound used a phase-based ACM to refine the rough segmentation results from a dilated FCN. Recent work by da Silva *et al* (2020) presented a superpixel-based CNN utilizing a manifold simple linear interactive clustering algorithm and a probabilistic atlas for coarse prostate segmentation in 3D MRI. A 3D ACWE model is applied later to obtain fine segmentation results. Feng *et al* (2020) also used an LS method for postprocessing to improve the performance of their pelvic floor structure segmentation network. The recently presented adaptive weighting and scalable distance regularized LS method (Li *et al* 2020c) also shows its strengths as postprocessing for DL methods. Some methods with LSs for postprocessing overlap with the methods mentioned in section 4.1.1.

Weakness, strength, and uncertainty: Postprocessing helps get fine or smooth segmentation results. The above works reported better performance after using their postprocessing techniques. But, as data-driven methods are the core algorithms, these methods have both the advantages and disadvantages of the initial neural networks. Similar to methods in section 4.1.1, the segmentation procedure is divided into multiple stages.

4.2.3. Loss function and regularization

The loss function is important for DL as it guides the training of the networks. Similar to the concept of the energy function of variational models, the loss function is designed to constrain the results and guide the optimization. With available anatomical information and related model-driven approaches, many loss functions show promising performance. The survey paper by Jurdia *et al* (2020) summarized high-level prior-based loss functions for medical image segmentation. Readers can obtain an overview of loss functions from this paper.

Since segmentation aims to find the optimized contour of the target objects to some extent, the information related to contours, edges, or boundaries was considered by many researchers. By incorporating boundary information directly into the loss function, Shen *et al* (2017) introduced a boundary-aware FCN for brain tumor segmentation. Another recent boundary-aware network by Chen *et al* (2019) for portrait segmentation utilizes not only a boundary loss but also a boundary feature mining branch to get boundary attention maps. Earlier work by Oktay *et al* (2017) mentioned the use of a shape regularization loss for cardiac image enhancement and segmentation. Recently an unsupervised microvascular image segmentation method by Gur *et al* (2019) employed a complex loss function with six terms. One of the terms is derived from the ACWE model. Similarly, Chen *et al* (2019) proposed a loss function inspired by the ACWE model. Similar to the energy terms in the ACWE model, the proposed loss function considers the length of the contour, area of the inside region, and area of the outside region. This method showed a promising performance on heart segmentation in MRI. Kim and Ye (2019) were inspired by another famous LS-related method, the Mumford–Shah model. The proposed Mumford–Shah loss function was demonstrated both on semi-supervised learning and unsupervised learning. Another cardiac segmentation in MRI method by Yue *et al* (2019) proposed a loss function with three terms: the segmentation loss (cross-entropy and Dice), the spatial constraint loss, and the shape reconstruction loss for shape regularization. Topological information can be applied to loss functions as well. In the work by Clough *et al* (2019), a topological loss is introduced by using PH to explicitly represent topological priors.

Distance transform maps (DTMs) are commonly used to design additional regularizers in loss functions. A recent study by Ma *et al* (2020) summarized the latest developments using DTM in the 3D medical segmentation field and evaluated five benchmark methods on two datasets. Classical ground truth label maps could be transferred into DTM as an alternative. For example, we could transform a binary mask into a gray-scale image by assigning the intensity of pixels according to their distance to the boundary. Signed distance function (SDF) is

one example of a transformation protocol that assigns negative or positive values inside or outside the objects. Two ways to using DTM for image segmentation with DL are: (1) designing new loss functions, (2) adding auxiliary tasks (Ma *et al* 2020). Here we only focus on loss functions, and the second way is discussed in section 4.2.5. The boundary loss proposed by Kervadec *et al* (2019) is designed for highly unbalanced segmentation problems. The widely used loss functions for segmentation like Dice loss, and cross-entropy loss are calculated by summing pixels over regions. If the number of positive pixels is much smaller than that of negative pixels in the ground truth labels, this kind of region-based loss may lead to networks that ignore positive pixels. The boundary loss is calculated by a non-symmetric L_2 distance on the space of shapes as a regional integral. Thus, the unbalanced data does not influence it. Hausdorff distance (HD) loss by Karimi and Salcudean (2019) is designed to minimize HD between segmentation and ground truth directly during training. In this work, three methods to estimate HD are described and one of them is based on distance transform. One more example is the signed distance function regression loss proposed by Xue *et al* (2020). During training, the network regresses the SDF of ground truth instead of calculating softmax. More details and explanations about these loss functions can be found in Ma *et al* (2020). For experiments, all the distance transform losses are coupled with dice loss to stabilize training, otherwise, training is hard to converge. The evaluation results show that distance transform losses have the potential to improve segmentation performance but the improvement is not consistent on different tasks. One drawback is the high computation cost of DTM. However, using DTM for performance improvement on image segmentation is still an open field.

Weakness, strength, and uncertainty: Modification of loss functions is an easy and effective way to employ anatomical information in neural networks. It does not change the network architecture or require complex implementation steps. However, many small things in loss functions could make big changes. In many cases, loss functions consist of more than one term. Any designed losses mentioned above could work as an additional regularizer to a pixel-wise loss. The weight parameters in such loss function could lead to large variation during training. Sometimes weight parameters that work on one dataset may fail on the other datasets.

4.2.4. Model as a module

A more integral combination of model-driven and data-driven approaches is to add a model as a module in segmentation networks. In the semi-supervised network for image segmentation proposed by Tang *et al* (2017), an LS model is incorporated within the training process to refine the contour from the predicted probability map and update the weights. Unlike using the LS for postprocessing, the LS model works interactively with the neural network to improve accuracy. Another semi-supervised network for 3D left atrium segmentation was proposed by Yu *et al* (2019). The highlight of that work is that they designed an uncertainty-aware scheme to enable the network to learn uncertainty using unlabeled data. In this framework, a teacher model is built and a student model learns from the teacher model when training with labeled data. When training with unlabeled data, the student model exploits the uncertainty from the teacher model, and the teacher model estimates the uncertainty as well. Recent work for vertebral bone segmentation (Rehman *et al* 2020) has a similar training strategy to Tang *et al* (2017) but in a supervised manner. An LS model is used to work interactively with a CNN of U-Net architecture type to refine the segmentation and updating weights in the network. Zhao *et al* (2019) proposed a knowledge-aided CNN (KaCNN) for small organ segmentation. Their KaCNN contains an information-fusion component that could combine the features from an additional model like multi-atlas models. In their work, they add LBP and BRIEF (Heinrich and Blendowski 2016) features as extra knowledge to boost the segmentation performance. Zhang *et al* (2020b) concatenate a morphological layer between two U-nets for epicardial fat segmentation. The proposed morphological layer refines the inside region of the pericardium where epicardial fat locates.

