The Journal of Supercomputing
https://doi.org/10.1007/s11227-021-03901-6

Deep learning techniques for tumor segmentation: a review

Huiyan Jiang[1,2] · **Zhaoshuo Diao**[1] · **Yu-Dong Yao**[3]

Accepted: 19 May 2021

Abstract
Recently, deep learning, especially convolutional neural networks, has achieved the remarkable results in natural image classification and segmentation. At the same time, in the field of medical image segmentation, researchers use deep learning techniques for tasks such as tumor segmentation, cell segmentation, and organ segmentation. Automatic tumor segmentation plays an important role in radiotherapy and clinical practice and is the basis for the implementation of follow-up treatment programs. This paper reviews the tumor segmentation methods based on deep learning in recent years. We first introduce the common medical image types and the evaluation criteria of segmentation results in tumor segmentation. Then, we review the tumor segmentation methods based on deep learning from technique view and tumor view, respectively. The technique view reviews the researches from the architecture of the deep learning and the tumor view reviews from the type of tumors.

Keywords Tumor segmentation · Deep learning · Medical image · Review

1 Introduction

At present, cancer has become the leading cause of death in China and even in the world, posing a huge threat to people's lives and health. The latest computer tomography (CT), magnetic resonance imaging (MRI), ultrasound (US), Positron emission tomography (PET) and other multimodal medical imaging technologies can achieve high-resolution imaging of tumor disease areas. The segmentation of

✉ Huiyan Jiang
hyjiang@mail.neu.edu.cn

[1] Software College, Northeastern University, Shenyang 110819, China

[2] Key Laboratory of Intelligent Computing in Medical Image, Ministry of Education, Northeastern University, Shenyang 110819, China

[3] Department of Electrical and Computer Engineering, Stevens Institute of Technology, Hoboken, NJ 07030, USA

Published online: 21 June 2021

tumor area from the above image is the basis of radiotherapy and clinical diagnosis. Relying on manual segmentation not only requires experienced radiologists, but also time-consuming and labor-intensive, so automated tumor segmentation is necessary. With the great achievements of deep learning in the field of natural image segmentation and the achievements of U-Net [1] in the field of medical image segmentation, tumor segmentation based on deep learning has become the actual standard of automatic tumor segmentation.

Although the characteristics of different tumors are not the same, the deep learning tumor segmentation methods are all based on the fully convolutional networks (FCN) [2]. Therefore, it is more important for researchers to understand the motivations for these improvements than to know what improvements have been made. Based on the understanding of the motivation of method improvement, researchers can better deal with the problem of tumor segmentation. Although the segmentation methods of different tumors are not the same, there is only one purpose of tumor segmentation, which is precise segmentation of tumor. It is consistent with the purpose of natural image segmentation. In addition, there is a special problem with tumor segmentation, which is insufficient data, especially the lack of labeled data. Therefore, all tumor segmentation methods focus on precision segmentation and lack of data. Precise segmentation of tumors can be achieved by three methods, which are to obtain more information for pixel classification, remove irrelevant areas in the image, and segmentation with clear boundaries. Based on the above motivation, this article first reviews the general methods of tumor segmentation from the perspective of network architecture and then reviews from the tumor perspective based on the characteristics of different tumors.

This paper reviews and combs the tumor segmentation algorithms based on deep learning. The algorithms are reviewed from two views as shown in Fig. 1: the technical architecture view and the tumor type view. From the perspective of technical architecture, the classic supervised learning architecture includes four parts: preprocessing, network structure, loss function and post-processing. In addition, there are advanced topics such as multimodal, semi-supervised and transfer learning. Compared with previous reviews of segmentation of medical and natural images [3–6] and review of segmentation of brain tumors [7–12], the main contributions of this paper as follows:

(1) This paper is the first review of tumor segmentation based on deep learning methods as we know.
(2) This paper first reviews from the perspective of technical architecture and divides it into six parts: network structure, loss function, preprocessing, post-processing, multimodality and semi-supervising. Then, from the view of tumor type, this paper selects brain tumor, breast tumor, colorectal cancer, kidney tumor, liver tumor, lung tumor and nasopharyngeal cancer, a total of seven common tumors are reviewed, which are convenient for researchers to consult.
(3) In addition to reviewing the previously published researches, we also summarized the problems that need to be solved in tumor segmentation based on the

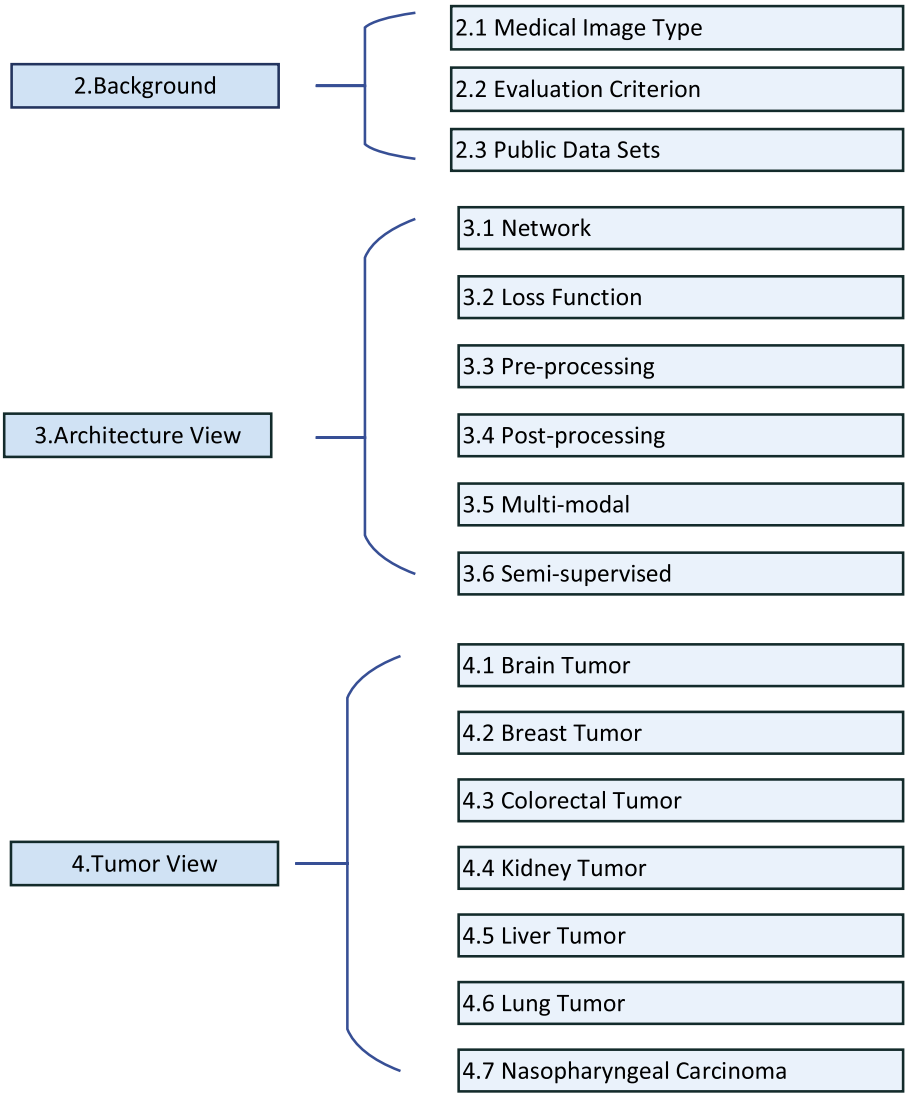

Fig. 1 Flow of the review. The tumor segmentation algorithms based on deep learning are classified from the perspective of deep learning architecture and tumor application, respectively. The detail of architecture view is shown in Fig. 2

difficulties of different tumors. Also, we use this as a motivation to analyze and compare the current published results. Finally, based on this, the paper summarizes some problems that have not been solved thoroughly and extends the possible research directions in the future.

This paper is organized as shown in Fig. 1. In Sect. 2, we introduce the background knowledge of tumor segmentation, including common biomedical image types, evaluation indicators and some public data set statistics. In Sect. 3, we review the method from the architecture view and introduce network structure, loss function, preprocessing, post-processing, multimodality and semi-supervising. We review the method from seven common tumors in Sect. 4. Section 5

introduces the future research direction of tumor segmentation. In Sect. 6, we conclude the whole paper.

2 Background knowledge of tumor segmentation

2.1 Medical images type

CT Images CT image is a cross-sectional scanning around a certain part of human body by using accurate and collimated X-rays, gamma rays, ultrasound, etc., together with highly sensitive detector [14]. Therefore, it has the characteristics of

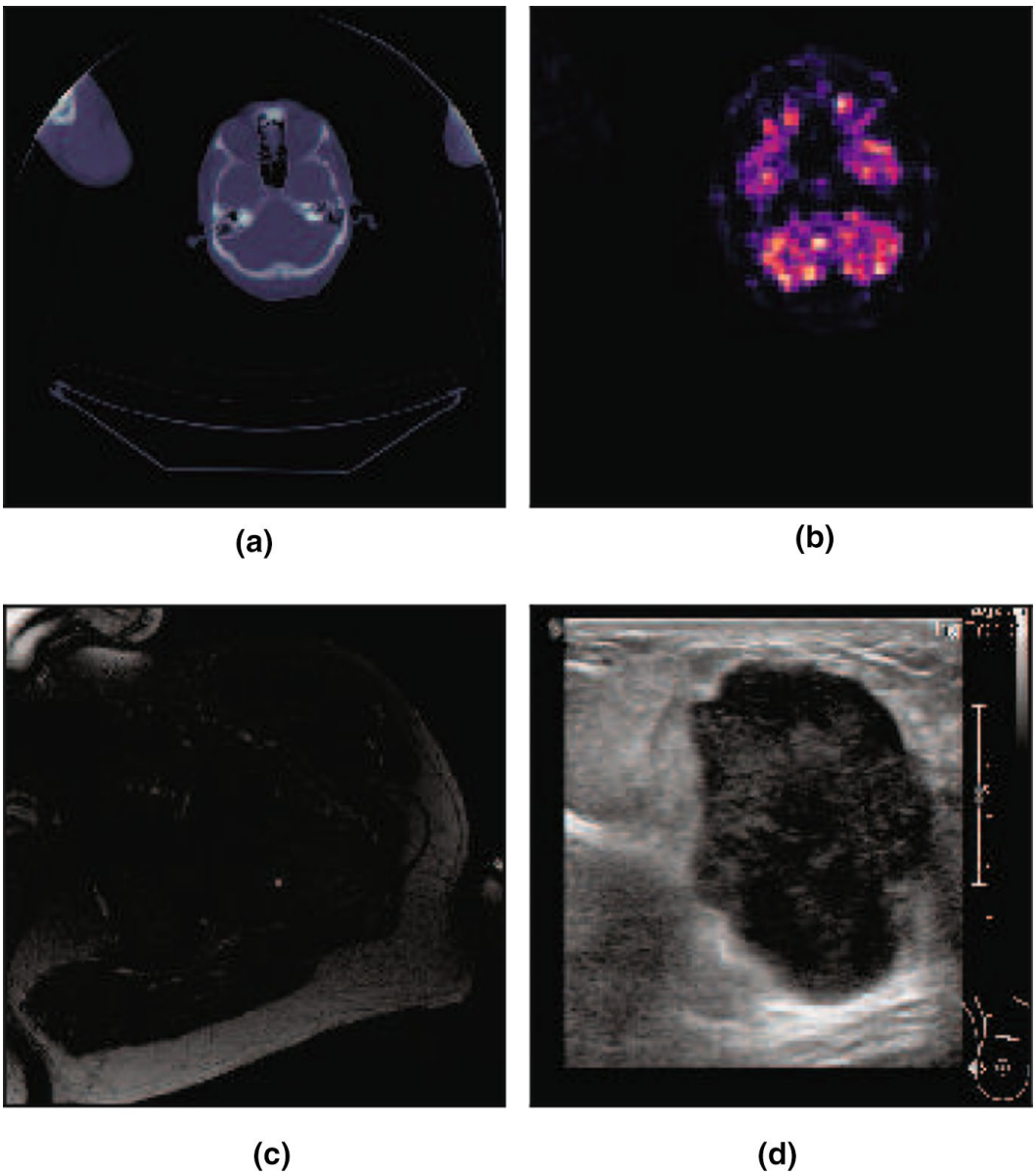

Fig. 2 Different types of medial image [13]. **a** CT image. **b** PET image. **c** MR image and **d** US image

fast scanning time and clear images. CT images play an important role in the diagnosis of neurological diseases, head and neck tumors and chest tumors [15]. The CT image of soft tissue sarcoma is shown in Fig. 2a. Low contrast is a hindrance during tumor diagnosis when using the CT images.

PET Images PET imaging is a relatively advanced clinical laboratory technology in the field of nuclear medicine. The general method is to label the short-lived radioactive elements with certain substances, which are generally necessary for the metabolism of biological life, such as glucose and protein [16–18]. After being injected into the human body, the metabolism of the substance is used to reflect the metabolism of life. Therefore, the PET image reflects the metabolic information, the position where metabolism is vigorous, and the density value in the PET image is large. Because the tumor area is generally rich in metabolism, PET images are generally used in combination with CT images to become PET-CT, which is mainly used as the standard for tumor diagnosis [19]. Figure 2b shows a PET image of soft tissue sarcoma. PET images are characterized by high contrast, but lack in low resolution.

MR Images MR image is a kind of tomographic imaging, as shown in Fig. 2c. It uses MRI phenomenon to obtain electromagnetic signal from human body and reconstruct human body information [20]. Compared with CT image, MR image can show the soft tissue structure clearly, and it is better than CT in the examination of central nervous system, bladder and rectum. MR images can be divided into T1 weighted image and T2 weighted image according to different weights [21].

US Images Ultrasound image is the use of ultrasonic beam scanning the human body, through the process of reflected signals to obtain an image of the body organs. Ultrasound imaging is widely used in ophthalmology, obstetrics, gynecology, cardiovascular system, digestive system and urinary system to determine the location, size and shape of organs, determine the scope and physical properties of lesions, provide anatomical maps of some gland tissues and identify the normal and abnormal fetus [22–24]. In tumor diagnosis, generally speaking, benign lesions have uniform texture and single interface, so the echo is uniform and regular. While in malignant lesions, the tumor tissue interface is complex and uneven, showing irregular echo structure by reason of the rapid growth, hemorrhage and degeneration [25]. Ultrasound image is shown in Fig. 2d.

2.2 Evaluation criterion

Tumor segmentation is actually a pixel-level classification problem. Then define image X, for all pixel in image X, we divide them into two classes which one is the tumor region and the other is the tissue area. The evaluation criterions include classification-based criterion, overlap-based criterion and surface-based criterion. Define V_p is the segmentation results of the method we proposed and V_g is the segmentation results from experienced radiologists or the ground-truth.

Classification-based Criterion For classification problems, the frequently used recall and precision indicators are also used in tumor segmentation, and their

indicator names have become positive predictive value (PPV) and sensitivity (SE) [26–28]. The calculation formulas are:

$$\mathrm{PPV} = \frac{V_p \cap V_g}{V_p} \tag{1}$$

$$\mathrm{SE} = \frac{V_p \cap V_g}{V_g} \tag{2}$$

where SE corresponds to the recall in the classification and PPV corresponds to the precision in the classification.

Overlap-based Criterion Overlap-based criterion compares the similarity and difference between the predicted tumor area and the ground-truth. As similarity coefficient (DSC) and Jaccard similarity coefficient (JSC) [29–31], they are calculated as:

$$\mathrm{DSC} = 2 \times \frac{\left|V_p \cap V_g\right|}{\left|V_p\right| + \left|V_g\right|} \tag{3}$$

$$\mathrm{JSC} = \frac{\left|V_p \cap V_g\right|}{\left|V_p \cup V_g\right|} \tag{4}$$

When the prediction result is completely consistent with the ground-truth, the DSC and JSC values are both 1. Also, JSC value corresponds to the Intersection Over Union (IoU) in the nature image segmentation [32]. Overlap-based criterions are widely used in image segmentation, especially medical image segmentation. In particular, the DCS value has actually become the most important criterion for evaluating the effect of medical image segmentation.

Surface-based Criterion In tumor segmentation, precise boundaries are also important. Therefore, for the prediction of tumor boundaries, surface-based evaluation criterions are generally used. A surface distance error (SDE) will be obtained for each pixel in the predicted tumor region V_p to measure the boundary prediction [33]. SDE is defined as the Euclidean distance between the pixel in V_p and its pixel closest pixel in V_g. Normally we use the average symmetric surface distance (ASSD) and it is defined as:

$$\mathrm{ASSD} = \frac{\sum_{a \in V_p} \min_{b \in V_g} d(a,b)}{2\left|V_p\right|} + \frac{\sum_{b \in V_g} \min_{a \in V_p} d(a,b)}{2\left|V_g\right|} \tag{5}$$

where $d(a, b)$ is the Euclidean distance between a and b. The smaller ASSD value indicates the smaller difference between the predicted tumor boundary and the ground-truth boundary. The smaller the ASSD value, the better the prediction result.

