Informatics Bulletin, Faculty of Computers and Artificial Intelligence, Helwan University
Published Online ***Vol 6 Issue 1, January 2024***
(https://fcihib.journals.ekb.eg)

Deep Learning Medical Image Segmentation Methods: A Survey

Mai Mokhtar[1*], Hala Abdel-Galil[1], Ghada Khoriba[1, 2]
[1]Computer Science Department, Faculty of Computers and Artificial Intelligence, Helwan University, Cairo, Egypt
[2]School of Information Technology and Computer Science ITCS, Nile University, Giza, Egypt
maimokhtar@fci.helwan.edu.eg, hala.nagy@fci.helwan.edu.eg,
ghada_khoriba@fci.helwan.edu.eg, ghadakhoriba@nu.edu.eg

***Abstract*—Medical image segmentation is essential for detecting and localizing tumors in medical image analysis. Image segmentation involves the identification of anatomical structures in images. Medical image segmentation starts with manual segmentation using Atlas methods, then auto-segmentation, facilitated by deep learning algorithms. Deep learning-based medical image segmentation retains a significant pledge in reducing treatment planning, radiation-related toxicities, and side effects. This study provides a complete overview of deep-learning medical image segmentation models. We review various deep-learning models and architectures applied to medical image segmentation, including fully convolutional networks, U-Net, and attention-based models. This literature review discusses using different loss functions, data augmentation techniques, and transfer learning in deep learning-based medical image segmentation and several types of medical image modality. Evaluation analysis encloses benchmark datasets for human body organs such as the brain, lungs, chest, and liver. Finally, we summarize the challenges and future directions of deep learning for medical image segmentation.**

***Index Terms*—Medical Image Segmentation, Computed Tomography (CT), Magnetic Resonance Imaging (MRI), Deep learning, CNN, U-Net**

I. INTRODUCTION

IMAGE segmentation divides an image into multiple segments or regions. It is an essential issue in computer vision, with applications ranging from object detection to medical imaging to autonomous driving. The traditional methods of image segmentation involve handcrafted features and supervised learning algorithms. However, with the introduction of deep learning, the segmentation of images is revolutionized. Deep learning algorithms outperform humans in segmenting image tasks, making it an active research area in recent years. FCNs are among the first algorithms to use a deep learning-based fully convolutional neural network for image segmentation. They can produce dense pixel-wise predictions, making them suitable for semantic segmentation tasks. One of the recent architectures is U-Net, a popular algorithm for deep learning-based medical image segmentation. It is an end- to-end architecture that comprises a network of encoders and decoders. The encoder network performs feature extraction, while the decoder network performs upsampling to produce a segmentation mask. The literature proposes various segmentation techniques or algorithms that overcome the limitations of traditional medical segmentation approaches. The appropriate approach or algorithm selection depends on the image issue's nature. As a result, examining recent developments in image segmentation methods is required to determine the best approach for a specific medical image problem [1].

The most common medical imaging technology types used in clinical diagnosis are computed tomography (CT), X-rays, magnetic resonance imaging (MRI), ultrasound imaging (UI), and positron emission tomography (PET). In addition, popular RGB images include microscope and fundus images of the retina. Medical imaging provides vast information, and physicians utilize CT scans alongside additional diagnostic images to evaluate the patient's health. Consequently, the research focus on computer vision shifts to medical image processing, as it is now the primary concern of medical professionals to extract as much information as possible from these images for diagnosis and treatment planning [2]. Deep learning-based image segmentation algorithms demonstrate promising results in image segmentation with the fast expansion of artificial intelligence, specifically deep learning [3]. Regarding segmentation accuracy and speed, deep learning provides several benefits over traditional machine learning and computer vision approaches. Clinicians can validate the size of tumor tissues, objectively assess the impact before and after therapy, and considerably lessen their burden thanks to effective medical image segmentation using deep learning.

The contribution of this paper is to present a study to provide insights into deep-learning medical image segmentation. It represents the challenges and open issues to promote interest in investigating and exploring medical image segmentation. This paper presents the approaches and deep learning models to give insights to the researchers who get started in medical image segmentation. It also illustrates the types of these images with the most frequently used dataset for different organs and their use with the deep learning models.

The remaining sections of the paper are structured as follows: Section II shows the image segmentation approaches, followed by an overview of deep learning segmentation models in Section III. We present various types of medical images related to the segmentation problem in Section IV. In Section V, medical image segmentation datasets are illustrated. Section VI focuses on image segmentation techniques for different organs. Section VII illustrates the challenges and open issues, and the paper concludes with a summary in Section VIII.

II. Image segmentation approaches

Image segmentation is a crucial subject in the study of computer vision, and in recent years, it has received attention in research on how to comprehend images. Segmentation is the process of breaking up an image into different parts depending on many features, including color, grayscale, spatial texture, and geometric forms [4]. The objective is to retain consistency across features within a single region while making it evident when one area differs. Image segmentation may be divided into semantic, instance, and panoramic segments depending on the level of feature depth. Semantic segmentation is one of these areas and includes medical image segmentation. The subfields of image segmentation research include autonomous driving, medical image segmentation, and satellite image segmentation, to name a few [5]. The image segmentation approach gradually improves in accuracy as the recommended network model architecture. Nevertheless, no unique segmentation method can be used for all images and segmentation situations [6].