Another way to involve model-driven approaches in DL is to transform the model into a network. For instance, the deep watershed transform network (WTN) segmentation proposed by Bai and Urtasun (2017) learns the energy of the watershed transform with a feed-forward neural network. PSPNet (Zhao *et al* 2017) is used to segment a rough ROI so that the WTN only focuses on relevant areas. This network combines the strengths of DL with the classical bottom-up grouping technique that can be trained end-to-end and be fully automatic. Gur *et al* (2019) introduced an end-to-end trainable ACM via differentiable rendering for image segmentation. In this model, an encoder-decoder architecture with U-Net skip connections is developed to produce a 2D displacement field J . The vertices of the polygon are updated by the value in J . In other words, the displacement field guides the polygon evolution, which is similar to the idea in Rupprecht *et al* (2016). Le *et al* (2018) reformulated LSs as RNNs for semantic segmentation. They call the reformulated module Recurrent LS. A very recent work by Actor *et al* (2020) looked at the similarity between CNN and LS methods for segmentation. They constructed a LS network with CNNs and compared it with common CNNs.

Weakness, strength, and uncertainty: Some methods in this category are multi-stage methods, which means either a pre-trained model is required or part of the network needs to be trained separately. This makes the implementation and training complex. One example of attempting to incorporate a model with DL as a one-

stage end-to-end trainable method is the brain tumor segmentation by Le *et al* (2018). In this work, a LevelSet layer is designed by incorporating the recurrent fully-convolutional network and LS framework. The proposed deep recurrent LS combines convolutional layers, deconvolutional layers, and LevelSet layers to obtain feature maps and refine contours for brain tumors. Comparing to other brain tumor segmentation methods, this method improves the speed but does not outperform all the other methods without LevelSet layers.

4.2.5. Multi-task network

The multi-task network or multi-task learning is referring to the networks trained for multiple purposes. Many researchers believe that learning complementary tasks in one network can improve the overall performance. In multi-task networks for segmentation, common auxiliary tasks are contour map learning, object detection, distance map learning, adjacency object detection/segmentation, etc.

The early multi-task network by Chen *et al* (2016) for gland segmentation won the 2015 MICCAI Gland Segmentation Challenge and learns gland objects and contours within one fully connected network. A multi-task network for multi-organ segmentation is published by Navarro *et al* (2019). In this network, an encoder-decoder network of U-Net architecture type is trained to learn segmentation maps, distance maps, and contour maps at the same time. In Wang *et al* (2020), a distance map is to refine the tubular structure. Similarly, in the shape and boundary-aware joint multi-task deep network proposed by Murugesan *et al* (2019), the network is trained to learn the segmentation, contour maps, and distance maps at the same time. Since there is only one segmentation mask in the dataset, the contour map is obtained by estimating the boundaries of connected components from a transformed distance map. And the distance map is estimated by applying a Euclidean distance transform to the mask. Another example is the cell segmentation model presented by Liu *et al* (2019). In this model, a U-Net is trained to learn centroids, regions, and contours of cells in adaptive optics retinal images. A LS model has followed to finally segment the cells. Myronenko and Hatamizadeh (2019) employed a similar idea in 3D kidney and kidney tumor segmentation. A boundary stream with an attention-driven decoder is deployed to emphasize the boundary features in the feature maps learned by the mainstream. Similar ideas were used for semantic segmentation too. Takikawa *et al* (2019) proposed a gated shape CNN for semantic segmentation in which a network branch of shape stream is developed to focus on processing the relevant edge-related information. A fusion module with atrous spatial pyramid pooling is used to combine the information from the shape stream and the regular stream. Recent work by Hatamizadeh *et al* (2020) proposed an edge-gated CNN module that can be integrated with any generic encoder-decoder architecture to enhance the edge representations in the learned feature maps. The idea of this module is to add an auxiliary task of learning edges to the original network. The ground truth is generated by applying the Sobel filter to the segmentation masks.

Unlikely, some researchers forced their networks to learn specific features like contour features in an inconspicuous multi-task way. A superpixel-based CNN for liver segmentation by Qin *et al* (2018) tried to learn liver boundary features. In this framework, with superpixels, segmentation is transformed into a classification problem. By labeling the superpixels into three classes (liver, liver boundary, and background), the network is enforced to learn boundary features and to be able to identify the liver boundary explicitly. Zhou *et al* (2019) developed a CIA-Net for nuclei instance segmentation with the contour-aware information aggregation. An information aggregation module is introduced for the bi-directional multi-level task-specific feature aggregation between two decoders. This work won the 2018 MICCAI challenge of multi-organ-nuclei-segmentation. Zhang *et al* (2019) proposed a generic medical segmentation network named edge-attention guidance network. In this network, a network branch is designed to learn the edge-attention representations in the encoding layers. Then, these representations are transferred to decoding layers to guide segmentation. A weighted aggregation module is used to fuse the former edge guidance module and the decoder. This work has experimented with good results on OD/cup, vessel, and lung segmentation in multiple modalities. Recent work by Li *et al* (2020b) used an autoencoder to learn low-dimensional anatomical features to constraint the segmentation results from the main U-Net stream. To reach a similar goal, Painchaud *et al* (2020) proposed a model with two variational autoencoders (VAEs) for cardiac segmentation. A constrained VAE is trained to learn anatomical features from valid cardiac shapes by reconstructing the ground truth, and the learned features are used to regulate predicted implausible segmentation to the closest correct shape in the latent space.

Weakness, strength, and uncertainty: Comparing to methods in the other sections, multi-task networks usually have a bigger network structure as some of them consist of multiple networks in parallel. As the network is trained for multiple tasks, multiple labels are fed into the network. Thus, naturally, there may be adjacency information, geometrical structure information, or other anatomical information in these labels. During training, as the network is learning more than one task, the correlation between tasks may contribute to all tasks and improve their performance. The strength of these networks is obvious as more information is learned. However, the drawback is obvious too. A bigger network leads to more computational costs. Multiple labels require both more objects for labeling and more annotation time. Multi-task networks do not apply to all segmentation tasks.

4.2.6. Graph neural network

GNNs are neural networks that learn and analyze graph data. Graph data is described with nodes (vertices) and edges between nodes (Zhou *et al* 2020). As graphs could represent a large number of systems, it is possible to use graphs to represent anatomical structures and shape priors. A recent survey paper of GNNs by Zhou *et al* (2020) provides a review of existing GNNs and their applications on text, image, science, etc. But medical image segmentation is not covered by this survey. In this section, we only focus on medical image segmentation.

For medical image segmentation, GNNs are known to extract trees like airways and vessels. Selvan *et al* (2018) proposed a graph auto-encoder (GAE) model based on GNNs (Kipf and Welling 2016a, 2016b) to extract airways from 3D chest CT. In this paper, CT scans are preprocessed into graph data that has N nodes with edges between nodes and node feature matrices consisting of 7-dimensional Gaussian density, local radius, position, orientation in 3D, and variances. GAEs are trained to learn node embeddings from the input feature matrices and a decoder is trained to predict edges between nodes. During training, the feature matrices keep updating and similar nodes in the graph are embedded more closely together. Later, this work and another work using mean field network (Selvan *et al* 2018) of the same authors are extended to a journal paper (Selvan *et al* 2020). Both models showed improved performance especially on detecting small branches and overcoming occlusions. One limitation is that preprocessing is necessary to obtain graphs from image data as input, which makes the method a two-stage model. Another work for airway segmentation by Juarez *et al* (2019) presented an end-to-end framework by replacing the deepest level of a 3D U-net with a GNN-based module with graph convolutions. In this model, preprocessing of images to graph inputs is avoided as the feature maps from downsampling layers are transformed into graph data for feeding to the GNN module. From their experimental results, the proposed model shows similar results as the baseline U-net model, but with a small improvement on airway completeness for a fixed volume leakage.