Classification-based criterion focuses on the classification of independent pixels and cannot well reflect the segmentation of the whole tumor area. Therefore, classification-based criterion is not generally used as a standard to evaluate the quality of the tumor segmentation results. Classification-based evaluation criterion is generally used as a supplement to other evaluation criterions. Overlap-based criterion is the most widely used evaluation criterion. It has actually become the primary evaluation criterion for tumor segmentation methods, and it is also the ranking basis for various competitions [34–39]. However, there are two main problems in overlap-based criterion. One is that it is not sensitive to the boundary, and the other is that the dice value of small objects fluctuates too much. For example, it is possible that the DSC value is high, but the tumor boundary of the predicted results is far from the ground-truth. In small object segmentation, the change of several pixels may cause the huge fluctuation of dice. Surface-based criterion circumvents the above two situations by establishing an evaluation method for the tumor surface. Therefore, in the task of tumor segmentation, if the tumor size is balanced, the overlap-based criterion is generally used as the standard to measure the quality of the method. If the tumor is small or the tumor fluctuates greatly, the surface-based criterion is better.

2.3 Public data sets

In the tumor segmentation task, there are also many publicly available data sets, which make comparison of some studies possible. In the public data set of tumor segmentation, most of the public data sets consist of a specific type of tumor, so this paper summarizes some of the more commonly used public data sets as shown in Table 1.

Table 1 Partial public data set in the field of medical image segmentation

Data set	Tumor type	Image modality	Year
KiTS19 [34]	Kidney tumor	CT	2019
BRATS2015 [35]	Brain tumor	MR	2015
BRATS2016 [36]	Brain tumor	MR	2016
BRATS2017 [37]	Brain tumor	MR	2017
BRATS2018 [38]	Brain tumor	MR	2018
BRATS2019 [39]	Brain tumor	MR	2019
Warwick-CRCHP [40]	Colorectal tumor	Histology images	2016
Warwick-CRAG [41]	Colorectal tumor	Histology images	2019
Warwick-ECCGD [42]	Colorectal tumor	Histology images	2020
Breast-Data set [13]	Breast tumor	US	2019
LiTS [43]	Liver tumor	CT	2019
3DIRCADb [44]	Liver tumor	CT	2018
LIDC-IDRI [28]	Lung tumor	CT	2010
TCIA-NSCLC [45]	Lung tumor	CT	2018

Including tumor type, image modality and release year

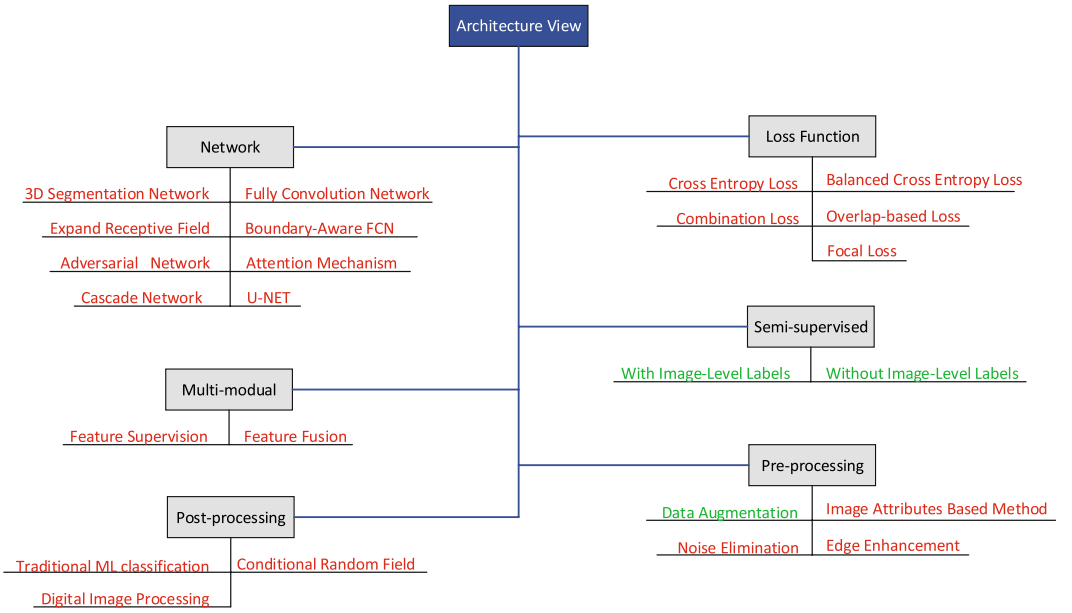

Fig. 3 Diagram of the review of architecture. The red font indicates that the purpose of this method is precise segmentation of tumor and the green indicates the purpose is the solution of lack data (color figure online)

In addition, researchers can also go to the TCIA official website (https://wiki.cancerimagingarchive.net) to search for tumor segmentation data sets published by the other researchers according to tumor type.

Table 2 Comparison of network

Motivation	Network	Parameters (MB)	Prediction time (s)	Year
Segmentation with clear boundaries	U-NET [1]	141	0.05	2015
	Boundary-Aware FCN [46]	204	0.07	2017
Remove irrelevant areas in the image	Adversarial network [47]	211	0.07	2018
	Cascaded Network [48]	280	0.11	2019
Obtain more information for pixel classification	3D U-NET [49]	150	0.32	2016
	PSPNet [50]	132	0.04	2017
	Attention U-Net [51]	110	0.03	2018
	H-DenseUnet [52]	180	0.14	2018
	Dilated convolution [53]	148	0.05	2019

The prediction time and parameters size are obtained by using a single 512×512 slice on two 11G 2080Ti machines

3 Review from architecture view

In this section, we will overview from the architecture view and review it from six aspects as shown in Fig. 3.

3.1 Network

As can be seen from Fig. 3, the purpose of all network structures is precise segmentation. However, each method has a different starting point to improve precise segmentation. Precise segmentation of tumors can be achieved by three methods, which are to obtain more information for pixel classification, remove irrelevant areas in the image, and segmentation with clear boundaries. Table 2 describes the comparison of various methods except FCN, because FCN is the backbone of all these methods. The table introduces the motivation of each method and compares their parameters and prediction time. The prediction time is obtained by using a single 512×512 slice on two 11G 2080Ti machines.

3.1.1 Fully convolution network (FCN)

Fully convolutional network (FCN) are first proposed by Shelhamer et al. [2], transforming convolutional neural networks [53–57] used in the field of image classification into image segmentation. The network structure diagram of FCN is shown in Fig. 4. The head of FCN is the same as the AlexNet [57], but removes

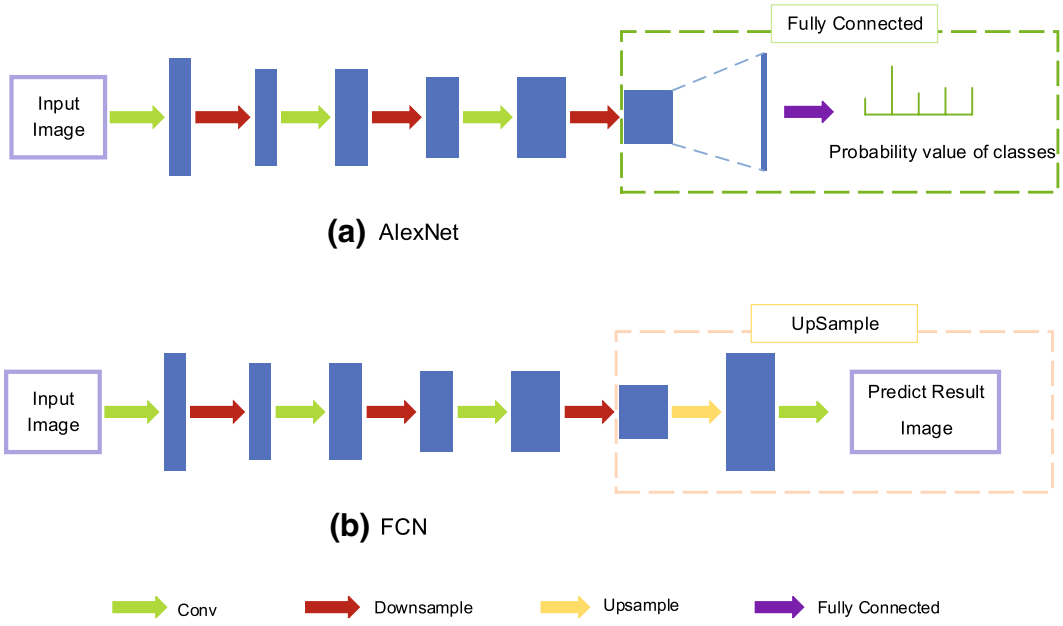

Fig. 4 Overview of the FCN [2]. FCN transformed the traditional AlexNet into a fully connected convolutional network, changing the original last fully connected layer to an upsampling and convolutional layer to obtain semantic segmentation results

the full connection layer of the classification network and replaces it with the upper sampling and convolution layer, so that it can get a prediction result for each pixel of the image, indicating what kind of image the pixel belongs to. Due to the downsampling process such as pooling, the image becomes blurred, especially the boundaries that are more important for segmentation become blurred. Therefore, how to use upsampling to recover the original information of the image is the key to the FCN network structure. The FCN uses deconvolution, while SegNet [58] saves the pixel index during the downsampling, and recovers pixel using index when upsampling. Meanwhile, SegNet proposes to refer to the downsampling stage of the full convolutional neural network as the encoder stage, the upsampling stage as the decoder stage, and the classical encoder–decoder network architecture. In the tumor segmentation task, some researchers directly use the FCN network as the backbone structure for tumor segmentation [59–62].

3.1.2 U-NET

The U-NET network, proposed by Ronneberger et al. [1], is one of the most successful networks based on the FCN structure and has now become a baseline standard network architecture in medical image segmentation. U-NET network architecture is shown in Fig. 5, which is based on the classic encoder–decoder architecture proposed by SegNet. It uses skip-connection to solve the problem of image edge information loss caused by downsampling. In the upsampling phase, the feature map generated in the corresponding downsampling process and the feature map after upsampling are fused, and this method is called skip-connection. Fusion methods include direct concatenation and pixel-wise summing. In addition to U-NET, 3D U-NET [49] and V-NET [63] adapting U-NET to 3D image environments are also

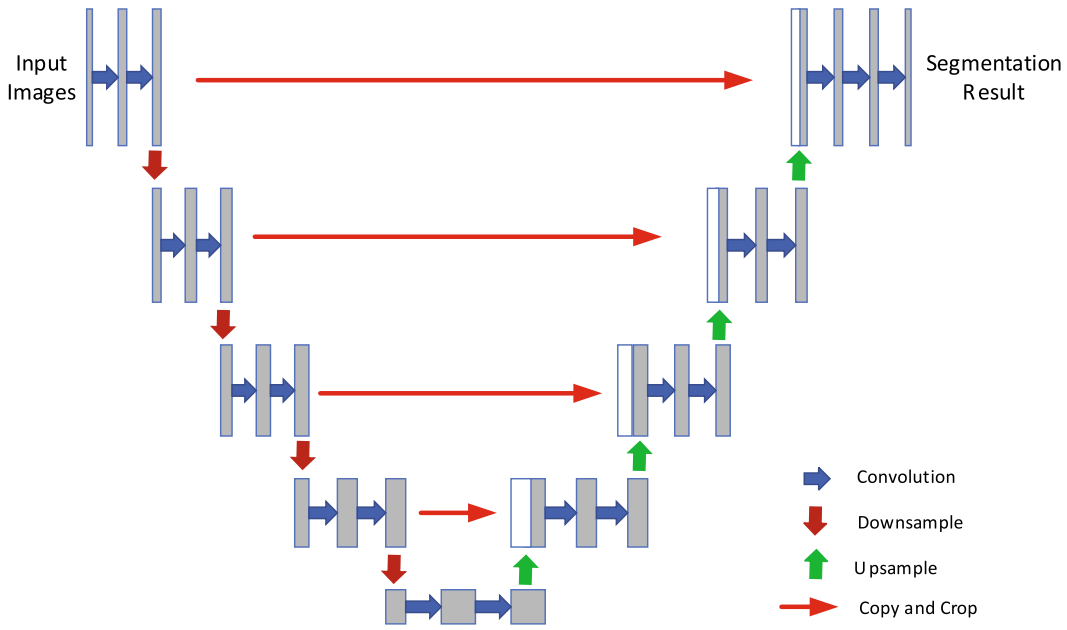

Fig. 5 U-NET [1] network architecture

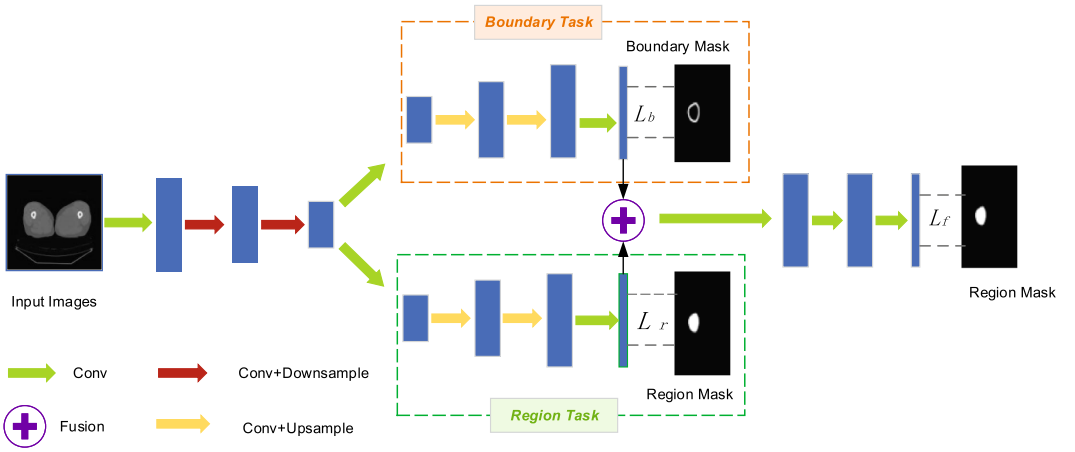

Fig. 6 Overview of the Boundary-Aware FCN [46]. Boundary-Aware transform the segmentation task into a multiple task. One task aiming at segmenting the boundary, called Boundary Task. And the other task aiming at segmenting tumor, called Region Task. Finally, the output from the two task are fused and pass the convolution layer and output the predicted result. Blue box is the layer shared by two task, orange box is the layer belong to Boundary task and green box is the layer belong to Region Task. L_b, L_r, L_f are the loss functions of the two branches and the final segmentation, respectively (color figure online)

widely used. U-NET is used as a baseline backbone in tumor segmentation due to its powerful effect and few training parameters [64–67].

3.1.3 Boundary-Aware FCN

Although U-NET solves the problem of how to restore the original information during the upsampling in FCN through skip-connection. But for the boundary, it still cannot get a good segmentation result, because the boundary pixels and the internal pixels are relatively blurred. The reason for this ambiguity is that even in the first convolution layer, convolution operators can cause similar values in the voxel feature map near the tumor boundary. Shen et al. [46] propose a Boundary-Aware FCN network for boundary segmentation, and the network structure is shown in Fig. 6. The Boundary-Aware FCN turns the segmentation problem from a single network into a multitasking network. The two FCN networks share the downsampling phase

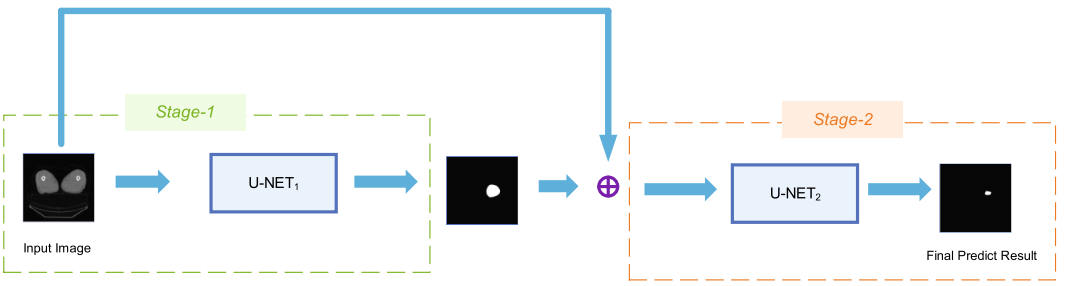

Fig. 7 Overview of the cascaded network architecture [48]. The network contains two stage. In the first stage, a rough segmentation result is made. In the second stage, the segmentation result of the first stage is merged with the original image and sent to the U-NET network

and have their own upsampling phase. The two upsampling correspond to different segmentation tasks, one segments the tumor region and the other segments the tumor boundary. The two segmentation results are then fused together, and after several convolutional layers, the final segmentation result is outputted.

3.1.4 Cascaded network

The relationship between tumor segmentation and organ segmentation is that tumor segmentation requires the location of tumor on the basis of organ segmentation. Based on this, a cascade network is proposed. The first neural network finds out the organ of the lesion, called rough segmentation, and second neural networks for accurate tumor segmentation. Figure 7 shows a classical two-stage cascaded tumor segmentation network proposed by Gruber et al. [48]. The entire network structure consists of two U-NET networks. First, the original image passes through the first U-NET network to obtain a binary image. Then, the image is multiplied with the original image and then sent to the second U-NET network. The output result is the final segmentation result [60, 62] all use similar methods for tumor segmentation. In the cascade network structure, there are still some studies using three networks for cascade, that is, the first two are rough segmentation, and the last one is precise segmentation [65, 67, 68] use a three-layer cascade network, and proves that the cascade tumor segmentation framework has a good ability to deal with the situation of positive and negative sample imbalance.