Traditional image segmentation approaches in computer vision and image processing are no longer adequate compared to deep learning-based approaches, although they are still valuable to learn [7, 8, 9]. Edge detection, threshold-based segmentation, and region-based segmentation are some of the classical approaches that segment images based on principles of mathematics and digital image processing [10, 11, 12]. These techniques are computationally efficient and have fast segmentation speeds. However, they need more specificity in terms of segmentation accuracy. Deep learning-based approaches have shown significant advances in image segmentation in recent years and demonstrated higher segmentation accuracy than conventional methods. The first successful use of deep learning to semantically segment images is achieved using a fully convolutional network. Many exceptional segmentation networks, such as RefineNet, Mask R-CNN, U-Net, and DeconvNet [13, 14, 15, 16, 17, 18], offer superior processing capability for fine edges.

III. Deep Learning Segmentation models

Convolutional neural networks (CNNs) possess robust feature expression and feature extraction abilities, which make them an ideal candidate for medical image segmentation tasks. They eliminate the need for extensive image preprocessing or manual feature extraction, resulting in more streamlined and efficient image segmentation procedures. Most recently, CNNs have been employed to segment medical images, achieving notable success. FCN, U-Net, and GAN are the categories currently existing in medical image segmentation methods based on deep learning. This section illustrates the results of the classical research studies and the benefits and limitations of each segmentation method.

A. Fully Convolutional Neural Networks

A fully convolutional network (FCN) is one of the first networks to implement this technique. This section analyzes the advantages and disadvantages of fully convolutional networks and highlights the various applications of a fully convolutional network and its variants.

1) FCN: Adding fully connected layers to convolutional neural networks, such as ResNet and VGG, provides category probability information after the softmax layer. However, only the overall image category may be determined, not the category of individual pixels, rendering it unsuitable while segmenting images. The Fully Convolutional Network (FCN) was suggested by Long et al. [19] as a solution to this problem with a 20% relative improvement to 62.2% mean IU in 2012. Convolutional layers comprise the first five levels of a standard CNN design, while fully connected (one-dimensional vector) layers of 4096 in length comprise the sixth and seventh layers. The eighth layer is a fully connected layer that has a length of 1000, signifying the likelihood of 1000 categories. Layers five through seven are turned into convolutional layers with different kernel sizes of 7 x 7, 1 x 1, and 1 x 1, respectively, to produce each pixel having a two-dimensional feature map. The softmax layer is used to obtain pixel categorization information, resulting in the resolution of the segmentation problem. FCN does not have any input image size restrictions. The final convolution layer's feature map is upsampled and restored to the original input image size using the deconvolution layer.

2) SegNet: SegNet is a pixel-level image segmentation model that builds upon the FCN semantic segmentation task and uses an encoder-decoder symmetric structure [20]. VGG16 is the encoder to extract object information from the input image. The decoder assigns each pixel a color or label associated with its object information, which generates the final image. Unlike FCN's deconvolution operation, which is used for upsampling lower-resolution feature maps, SegNet upsamples its input in a non-linear manner using a more extensive pooling index from the encoder rather than learning how to do so. This approach produces a dense feature map by generating a trainable convolution kernel applied to a sparse feature map. Eventually, the softmax classifier categorizes pixels in the feature maps, which have been restored to their original resolution. Depooling the low-resolution feature maps helps preserve the information with high frequency, enhances edge detection, and decreases the number of training parameters.

B. U-Net

1) 2D U-Net: To improve the widely-used FCN model, Ronneberger et al. [21] proposed a U-Net structure for biological images. U-Net and its variants have shown remarkable success in several areas of computer vision, which has resulted in over 4000 citations since its introduction at the 2015 MICCAI conference. Although many recent innovations in convolutional neural network design have occurred, U-Net's concept remains an essential reference, with many researchers incorporating additional modules or design principles.

The U-Net network is comprised of a U-shaped architecture with skip connections. It has a structure that is similar to the encoder-decoder structure of SegNet. Each of the four sub-modules in the encoder contains two convolutional layers used for downsampling, followed by max pooling. Similarly, the decoder includes four sub-modules, with incremental upsampling to increase the resolution. Finally, pixel-wise predictions are made. Figure 1 displays the network architecture, which results in an output of 388 x 388 from a 572 x 572 input due to the medical profession's need for greater segmentation precision. The network relies solely on convolution and downsampling without any fully connected layers. Each upsampling layer's output is connected to the appropriate submodule in the encoder with the exact resolution using a skip connection.

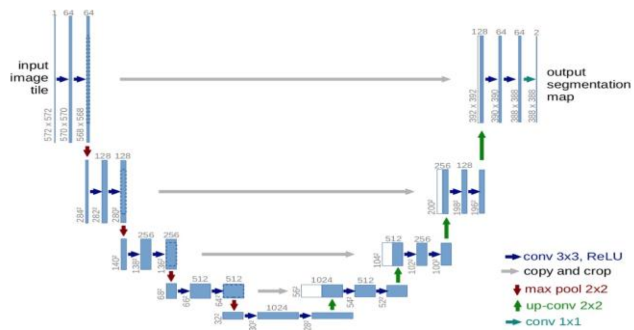

Fig. 1: 2D U-Net Structure [22].