The first method applying GNN to blood vessel segmentation is proposed by Shin *et al* (2019). They combine a GNN module for learning the global structure of vessels and a U-net-like CNN with downsampling and upsampling for learning local appearances. The model was evaluated on four retinal image datasets and a coronary artery x-ray angiography dataset. Their experiments show that the vessel graph network (VGN) has better performance in terms of average precision and area under the curve, which means the VGN has a better ability to detect both vessels and background. Another work published by Wolterink *et al* (2019) for coronary artery segmentation in cardiac CT angiography (CCTA) utilizes GNNs as well. In this work, a network with five graph convolutional network (GCN) layers is used to optimize the location of nodes in a tubular surface mesh graph by learning the local features and neighbor features. The model requires both CCTA and a coronary artery centerline as input. The centerlines are extracted automatically from their previous CNN-based method Wolterink *et al* (2019). The paper shows that GCNs improve segmentation accuracy and produce regular and better meshes directly. The same group of researchers proposed graph attention networks (GAT) for coronary artery segment labeling (Hampe *et al* 2021) of a similar framework. This approach shows similar performance to previous approaches on most branches, and better performance on small leaf branches. Yao *et al* (2020) proposed a GCN-based point cloud approach to improve head and neck vessel segmentation in CT angiography. In their model, the rough segmentation results from V-net are refined by the proposed approach in two steps. First, a point cloud network takes the rough segmentation and refines the initial voxels of vessels. Then, a GCN is applied to the point cloud to classify vessels into 13 categories. More recent work by Yang *et al* (2020) includes residual connections and a condition extractor in GCN for coronary artery labeling in CCTA. A partial-residual GCN takes centerlines of coronary arteries as input, while a condition extractor with 3D CNN and bi-directional long short-term memory takes the images and centerlines as input and extracts features along vessel branches as the conditions for the GCN. The two parts are trained end-to-end taking both position features and spatial features.

Apart from tree structure segmentation, GNNs are applied to other medical image segmentation tasks. Cucurull *et al* (2018) approached cerebral cortex parcellation as a graph segmentation task in 3D MRI. Both GCN (Kipf and Welling 2016a) and GAT (Veličković *et al* 2018) were trained and evaluated on the Human Connectome Project dataset for Broca's area parcellation. Their work showed improving performance compared to other alternatives and baselines. Tian *et al* (2020) presented a framework using a multi-scale CNN to generate feature maps and then a GCN takes the feature maps and graph nodes as its input to segment OD and cup. Their method was evaluated on the REFUGE and Drishti-GSI datasets and outperformed the state-of-the-art methods. A similar idea was applied to an interactive prostate segmentation method by Tian *et al* (2020) as well. Another example is the uncertainty-based GCNs for organ segmentation refinement published by Soberanis-Mukul *et al* (2020). They employ uncertainty levels from the output of a CNN (2D U-net in their experiments) as input for their GCN to formulate a semi-supervised learning strategy. The approach was tested on the NIH pancreas dataset and spleen dataset of the medical segmentation decathlon. Wickramasinghe *et al* (2020) introduced a GNN structure that can segment voxels into 3D surface meshes directly. This model was

evaluated on electron microscopy and MRI of brain and CT of liver and outperformed all state-of-the-art methods.

Weakness, strength, and uncertainty: Though many works reported improved performance of GNNs in medical image segmentation, GNNs have many limitations. Tian *et al* (2020) pointed out that their GNNs failed when the objects have irregular edges or different structures. For tree segmentation, many researchers reported their GNNs performed worse on small branches, sometimes even missed them. As the number of publications using GNNs for medical image segmentation is much smaller than for other methods, many possibilities but also problems exist.

5. Discussion and future work

By reviewing the literature of anatomy-aided DL for medical image segmentation, we observed that the common goal for all these methods is to incorporate the advantages of both model-driven methods and data-driven methods. DL methods have shown outstanding performance on many medical image segmentation tasks. But it requires a large number of labeled data. In the meantime, a neural network trained to segment a specific object can hardly segment other objects. Many model-driven methods are on the opposite side. They may be able to segment many objects after tuning the parameters. But some complex tasks can be difficult for them. Model-driven methods are commonly robust against noise which is an outstanding advantage compared to DL methods. Many trained networks are noise-sensitive. Adding a simple Gaussian noise to original images could confuse the network and could result in wrong segments. In addition, interpretability is crucial for medical image analysis since it is related to medical data and patients.

The advantages of anatomy-aided DL include but are not limited to: (1) reducing computational cost, (2) decreasing human interaction, (3) guiding neural networks, (4) improving segmentation accuracy, (5) higher interpretability, (6) higher robustness. However, it has drawbacks as well. Segmentation targets have various anatomical information and they have different visual effects or features in different image modalities. For a medical image segmentation task, usually, the segmentation target and the image modality are known. The selection of proper anatomical information and corresponding model-driven methods can be difficult. It is hard to represent or apply some complex anatomical information to segmentation methods. Besides, the way of combining anatomical information and data-driven methods varies. It is not clear which way is better. Considering that some segmentation tasks are based on more than one image modality, sometimes even video sequences, finding the right segmentation methods with the right anatomical information can be more challenging. Another drawback is that sometimes more labels are required for representing anatomical features. Similar to the acquisition of segmentation labels, this can be time-consuming and labor-intensive.

5.1. Future work

With the development of computer-aided diagnosis systems, there are diverse medical image segmentation tasks with various segmentation targets and image modalities. Anatomy-aided DL for medical image segmentation has bright prospects due to its flexibility and adaptability. Here we introduce some future research directions that we believe will be helpful in this field.

- *End-to-end framework:* Most works reviewed in this paper are multi-stage methods, which require extra preprocessing, pretraining, postprocessing, or other steps. With the maturity of DL frameworks like PyTorch, much complex design of anatomy-aided DL methods can be implemented in an end-to-end framework Bohlender *et al* (2021). The trend of end-to-end approaches for medical image segmentation could lead to better practicability for clinical use.
- *Learning with less data:* Data and label acquisition is one of the biggest challenges when applying DL to medical data. Thus, the ability of anatomy-aided DL to reduce the required amount of labeled data in obtaining good segmentation is significant. It could be semi-supervised learning, self-supervised learning, unsupervised learning, or other learning strategies that require less labeled data. Or it could be PINN which does not require plentiful training data.
- *Explainable model:* DL models are known as 'black boxes' with low interpretability. Thus, once a model is trained, the predictions from it are hardly interpretable. In practice, these 'black boxes' can be attacked by imperceptible changes in images Finlayson *et al* (2019). Thus, opening these 'black boxes' and building resilient models robust to potential attacks is significant.