3.1.5 Expand the receptive field

In convolutional neural networks, the purpose of pooling or downsampling is to increase the receptive field of a single pixel by reducing the resolution of the image. If the receptive field of a single pixel increases, the richer the contextual information contained in a single pixel, the more robust the final segmentation result is and the better the effect is. Since the task of tumor segmentation is to predict the class of each pixel, enlarging the receptive field by adding downsampling will result in the

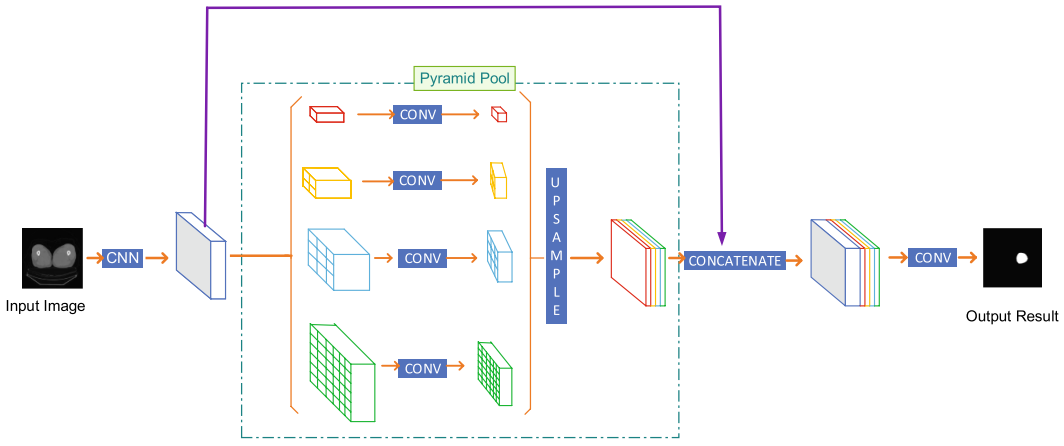

Fig. 8 PSPNet architecture [50]

inability to restore the original information of the image in the process of upsampling, resulting in the loss of information, and ultimately causing unsatisfactory segmentation effect. Gauss pyramid and dilated convolution is two common methods to increase the receptive field in image segmentation.

Applying Gauss pyramids in convolutional neural networks is first proposed in PSPNet by Zhao et al. [50] and PSPNet network architecture is shown in Fig. 8. After the original image has undergone convolution operation, the feature maps obtained are then subjected to several convolution operations with different convolution kernels and strides to obtain several feature maps of different sizes. Then, these feature maps are converted into the same size by upsampling, connected by pixel addition or concatenation and entered into subsequent convolution operations. The receptive fields of the feature maps obtained through different convolution kernels and stride sizes are different. The feature maps of the smaller receptive fields can well retain the original information and the feature maps of the bigger receptive fields can contain more information. Yang et al. [60] and Sarker et al. [69] used Gaussian pyramids to expand the receptive field in tumor segmentation.

Dilated convolution is another way to increase the receptive field. Dilated convolution is first proposed by Li et al. [70]. The introduction of dilated convolution is shown in Fig. 9. Dilated convolution achieves the effect of obtaining different receptive fields by setting some 0 in the convolution kernel. The feature map after dilated convolution is consistent with the original image size, so it will not cause information loss. Qin et al. [71] use dilated convolution to expand the receptive field during brain tumor segmentation. Hu et al. [59] also use dilated convolution in breast tumor segmentation. In addition to Gaussian pyramids and dilated convolutions, Mlynarski et al. [72] use slices from different angles to simultaneously extract features to increase the receptive field from the physical level. Jiang et al. [73] use multiple residual networks ResNet [53] to extract different resolution features to increase the receptive field.

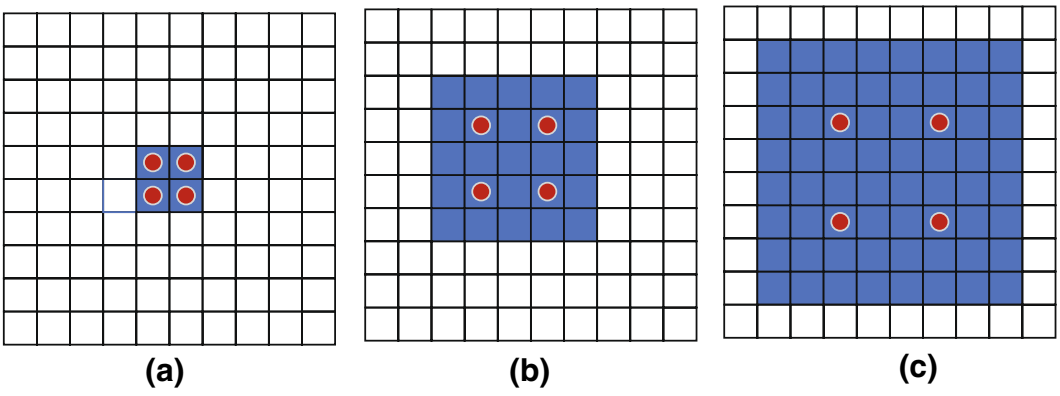

Fig. 9 Dilated convolution [70] diagram. The abc all represent the 2×2 convolution kernel. **a** means the dilation rate is 0. **b** means the dilation rate is 1 and **c** means the dilation rate is 2. The blue area represents the receptive filed of convolution kernel and read point represents the pixel participate in the convolution operation

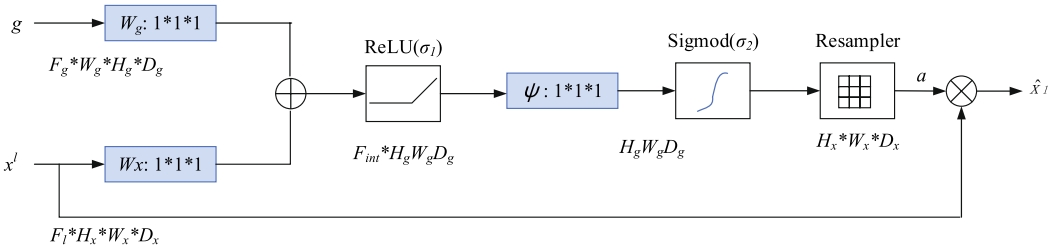

Fig. 10 Attention gate diagram [51]. g is the feature map from the upsampling and x^l is the corresponding feature map from the downsampling. The attention replaces the direct fusion of the upsampling feature map and downsampling feature map

3.1.6 Attention mechanism

Attention mechanism is based on human vision, which is first proposed in machine translation [74]. Oktay et al. [51] add the Attention mechanism to U-NET and propose the attention U-NET. At the end of each skip-connection, for the fusion of two feature maps, uses the attention gate to replace the original per-pixel summation and direct concatenation. The network structure of attention U-NET is shown in Fig. 10, where g is the upsampling feature map and x^l is the corresponding downsampling feature map. The two feature map first go through a 1×1 convolution to get W_g and W_x. After the ReLU and Sigmod functions, the attention coefficient is obtained, and the attention coefficient is then multiplied by x^l to obtain the final result. Jiang et al. [75] apply the attention mechanism to the segmentation of liver tumors and proposed the AHCNet network. Xu et al. [76] apply the attention mechanism to a cascaded network structure and propose a DCAN network for brain tumor segmentation. Sabarinathan et al. [77] fuse attention mechanism and coordinated convolution to process kidney tumor segmentation.

3.1.7 Adversarial tumor segmentation network

Adversarial learning proposed by Goodfellow et al. [78] is to learn generative networks using adversarial learning. The adversarial network generally consists of a generator and a discriminator. The purpose of the generator is to generate samples from a bunch of initial noise. Then, these generated samples and original samples pass through the discriminator and output whether these samples are outputted by the generator. Therefore, the purpose of the generator is to fool the discriminator as much as possible. The purpose of the discriminator is to identify the samples outputted by the generator as much as possible. Finally, the network can get a satisfactory generator in the process of adversarial learning. [47, 79–81] introduce adversarial networks into tumor segmentation tasks. A typical adversarial tumor segmentation network is shown in Fig. 11. The U-NET segmentation network is used as a generator in the adversarial network, and the segmentation result after U-NET is put into a discriminator at the same time as the ground-truth image. The discriminator is a common classification network.

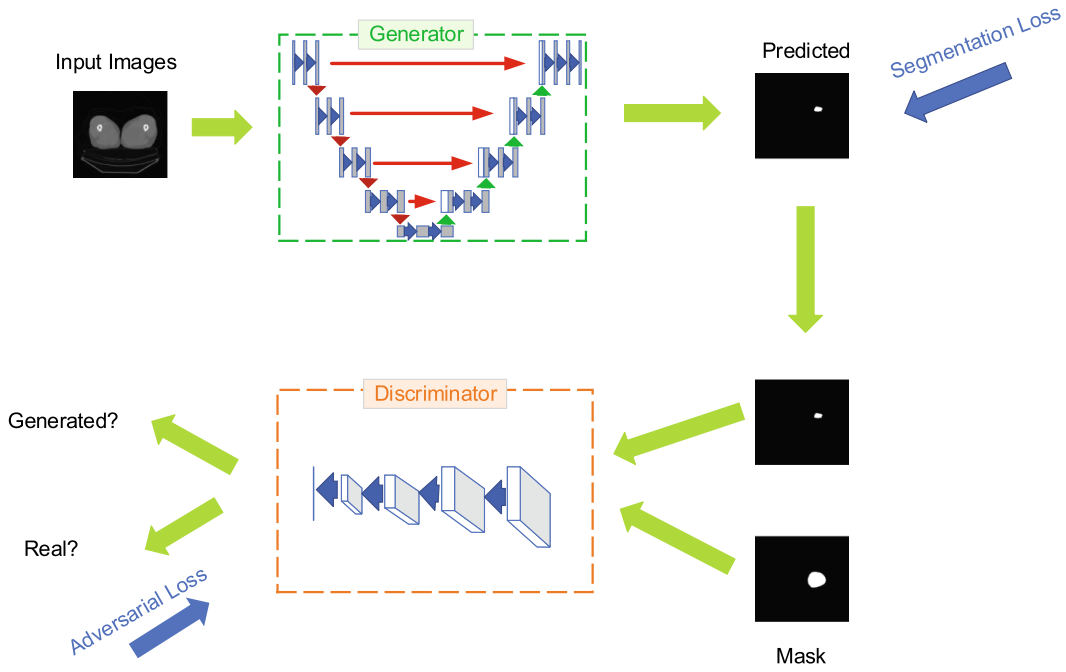

Fig. 11 Adversarial network [47]. In the adversarial network, the U-NET is used as a generator to output the prediction segmentation result, and then the prediction result and the ground-truth are input into the discriminator network. The loss function of the entire network includes two parts, one is the segmentation loss, and the other is the adversarial loss of the discriminator

3.1.8 3D Tumor segmentation network

A main difference between medical images and natural images is that during a medical imaging process, the whole body is scanned to generate consecutive slices, especially in the clinical practice of tumor diagnosis. Therefore, the medical image is naturally 3D, and most of the natural image processing methods focus on processing a single image, lacking the processing of context information for continuous slices. Although 3D U-NET [49], V-NET [63] and DeepMedic [82] directly process consecutive 3D slices, 3D convolution has many problems, such as many parameters, high computational resources and easy overfitting, so it has not become the first

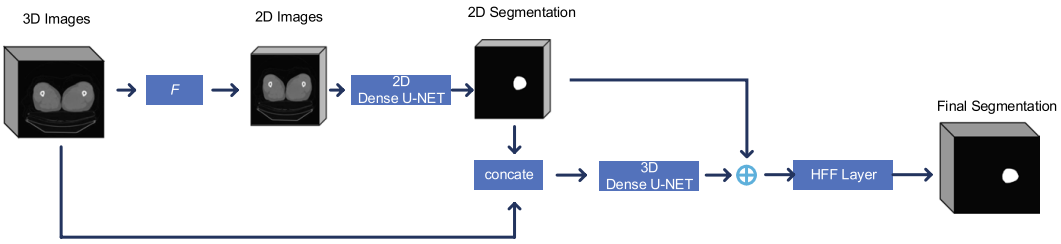

Fig. 12 H-DenseUNet [52] architecture. First, the 3D image is cut into 2.5D continuous slices through the F operation, and the segmentation results are obtained through 2D U-NET. Second, the segmentation results are concatenated with 3D input, and then the features between the slices are extracted through 3D U-NET. Finally, the 2D segmentation result and former result pass through the HFF layer, and the final segmentation result is output

choice for researchers to deal with 3D slices. There are two main solutions to deal with 3D slices, one is to convert them into 2.5D which contains one image along with its front and back images for processing [83], and the other adds a network structure to deal with information between slices directly after processing the 2D slice [52, 84].

The method of converting 3D slices into 2.5D slices still uses the 2D network structure, but the slices sent to the network are no longer only the current slice, but include several previous and after slices of the current slice, which are added to the channel of the current slice to handles the context information between slices. Han [83] uses 2.5D method to treat liver tumor segmentation. The typical network of the second processing method is the H-DenseUNet [52] network. The network structure is shown in Fig. 12. For each 3D input, the 3D volume block is transformed into a 2.5D slice with adjacent slices through the transformation processing function F. These 2.5D slices are then sent to 2D DenseUNet to extract inner-slice features. The 3D raw input and the predicted result after the 2D DenseUNet transformation are concatenated and sent together into the 3D network to extract the inter-slice features. Finally, the two features are fused and the final result is predicted through the HFF layer. Tseng et al. [84] use LSTM [85] to process context information between slices.

3.2 Loss function

Before summarizing the deep learning-based loss function, we first make some definitions. Let $x_i \in X$ represent the pixel in image, y_i is the class of the pixel x_i, p_i is the predicted class of the pixel x_i.

3.2.1 Cross-entropy loss

In the field of image segmentation, including natural image segmentation and medical image segmentation, cross-entropy loss is the most commonly used loss function. The cross-entropy loss of the binary classification is defined as:

$$L_{\mathrm{CE}}^{\mathrm{binary}} = \sum_{i=1}^{N} -\left(y_i \log p_i + \left(1 - y_i\right) \log\left(1 - p_i\right)\right) \tag{6}$$

Two classifications is that y_i have only two values, 0 or 1. Also, $P(y_i = 1) = p_i$ and $P(y_i = 0) = 1 - p_i$. For general multi-classification problems, the definition of cross-entropy loss is as:

$$L_{\mathrm{CE}}^{\mathrm{multi}} = \sum_{c=1}^{C} \sum_{i=1}^{N} -y_i^c p_i^c \tag{7}$$

where c is the class of the pixel, y_i^c means that the class of the pixel x_i is c and p_i^c is the predict probability value that x_i belong to class c.

3.2.2 Balanced cross-entropy loss

In the field of medical segmentation, especially tumor segmentation, the proportion of tumor area is very low in one image, and most pixels are background or normal tissue, that is, the proportion of positive and negative samples is seriously imbalanced. Therefore, the loss predicting the tumor area into the normal area should be different from the loss predicting the normal area into the tumor area, and the loss of predicting the tumor area into the normal area should be increased [2]. The balanced cross-entropy is defined as:

$$L_{\mathrm{WCE}} = \sum_{i=1}^{N} -\left(\beta y_i \log p_i + \left(1 - y_i\right) \log\left(1 - p_i\right)\right) \tag{8}$$

where β is the loss weight of the misclassification of the tumor area. Extending it from two-class to multi-class problems, the loss of balanced cross-entropy is defined as follows [86]:

$$L_{\mathrm{BCE}} = \sum_{c=1}^{C} \sum_{i=1}^{N} -w^c y_i^c p_i^c \tag{9}$$

where w^c is the loss weight of misclassification of the pixel belong to class c, called balanced weight. Wang et al. [87] propose to use the inverse of the class proportion as the balance weight value.

3.2.3 Focal loss

In addition to the problem of sample imbalance, some pixels are easy to be classified and some pixels are difficult to be classified in image segmentation. Lin et al. [88] propose focal loss to give greater weight to samples that are easy to error. The form of focal loss is:

$$L_{\mathrm{Focal}} = \sum_{i=1}^{N} -\left(\beta\left(1 - p_i\right)^{\gamma} y_i \log\left(p_i\right) + p_i^{\gamma}\left(1 - y_i\right) \log\left(1 - p_i\right)\right) \tag{10}$$

where γ is a parameter that controls whether the sample is difficult to distinguish. If $\gamma = 0$, the focal loss is equal with cross-entropy loss.

3.2.4 Overlap-based Loss

Dice loss is overlap-based loss function often used in medical image segmentation. The formula (3) defines the Dice criterion in tumor segmentation. The larger the Dice criterion, the better the segmentation effect. The Dice value is in [0, 1]. Dice loss is defined as 1 minus Dice value, the formula is:

$$L_{\text{Dice}} = 1 - 2 \times \frac{\left|V_p \cap V_g\right|}{\left|V_p\right| + \left|V_g\right|} \tag{11}$$

where V_p is the predicted segmentation result and V_g is the ground-truth segmentation result. In addition to Dice, JSC is often used as an overlap-based evaluation criterion. Yuan et al. [89] proposed a loss function based on Jaccard distance, whose logic is consistent with Dice loss and also subtracts JSC value with 1. The formula is:

$$L_{\text{Jaccard}} = 1 - \frac{\left|V_p \cap V_g\right|}{\left|V_p \cup V_g\right|} \tag{12}$$

Loss function based on overlap criterion will make the prediction result coincide with ground-truth higher and effectively reduce the occurrence of isolated false positives.

3.2.5 Combination of multiple loss

The main problem with the loss function in the field of tumor segmentation is the extreme imbalance of the sample, that is, the proportion of tumor areas in the entire image is much lower than the background and normal tissue areas. This problem is called the small target segmentation problem in the field of natural image processing. Miletari et al. [63] find that the Dice loss is much better than the cross-entropy loss, or even the balanced cross-entropy loss, in small target segmentation. However, Taghanaki et al. [3] conclude after analyzing multiple loss functions that overlap-based loss is not stable as cross-entropy loss in both large and small target segmentation tasks, making it more difficult to optimize. Therefore, most of the current studies use function that combines the two, that is, based on the balanced cross-entropy loss to ensure the stability of the results, and supplemented by the Dice loss to optimize the segmentation results of small targets. The final loss function is:

$$L = L_{\text{BCE}} + \alpha L_{\text{Dice}} \tag{13}$$

where α is the balance factor between the two loss. [90–92] all use the above combination loss for tumor segmentation.

3.3 Preprocessing

Before the data are sent into the deep learning network, it usually undergoes preprocessing. There are generally four kinds of preprocessing, which are image attributes based method, data augmentation, noise elimination and edge enhancement. We have basically reached a consensus that if the image is directly sent into the deep neural network without preprocessing, the effect will be greatly discounted, and even sometimes, good preprocessing is the key to the success of the model [93].

Image attributes based method Image attributes based method is mostly based on the characteristics of the image to do some processing. For example, Zhong et al. [94] propose that when processing CT images and PET images, it is necessary to truncate the pixel density values, because some too large or too small pixel density values have no practical significance and will affect our final processing results, while increasing the complexity of image information.

Data augmentation Data augmentation, that is increasing the number of training sets, aims to increase the data set using graphical transformation, making the model more robust and less likely to overfit. Common methods to increase the number of data sets including Flip, Rotation, Shift, Shear, Zoom, Brightness and Elastic distortion [26]. [33, 95] find through comparative experiments that using MR images of the brain provided by different institutions yields better results than those provided by a single institution, validating the urgency of improving data diversity through data augmentation. The experimental results of the [90, 92] point out that in the case of a small data set, the benefit of using data augmentation is not less than the benefit of model updating.

Noise elimination The purpose of noise elimination is to eliminate noise in images, making deep learning easier. Wavelet transform is basically used as a preprocessing method for noise elimination. Before deep learning is applied to tumor segmentation, traditional machine learning segmentation generally uses wavelet transform as a preprocessing method for tumor segmentation. Median filter with large kernel can remove noise well and has less loss for edge. Husham et al. [96] extend the median filtering and proposed a wavelet denoising method. The wavelet denoising algorithm consists of three steps: wavelet transform, threshold filtering and inverse wavelet transform. Mittal et al. [97] use Wiener filter as the basis of stationary wavelet transform to do image preprocessing because Wiener filter can smooth the boundary and preserve the internal information of the image well. Meanwhile, the authors point out that although anisotropic filter [98] can achieve the same function, its peak signal-to-noise ratio is too low.

Edge enhancement In the process of tumor segmentation, the blurry edge between tumor tissue and normal tissue is an important obstacle to the accuracy of tumor segmentation, which is why the edge enhancement preprocessing should be done. Histogram equalization uses cumulative distribution function to normalize image intensity, so as to improve image brightness. In some cases, histogram equalization produces a scour effect by averaging the intensity to the middle [99]. Laplace edge enhancement or non-sharp masking technology is to enhance the high-frequency components by filtering the image and scaling it into the original image, so as to achieve edge enhancement [100]. In the practice of edge enhancement, the trade-off between noise enhancement and edge enhancement needs to be considered.

3.4 Post-processing

After passing through the deep neural network, the output can be directly used as the result of tumor segmentation, but in general, the output of the deep neural network

may not directly meet the needs and it is not interpretable. Therefore, some researchers do some post-processing operations after the output of deep learning to achieve better results.

3.4.1 Conditional random field (CRF)

Conditional random field (CRF) is a conditional probability distribution model for a set of output sequences by given a set of input sequences, which is widely used in natural language processing. In the field of image segmentation, the significance of the application of CRF is that the class of each pixel is determined by the pixel and the surrounding pixels. Krähenbühl et al. [101] extend CRF and the class of each pixel is related to all pixels in the whole picture, and propose DenseCRF. The role of CRF and its variant DenseCRF in tumor segmentation has been replaced by deep learning methods led by FCN. But in post-processing, in order to make the final segmentation effect more reasonable, it is still widely used. The energy function based on Gibbs distribution to be optimized by DenseCRF is defined as:

$$E(x) = \sum_{i=1}^{N} \varphi_i(x_i) + \sum_{j<i} \psi_{i,j}(x_i, x_j) \tag{14}$$

where $x_i \in X$ is the pixel in image X, $\varphi_i(x_i)$ is the characteristic function of each pixel and $\psi_{i,j}(x_i, x_j)$ is the characteristic function of the pixel x_i and pixel x_j. In the method of using DenseCRF as post-processing, the value of $\varphi_i(x_i)$ is generally modified to the output of the previous deep convolution neural network [62, 102].

3.4.2 Machine learning on features using deep learning

In the field of post-processing research, some researchers believe that the advantage of deep networks represented by FCN lies in feature extraction, but it is not as good as traditional machine learning methods in subsequent classification problems. Therefore, some researchers use FCN as a feature extraction tool. After FCN, they use the features extracted by FCN to follow other machine learning methods [103]. Qaiser et al. [29] propose that because of the morphological difference between the normal nucleus and the nucleus of cancer cells, after deep learning, the method of persistent homology is used to deal with it. Persistent homology is an algebraic tool. Given a topological space, some algebraic invariants are calculated by using the structure of the space [104]. Hu et al. [59] believe that during breast tumor segmentation, the FCN output result is very rough, because the boundaries of the tumor area are blurred and diverse, so after FCN, the author added phase-based active contour (PBAC) model. Chlebus et al. [105] propose that using random forest as post-processing can eliminate false positives. After the deep network is executed, the features of the last few layers are used to send them into the random forest model, and finally the pixel category is obtained. Chen et al. [106] propose using multilayer perceptron for post-processing, the effect is better than CRF and other methods.

3.5 Multimodal tumor segmentation

More and more researchers have begun to use multiple modal images for tumor segmentation to compensate for shortcomings in single-modal imaging. The mainstream multimodal joint segmentation mainly includes CT and MR images, T1 and T2 weight images of MR, CT and PET images. The main problem of multimodal image segmentation is how to effectively use the information of multiple modalities to improve the segmentation effect. At present, there are two ways to do multimodal tumor segmentation. One is that different modes contain the same information, so the same information has the highest confidence and should be used most. Then how the features of the two modes supervise each other to extract useful features were discussed, which is called feature supervision [107–109]. The second way considers that the images of different modalities contain different information, and discusses how to effectively use this complementary information, which is called feature fusion [84, 90, 92].

A typical network architecture for feature supervision is shown in Fig. 13. The two modal images pass through a separate U-NET module. A complete U-NET module is divided into encoder and decoder modules, where the role of the encoder module is for feature extraction, and the role of the decoder module is for image recovery. Feature supervision is based on the assumption that the deep semantic information contained in the two modal images is consistent, so after sufficient feature abstraction, the resulting features should be similar. Therefore, after encoder, two independent U-Net continue to extract features after several layers of convolution network,

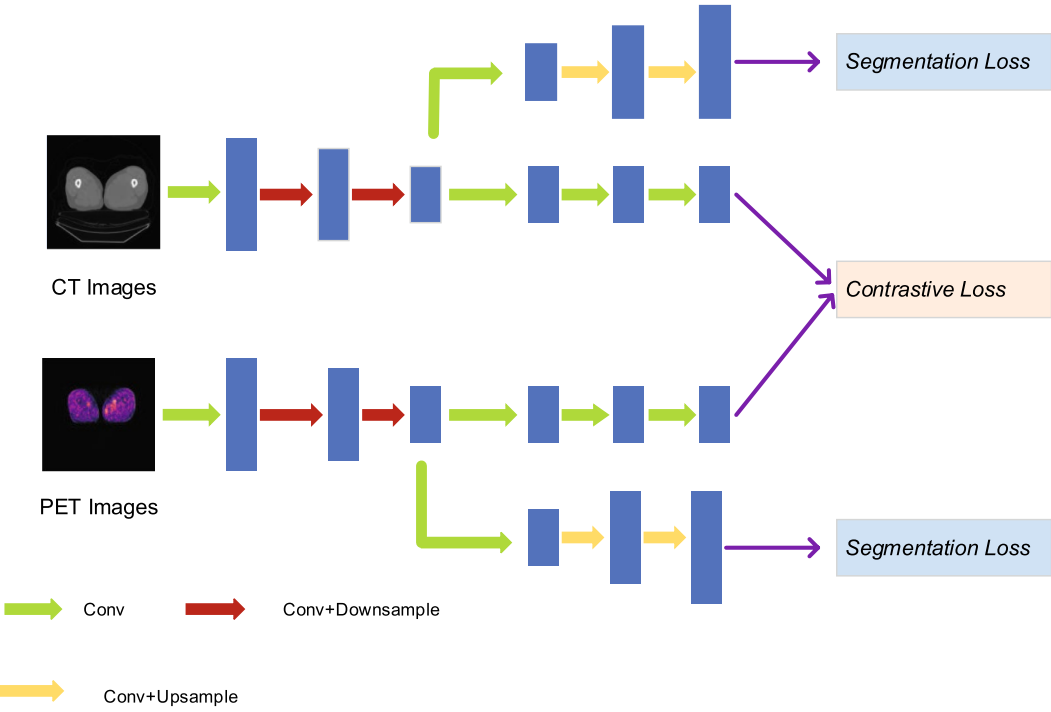

Fig. 13 Feature supervision [107]. The images of two modes pass through different u-net networks, respectively, and then do feature extraction after upsampling, then the features of two different modes should be similar, that is contrastive loss

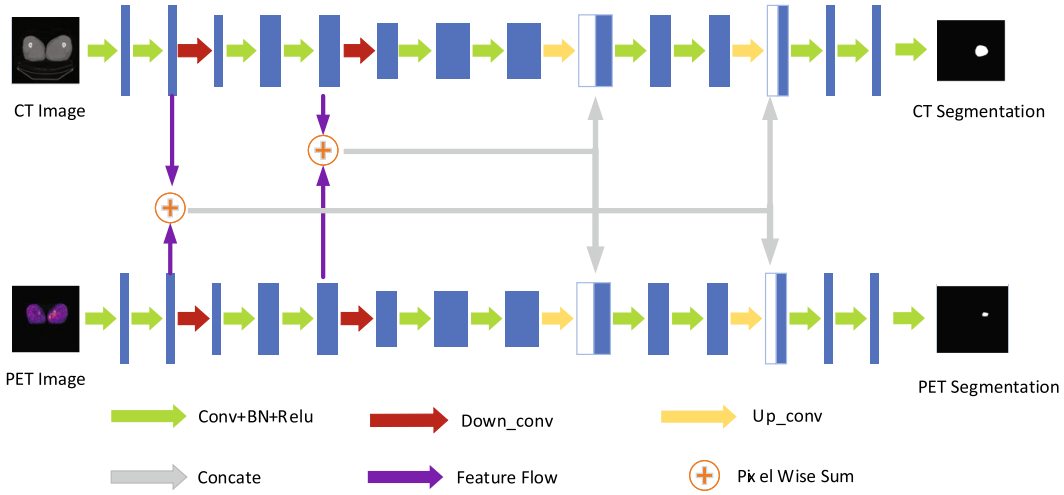

Fig. 14 Feature fusion [84]. After the two modal images have passed through the U-NET encoder, feature fusion will be performed at the decoder stage. In the decoder layers of the two modalities, when skip-connection, not only the down-sampled features of the current modal image are concatenated in, but also the down-sampled features of the corresponding modal image are concatenated in to perform feature fusion

and finally get semantic features, which should be similar. Using this similarity as a loss function to guide the network structure of the encoder layer of both u-net.

The typical network of feature fusion is shown in Fig. 14. The idea of feature fusion is based on the use of the complementary information between the two to improve the segmentation results. After the two modal images have passed through the U-NET encoder, feature fusion will be performed at the decoder stage. In the decoder layers of the two modalities, when skip-connection, not only the down-sampled features of the current modal image are concatenated in, but also the down-sampled features of the corresponding modal image are concatenated in to perform feature fusion. Moreover, Jin et al. [91] use the attention mechanism to complete the feature fusion of the two modalities.

3.6 Semi-supervised tumor segmentation

In tumor segmentation based on medical images, the main problem is insufficient labeled samples. For tumor segmentation tasks, it is difficult to obtain enough finely labeled training and test images. So researchers started using semi-supervised segmentation to solve the problem of lacking labeled samples. Similarly, semi-supervised segmentation can also be divided into two ways at present. One is that although fine pixel-level labels are difficult to obtain, image-level labels, that is, whether the image contains tumors, are easier to get. Therefore, this kind of method is trained for the part containing pixel-level labels and part containing image-level labels [64, 110]. The second kind of method is training for completely missing label data [111, 112].

The structure of the first type of semi-supervised segmentation method is shown in Fig. 15. Part of the input data of this type of semi-supervised method have

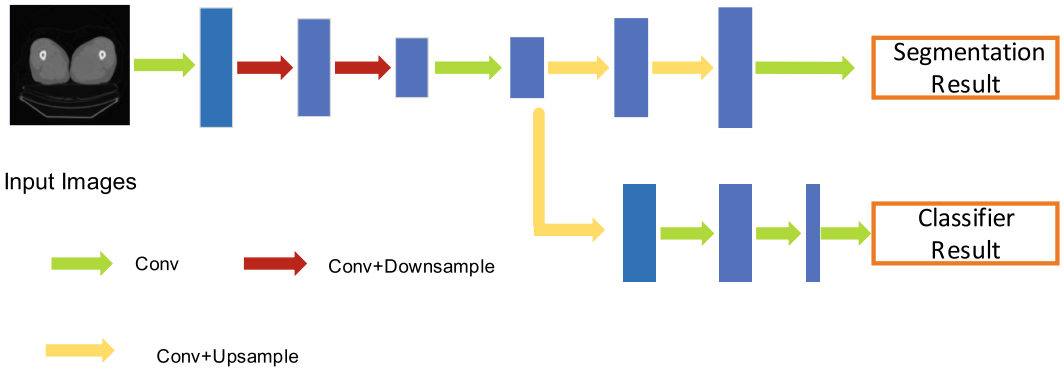

Fig. 15 First way of semi-supervised segmentation [110]. Train classification network and segmentation network at the same time according to different levels of labeling

pixel-level labels, and some data have only image-level labels. Therefore, training segmentation network uses data with pixel-level label and training classification networks uses data only with image-level label. In order to share parameters, all the network structures of the segmented network are shared by the classified network. Also, on the basis of the segmented network, additional layers of convolution network are added to form a complete classified network.

The second kind of semi-supervised method faces the problem of insufficient samples, but all the samples are pixel-level labeled. The method proposed by Wang et al. [111] is to use labeled data to train the model, then use unlabeled data to predict and see the probability of the final result. If the probability values are close to 1 or 0, that is, the accuracy is high, then the information contained in the image is considered to have been learned. If the probability value approaches 0.5, it is considered that the tumor information contained in the image has not been learned before, so the image is manually labeled at the pixel level and then learned. Through the above method, the number of the data that should be labeled is reduced. Jiang et al. [112] use adversarial learning to generate samples to increase the number of samples for learning.

4 Review from tumor view

Although tumor segmentation can extract a general paradigm, different types of tumors face some difficulties because of their different characteristics and different imaging techniques. Table 3 analyzes the characteristics of each tumor and the difficulty of segmentation, as well as the current mainstream solutions. The shape of the tumor is almost variation. In terms of tumor size, we analyze the public data set. If the ratio of the largest tumor to the smallest tumor size in a slice is less than 10, it is considered small variation, and if it is between 10 and 100, it is considered variation, and if it is greater than 100, it is considered big variation.

Table 3 Summary of different types of tumors

Tumor type	Size	Difficulties	Skills
Brain tumor	Big variation	Different modalities have different enhancement points for the image	Multi-modual [79, 84, 113]
		The distinction between tumor and non-tumor areas is blurred	Obtain more information for pixel classification [71, 114, 115]
		Tumor size fluctuates too much in the same image	Soft Dice Loss [26]
Breast tumor	Small variation	The distinction between tumor and non-tumor areas is blurred	Obtain more information for pixel classification [32, 59]
		High degree of speckle noise	Attention mechanism and adversarial learning [81]
Colorectal tumor	Variation	The distinction between tumor and non-tumor areas is blurred	Obtain more information for pixel classification [116, 117]
		Hard mimics from the complex peritumoral background	Hybrid Loss [118]
		Polyp's malignancy degree is only known after completing the histological analysis.	Segment the polyab directly [27, 119]
Kidney tumor	Variation	The distinction between tumor and non-tumor areas is blurred	Obtain more information for pixel classification [60, 68, 120–122]
Liver tumor	Big variation	The distinction between tumor and non-tumor areas is blurred	Obtain more information for pixel classification [48, 62, 67, 123]
		Different phase have different enhancement points for the different tumors	Multi-phase [61]
Lung tumor	Small variation	The distinction between tumor and non-tumor areas is blurred	Obtain more information for pixel classification [30, 124]
Nasopharyngeal carcinoma	Variation	Complex anatomical structure	Multimodality fusion [125]

For different types of tumors, the shape of the tumor, the difficulties in segmentation and the corresponding solutions are shown. All tumors are variable in shape

4.1 Brain tumor

Brain tumors are caused by uncontrolled brain conditions, resulting in abnormal cells appearing in the brain regions. The emergence of these tumors breaks the normal work of the brain and has a negative impact on patient health [126]. In medical practice of brain tumor segmentation, brain MR images are most commonly used in order to separate tumor regions from normal tissue regions, such as gray matter, white matter and cerebrospinal fluid. Havaei et al. [127] use deep learning for the first time to segment brain tumors and validate its results faster than traditional CRF and better at dealing with indistinguishable brain tumor segmentation (such as glioma and malignant glioma). Although the FCN network was not proposed at the time, the paper realized the function of Gaussian pyramid by setting different convolution kernels and implemented the skip-connection function by concatenating the results after multiple convolution kernels.