- *Physics-informed neural networks*: PINNs could combine powerful DL and classical physical models for image segmentation. They have better interpretability than black-box DL methods. However, they have not been applied to many medical image segmentation tasks yet. Thus, both potential and challenges exist here.
- *Graph neural networks*: GNN has shown its ability in many applications but not much in medical image segmentation. As it naturally includes graphical information in the network, it has the potential to outperform CNN in some medical image segmentation tasks requiring more contextual information.

5.2. Conclusion

In this review paper, we provided a comprehensive overview of these anatomy-aided DL methods for medical image segmentation. In particular, we presented the anatomical information of four types (shape, appearance, motion, and context), and challenges in medical data of six categories. To overcome the challenges, we discussed methodologies in two directions (model-driven assisted by data-driven and data-driven assisted by model-driven) of incorporating anatomical information into DL for segmentation. We hope that this review can provide an intelligible understanding of those methods that have combined advantages of both DL and anatomy, and bring awareness of challenges and possible future works.

Acknowledgments

The authors would like to thank the scientific advice provided by Prof R Vliegenthart from the Department of Radiology, University Medical Center Groningen. This work is supported by ZonMw under project B3CARE (project number:104006003). This project has received funding from the EU Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 777 826. JW and CB acknowledge support by the Dutch 4TU HTSF program Precision Medicine.

ORCID iDs

Lu Liu  <https://orcid.org/0000-0003-1352-421X>

Jelmer M Wolterink  <https://orcid.org/0000-0001-5505-475X>

Christoph Brune  <https://orcid.org/0000-0003-0145-5069>

Raymond N J Veldhuis  <https://orcid.org/0000-0002-0381-5235>

References

- Actor J A, Fuentes D T and Rivière B 2020 Identification of kernels in a convolutional neural network: connections between the level set equation and deep learning for image segmentation *Proc. SPIE* **11313** 1131317
- Bai M and Urtasun R 2017 Deep watershed transform for instance segmentation *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.* **2017** 5221–9
- Bigün J, Granlund G H and Wiklund J 1991 Multidimensional orientation estimation with applications to texture analysis and optical flow *IEEE Trans. Pattern Anal. Mach. Intell.* **13** 775–90
- Bohlender S, Oksuz I and Mukhopadhyay A 2021 A survey on shape-constraint deep learning for medical image segmentation arXiv:2101.07721 19 January 2021
- Boink Y E, Manohar S and Brune C 2019 A partially-learned algorithm for joint photo-acoustic reconstruction and segmentation *IEEE Trans. Med. Imaging* **39** 129–39
- Cai Q, Liu H, Qian Y, Zhou S, Duan X and Yang Y-H 2019 Saliency-guided level set model for automatic object segmentation *Pattern Recognit.* **93** 147–63
- Caselles V, Kimmel R and Sapiro G 1997 Geodesic active contours *Int. J. Comput. Vis.* **22** 61–79
- Cha K H, Hadjiiski L M, Samala R K, Chan H-P, Cohan R H, Caoili E M, Paramagul C, Alva A and Weizer A Z 2016a Bladder cancer segmentation in ct for treatment response assessment: application of deep-learning convolution neural network a pilot study *Tomography* **2** 421–29
- Cha K H, Hadjiiski L, Samala R K, Chan H-P, Caoili E M and Cohan R H 2016b Urinary bladder segmentation in ct urography using deep-learning convolutional neural network and level sets *Med. Phys.* **43** 1882–96
- Chakravarty A and Sivaswamy J 2018 Race-net: a recurrent neural network for biomedical image segmentation *IEEE J. Biomed. Health Inform.* **23** 1151–62
- Chan T F and Vese L A 2001 Active contours without edges *IEEE Trans. Image Process.* **10** 266–77
- Chazal F and Michel B 2021 An introduction to topological data analysis: fundamental and practical aspects for data scientists arXiv:1710.04019v2 25 February 2021
- Chen C, Qin C, Qiu H, Tarroni G, Duan J, Bai W and Rueckert D 2020a Deep learning for cardiac image segmentation: a review *Front. Cardiovascular Med.* **7** 25
- Chen H, Qi X, Yu L and Heng P-A 2016 Dcan: deep contour-aware networks for accurate gland segmentation *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.* **2016** 2487–96
- Chen J, You H and Li K 2020b A review of thyroid gland segmentation and thyroid nodule segmentation methods for medical ultrasound images *Comput. Methods Programs Biomed.* **185** 105329