With [26, 97, 128, 129] introducing FCN and U-NET into the brain tumor segmentation challenge, the brain tumor segmentation method based on FCN has become the backbone of the brain tumor segmentation network. Because the basic deep network has too small receptive fields to affect the final segmentation results, [71, 114, 115] have adopted dilated convolution to expand receptive fields. Shen et al. [46] aim at the problem that the convolutional neural network is not clear for the segmentation of tumor boundary, in addition to the segmentation of tumor region, a sub-task with tumor boundary as the ground-truth was added, and both are trained simultaneously to improve the accuracy of boundary segmentation. Dong et al. [65] use cascaded FCN network for brain tumor segmentation, which improved the accuracy, but Zhou et al. [130] believe that the cascaded segmentation network would lead to extremely complex system, and ignore the relationship between various deep models. In view of this, Zhou et al. [130] divide the cascaded network into three independent but interconnected networks, and the three networks are trained together to explore their relationship.

In addition to adding a branch structure that is conducive to brain tumor segmentation on the basis of FCN, some researchers have also updated the basic structure of FCN. Pereira et al. [131] believe that not all the feature maps generated by convolutional networks are related to the final segmentation results, so they could not be simply added up. Therefore, Squeeze-and-Excitation theory in image classification theory is introduced into FCN to make the combination of feature maps. Chen et al. [106] hold the opinion that existing deep learning frameworks cannot guide models to learn high-quality hierarchical features, and proposes a multi-level Deep Medic to solve this problem. Chang et al. [102] consider that the widely used maximal pooling lost a lot of context information because it only preserves the maximum value in the receptive field. In order to compensate for this, in deep learning, two models are trained, using maximal pooling and average pooling, respectively.

Xu et al. [76] use 3D convolutional neural networks to capture information between slices, but 3D U-NET has many parameters, consumes many resources and is prone to overfitting. Mlynarski et al. [72] use a lighter 2D network for feature extraction within a single image and 3D network is only used for inter-slice feature processing. To solve the problem that 3D U-NET is easy to overfit, Myronenko et al.

[132] add a self-coding branch to reconstruct the image in the encoder phase of 3D U-NET, trained with segmentation during training, and regularize 3D U-NET by this method. [47, 80] introduce an adversarial neural network into the segmentation of brain tumors, which is basically the same as introduced in Sect. 3.1.7. [79, 84, 113] begin to explore the use of multiple modal MR images to improve the effect.

4.2 Breast tumor

Breast cancer is the most common tumor disease in women and it has now become the second leading cause of death [33]. So automatic segmentation for breast tumors is especially important. Breast tumor segmentation based on deep learning is mostly focused on ultrasound images, and few researchers focus on breast tumor segmentation from X-ray and MR images. So far, ultrasound images are still the most commonly used imaging format for radiologists to diagnose breast tumors. For breast tumor segmentation using ultrasound images, the traditional machine learning methods still occupy the mainstream [133–137]. Even if deep learning is used for tumor segmentation based on ultrasound images, unlike brain tumor segmentation that directly use FCN or U-NET, most of them have complicated deformations. This is because breast tumors in ultrasound images have different morphologies and different sizes.

Hu et al. [59] transplant FCN to ultrasound-based breast tumor segmentation. In addition to using dilated convolution to expand the receptive field, the phase-based active contour (PBAC) model is used for precise segmentation of tumors in the post-processing stage. Vakanski et al. [138] and Liang et al. [139] also use dilated convolution to expand the receptive field of deep learning networks. Vakanski et al. [138] argue that it is necessary and challenging to incorporate prior knowledge into FCN because breast tumors exhibit different occurrence locations, boundaries, curvatures and intensities in different patients. The prior knowledge selected by the author is a visual saliency image. Visual saliency estimation is an important method for automatic detection of tumors in medical images. The purpose is to establish a model between the saliency level of the image area and the ability to attract the visual attention of the radiologist [140, 141]. The author integrates the information expressed by visual saliency into the FCN architecture through the attention mechanism. Liang et al. [139] use instance segmentation network Mask R-CNN [142] for breast tumor segmentation in ultrasound images. Due to the high complexity of Mask R-CNN model and the large scale of the required training data set, in order to solve this problem, a new elastic transformation is proposed for data augmentation. Singh et al. [81] introduce adversarial learning networks to tumor segmentation tasks on ultrasound images.

In addition to ultrasound images, some studies have focused on X-ray and MR images, but their difficulty is lower than that of ultrasound images. Hai et al. [32] use fully connected DenseNet [143] to process X-ray images, while using dilated convolution to improve the receptive field. Zhang et al. [144] train a FCN with ROI region as the ground-truth in dealing with breast tumor segmentation from MR images, and then, the result passes through a cascaded network. Adoui et al. [66] use

two networks, U-NET and SegNet, to perform MR-based breast tumor segmentation to improve segmentation accuracy.

4.3 Colorectal tumor

Colorectal tumors are the second most common tumor in women and the third most common tumor in men. Segmentation of colorectal tumors is mainly focused on three aspects, one is based on MR image segmentation, the second is based on pathological images, do the segmentation of colorectal tumor cells and normal cells, and the last is based on CT image of tumor segmentation. The segmentation based on MR images and pathological images is more concerned by researchers.

For colorectal tumor segmentation based on CT images, Men et al. [145] first use FCN network, meanwhile use dilated convolution to improve the receptive field, and use patches of different sizes as input to obtain different resolution features. Liu et al. [146] use the adversarial network to segment colorectal tumors from CT images, which improves the segmentation effect. Trebeschi et al. [147] use deep learning network for colorectal tumor segmentation based on MR images for the first time. Jian et al. [148] introduce FCN into MR image-based colorectal tumor segmentation, while aiming at the defect that FCN could not obtain multi-level semantic information, the segmentation results are output after each workflow ended. Soomro et al. [149] propose different solutions to the same problem. In order to obtain different levels of semantic information, different resolution images are input and then the segmentation results obtained at different resolutions are integrated. Huang et al. [118] propose a mixed loss function based on Dice loss to solve the problem of sample imbalance in the segmentation of colorectal tumors. Meanwhile, different resolution images are input to eliminate false positives. Different from previous sliding window or single training model, Huang et al. [150] propose a joint training model based on the encoder–decoder framework, which includes the ROI acquisition branch and the ROI-based tumor segmentation branch, respectively. The two branches share the encoder stage. In the study of tumor segmentation in colorectal tissue images, Tang et al. [151] introduce SegNet to the segmentation task for the first time and achieve better results than traditional methods. Qaiser et al. [29] do not blindly use the deep learning method in tumor segmentation based on tissue images. He believes that there were obvious morphological differences between tumor cell nuclei and non-tumor cell nuclei. Based on this, a segmentation algorithm combining continuous coherence method and deep learning is proposed.

About 80%–95% of colorectal tumors develop from colorectal polyps. Therefore, the diagnosis of colorectal polyps can prevent the occurrence of colorectal tumors. Some researchers focus on the automatic segmentation of colorectal polyps. The diagnosis of colorectal polyps is mostly based on images generated by colonoscopy. Li et al. [27] introduce FCN to the segmentation of colorectal polyps for the first time. Akbari et al. [116] divide the colorectal segmentation into two stages. The first stage is to obtain the ROI region through FCN. The second stage uses the multi-level Ostu algorithm to segment the tumor region based on the ROI obtained in the first stage. Nguyen et al. [119] introduce the encoder–decoder architecture to the

colorectal polyp segmentation task. Kang et al. [117] use Mask R-CNN to segment colorectal polyps and a Mask R-CNN network based on two different backbones (ResNet-50 and ResNet-101) was integrated.

4.4 Kidney tumor

Kidney tumors caused approximately 175,000 deaths in 2018 [152] and is expected to cause 14,770 deaths across the United States in 2019 [153]. Currently, CT images are commonly used in the diagnosis of kidney tumors, and kidney tumor segmentation based on deep learning is also focused on CT images. Efremova et al. [154], Zavala-Romero et al. [155] and Shen et al. [156] introduce U-NET and 3D U-NET into the task of kidney tumor segmentation, all of which have achieved far better results than traditional methods. Yang et al. [60], Vu et al. [68], Lv et al. [120], Mu et al. [121] and Wei et al. [122] all use cascaded network structures to distinguish kidney tumors. The difference between the methods is that Yang et al. [60] uses the Gaussian pyramid to expand the receptive field in the network structure of the first stage; Vu et al. [68] expand the cascade network to three layers, the first layer directly obtains the results, the second layer obtains the tumor and kidney regions, the third layer obtains the tumor segmentation results with the input of the second layer cascade, and the final results are fused by the first layer cascade results and the third layer cascade results. In order to obtain different features in the cascade stage, the two cascade frameworks are composed of U-NET and V-NET, respectively, in [120]; Mu et al. [121] re-randomize sampling and regularization when the output of the first layer of the cascade is input to the second layer of the cascade. Yu et al. [157] process images from two different perspectives and then cross over two networks which is connected each other. Myronenko et al. [158] propose a Boundary-Aware architecture for kidney tumor segmentation, where one task branch segments the tumor boundary and the other branch segments the tumor region. Sabarinathan et al. [159] and Zhang et al. [160] use a multi-level semantic information fusion architecture to obtain different levels of semantic information. In addition, Zhang et al. [160] use a cascade architecture to reduce sample imbalance before fusion. Chen et al. [161] extend Mask R-CNN to 3D Mask R-CNN and applies it to kidney tumor segmentation. Sun et al. [162] introduce attention mechanism to kidney tumor segmentation.

4.5 Liver tumor

According to the WHO report, liver tumors are the fifth most deadly tumor in the world [163]. The segmentation of deep learning methods for liver tumors is focused on CT images. Chlebus et al. [105, 164] introduce 2D U-NET to liver tumor segmentation for the first time and optimize post-processing to reduce false positives by using random forests. Vorontsov et al. [165] use two different architectures to complete the encoder–decoder network structure, in which the encoder phase uses

ResNet and the decoder phase uses U-NET, and uses the skip-connection mechanism to connect the high-channel information of ResNet with the low-channel information of U-NET.

In addition to 2D U-NET, a large number of researchers have focused on how to process 3D CT slices to perform liver tumor segmentation tasks. Because the direct use of 3D U-NET has the problems of difficult training, easy overfitting and large calculation, researchers focus on how to coordinate the processing of information within slices and information between slices. Han et al. [83] use a 2.5D network to process, that is concatenating together several adjacent slices of the current slice into a 2D U-NET. Although this method can obtain the connection between adjacent slices, there are faults, and the receptive field is not large enough. Li et al. [52] use 2D U-NET to process the information within the slices and then selected some connected slices to send to the 3D U-NET network. Liu et al. [166] point out that the network for processing 3D information between slices such as 3D U-NET and LSTM is difficult to learn. Therefore, the authors propose a 3D anisotropic hybrid model to convert 2D features into 3D features.

In the segmentation of liver tumors, another widely used network is the cascade network, but different researchers have made corresponding updates for the cascade network. Gruber et al. [48] propose a two-stage cascade structure. The first stage cuts out the position of the liver, and the second stage does fine tumor segmentation. Christ et al. [62] add CRF to do post-processing on this basis, which further improved the accuracy and interpretability of the final segmentation results. Yuan et al. [67] extend the cascade structure to three stages. The first stage is used to locate liver organs, the second stage is used to segment liver organs, and the third stage is used to segment liver tumors. Jin et al. [123] introduce 3D U-NET into cascade network, and in order to solve the computational cost problem of 3D U-NET, first use 2D U-NET to segment the ROI area and then put this ROI area into cascade 3D U-NET network, which reduces the computational cost by reducing the size of the 3D input image. Jiang et al. [75] introduce soft and hard attention mechanism and long and short skip mechanism into the cascade network. Bai et al. [167] cascade the deep learning method with the traditional method. First, the liver region that may contain tumors is obtained using 3D U-NET, then the ROI region in the liver was obtained using 3D fractal residual network, and finally the final tumor segmentation is done using active contour model.

In addition, Sun et al. [61] believe that RTCT information acquired at different stages has different information and characteristics, so one FCN extraction feature is trained for three different stages of RTCT [arterial (ART) phase, portal venous (PV) phase, and delayed (DL) phase], and then fused together to make final segmentation. Seo et al. [168] argue that U-NET has three defects when applied to liver tumor segmentation. First, although the skip mechanism is useful, it will cause some information duplication. Second, although the high-level features have good expressive ability, the boundary information of the input pictures has long been blurred. Third, the number of pooling layers is difficult to determine due to the different size of the input image. For this reason, the solution given by the author is to use residual deconvolution and activation function in the skip to avoid information duplication. For the features of small objects, it is not in the residual deconvolution calculation

in the skip, while adding additional network in the skip to obtain higher-resolution boundary information in the high-level feature stage.

4.6 Lung tumor

Lung cancer is one of the leading lethal cancers with a 5-year survival rate of only 18% [169]. Current lung tumor segmentation methods based on deep learning are mainly focused on CT images. Wang et al. [170] introduce deep learning into the segmentation of lung tumors. Meanwhile, the authors believe that CT images from different perspectives represented different information. Therefore, each CNN branch is trained for different perspectives. Finally, these results are fused into a full-connected layer to output the probability that the central pixel point of the branch belonged to the tumor area or not. Zhou et al. [30] use deep learning to process both 2D and 3D images to acquire features separately, and in addition, change the pooling into the central pooling which pays more attention to the central pixels. Rocha et al. [171] introduce U-NET to lung tumor segmentation and make detailed comparisons with traditional methods in various aspects, and conclude that deep learning methods are far superior to traditional methods. Tong et al. [172] update U-NET, add a residual network to speed up model training and use batch normalization to make the model more robust. Qin et al. [124] use adversarial network to generate more positive samples in order to solve the problem of data sample imbalance. Jiang et al. [73] use multiple ResNet networks to extract features of CT images with different resolution. Roy et al. [173] first obtain the rough segmentation results using U-NET, then use morphology-driven level set method to do fine segmentation, and the initial seed points of the level set are provided by the rough segmentation results of U-NET. Feng et al. [174], Tang et al. [31] and Wu et al. [175] all use the combined framework of classification and segmentation of pulmonary nodules and complete the classification and segmentation tasks of pulmonary nodules in one deep learning process. Among them, Feng et al. [174] use the simple linear iterative clustering (SLIC) method [176] for preprocessing. Tang et al. [31] use decoupled feature maps to reduce false positives.

4.7 Nasopharyngeal carcinoma

Nasopharyngeal carcinoma is a neoplasm originating in the nasopharynx and is a malignant tumor that occurs frequently in southern China, Southeast Asia, the Arctic, the Middle East and North Africa [177]. Most of the current research on automated segmentation of nasopharyngeal carcinoma has focused on CT and MR images. In the CT based automation methods, Men et al. [178] introduce the decoder and encoder architecture to the segmentation of nasopharyngeal carcinoma, but the upsampling and downsampling methods are rough. Liang et al. [179] use a cascaded network for processing, which includes a total of two layers of networks, one layer obtains the ROI area, and the other layer performs segmentation based

on ROI. Zhong et al. [180] extend the cascade network structure to three layers. Li et al. [181] introduce U-NET to automated segmentation of nasopharyngeal carcinoma. MR image-based segmentation of nasopharyngeal carcinoma has been studied slightly more than CT images. Wang et al. [182] introduce CNN into MR image-based segmentation of nasopharyngeal carcinoma and use morphological methods in post-processing. Ma et al. [183] believe that MR images from different perspectives store different information, so a neural network is trained separately for MR images based on different perspectives. Chen et al. [125] similarly process the MR images of different modes and achieve better results. Wu et al. [184] use graph segmentation as post-processing, and the output of the neural network is used as input for graph segmentation. Li et al. [185] use U-NET as a MR image-based nasopharyngeal cancer segmentation framework. Huang et al. [186] add an attention mechanism to the U-NET structure, which allows high-level semantic features to guide the learning of low-level features.

5 Discussion and future direction

We review the tumor segmentation algorithms based on deep learning from two perspectives, which are the architecture view and the tumor view. But in this section, we integrate two perspectives together to summarize. As shown in Table 4, we summarize the research topics of different types of tumors according to the architecture view introduced in Sect. 3. Because a large number of research topics are concerned about the innovation of network structure, we split the topic of network into specific technical details.

Although the tumor segmentation based on deep learning has achieved far better results than traditional methods, there are still some problems that have not been solved or lack reasonable and widely accepted methods. As shown in Table 2, with the improvement in network capabilities, the amount of network parameters has become larger and larger, especially the amount of network parameters that process information between 3D slices has expanded particularly severely. Therefore, how to effectively process 3D slice information, and how to compress the model as the amount of network parameters increases, are issues that need to be addressed in the future. And, due to the small amount of medical imaging data, how to prevent overfitting also requires researchers to think. It can be seen from Table 3 that although the characteristics of different tumors are not the same, there are still many commonalities between tumors, but there is currently no effective research aimed at unified segmentation models and transfer learning between different tumors. Based on this, we summarize the following four future research directions of tumor segmentation based on deep learning.

5.1 Processing of 3D medical images

Although there are currently 3D U-NET [49] and V-NET [63] for 3D image processing, the computational complexity of 3D convolutional neural network is an

Table 4 Summary of research topics of different types of tumors, including the main imaging modalities of each type of tumor, the main focus of the current research and the corresponding references

Tumor type	Major image type	Major research topic	References
Brain tumor	MR	Multi-modual	[79, 84, 113]
		FCN or U-NET	[128]△, [26]△, [97, 129]
		Cascade network	[46]△, [65]△, [130]△
		Adversarial network	[47]△, [80]△
		Expand receptive field	[114]△, [115]△, [71]△
		3D segmentation network	[76]△, [72]△, [132]△
Breast tumor	X-ray, MR, US	Preprocessing	[141]
		FCN or U-NET	[59]
		Cascade network	[144]△, [66]
		Adversarial network	[81]△
		Attention mechanism	[138]
		Expand receptive field	[32, 59]
Colorectal tumor	CT, MR, tissue image	Post-processing	[29]△
		Loss function	[148]
		FCN or U-NET	[145, 147, 149]△, [27]△, [119]△
		Cascade network	[116]△, [117]△
Kidney tumor	CT	FCN or U-NET	[154]△, [155]△, [156]△
		Cascade network	[60]△, [68]△, [120]△, [121]△, [122]△
		Attention mechanism	[162]△
		Boundary-Aware FCN	[158]△
		3D segmentation network	[161]△

Table 4 (continued)

Tumor type	Major image type	Major research topic	References
Liver tumor	CT	Post-processing	[105, 164]△
		FCN or U-NET	[165]△
		Cascade network	[48]△, [62]△, [67]△, [123]△, [75]△, [167]△, [61]△
		3D segmentation network	[83]△, [52]△, [166]△
Lung tumor	CT	Preprocessing	[174]△
		Post-processing	[173]△
		FCN or U-NET	[170]△, [171]△, [172]
		Semi-supervised	[31, 174, 175]△
		Adversarial network	[124]△
		3D segmentation network	[30]△
Nasopharyngeal carcinoma	MR, CT	Post-processing	[182, 184]
		Multi-modual	[125]
		FCN or U-NET	[178, 181, 183, 185]
		Cascade network	[179]
		Attention mechanism	[186]

△ represents using public data set

unavoidable problem. Because the number of parameters is extremely inflated and the number of trainable images from 2D to 3D is rapidly reduced, overfitting is inevitable. However, current studies have proven that 3D U-ET is better than 2D U-NET [a20], which also proves that feature extraction between slices is necessary. Although some studies have addressed the above problems and used LSTM-like methods to deal with feature extraction between slices, the network of such methods is extremely complex and difficult to reproduce [52, 84]. Therefore, how to reasonably and effectively solve the problem of 3D medical image processing, which not only applies the inter-slice features of 3D image, but also has less parameters and less computational amount than 3D U-NET, still needs further research.

5.2 Transfer learning and unified segmentation model

At present, there are few studies on transfer learning in tumor segmentation. Most researchers have proposed a new network structure and then verified the effectiveness of tumor segmentation on two different data sets to prove the effectiveness of their own effects. However, in natural image segmentation or organ cell segmentation in medical image processing, there is still a large amount of research on transfer learning. Research on transfer learning mainly focuses on unsupervised domain adaptation [187–191], that is, the data in the source domain are labeled, and the data in the target domain are unlabeled. Transforming to tumor segmentation is that the segmentation model trained on labeled data of brain tumors is adapted to the segmentation of unlabeled lung tumor data. In addition, since each tumor segmentation is trained with an independent model, in clinical practice, a unified system or model is required to complete multiple tumor segmentation. Huang et al. [192] propose a unified U-NET model for multi-organ segmentation, but research on a unified segmentation model for multi-organ tumors is still lacking.

5.3 Model compression

Current research on model compression has focused on natural image segmentation [193–196]. Due to the increasing complexity of natural image segmentation networks and the current demand for transplanting natural image classification or segmentation models to portable devices such as mobile phones, research on the compression of natural image segmentation models is in full swing. Although the medical image segmentation model is slightly more complex than the natural image, especially the current multimodal, 3D image processing model is not less complex than the natural image. It is just less than the training data of natural images, and the training difficulty is a bit lower, but the model parameter magnitude is not low. With the future demand for telemedicine technology, the need to run medical tumor segmentation models on mobile devices will also exist, so the compression research of medical segmentation models is also very meaningful.

5.4 Overfitting problem

Overfitting is a fundamental problem in training deep learning. A series of solutions have been developed. For example, increase the sample amount, regularization, dropout, etc. But the effectiveness of all the above methods is mostly verified based on natural image classification and natural image segmentation problems. It should be noted that the biggest difference between medical image segmentation and natural image classification segmentation is that, first, the segmentation of medical images is the segmentation of small targets. Second, the sample number is smaller than that of natural images. Therefore, in the medical image segmentation problem, it is more meaningful to solve the problem of over-fitting when training with a small amount of data than to propose a more complex network architecture.

6 Conclusion

This paper reviews the methods of tumor segmentation based on deep learning. We review from two perspectives, one is from the technical perspective of deep learning and the other is from the perspective of tumor types. From the technical view of deep learning, we review the network architecture, preprocessing, post-processing, loss function, multimodal segmentation and semi-supervised segmentation. Based on the types of tumors, we review the seven tumors that are the main focus of current research, namely brain tumors, breast tumors, colorectal tumors, kidney tumors, liver tumors, lung tumors and nasopharyngeal cancer.

The tumor segmentation method based on deep learning is summarized from the two perspectives of tumor type and technical architecture. The current methods basically have two purposes, accurately segmenting tumors and making up for lack of training data. Deep learning is good at accurately segmenting tumors based on sufficient training data, and all the methods are basically from the following three perspectives: segmentation with clear boundaries, remove infrared areas in the image and obtain more information for pixel classification. Based on this, a large number of networks have been proposed, and detailed comparative introductions can be found in the article. However, due to the natural demand of neural network for the amount of data, the current methods to make up for the lack of data are relatively few, and most of them are based on modifying the training strategy. Based on the above situation, we have summarized four possible future research directions, that is 3D image segmentation, transfer learning, model compression and reasonable solution to overfitting.

The tumor segmentation method based on deep learning has made remarkable achievements so far, and the related research directions and research points are relatively scattered. Based on the motivation of the method, this paper reviews it from the perspectives of tumor type and network architecture. For related researchers and people who want to quickly understand this field, you can find some helpful content through this review.

Acknowledgements This work was supported by the National Natural Science Foundation of China under Grant 61872075.

References

1. Ronneberger O, Fischer P, Brox T (2015) U-net: convolutional networks for biomedical image segmentation. Medical Image Computing and Computer-Assisted intervention—MICCAI 2015. Springer, Cham, pp 234–241
2. Shelhamer E, Long J, Darrell T (2017) Fully convolutional networks for semantic segmentation. IEEE Trans Pattern Anal Mach Intell 39(4):640–651
3. Taghanaki SA, Abhishek K, Cohen JP, Cohenadad J, Hamarneh G (2021) Deep semantic segmentation of natural and medical images: a review. Artif Intell Rev 54(1):137–178
4. Seo H, Badiei Khuzani M, Vasudevan V, Huang C, Ren H, Xiao R, Jia X, Xing L (2020) Machine learning techniques for biomedical image segmentation: an overview of technical aspects and introduction to state-of-art applications. Med Phys 47(5):e148–e167
5. Ghosh S, Das N, Das I, Maulik U (2019) Understanding deep learning techniques for image segmentation. ACM Comput Surv 52(4):73
6. Guo Y, Liu Y, Georgiou T, Lew MS (2018) A review of semantic segmentation using deep neural networks. Int J Multimed Inf Retr 7(2):87–93
7. Kothari A, Indira B (2016) An overview on automated brain tumor segmentation techniques. Int J Comput Trends Technol 40(1):45–48
8. Sharma M, Miglani N (2020) Automated brain tumor segmentation in MRI images using deep learning: overview, challenges and future. Springer, Cham, pp 347–383
9. Isin A, Direkoglu C, Şah M (2016) Review of MRI-based brain tumor image segmentation using deep learning methods. Procedia Comput Sci 102:317–324
10. At A, Ss B, Mp A (2020) Brain tumor segmentation and classification from magnetic resonance images: review of selected methods from 2014 to 2019. Pattern Recogn Lett 131:244–260
11. Chihati S, Gaceb D (2020) A review of recent progress in deep learning-based methods for MRI brain tumor segmentation. In: 2020 11th International Conference on Information and Communication Systems (ICICS), pp 149–154
12. Nadeem MW, Ghamdi MAA, Hussain M, Khan MA, Khan KM, Almotiri SH, Butt SA (2020) Brain tumor analysis empowered with deep learning: a review, taxonomy, and future challenges. Brain Sci 10(2):118
13. A R, K K, W L, SS M (2018) Medical image database. https://www.med.upenn.edu/cbica/brats2019/data.html
14. Mo J, Zhang L, Wang Y, Huang H (2020) Iterative 3D feature enhancement network for pancreas segmentation from CT images. Neural Comput Appl 32(16):12535–12546
15. Dehmeshki J, Amin H, Valdivieso M, Ye X (2008) Segmentation of pulmonary nodules in thoracic CT scans: a region growing approach. IEEE Trans Med Imaging 27(4):467–480
16. Foster B, Bagci U, Mansoor A, Xu Z, Mollura DJ (2014) A review on segmentation of positron emission tomography images. Comput Biol Med 50:76–96
17. Kumar A, Fulham MJ, Feng D, Kim J (2020) Co-learning feature fusion maps from PET-CT images of lung cancer. IEEE Trans Med Imaging 39(1):204–217
18. Gsaxner C, Roth PM, Wallner J, Egger J (2019) Exploit fully automatic low-level segmented pet data for training high-level deep learning algorithms for the corresponding CT data. PLoS ONE 14(3):1–20
19. Bagci U, Udupa JK, Mendhiratta N, Foster B, Xu Z, Yao J, Chen X, Mollura DJ (2013) Joint segmentation of anatomical and functional images: applications in quantification of lesions from PET, PET-CT, MRI-PET, and MRI-PET-CT images. Med Image Anal 17(8):929–945
20. Milshteyn E, Guryev G, Torrado-Carvajal A, Adalsteinsson E, White JK, Wald LL, Guerin B (2021) Individualized SAR calculations using computer vision-based MR segmentation and a fast electromagnetic solver. Magn Reson Med 85(1):429–443
21. Zhu Q, Du B, Yan P (2020) Boundary-weighted domain adaptive neural network for prostate MR image segmentation. IEEE Trans Med Imaging 39(3):753–763

22. Rodrigues PS, Lopes GAW, Giraldi GA, Barcelos CAZ, Vieira L, Guliato D, Singh BK (2019) Cad system for breast us images with speckle noise reduction and bio-inspired segmentation, pp 68–75
23. Martin M, Sciolla B, Sdika M, Quétin P, Delachartre P (2019) Segmentation of neonates cerebral ventricles with 2D CNN in 3D US data: suitable training-set size and data augmentation strategies. In: 2019 IEEE International Ultrasonics Symposium (IUS), pp 2122–2125
24. Pauly PS, Rajan BK (2019) Automated feature extraction of hepatic US image of dog using image processing. In: 2019 2nd International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT), pp 456–460
25. Egger J, Voglreiter P, Dokter M, Hofmann M, Chen X, Zoller WG, Schmalstieg D, Hann A (2016) US-Cut: interactive algorithm for rapid detection and segmentation of liver tumors in ultrasound acquisitions. In: Duric N, Heyde B (eds) Medical Imaging 2016: Ultrasonic Imaging and Tomography, International Society for Optics and Photonics, vol 9790. SPIE, pp 372–377
26. Dong H, Yang G, Liu F, Mo Y, Guo Y (2017) Automatic brain tumor detection and segmentation using U-Net based fully convolutional networks. Medical image understanding and analysis. Springer, Cham, pp 506–517
27. Li Q, Yang G, Chen Z, Huang B, Chen L, Xu D, Zhou X, Zhong S, Zhang H, Wang T (2017) Colorectal polyp segmentation using a fully convolutional neural network. In: 2017 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), pp 1–5
28. Armato S, McLennan G, McNitt-Gray M, Meyer C, Reeves A, Bidaut L, Zhao B, Croft B, Clarke L (2010) WE-B-201B-02: the lung image database consortium (LIDC) and image database resource initiative (IDRI): a completed public database of CT scans for lung nodule analysis. Med Phys 37(6Part6):3416–3417
29. Qaiser T, Tsang YW, Taniyama D, Sakamoto N, Nakane K, Epstein D, Rajpoot N (2019) Fast and accurate tumor segmentation of histology images using persistent homology and deep convolutional features. Med Image Anal 55:1–14
30. Wang S, Zhou M, Liu Z, Liu Z, Gu D, Zang Y, Dong D, Gevaert O, Tian J (2017) Central focused convolutional neural networks: developing a data-driven model for lung nodule segmentation. Med Image Anal 40:172–183
31. Tang H, Zhang C, Xie X (2019) Nodulenet: decoupled false positive reduction for pulmonary nodule detection and segmentation. Medical Image Computing and Computer Assisted Intervention—ICCAI 2019. Springer, Cham, pp 266–274
32. Hai J, Qiao K, Chen J, Tan H, Xu J, Zeng L, Shi D, Yan B (2019) Fully convolutional densenet with multiscale context for automated breast tumor segmentation. J Healthc Eng 2019:8415485
33. Albadawy E, Saha A, Mazurowski MA (2018) Deep learning for segmentation of brain tumors: impact of cross-institutional training and testing. Med Phys 45(3):1150–1158
34. Heller N, Isensee F, Maier-Hein KH, Hou X, Xie C, Li F, Nan Y, Mu G, Lin Z, Han M, Yao G, Gao Y, Zhang Y, Wang Y, Hou F, Yang J, Xiong G, Tian J, Zhong C, Ma J, Rickman J, Dean J, Stai B, Tejpaul R, Oestreich M, Blake P, Kaluzniak H, Raza S, Rosenberg J, Moore K, Walczak E, Rengel Z, Edgerton Z, Vasdev R, Peterson M, McSweeney S, Peterson S, Kalapara A, Sathianathen N, Papanikolopoulos N, Weight C (2021) The state of the art in kidney and kidney tumor segmentation in contrast-enhanced CT imaging: results of the kits19 challenge. Med Image Anal 67:101821
35. Menze BH, Jakab A, Bauer S, Kalpathy-Cramer J, Farahani K, Kirby J, Burren Y, Porz N, Slotboom J, Wiest R, Lanczi L, Gerstner E, Weber M, Arbel T, Avants BB, Ayache N, Buendia P, Collins DL, Cordier N, Corso JJ, Criminisi A, Das T, Delingette H, Demiralp C, Durst CR, Dojat M, Doyle S, Festa J, Forbes F, Geremia E, Glocker B, Golland P, Guo X, Hamamci A, Iftekharuddin KM, Jena R, John NM, Konukoglu E, Lashkari D, Mariz JA, Meier R, Pereira S, Precup D, Price SJ, Raviv TR, Reza SMS, Ryan M, Sarikaya D, Schwartz L, Shin H, Shotton J, Silva CA, Sousa N, Subbanna NK, Szekely G, Taylor TJ, Thomas OM, Tustison NJ, Unal G, Vasseur F, Wintermark M, Ye DH, Zhao L, Zhao B, Zikic D, Prastawa M, Reyes M, Van Leemput K (2015) The multimodal brain tumor image segmentation benchmark (BRATS). IEEE Trans Med Imaging 34(10):1993–2024
36. Randhawa RS, Modi A, Jain P, Warier P (2016) Improving boundary classification for brain tumor segmentation and longitudinal disease progression. In: Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. Springer, Cham, pp 65–74
37. Isensee F, Kickingereder P, Wick W, Bendszus M, Maier-Hein KH (2018) Brain tumor segmentation and radiomics survival prediction: contribution to the brats 2017 challenge. In: Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. Springer, Cham, pp 287–297