- Chen L-C, Papandreou G, Kokkinos I, Murphy K and Yuille A L 2017 Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs *IEEE Trans. Pattern Anal. Mach. Intell.* **40** 834–48
- Chen X, Qi D and Shen J 2019 Boundary-aware network for fast and high-accuracy portrait segmentation arXiv:1901.03814 12 January 2019
- Chen X, Williams B M, Vallabhaneni S R, Czanner G, Williams R and Zheng Y 2019 Learning active contour models for medical image segmentation *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.* **2019** 11632–40
- Cheplygina V, de Bruijne M and Pluim J P 2019 Not-so-supervised: a survey of semi-supervised, multi-instance, and transfer learning in medical image analysis *Med. Image Anal.* **54** 280–96
- Clough J R, Oksuz I, Byrne N, Schnabel J A and King A P 2019 Explicit topological priors for deep-learning based image segmentation using persistent homology *Information Processing in Medical Imaging* ed A Chung et al vol 11492 (Berlin: Springer) pp 16–28
- Commandeur F, Goeller M, Betancur J, Cadet S, Doris M, Chen X, Berman D S, Slomka P J, Tamarappoo B K and Dey D 2018 Deep learning for quantification of epicardial and thoracic adipose tissue from non-contrast ct *IEEE Trans. Med. Imaging* **37** 1835–46
- Cucurull G, Wagstyl K, Casanova A, Veličković P, Jakobsen E, Drozdal M, Romero A, Evans A and Bengio Y 2018 Convolutional neural networks for mesh-based parcellation of the cerebral cortex *Medical Imaging with Deep Learning (Amsterdam, 4–6 July 2018)* (<https://openreview.net/pdf?id=rkKvBAiiz>)
- da Silva G L, Diniz P S, Ferreira J L, França J V, Silva A C, de Paiva A C and de Cavalcanti E A 2020 Superpixel-based deep convolutional neural networks and active contour model for automatic prostate segmentation on 3d mri scans *Med. Biol. Eng. Comput.* **58** 1947–64
- Deng Z, Guo Q and Zhu Z 2019 Dynamic regulation of level set parameters using 3d convolutional neural network for liver tumor segmentation *J. Healthc. Eng.* **2019** 4321645
- Desjardins B and Kazerooni E A 2004 Ecg-gated cardiac ct *Am. J. Roentgenol.* **182** 993–1010
- Ding W, Li L, Zhuang X and Huang L 2020 Cross-modality multi-atlas segmentation using deep neural networks *Medical Image Computing and Computer Assisted Intervention—MICCAI 2020* ed A L Martel vol 12263 (Berlin: Springer) pp 233–42
- Duan J, Schlemper J, Bai W, Dawes T J, Bello G, Doumou G, De Marva A, O'Regan D P and Rueckert D 2018 Deep nested level sets: fully automated segmentation of cardiac mr images in patients with pulmonary hypertension *Medical Image Computing and Computer Assisted Intervention—MICCAI 2018* ed A Frangi et al vol 11073 (Berlin: Springer) pp 595–603
- Ebrahimkhani S, Jaward M H, Cicutini F M, Dharmaratne A, Wang Y and de Herrera A G S 2020 A review on segmentation of knee articular cartilage: from conventional methods towards deep learning *Artif. Intell. Med.* **106** 101851
- Egger J, Pfarrkirchner B, Gsaxner C, Lindner L, Schmalstieg D and Wallner J 2018 Fully convolutional mandible segmentation on a valid ground-truth dataset *40th Annual Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)* **2018** 656–60
- Feng F, Ashton-Miller J A, DeLancey J O L and Luo J 2020 Convolutional neural network-based pelvic floor structure segmentation using magnetic resonance imaging in pelvic organ prolapse *Med. Phys.* **47** 4281–93
- Feo R and Giove F 2019 Towards an efficient segmentation of small rodents brain: a short critical review *J. Neurosci. Methods* **323** 82–9
- Finlayson S G, Bowers J D, Ito J, Zittrain J L, Beam A L and Kohane I S 2019 Adversarial attacks on medical machine learning *Science* **363** 1287–9
- Gordon M N, Hadjiiski L M, Cha K H, Samala R K, Chan H-P, Cohan R H and Caoili E M 2019 Deep-learning convolutional neural network: Inner and outer bladder wall segmentation in ct urography *Med. Phys.* **46** 634–48
- Guo X, Schwartz L H and Zhao B 2019 Automatic liver segmentation by integrating fully convolutional networks into active contour models *Med. Phys.* **46** 4455–69
- Gur S, Shaharabany T and Wolf L 2019 End to end trainable active contours via differentiable rendering *Int. Conf. on Learning Representations 2020 (Addis Ababa, Ethiopia, 30 April 2020)* (<https://openreview.net/pdf?id=rkxawHKDr>)
- Gur S, Wolf L, Golgher L and Blinder P 2019 Unsupervised microvascular image segmentation using an active contours mimicking neural network *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)* **2019** 10722–31
- Hameurlaine M and Moussaoui A 2019 Survey of brain tumor segmentation techniques on magnetic resonance imaging *Nano Biomed. Eng.* **11** 178–91
- Hampe N, Wolterink J M, Collet C, Planken N and Išgum I 2021 Graph attention networks for segment labeling in coronary artery trees *Proc SPIE*, **11596** 1159611
- Han Y, Zhang S, Geng Z, Wei Q and Ouyang Z 2019 Level set based shape prior and deep learning for image segmentation *IET Image Proc.* **14** 183–91
- Haque I R I and Neubert J 2020 Deep learning approaches to biomedical image segmentation *Inform. Med. Unlocked* **18** 100297
- Hatamizadeh A, Hoogi A, Sengupta D, Lu W, Wilcox B, Rubin D and Terzopoulos D 2019 Deep active lesion segmentation *Machine Learning in Medical Imaging* ed H I Suk et al vol 11861 (Berlin: Springer) pp 98–105
- Hatamizadeh A, Terzopoulos D and Myronenko A 2020 Edge-gated cnns for volumetric semantic segmentation of medical images arXiv:2002.04207
- Heinrich M P and Blendowski M 2016 Multi-organ segmentation using vantage point forests and binary context features *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2016* ed S Ourselin et al vol 9901 (Berlin: Springer) pp 598–606
- Hesamian M H, Jia W, He X and Kennedy P 2019 Deep learning techniques for medical image segmentation: achievements and challenges *J. Digit. Imaging* **32** 582–96
- Homayounfar N, Xiong Y, Liang J, Ma W-C and Urtasun R 2020 Levelset r-cnn: a deep variational method for instance segmentation *Computer Vision—ECCV 2020* ed A Vedaldi et al vol 12368 (Berlin: Springer) pp 555–71
- Hoogi A, Subramaniam A, Veerapaneni R and Rubin D L 2016 Adaptive estimation of active contour parameters using convolutional neural networks and texture analysis *IEEE Trans. Med. Imaging* **36** 781–91
- Hu P, Wu F, Peng J, Bao Y, Chen F and Kong D 2017 Automatic abdominal multi-organ segmentation using deep convolutional neural network and time-implicit level sets *Int. J. Comput. Assist. Radiol. Surg.* **12** 399–411
- Hu Y, Guo Y, Wang Y, Yu J, Li J, Zhou S and Chang C 2019 Automatic tumor segmentation in breast ultrasound images using a dilated fully convolutional network combined with an active contour model *Med. Phys.* **46** 215–28
- Humeau-Heurtier A 2019 Texture feature extraction methods: a survey *IEEE Access* **7** 8975–9000
- Ibragimov B, Toesca D, Chang D, Koong A and Xing L 2017 Combining deep learning with anatomical analysis for segmentation of the portal vein for liver sbrrt planning *Phys. Med. Biol.* **62** 8943–8958
- Iglesias J E and Sabuncu M R 2015 Multi-atlas segmentation of biomedical images: a survey *Med. Image Anal.* **24** 205–19
- Jayadevappa D, Srinivas Kumar S and Murty D 2011 Medical image segmentation algorithms using deformable models: a review *IETE Tech. Rev.* **28** 248–55
- Jiang Z, Lyu X, Zhang J, Zhang Q and Wei X 2020 Review of deep learning methods for mri brain tumor image segmentation *J. Image Graph.* **25** 215–28

- Juarez A G-U, Selvan R, Saghir Z and de Bruijne M 2019 A joint 3d unet-graph neural network-based method for airway segmentation from chest cts *Machine Learning in Medical Imaging* ed H I Suk et al vol 11861 (Berlin: Springer) pp 583–91
- Jurdia R E, Petitjean C, Honeine P, Cheplygina V and Abdallah F 2020 High-level prior-based loss functions for medical image segmentation: a survey arXiv:2011.08018v2 22 November 2020
- Kainz P, Pfeiffer M and Urschler M 2017 Segmentation and classification of colon glands with deep convolutional neural networks and total variation regularization *PeerJ* **5** e3874
- Karimi D, Dou H, Warfield S K and Gholipour A 2020 Deep learning with noisy labels: exploring techniques and remedies in medical image analysis *Med. Image Anal.* **65** 101759
- Karimi D and Salcudean S E 2019 Reducing the Hausdorff distance in medical image segmentation with convolutional neural networks *IEEE Trans. Med. Imaging* **39** 499–513
- Kass M, Witkin A and Terzopoulos D 1988 Snakes: active contour models *Int. J. Comput. Vis.* **1** 321–31
- Kervadec H, Bouchtiba J, Desrosiers C, Granger E, Dolz J and Ayed I B 2021 Boundary loss for highly unbalanced segmentation *Med. Image Anal.* **67** 101851
- Kim B and Ye J C 2019 Mumford–Shah loss functional for image segmentation with deep learning *IEEE Trans. Image Process.* **29** 1856–66
- Kipf T N and Welling M 2016a Semi-supervised classification with graph convolutional networks *Int. Conf. on Learning Representations 2017 (Toulon, France)* <https://openreview.net/pdf?id=SJU4ayYgl>
- Kipf T N and Welling M 2016b Variational graph auto-encoders arXiv:1611.07308 Bayesian Deep Learning Workshop (NIPS 2016)
- Le T H N, Gummadi R and Savvides M 2018 Deep recurrent level set for segmenting brain tumors *Medical Image Computing and Computer Assisted Intervention—MICCAI 2018* ed A Frangi et al vol 11072 (Berlin: Springer) pp 646–53
- Le T H N, Luu K, Duong C N, Quach K G, Truong T D, Sadler K and Savvides M 2020 Active contour model in deep learning era: a revise and review *Applications of Hybrid Metaheuristic Algorithms for Image Processing* ed D Oliva and S Hinojosa vol 890 (Berlin: Springer) pp 231–60
- Le T H N, Quach K G, Luu K, Duong C N and Savvides M 2018 Reformulating level sets as deep recurrent neural network approach to semantic segmentation *IEEE Trans. Image Process.* **27** 2393–407
- Lee L K, Liew S C and Thong W J 2015 A review of image segmentation methodologies in medical image *Advanced Computer and Communication Engineering Technology* ed H Sulaiman et al vol 315 (Berlin: Springer) pp 1069–80
- Li Q, Bai K, Zhao L and Guan X 2020a Progress and challenges of mri brain tumor image segmentation *J. Image Graph.* **25** 419–31
- Li R, Chen H, Gong G and Wang L 2020b Bladder wall segmentation in mri images via deep learning and anatomical constraints *Proc. Annual Int. Conf. IEEE Eng. Med. Biol. Soci. EMBS 2020* 1629–32
- Li S, Jiang H, Li H and Yao Y-D 2020c Aw-sdrlse: Adaptive weighting and scalable distance regularized level set evolution for lymphoma segmentation on pet images *IEEE J. Biomed. Health Inform.* **25** 1173–84
- Li W et al 2015 Automatic segmentation of liver tumor in ct images with deep convolutional neural networks *J. Comput. Commun.* **3** 146
- Li Z, Wang Y, Yu J, Shi Z, Guo Y, Chen L and Mao Y 2017 Low-grade glioma segmentation based on cnn with fully connected crf *J. Healthc. Eng.* **2017** 9283480
- Lin X and Li X 2019 Image based brain segmentation: From multi-atlas fusion to deep learning *Curr. Med. Imaging* **15** 443–52
- Litjens G, Kooi T, Bejnordi B E, Setio A A A, Ciompi F, Ghafoorian M, Van Der Laak J A, Van Ginneken B and Sánchez C I 2017 A survey on deep learning in medical image analysis *Med. Image Anal.* **42** 60–88
- Liu J, Shen C, Liu T, Aguilera N and Tam J 2019 Deriving visual cues from deep learning to achieve subpixel cell segmentation in adaptive optics retinal images *Ophthalmic Medical Image Analysis* ed H Fu et al vol 11855 (Berlin: Springer) pp 86–94
- Liu N and Han J 2016 Dhsnet: Deep hierarchical saliency network for salient object detection *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.* **2016** 678–86
- López-Linares K, Aranjuelo N, Kabongo L, Maclair G, Lete N, Ceresa M, García-Familiar A, Macía I and Ballester M A G 2018 Fully automatic detection and segmentation of abdominal aortic thrombus in post-operative cta images using deep convolutional neural networks *Med. Image Anal.* **46** 202–14
- Ma J, Deng Y and Ma Z 2020 Review of deep learning segmentation methods for ct images of liver tumors *J. Image Graph.* **25** 2024–46
- Ma J et al 2020 How distance transform maps boost segmentation cnns: an empirical study *Proc. 3rd Conf. on Medical Imaging with Deep Learning* vol 121 (PMLR) pp 479–92
- Ma Z, Wu X, Song Q, Luo Y, Wang Y and Zhou J 2018 Automated nasopharyngeal carcinoma segmentation in magnetic resonance images by combination of convolutional neural networks and graph cut *Exp. Therapeutic Med.* **16** 2511–21
- McInerney T and Terzopoulos D 1996 Deformable models in medical image analysis: a survey *Med. Image Anal.* **1** 91–108
- Millatary F, Navab N and Ahmadi S-A 2016 V-net: Fully convolutional neural networks for volumetric medical image segmentation *2016 4th Int. Conf. on 3D Vision (3DV) (Stanford, CA, 25–28 October 2016)* (Piscataway, NJ: IEEE) pp 565–71
- Močnik D, Ibragimov B, Xing L, Strojani P, Likar B, Pernuš F and Vrtovec T 2018 Segmentation of parotid glands from registered ct and mr images *Phys. Med.* **52** 33–41
- Morar A, Moldoveanu F and Gröller E 2012 Image segmentation based on active contours without edges *IEEE 8th Int. Conf. Intell. Comput. Commun. Process.* **2012** 213–20
- Mumford D B and Shah J 1989 Optimal approximations by piecewise smooth functions and associated variational problems *Commun. Pure Appl. Math.* **42** 577–685
- Murugesan B, Sarveswaran K, Shankaranarayana S M, Ram K, Joseph J and Sivaprakasam M 2019 Psi-net: Shape and boundary aware joint multi-task deep network for medical image segmentation *41st Annual Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)* **2019** 7223–6
- Myronenko A and Hatamizadeh A 2019 3d kidneys and kidney tumor semantic segmentation using boundary-aware networks arXiv:1909.06684 14 September 2019
- Navarro F, Shit S, Ezhov I, Paetzold J, Gafita A, Peeken J C, Combs S E and Menze B H 2019 Shape-aware complementary-task learning for multi-organ segmentation *Machine Learning in Medical Imaging* ed H I Suk et al (Berlin: Springer) pp 620–7
- Ngo T A, Lu Z and Carneiro G 2017 Combining deep learning and level set for the automated segmentation of the left ventricle of the heart from cardiac cine magnetic resonance *Med. Image Anal.* **35** 159–71
- Nosrati M S and Hamarneh G 2016 Incorporating prior knowledge in medical image segmentation: a survey arXiv:1607.01092 5 July 2016
- Nosrati M S, Peyrat J-M, Abinshed J, Al-Alao O, Al-Ansari A, Abugharbieh R and Hamarneh G 2014 Efficient multi-organ segmentation in multi-view endoscopic videos using pre-operative priors *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2014* ed P Golland et al vol 8674 (Berlin: Springer) pp 324–31
- Oktay O et al 2017 Anatomically constrained neural networks (acnns): application to cardiac image enhancement and segmentation *IEEE Trans. Med. Imaging* **37** 384–95