38. Weninger L, Rippel O, Koppers S, Merhof D (2019) Segmentation of brain tumors and patient survival prediction: Methods for the brats 2018 challenge. In: Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. Springer, Cham, pp 3–12
39. CBICA (2019) Multimodal brain tumor segmentation challenge 2019. http://www.onlinemedicalimages.com
40. Sirinukunwattana K, Raza SEA, Tsang Y, Snead D, Cree IA, Rajpoot NM (2016) Locality sensitive deep learning for detection and classification of nuclei in routine colon cancer histology images. IEEE Trans Med Imaging 35(5):1196–1206
41. Graham S, Chen H, Gamper J, Dou Q, Heng PA, Snead D, Tsang YW, Rajpoot N (2019) Mild-net: minimal information loss dilated network for gland instance segmentation in colon histology images. Med Image Anal 52:199–211
42. Shaban M, Awan R, Fraz MM, Azam A, Tsang YW, Snead D, Rajpoot NM (2020) Context-aware convolutional neural network for grading of colorectal cancer histology images. IEEE Trans Med Imaging 39(7):2395–2405
43. Bilic P, Christ PF, Vorontsov E, Chlebus G, Chen H, Dou Q, Fu CW, Han X, Heng PA, Hesser J, Kadoury S, Konopczynski T, Le M, Li C, Li X, Lipkovà J, Lowengrub J, Meine H, Moltz JH, Pal C, Piraud M, Qi X, Qi J, Rempfler M, Roth K, Schenk A, Sekuboyina A, Vorontsov E, Zhou P, Hülsemeyer C, Beetz M, Ettlinger F, Gruen F, Kaissis G, Lohöfer F, Braren R, Holch J, Hofmann F, Sommer W, Heinemann V, Jacobs C, Mamani GEH, van Ginneken B, Chartrand G, Tang A, Drozdzal M, Ben-Cohen A, Klang E, Amitai MM, Konen E, Greenspan H, Moreau J, Hostettler A, Soler L, Vivanti R, Szeskin A, Lev-Cohain N, Sosna J, Joskowicz L, Menze BH (2019) The liver tumor segmentation benchmark (liTS). arXiv:1901.04056
44. Universitaires H (2018) 3Dircadb. https://www.ircad.fr/research/3dircadb/
45. TCIA (2018) NSCLC-Radiomics. https://wiki.cancerimagingarchive.net/display/Public/NSCLC-Radiomics
46. Shen H, Wang R, Zhang J, McKenna SJ (2017) Boundary-aware fully convolutional network for brain tumor segmentation. Medical Image Computing and Computer-Assisted Intervention—MICCAI 2017. Springer, Cham, pp 433–441
47. Rezaei M, Harmuth K, Gierke W, Kellermeier T, Fischer M, Yang H, Meinel C (2018) A conditional adversarial network for semantic segmentation of brain tumor. In: Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. Springer, Cham, pp 241–252
48. Gruber N, Antholzer S, Jaschke W, Kremser C, Haltmeier M (2019) A joint deep learning approach for automated liver and tumor segmentation. In: 2019 13th International Conference on Sampling Theory and Applications (SampTA), vol 1, pp 1–5
49. Çiçek Ö, Abdulkadir A, Lienkamp SS, Brox T, Ronneberger O (2016) 3D U-Net: learning dense volumetric segmentation from sparse annotation. Medical Image Computing and Computer-Assisted Intervention—MICCAI 2016. Springer, Cham, pp 424–432
50. Zhao H, Shi J, Qi X, Wang X, Jia J (2017) Pyramid scene parsing network. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), vol 1, pp 6230–6239
51. Oktay O, Schlemper J, Folgoc LL, Lee M, Heinrich M, Misawa K, Mori K, McDonagh S, Hammerla NY, Kainz B, Glocker B, Rueckert D (2018) Attention U-Net: learning where to look for the pancreas. arXiv:1804.03999
52. Li X, Chen H, Qi X, Dou Q, Fu C, Heng P (2018) H-DenseUNet: Hybrid densely connected UNet for liver and tumor segmentation from CT volumes. IEEE Trans Med Imaging 37(12):2663–2674
53. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE Computer Society, Los Alamitos, CA, USA, vol 1, pp 770–778
54. Giusti A, Cireşan DC, Masci J, Gambardella LM, Schmidhuber J (2013) Fast image scanning with deep max-pooling convolutional neural networks. In: 2013 IEEE International Conference on Image Processing, vol 1, pp 4034–4038
55. Zhang H, Dana K, Shi J, Zhang Z, Wang X, Tyagi A, Agrawal A (2018) Context encoding for semantic segmentation. In: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, vol 1, pp 7151–7160
56. Zhang C, Lin G, Liu F, Yao R, Shen C (2019) Canet: class-agnostic segmentation networks with iterative refinement and attentive few-shot learning. In: 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), vol 1, pp 5212–5221
57. Krizhevsky A, Sutskever I, Hinton GE (2017) Imagenet classification with deep convolutional neural networks. Commun ACM 60(6):84–90

58. Badrinarayanan V, Kendall A, Cipolla R (2017) Segnet: a deep convolutional encoder-decoder architecture for image segmentation. IEEE Trans Pattern Anal Mach Intell 39(12):2481–2495
59. Hu Y, Guo Y, Wang Y, Yu J, Li J, Zhou S, 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(1):215–228
60. Yang G, Li G, Pan T, Kong Y, Wu J, Shu H, Luo L, Dillenseger J, Coatrieux J, Tang L, Zhu X (2018) Automatic segmentation of kidney and renal tumor in CT images based on 3D fully convolutional neural network with pyramid pooling module. In: 2018 24th International Conference on Pattern Recognition (ICPR), vol 1, pp 3790–3795
61. Sun C, Guo S, Zhang H, Li J, Chen M, Ma S, Jin L, Liu X, Li X, Qian X (2017) Automatic segmentation of liver tumors from multiphase contrast-enhanced CT images based on FCNs. Artif Intell Med 83:58–66
62. Christ PF, Ettlinger F, Grün F, Elshaera MEA, Lipkova J, Schlecht S, Ahmaddy F, Tatavarty S, Bickel M, Bilic P, Rempfler M, Hofmann F, Anastasi MD, Ahmadi SA, Kaissis G, Holch J, Sommer W, Braren R, Heinemann V, Menze B (2017) Automatic liver and tumor segmentation of CT and MRI volumes using cascaded fully convolutional neural networks. arXiv:1702.05970
63. Milletari F, Navab N, Ahmadi S (2016) V-net: fully convolutional neural networks for volumetric medical image segmentation. In: 2016 Fourth International Conference on 3D Vision (3DV), vol 1, pp 565–571
64. Mlynarski P, Delingette H, Criminisi A, Ayache N (2019) Deep learning with mixed supervision for brain tumor segmentation. J Med Imaging 6(3):1–13
65. Wang G, Li W, Ourselin S, Vercauteren T (2018) Automatic brain tumor segmentation using cascaded anisotropic convolutional neural networks. In: Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. Springer, Cham, pp 178–190
66. Adoui ME, Mahmoudi SA, Larhmam MA, Benjelloun M (2019) MRI breast tumor segmentation using different encoder and decoder CNN architectures. Computers 8(3)
67. Zhang J, Saha A, Zhu Z, Mazurowski MA (2019) Hierarchical convolutional neural networks for segmentation of breast tumors in MRI with application to radiogenomics. IEEE Trans Med Imaging 38(2):435–447. https://doi.org/10.1109/TMI.2018.2865671
68. Vu MH, Grimbergen G, Simkó A, Nyholm T, Löfstedt T (2019) Localization network and end-to-end cascaded U-Nets for kidney tumor segmentation. Submissions to the 2019 Kidney Tumor Segmentation Challenge: KiTS19. https://doi.org/10.24926/548719.073
69. Sarker MMK, Rashwan HA, Akram F, Banu SF, Saleh A, Singh VK, Chowdhury FUH, Abdulwahab S, Romani S, Radeva P, Puig D (2018) SLSDeep: skin lesion segmentation based on dilated residual and pyramid pooling networks. Medical Image Computing and Computer Assisted Intervention—MICCAI 2018. Springer, Cham, pp 21–29
70. Li X, Yang D, Wang Y, Yang S, Qi L, Li F, Gan Z, Zhang W (2019) Automatic tongue image segmentation for real-time remote diagnosis. In: 2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), vol 1, pp 409–414
71. Qin Y, Kamnitsas K, Ancha S, Nanavati J, Cottrell G, Criminisi A, Nori A (2018) Autofocus layer for semantic segmentation. In: Medical Image Computing and Computer Assisted Intervention—MICCAI 2018. Springer, Cham, pp 603–611
72. Mlynarski P, Delingette H, Criminisi A, Ayache N (2019) 3D convolutional neural networks for tumor segmentation using long-range 2D context. Comput Med Imaging Graph 73:60–72
73. Jiang J, Hu Y, Liu C, Halpenny D, Hellmann MD, Deasy JO, Mageras GS, Veeraraghavan H (2019) Multiple resolution residually connected feature streams for automatic lung tumor segmentation from CT images. IEEE Trans Med Imaging 38(1):134–144
74. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser U, Polosukhin I (2017) Attention is all you need. In: Proceedings of the 31st International Conference on Neural Information Processing Systems, Curran Associates Inc., Red Hook, NY, USA, NIPS'17, pp 6000–6010
75. Jiang H, Shi T, Bai Z, Huang L (2019) AHCNet: an application of attention mechanism and hybrid connection for liver tumor segmentation in CT volumes. IEEE Access 7:24898–24909
76. Xu H, Xie H, Liu Y, Cheng C, Niu C, Zhang Y (2019) Deep cascaded attention network for multi-task brain tumor segmentation. In: Medical Image Computing and Computer Assisted Intervention—MICCAI 2019. Springer, Cham, pp 420–428
77. Sabarinathan D, Beham MP, Roomi SMMM (2019) Hyper vision net: kidney tumor segmentation using coordinate convolutional layer and attention unit

78. Goodfellow IJ, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, Bengio Y (2014) Generative adversarial nets. In: Proceedings of the 27th International Conference on Neural Information Processing Systems—Volume 2, MIT Press, Cambridge, MA, USA, NIPS'14, pp 2672–2680
79. Yu B, Zhou L, Wang L, Fripp J, Bourgeat P (2018) 3D cGAN based cross-modality MR image synthesis for brain tumor segmentation. In: 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), vol 1, pp 626–630
80. Li Z, Wang Y, Yu J (2018) Brain tumor segmentation using an adversarial network. In: Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. Springer, Cham, pp 123–132
81. Singh VK, Rashwan HA, Abdel-Nasser M, Sarker MMK, Akram F, Pandey N, Romani S, Puig D (2019) An efficient solution for breast tumor segmentation and classification in ultrasound images using deep adversarial learning. arXiv:Image and Video Processing
82. Kamnitsas K, Ledig C, Newcombe V, Simpson JP, Kane AD, Menon DK, Rueckert D, Glocker B (2017) Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. Med Image Anal 36:61–78
83. Han X (2017) MR-based synthetic CT generation using a deep convolutional neural network method. Med Phys 44(4):1408–1419
84. Tseng K, Lin Y, Hsu W, Huang C (2017) Joint sequence learning and cross-modality convolution for 3d biomedical segmentation. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), vol 1, pp 3739–3746
85. Hochreiter S, Schmidhuber J (1997) Long short-term memory. Neural Comput 9(8):1735–1780
86. Xie S, Tu Z (2017) Holistically-nested edge detection. Int J Comput Vis 125(1):3–18
87. Wang X, Peng Y, Lu L, Lu Z, Bagheri M, Summers RM (2019) ChestX-ray: hospital-scale chest X-ray database and benchmarks on weakly supervised classification and localization of common thorax diseases. Springer, Cham, pp 369–392
88. Lin T, Goyal P, Girshick R, He K, Dollár P (2017) Focal loss for dense object detection. In: 2017 IEEE International Conference on Computer Vision (ICCV), vol 1, pp 2999–3007
89. Yuan Y, Chao M, Lo Y (2017) Automatic skin lesion segmentation using deep fully convolutional networks with Jaccard distance. IEEE Trans Med Imaging 36(9):1876–1886
90. Zhong Z, Kim Y, Plichta KA, Allen BG, Zhou L, Buatti JM, Wu X (2018) Simultaneous cosegmentation of tumors in PET-CT images using deep fully convolutional networks. Med Phys 46(2):619–633
91. Jin D, Guo D, Ho TY, Harrison AP, Xiao J, Tseng CK, Lu L (2019) Accurate esophageal gross tumor volume segmentation in PET/CT using two-stream chained 3D deep network fusion. In: Medical Image Computing and Computer Assisted Intervention—MICCAI 2019. Springer, Cham, pp 182–191
92. Zhao X, Li L, Lu W, Tan S (2018) Tumor co-segmentation in PET/CT using multi-modality fully convolutional neural network. Phys Med Biol 64(1):015011
93. Tahir B, Iqbal S, Khan MUG, Saba T, Mehmood Z, Anjum A, Mahmood T (2019) Feature enhancement framework for brain tumor segmentation and classification. Microsc Res Tech 82(6):803–811
94. Zhong Z, Kim Y, Buatti J, Wu X (2017) 3D alpha matting based co-segmentation of tumors on PET-CT images. In: Molecular Imaging, Reconstruction and Analysis of Moving Body Organs, and Stroke Imaging and Treatment. Springer, Cham, pp 31–42
95. Sheller MJ, Reina GA, Edwards B, Martin J, Bakas S (2019) Multi-institutional deep learning modeling without sharing patient data: a feasibility study on brain tumor segmentation. In: Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. Springer, Cham
96. Husham A, Alkawaz MH, Saba T, Rehman A, Alghamdi J (2016) Automated nuclei segmentation of malignant using level sets. Microsc Res Tech 79(10):993–997
97. Mittal M, Goyal LM, Kaur S, Kaur I, Verma A, Hemanth DJ (2019) Deep learning based enhanced tumor segmentation approach for MR brain images. Appl Soft Comput 78:346–354
98. Sampergonzalez J, Burgos N, Bottani S, Fontanella S, Lu P, Marcoux A, Routier A, Guillon J, Bacci M, Wen J et al (2018) Reproducible evaluation of classification methods in Alzheimer's disease: framework and application to MRI and pet data. Neuroimage 183:504–521
99. Zuiderveld KJ (1994) Contrast limited adaptive histogram equalization. Graphics Gems, pp 474–485
100. Polesel A, Ramponi G, Mathews VJ (2000) Image enhancement via adaptive unsharp masking. IEEE Trans Image Process 9(3):505–510

101. Krähenbühl P, Koltun V (2011) Efficient inference in fully connected crfs with gaussian edge potentials. In: Shawe-Taylor J, Zemel R, Bartlett P, Pereira F, Weinberger KQ (eds) Advances in Neural Information Processing Systems, vol 24. Curran Associates Inc
102. Chang J, Zhang L, Gu N, Zhang X, Ye M, Yin R, Meng Q (2019) A mix-pooling CNN architecture with FCRF for brain tumor segmentation. J Vis Commun Image Represent 58:316–322
103. Soltaninejad M, Zhang L, Lambrou T, Yang G, Allinson N, Ye X (2018) MRI brain tumor segmentation and patient survival prediction using random forests and fully convolutional networks. In: Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. Springer, Cham, pp 204–215
104. Xia K, Wei G (2014) Persistent homology analysis of protein structure, flexibility, and folding. Int J Numer Methods Biomed Eng 30(8):814–844
105. Chlebus G, Meine H, Moltz JH, Schenk A (2017) Neural network-based automatic liver tumor segmentation with random forest-based candidate filtering. arXiv:1706.00842
106. Chen S, Ding C, Liu M (2019) Dual-force convolutional neural networks for accurate brain tumor segmentation. Pattern Recogn 88:90–100
107. Ma Z, Wu X, Sun S, Xia C, Yang Z, Li S, Zhou J (2018) A discriminative learning based approach for automated nasopharyngeal carcinoma segmentation leveraging multi-modality similarity metric learning. In: 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), vol 1, pp 813–816
108. Zhao L, Lu Z, Jiang J, Zhou Y, Wu Y, Feng Q (2019) Automatic nasopharyngeal carcinoma segmentation using fully convolutional networks with auxiliary paths on dual-modality PET-CT images. J Digit Imaging 32(3):462–470
109. Ma Z, Zhou S, Wu X, Zhang H, Yan W, Sun S, Zhou J (2019) Nasopharyngeal carcinoma segmentation based on enhanced convolutional neural networks using multi-modal metric learning. Phys Med Biol 64(2):025005
110. Feng X, Yang J, Laine AF, Angelini ED (2017) Discriminative localization in CNNs for weakly-supervised segmentation of pulmonary nodules. In: Medical Image Computing and Computer Assisted Intervention—MICCAI 2017. Springer, Cham, pp 568–576
111. Wang W, Lu Y, Wu B, Chen T, Chen DZ, Wu J (2018) Deep active self-paced learning for accurate pulmonary nodule segmentation. In: Medical Image Computing and Computer Assisted Intervention—MICCAI 2018. Springer, Cham, pp 723–731
112. Jiang J, Hu YC, Tyagi N, Zhang P, Rimner A, Mageras GS, Deasy JO, Veeraraghavan H (2018) Tumor-aware, adversarial domain adaptation from CT to MRI for lung cancer segmentation. In: Medical Image Computing and Computer Assisted Intervention—MICCAI 2018. Springer, Cham, pp 777–785
113. Shen H, Zhang J, Zheng W (2017) Efficient symmetry-driven fully convolutional network for multimodal brain tumor segmentation. In: 2017 IEEE International Conference on Image Processing (ICIP), vol 1, pp 3864–3868
114. Moreno Lopez M, Ventura J (2018) Dilated convolutions for brain tumor segmentation in MRI scans. In: Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. Springer, Cham, pp 253–262
115. McKinley R, Meier R, Wiest R (2019) Ensembles of densely-connected CNNs with label-uncertainty for brain tumor segmentation. In: Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. Springer, Cham, pp 456–465
116. Akbari M, Mohrekesh M, Nasr-Esfahani E, Soroushmehr SMR, Karimi N, Samavi S, Najarian K (2018) Polyp segmentation in colonoscopy images using fully convolutional network. In: 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), vol 1, pp 69–72
117. Kang J, Gwak J (2019) Ensemble of instance segmentation models for polyp segmentation in colonoscopy images. IEEE Access 7:26440–26447
118. Huang Y, Dou Q, Wang Z, Liu L, Wang L, Chen H, Heng P, Xu R (2018) HL-FCN: Hybrid loss guided FCN for colorectal cancer segmentation. In: 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), vol 1, pp 195–198
119. Nguyen Q, Lee S (2018) Colorectal segmentation using multiple encoder-decoder network in colonoscopy images. In: 2018 IEEE First International Conference on Artificial Intelligence and Knowledge Engineering (AIKE), vol 1, pp 208–211
120. Lv Y, Wang J (2019) Kidney tumor segmentation based on U-Net and V-Net with double loss function training. https://doi.org/10.24926/548719.054