- Painchaud N, Skandarani Y, Judge T, Bernard O, Lalande A and Jodoin P-M 2020 Cardiac segmentation with strong anatomical guarantees *IEEE Trans. Med. Imaging* **39** 3703–13
- Qin W, Wu J, Han F, Yuan Y, Zhao W, Ibragimov B, Gu J and Xing L 2018 Superpixel-based and boundary-sensitive convolutional neural network for automated liver segmentation *Phys. Med. Biol.* **63** 095017
- Raissi M, Perdikaris P and Karniadakis G E 2019 Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations *J. Comput. Phys.* **378** 686–707
- Ramírez I, Martín A and Schiavi E 2018 Optimization of a variational model using deep learning: An application to brain tumor segmentation *IEEE 15th Int. Symp. Biomed. Imaging (ISBI 2018)* **2018** 631–4
- Rehman F, Shah S I A, Riaz M N, Gilani S O and Faiza R 2020 A region-based deep level set formulation for vertebral bone segmentation of osteoporotic fractures *J. Digit. Imaging* **33** 191–203
- Ronneberger O, Fischer P and Brox T 2015 U-net: convolutional networks for biomedical image segmentation *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015* ed N Navab et al vol 9351 (Berlin: Springer) pp 234–41
- Roy R, Chakraborti T and Chowdhury A S 2019 A deep learning-shape driven level set synergism for pulmonary nodule segmentation *Pattern Recognit. Lett.* **123** 31–8
- Rupprecht C, Huaroc E, Baust M and Navab N 2016 Deep active contours arXiv:1607.05074 18 July 2016
- Saha A, Hosseinzadeh M and Huisman H 2020 Encoding clinical priori in 3d convolutional neural networks for prostate cancer detection in bpmri arXiv:2011.00263v3
- Saman S and Narayanan S J 2019 Survey on brain tumor segmentation and feature extraction of mr images *Int. J. Multimedia Inf. Retr.* **8** 79–99
- Samuel P M and Veeramalai T 2020 Review on retinal blood vessel segmentation—an algorithmic perspective *Int. J. Biomed. Eng. Technol.* **34** 75–105
- Santner J, Unger M, Pock T, Leistner C, Saffari A and Bischof H 2009 Interactive texture segmentation using random forests and total variation *British Machine Vision Conference (London, United Kingdom, 7/September/09–10/September/09)* (BMVA Press) pp 66.1–12
- Scarselli F, Gori M, Tsoi A C, Hagenbuchner M and Monfardini G 2008 The graph neural network model *IEEE Trans. Neural Netw.* **20** 61–80
- Selvan R, Kipf T, Welling M, Juarez A G-U, Pedersen J H, Petersen J and de Bruijne M 2020 Graph refinement based airway extraction using mean-field networks and graph neural networks *Med. Image Anal.* **64** 101751
- Selvan R, Kipf T, Welling M, Pedersen J H, Petersen J and de Bruijne M 2018 Extraction of airways using graph neural networks *MIDL 2018 (Amsterdam)* <https://openreview.net/pdf?id=rkn2fjjjG>
- Selvan R, Welling M, Pedersen J H, Petersen J and de Bruijne M 2018 Mean field network based graph refinement with application to airway tree extraction *Medical Image Computing and Computer Assisted Intervention—MICCAI 2018* ed A Frangi et al vol 11071 (Berlin: Springer) pp 750–8
- Shen D, Wu G and Suk H-I 2017 Deep learning in medical image analysis *Annu. Rev. Biomed. Eng.* **19** 221–48
- Shen H, Wang R, Zhang J and McKenna S J 2017 Boundary-aware fully convolutional network for brain tumor segmentation *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2017* ed M Descoteaux et al vol 10434 (Berlin: Springer) pp 433–41
- Shin S Y, Lee S, Yun I D and Lee K M 2019 Deep vessel segmentation by learning graphical connectivity *Med. Image Anal.* **58** 101556
- Soberanis-Mukul R D, Navab N and Albarqouni S 2020 Uncertainty-based graph convolutional networks for organ segmentation refinement *Proc. 3rd Conf. on Medical Imaging with Deep Learning* vol 121 (PMLR) pp 755–69 <http://proceedings.mlr.press/v121/soberanis-mukul20a.html>
- Song B, Chou C-R, Chen X, Huang A and Liu M-C 2016 Anatomy-guided brain tumor segmentation and classification *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries* ed A Crimi et al vol 10154 (Berlin: Springer) pp 162–70
- Soomro T A, Afifi A J, Zheng L, Soomro S, Gao J, Hellwich O and Paul M 2019 Deep learning models for retinal blood vessels segmentation: a review *IEEE Access* **7** 71696–717
- Šprem J, De Vos B D, Lessmann N, Van Hamersvelt R W, Greuter M J, De Jong P A, Leiner T, Viergever M A and Išgum I 2018 Coronary calcium scoring with partial volume correction in anthropomorphic thorax phantom and screening chest ct images *PLoS One* **13** e0209318
- Sun L, Zhang L and Zhang D 2019 Multi-atlas based methods in brain mr image segmentation *Chin. Med. Sci. J.* **34** 110–9
- Szabo T L 2014 *Diagnostic Ultrasound Imaging: Inside Out* 2nd edn (New York: Academic) (<https://doi.org/10.1016/c2011-0-07261-7>)
- Szeliski R 2010 *Computer Vision: Algorithms and Applications* 1st edn (Berlin: Springer) (<https://doi.org/10.1007/978-1-84882-935-0>)
- Taghanaki S A, Abhishek K, Cohen J P, Cohen-Adad J and Hamarneh G 2021 Deep semantic segmentation of natural and medical images: a review *Artif. Intell. Rev.* **54** 1–42
- Tajbakhsh N, Jeyaseelan L, Li Q, Chiang J N, Wu Z and Ding X 2020 Embracing imperfect datasets: a review of deep learning solutions for medical image segmentation *Med. Image Anal.* **63** 101693
- Takikawa T, Acuna D, Jampani V and Fidler S 2019 Gated-scnn: gated shape cnns for semantic segmentation *Proc. IEEE Int. Conf. Comput. Vis.* **2019** 5228–37
- Tang M, Valipour S, Zhang Z, Cobzas D and Jagersand M 2017 A deep level set method for image segmentation *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support* ed M Cardoso (Berlin: Springer) pp 126–34
- Tian Z, Li X, Zheng Y, Chen Z, Shi Z, Liu L and Fei B 2020 Graph-convolutional-network-based interactive prostate segmentation in mr images *Med. Phys.* **47** 4164–76
- Tian Z, Zheng Y, Li X, Du S and Xu X 2020 Graph convolutional network based optic disc and cup segmentation on fundus images *Biomed. Opt. Express* **11** 3043–57
- Veličković P, Cucurull G, Casanova A, Romero A, Lio P and Bengio Y 2018 Graph attention networks *Int. Conf. on Learning Representations* https://openreview.net/references/pdf?id=HkKE_oELM
- Vrtovec T, Močnik D, Strojani P, Pernuš F and Ibragimov B 2020 Auto-segmentation of organs at risk for head and neck radiotherapy planning: from atlas-based to deep learning methods *Med. Phys.* **47** e929–50
- Wang Y, Wei X, Liu F, Chen J, Zhou Y, Shen W, Fishman E K and Yuille A L 2020 Deep distance transform for tubular structure segmentation in ct scans *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.* **2020** 3833–42
- Wasserman L 2018 Topological data analysis *Annu. Rev. Stat. Appl.* **5** 501–32
- Wickramasinghe U, Remelli E, Knott G and Fua P 2020 Voxel2mesh: 3d mesh model generation from volumetric data *Medical Image Computing and Computer Assisted Intervention—MICCAI 2020* ed A L Martel vol 12264 (Springer: Berlin) pp 299–308
- Wolterink J M, Leiner T and Išgum I 2019 Graph convolutional networks for coronary artery segmentation in cardiac ct angiography *Graph Learning in Medical Imaging* ed D Zhang et al vol 11849 (Berlin: Springer) pp 62–9
- Wolterink J M, van Hamersvelt R W, Viergever M A, Leiner T and Išgum I 2019 Coronary artery centerline extraction in cardiac ct angiography using a cnn-based orientation classifier *Med. Image Anal.* **51** 46–60

- Xie H, Pan Z, Zhou L, Zaman F A, Chen D, Jonas J B, Wang Y and Wu X 2020 Globally optimal segmentation of mutually interacting surfaces using deep learning arXiv:2007.01259v3 21 July 2020
- Xie L, Song Y and Chen Q 2020 Automatic left ventricle segmentation in short-axis mri using deep convolutional neural networks and central-line guided level set approach *Comput. Biol. Med.* **122** 103877
- Xu J, Gong L, Wang G, Lu C, Gilmore H, Zhang S and Madabhushi A 2019 Convolutional neural network initialized active contour model with adaptive ellipse fitting for nuclear segmentation on breast histopathological images *J. Med. Imaging* **6** 017501
- Xu X, Zhou F and Liu B 2018 Automatic bladder segmentation from ct images using deep cnn and 3d fully connected crf-rnn *Int. J. Comput. Assist. Radiol. Surg.* **13** 967–75
- Xue Y, Tang H, Qiao Z, Gong G, Yin Y, Qian Z, Huang C, Fan W and Huang X 2020 Shape-aware organ segmentation by predicting signed distance maps *Proc. AAAI Conf. on Artificial Intelligence* vol 34, pp 12565–72
- Yang H, Sun J, Li H, Wang L and Xu Z 2016 Deep fusion net for multi-atlas segmentation: application to cardiac mr images *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2016* ed S Ourselin et al vol 9901 (Berlin: Springer) pp 521–8
- Yang H, Zhen X, Chi Y, Zhang L and Hua X-S 2020 Cpr-gcn: Conditional partial-residual graph convolutional network in automated anatomical labeling of coronary arteries *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.* **2020** 3803–11
- Yang X, Wu N, Cheng G, Zhou Z, David S Y, Beitler J J, Curran W J and Liu T 2014 Automated segmentation of the parotid gland based on atlas registration and machine learning: a longitudinal mri study in head-and-neck radiation therapy *Int. J. Radiat. Oncol. Biol. Phys.* **90** 1225–33
- Yang X, Yu L, Li S, Wang X, Wang N, Qin J, Ni D and Heng P-A 2017 Towards automatic semantic segmentation in volumetric ultrasound *Medical Image Computing and Computer Assisted Intervention—MICCAI 2017* ed M Descoteaux et al vol 10433 (Berlin: Springer) pp 711–9
- Yao L, Jiang P, Xue Z, Zhan Y, Wu D, Zhang L, Wang Q, Shi F and Shen D 2020 Graph convolutional network based point cloud for head and neck vessel labeling *Machine Learning in Medical Imaging* ed M Liu et al vol 12436 (Berlin: Springer) pp 474–83
- Yedavalli V, Tong E, Martin D, Yeom K and Forkert N 2020 Artificial intelligence in stroke imaging: Current and future perspectives *Clin. Imaging* **69** 246–54
- Yi X, Walia E and Babyn P 2019 Generative adversarial network in medical imaging: a review *Med. Image Anal.* **58** 101552
- Yu L, Wang S, Li X, Fu C-W and Heng P-A 2019 Uncertainty-aware self-ensembling model for semi-supervised 3d left atrium segmentation *Medical Image Computing and Computer Assisted Intervention—MICCAI 2019* ed D Shen vol 11765 (Berlin: Springer) pp 605–13
- Yue Q, Luo X, Ye Q, Xu L and Zhuang X 2019 Cardiac segmentation from lge mri using deep neural network incorporating shape and spatial priors *Medical Image Computing and Computer Assisted Intervention—MICCAI 2019* ed D Shen vol 11765 (Berlin: Springer) pp 559–67
- Zabihollahy F, White J A and Ukwatta E 2019 Convolutional neural network-based approach for segmentation of left ventricle myocardial scar from 3d late gadolinium enhancement mr images *Med. Phys.* **46** 1740–51
- Zeng X, Staib L H, Schultz R T and Duncan J S 1998 Volumetric layer segmentation using coupled surfaces propagation *Proc. 1998 IEEE Comput. Soc. Conf. on Comput. Vis. Pattern Recognit. (Cat. No. 98CB36231)* **1998** 708–15
- Zhang D, Zhu W, Zhao H, Shi F and Chen X 2018 Automatic localization and segmentation of optical disk based on faster r-cnn and level set in fundus image *Proc. SPIE* **10574** 105741U
- Zhang H, Wang Y, Geng X and Fu P 2020a Review of deep learning methods for iso-intense infant brain mr image segmentation *J. Image Graph.* **25** 2068–78
- Zhang P, Zhong Y, Deng Y, Tang X and Li X 2019 A survey on deep learning of small sample in biomedical image analysis arXiv:1908.00473
- Zhang Q, Zhou J, Zhang B, Jia W and Wu E 2020b Automatic epicardial fat segmentation and quantification of ct scans using dual u-nets with a morphological processing layer *IEEE Access* **8** 128032
- Zhang Z, Fu H, Dai H, Shen J, Pang Y and Shao L 2019 Et-net: A generic edge-attention guidance network for medical image segmentation *Medical Image Computing and Computer Assisted Intervention—MICCAI 2019* ed D Shen (Springer: Berlin) pp 442–50
- Zhao H, Shi J, Qi X, Wang X and Jia J 2017 Pyramid scene parsing network *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)* **2017** 2881–90
- Zhao X, Wu Y, Song G, Li Z, Zhang Y and Fan Y 2018 A deep learning model integrating fcnn and crfs for brain tumor segmentation *Med. Image Anal.* **43** 98–111
- Zhao Y, Li H, Wan S, Sekuboyina A, Hu X, Tetteh G, Piraud M and Menze B 2019 Knowledge-aided convolutional neural network for small organ segmentation *IEEE J. Biomed. Health Inform.* **23** 1363–73
- Zhou H, Li X, Schaefer G, Celebi M E and Miller P 2013 Mean shift based gradient vector flow for image segmentation *Comput. Vis. Image Understand.* **117** 1004–16
- Zhou Y, Onder O F, Dou Q, Tsougenis E, Chen H and Heng P-A 2019 Cia-net: robust nuclei instance segmentation with contour-aware information aggregation *Information Processing in Medical Imaging* ed A Chung et al vol 11492 (Berlin: Springer) pp 682–93
- Zhou J, Cui G, Hu S, Zhang Z, Yang C, Liu Z, Wang L, Li C and Sun M 2020 Graph neural networks: a review of methods and applications *AI Open* **1** 57–81
- Zhuang X et al 2019 Evaluation of algorithms for multi-modality whole heart segmentation: an open-access grand challenge *Med. Image Anal.* **58** 101537