121. Mu G, Lin Z, Han M, Yao G, Gao Y (2019) Segmentation of kidney tumor by multi-resolution VB-nets. https://doi.org/10.24926/548719.003
122. Wei H, Wang Q, Zhao W, Zhang M, Yuan K, Li Z (2019) Two-phase framework for automatic kidney and kidney tumor segmentation. https://doi.org/10.24926/548719.043
123. Jin Q, Meng Z, Sun C, Cui H, Su R (2020) RA-UNet: a hybrid deep attention-aware network to extract liver and tumor in CT scans. Front Bioeng Biotechnol 8:1471
124. Qin Y, Zheng H, Huang X, Yang J, Zhu Y (2019) Pulmonary nodule segmentation with CT sample synthesis using adversarial networks. Med Phys 46(3):1218–1229
125. Chen H, Qi Y, Yin Y, Li T, Liu X, Li X, Gong G, Wang L (2020) MMFNet: a multi-modality MRI fusion network for segmentation of nasopharyngeal carcinoma. Neurocomputing 394:27–40
126. Singh L, Chetty G, Sharma D (2012) A novel machine learning approach for detecting the brain abnormalities from MRI structural images. In: Pattern Recognition in Bioinformatics. Springer, Berlin, pp 94–105
127. Havaei M, Davy A, Wardefarley D, Biard A, Courville A, Bengio Y, Pal C, Jodoin P, Larochelle H (2017) Brain tumor segmentation with deep neural networks. Med Image Anal 35:18–31
128. Pereira S, Oliveira A, Alves V, Silva CA (2017) On hierarchical brain tumor segmentation in MRI using fully convolutional neural networks: a preliminary study. In: 2017 IEEE 5th Portuguese Meeting on Bioengineering (ENBENG), vol 1, pp 1–4
129. Isensee F, Kickingereder P, Bonekamp D, Bendszus M, Wick W, Schlemmer HP, Maier-Hein K (2017) Brain tumor segmentation using large receptive field deep convolutional neural networks. In: Bildverarbeitung für die Medizin 2017. Springer, Berlin, pp 86–91
130. Zhou C, Ding C, Lu Z, Wang X, Tao D (2018) One-pass multi-task convolutional neural networks for efficient brain tumor segmentation. In: Medical Image Computing and Computer Assisted Intervention—MICCAI 2018. Springer, Cham, pp 637–645
131. Pereira S, Alves V, Silva CA (2018) Adaptive feature recombination and recalibration for semantic segmentation: application to brain tumor segmentation in MRI. In: Medical Image Computing and Computer Assisted Intervention—MICCAI 2018. Springer, Cham, pp 706–714
132. Myronenko A (2019) 3D MRI brain tumor segmentation using autoencoder regularization. In: Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. Springer, Cham, pp 311–320
133. Huang Y, Chen D, Lin Y (2019) 3D contouring for breast tumor in sonography. arXiv: Computer Vision and Pattern Recognition
134. Arjmand A, Meshgini S, Afrouzian R, Farzamnia A (2019) Breast tumor segmentation using k-means clustering and cuckoo search optimization. In: 2019 9th International Conference on Computer and Knowledge Engineering (ICCKE), vol 1, pp 305–308
135. Liu L, Li K, Qin W, Wen T, Li L, Wu J, Gu J (2018) Automated breast tumor detection and segmentation with a novel computational framework of whole ultrasound images. Med Biol Eng Comput 56(2):183–199
136. Ying T, Li R, Yangyang L, Xuehong C (2019) Micro calcification point automatic detection method based on ultrasonic breast tumor image. Pat. Appl. https://lens.org/053-363-278-817-556
137. El-Azizy ARM, Salaheldien M, Rushdi MA, Gewefel H, Mahmoud AM (2019) Morphological characterization of breast tumors using conventional b-mode ultrasound images. In: 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), vol 1, pp 6620–6623
138. Vakanski A, Xian M, Freer P (2019) Attention enriched deep learning model for breast tumor segmentation in ultrasound images. Ultrasound Med Biol 46(10):2819–2833
139. Liang Y, He R, Li Y, Wang Z (2019) Simultaneous segmentation and classification of breast lesions from ultrasound images using mask R-CNN. In: 2019 IEEE International Ultrasonics Symposium (IUS), vol 1, pp 1470–1472
140. Xie Y, Chen K, Lin J (2017) An automatic localization algorithm for ultrasound breast tumors based on human visual mechanism. Sensors 17(5):1101
141. Shao H, Zhang Y, Xian M, Cheng HD, Xu F, Ding J (2015) A saliency model for automated tumor detection in breast ultrasound images. In: 2015 IEEE International Conference on Image Processing (ICIP), vol 1, pp 1424–1428
142. He K, Gkioxari G, Dollar P, Girshick R (2020) Mask R-CNN. IEEE Trans Pattern Anal Mach Intell 42(2):386–397

143. Huang G, Liu Z, Van Der Maaten L, Weinberger KQ (2017) Densely connected convolutional networks. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), vol 1, pp 2261–2269
144. Zhang J, Saha A, Zhu Z, Mazurowski MA (2019) Hierarchical convolutional neural networks for segmentation of breast tumors in MRI with application to radiogenomics. IEEE Trans Med Imaging 38(2):435–447
145. Men K, Dai J, Li Y (2017) Automatic segmentation of the clinical target volume and organs at risk in the planning CT for rectal cancer using deep dilated convolutional neural networks. Med Phys 44(12):6377–6389
146. Liu X, Guo S, Zhang H, He K, Mu S, Guo Y, Li X (2019) Accurate colorectal tumor segmentation for CT scans based on the label assignment generative adversarial network. Med Phys 46(8):3532–3542
147. Trebeschi S, Van Griethuysen JJM, Lambregts DMJ, Lahaye MJ, Parmer C, Bakers FCH, Peters NHGM, Beetstan RGH, Aerts HJWL (2017) Deep learning for fully-automated localization and segmentation of rectal cancer on multiparametric MR. Sci Rep 7(1):5301
148. Jian J, Xiong F, Xia W, Zhang R, Gu J, Wu X, Meng X, Gao X (2018) Fully convolutional networks (FCNs)-based segmentation method for colorectal tumors on T2-weighted magnetic resonance images. Australas Phys Eng Sci Med 41:393–401
149. Soomro MH, Coppotelli M, Conforto S, Schmid M, Giunta G, Secco LD, Neri E, Caruso D, Rengo M, Laghi A (2019) Automated segmentation of colorectal tumor in 3D MRI using 3D multiscale densely connected convolutional neural network. J Healthc Eng 2019:1–11
150. Huang YJ, Dou Q, Wang ZX, Liu LZ, Jin Y, Li CF, Wang L, Chen H, Xu RH (2020) 3-D RoI-aware U-Net for accurate and efficient colorectal tumor segmentation. IEEE Trans Cybern. https://doi.org/10.1109/TCYB.2020.2980145
151. Tang J, Li J, Xu X (2018) Segnet-based gland segmentation from colon cancer histology images. In: 2018 33rd Youth Academic Annual Conference of Chinese Association of Automation (YAC), vol 1, pp 1078–1082
152. Bray FI, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A (2018) Global cancer statistics 2018: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 68(6):394–424
153. Siegel RL, Miller KD, Jemal A (2019) Cancer statistics, 2019. Ca A Cancer J Clin 69(1)
154. Efremova DB, Konovalov DA, Siriapisith T, Kusakunniran W, Haddawy P (2019) Automatic segmentation of kidney and liver tumors in CT images. arXiv:1908.01279
155. Zavala-Romero O, Matos J, Illan I, Stoyanova R, Breto A, Xu I, Zavala-Hidalgo J, Romero-Centeno R (2019) Nested 3D neural networks for kidney and tumor segmentation. https://doi.org/10.24926/548719.084
156. Shen C, Wang C, Oda M, Mori K (2019) Coarse-to-fine kidney and tumor segmentation with fully convolutional networks https://doi.org/10.24926/548719.072
157. Yu Q, Shi Y, Sun J, Gao Y, Zhu J, Dai Y (2019) Crossbar-net: a novel convolutional neural network for kidney tumor segmentation in CT images. IEEE Trans Image Process 28(8):4060–4074
158. Myronenko A, Hatamizadeh A (2019) 3D kidneys and kidney tumor semantic segmentation using boundary-aware networks. arXiv:1909.06684
159. Sabarinathan D, Beham D, Roomi S (2020) Hyper Vision Net: kidney tumor segmentation using coordinate convolutional layer and attention unit, pp 609–618. https://doi.org/10.1007/978-981-15-8697-2_57
160. Zhang Y, Wang Y, Hou F, Yang J, Xiong G, Tian J, Zhong C (2019) Cascaded volumetric convolutional network for kidney tumor segmentation from CT volumes. https://doi.org/10.24926/548719.004. arXiv:Image and Video Processing
161. Chen C, Ma L, Jia Y, Zuo P (2019) Kidney and tumor segmentation using modified 3D mask RCNN. https://doi.org/10.24926/548719.061
162. Sun L, Zeng W, Ding X, Huang Y (2019) A multi-scale attention network for kidney tumor segmentation on CT scans. https://doi.org/10.24926/548719.055
163. Forner A (2015) Hepatocellular carcinoma surveillance with miRNAs. Lancet Oncol 16(7):743–745
164. Roy S (2013) A review on automated brain tumor detection and segmentation from MRI of brain. Int J Adv Res Comput Sci Softw Eng 3:1706–1746

165. Vorontsov E, Tang A, Pal C, Kadoury S (2018) Liver lesion segmentation informed by joint liver segmentation. In: 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), vol 1, pp 1332–1335
166. Liu S, Xu D, Zhou SK, Pauly O, Grbic S, Mertelmeier T, Wicklein J, Jerebko A, Cai W, Comaniciu D (2018) 3D anisotropic hybrid network: transferring convolutional features from 2D images to 3D anisotropic volumes. In: Medical Image Computing and Computer Assisted Intervention—MICCAI 2018. Springer, Cham, pp 851–858
167. Bai Z, Jiang H, Li S, Yao Y (2019) Liver tumor segmentation based on multi-scale candidate generation and fractal residual network. IEEE Access 7:82122–82133
168. Seo H, Huang C, Bassenne M, Xiao R, Xing L (2020) Modified U-Net (mU-Net) with incorporation of object-dependent high level features for improved liver and liver-tumor segmentation in CT images. IEEE Trans Med Imaging 39(5):1316–1325
169. Siegel RL, Miller KD, Jemal A (2016) Cancer statistics, 2016. CA Cancer J Clin 66(1):7–30
170. Wang S, Zhou M, Gevaert O, Tang Z, Dong D, Liu Z, Jie T (2017) A multi-view deep convolutional neural networks for lung nodule segmentation. In: 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), vol 1, pp 1752–1755
171. Rocha J, Cunha A, Maria Mendonça A (2019) Comparison of conventional and deep learning based methods for pulmonary nodule segmentation in CT images. In: Progress in Artificial Intelligence. Springer, Cham, pp 361–371
172. Tong G, Li Y, Chen H, Zhang Q, Jiang H (2018) Improved U-NET network for pulmonary nodules segmentation. Optik 174:460–469
173. Roy R, Chakraborti T, Chowdhury AS (2019) A deep learning-shape driven level set synergism for pulmonary nodule segmentation. Pattern Recogn Lett 123:31–38
174. Feng Y, Hao P, Zhang P, Liu X, Wu F, Wang H (2019) Supervoxel based weakly-supervised multi-level 3D CNNs for lung nodule detection and segmentation. J Ambient Intell Humaniz Comput 1–11
175. Wu B, Zhou Z, Wang J, Wang Y (2018) Joint learning for pulmonary nodule segmentation, attributes and malignancy prediction. In: 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), vol 1, pp 1109–1113
176. Achanta R, Shaji A, Smith K, Lucchi A, Fua P, Susstrunk S (2012) Slic superpixels compared to state-of-the-art superpixel methods. IEEE Trans Pattern Anal Mach Intell 34(11):2274–2282
177. Chang ET, Adami H (2006) The enigmatic epidemiology of nasopharyngeal carcinoma. Cancer Epidemiol Biomark Prev 15(10):1765–1777
178. Men K, Chen X, Zhang Y, Zhang T, Dai J, Yi J, Li Y (2017) Deep deconvolutional neural network for target segmentation of nasopharyngeal cancer in planning computed tomography images. Front Oncol 7:315
179. Liang S, Tang F, Huang X, Yang K, Zhong T, Hu R, Liu S, Yuan X, Zhang Y (2019) Deep-learning-based detection and segmentation of organs at risk in nasopharyngeal carcinoma computed tomographic images for radiotherapy planning. Eur Radiol 29:1961–1967
180. Zhong T, Huang X, Tang F, Liang S, Deng X, Zhang Y (2019) Boosting-based cascaded convolutional neural networks for the segmentation of CT organs-at-risk in nasopharyngeal carcinoma. Med Phys 46(12):5602–5611
181. Li S, Xiao J, He L, Peng X, Yuan X (2019) The tumor target segmentation of nasopharyngeal cancer in CT images based on deep learning methods. Technol Cancer Res Treat 18:1533033819884561
182. Wang Y, Zu C, Hu G, Luo Y, Ma Z, He K, Wu X, Zhou J (2018) Automatic tumor segmentation with deep convolutional neural networks for radiotherapy applications. Neural Process Lett 48(3):1323–1334
183. Ma Z, Wu X, Zhou J (2017) Automatic nasopharyngeal carcinoma segmentation in MR images with convolutional neural networks. In: 2017 International Conference on the Frontiers and Advances in Data Science (FADS), vol 1, pp 147–150
184. Ma Z, Wu X, Song Q, Luo Y, Wang Y, Zhou J (2018) Automated nasopharyngeal carcinoma segmentation in magnetic resonance images by combination of convolutional neural networks and graph cut. Exp Ther Med 16(3):2511–2521
185. Li Q, Xu Y, Chen Z, Liu D, Feng S, Law M, Ye Y, Huang B (2018) Tumor segmentation in contrast-enhanced magnetic resonance imaging for nasopharyngeal carcinoma: deep learning with convolutional neural network. Biomed Res Int 2018:9128527

186. Jb Huang, Zhuo E, Li H, Liu L, Cai H, Ou Y (2019) Achieving accurate segmentation of nasopharyngeal carcinoma in MR images through recurrent attention. In: Medical Image Computing and Computer Assisted Intervention—MICCAI 2019. Springer, Cham, pp 494–502
187. Xing F, Bennett T, Ghosh D (2019) Adversarial domain adaptation and pseudo-labeling for cross-modality microscopy image quantification. In: Medical Image Computing and Computer Assisted Intervention—MICCAI 2019. Springer, Cham, pp 740–749
188. Chen C, Dou Q, Chen H, Qin J, Heng PA (2020) Unsupervised bidirectional cross-modality adaptation via deeply synergistic image and feature alignment for medical image segmentation. IEEE Trans Med Imaging 39(7):2494–2505
189. Dou Q, Ouyang C, Chen C, Chen H, Glocker B, Zhuang X, Heng P (2019) PnP-AdaNet: plug-and-play adversarial domain adaptation network at unpaired cross-modality cardiac segmentation. IEEE Access 7(1):99065–99076
190. Yan W, Wang Y, Gu S, Huang L, Yan F, Xia L, Tao Q (2019) The domain shift problem of medical image segmentation and vendor-adaptation by Unet-GAN. In: Medical Image Computing and Computer Assisted Intervention—MICCAI 2019. Springer, Cham, pp 623–631
191. Ouyang C, Kamnitsas K, Biffi C, Duan J, Rueckert D (2019) Data efficient unsupervised domain adaptation for cross-modality image segmentation. In: Medical Image Computing and Computer Assisted Intervention—MICCAI 2019. Springer, Cham, pp 669–677
192. Huang C, Han H, Yao Q, Zhu S, Zhou SK (2019) 3D U^2-Net: a 3D universal U-Net for multi-domain medical image segmentation. In: Medical Image Computing and Computer Assisted Intervention—MICCAI 2019. Springer, Cham, pp 291–299
193. Wang T, Yuan L, Zhang X, Feng J (2019) Distilling object detectors with fine-grained feature imitation. In: 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), vol 1, pp 4928–4937
194. Gotmare A, Keskar NS, Xiong C, Socher R (2019) A closer look at deep learning heuristics: Learning rate restarts, warmup and distillation. In: International Conference on Learning Representations. https://openreview.net/forum?id=r14EOsCqKX
195. Pilzer A, Lathuilière S, Sebe N, Ricci E (2019) Refine and distill: exploiting cycle-inconsistency and knowledge distillation for unsupervised monocular depth estimation. In: 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), vol 1, pp 9760–9769
196. Bhardwaj K, Suda N, Marculescu R (2019) Dream distillation: a data-independent model compression framework. CoRR. arXiv:1905.07072