2) 3D U-Net: Current medical image segmentation research is focused on developing the U-Net model, and several versions of it have been created. One of the versions is the 3D U-Net model presented by içek et al. [23]. This approach provides the U-Net structure with more geographic information. The network architecture is shown in Figure 2; it is the same as U-Net in that there is only one path for encoding and one for decoding. Each course in the network may be adjusted to one of four different resolutions. Two 3 × 3 convolutions and a ReLU layer comprise each encoding route layer; a maximum pooling layer is used for dimensionality reduction. Each convolution layer in the decoding pipeline consists of two 3 × 3 × 3 convolution layers, followed by a 2 × 2 × 2 deconvolution layer with a stride of 2, and finally, a ReLU layer. The layer in the encoding route with the exact resolution is given to the decoding path via a shortcut, which grants it access to high-resolution versions of the original features. The network can segment 3D images if it is got a sequence of images in which each slice is a 2D representation of the original 3D image. In addition to training on a sparsely labeled data set and different unlabeled places on that data set, this network may be trained on several sparsely labeled data sets and then used to predict new data. Compared to U-Net input, the stereo image input has three channels and is more significant at 132 × 132 × 116 pixels. The final image is 44 × 44 × 28 pixels. The outstanding original features of FCN and U-Net are still present in 3D U-Net, and volumetric images can benefit from its introduction.

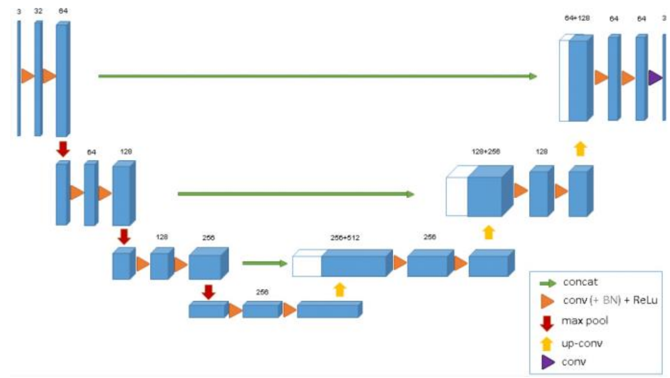

Fig. 2: 3D U-Net Structure [23].

3) Segmentation Adversarial Network (SegAN): Using the U-Net architecture, Xue et al. [24] created an adversarial segmentation network (SegAN) for generating segments. Successful segmentation of medical images requires overcoming a number of obstacles, one of the most critical of which is the issue of unbalanced pixel categories. To optimize the segmentation network, the authors introduced a novel GAN-based segmentation network and a multiscale $L1loss$ function. SegAN consists of segment network S and critic network *C*. A model with excellent performance is developed by alternatingly training a segmenter and a critic network in a min-max game. S adopts a typical U-Net topology for the segmented network. By using convolutional layers with a stride of 2 and a kernel size of 4x4, downsampling may be accomplished. Instead, 3x3 kernels and a stride of 1 are used in convolutional layers to do the upsampling. The original images and those masked by ground-truth-based or S-based label maps are sent to the critic network. This study uses the brain tumor segmentation (BRATS) dataset, which is more reliable and efficient for segmentation tasks. The authors suggested that the multiscale $L1loss$ function optimizes the whole segmentation network compared to a single-scale loss function.

4) Structure Correcting Adversarial Network (SCAN): In clinical practice, the most frequent kind of imaging is a chest X-ray (CXR) which is used as an imaging modality for detecting various cardiopulmonary pathologies due to its low radiation exposure and affordability, which account for more than 55% of all medical imaging. Therefore, developing computer-aided detection (CAD) techniques that can complement CXRs is essential for medical professionals. Dai et al. [25] suggested a structural correction adversarial network (SCAN) to achieve as segment the heart and lung regions in CXR images. Unlike previous works, where generative adversarial networks (GANs) were employed for image segmentation, this network architecture utilizes both discriminative and segmentation fully convolutional networks (FCNs) to process grayscale CXR images due to the limitations of a relatively small training dataset consisting of 247 images.

In this study, the Fully Convolutional Network (FCN) is modified to be trained without using previously-established model information. Human physiological regularities are used by the critic network to place constraints on the convolutional segmentation network. The critic network can improve its accuracy by comparing the mask produced by the segmentation network during training with the ground truth organ annotations. This allows the critic network to acquire higher-order structures during the confrontation phase, which helps direct the segmentation model toward more precise findings. The proposed method, SCAN, also incorporates a downsampling module tailored to the unique characteristics of CXR images. As results summary, Table I shows different segmentation models with their results and the datasets used.

IV. Types of Medical Images

The choice of the input data is an important part of making accurate and useful deep-learning models for medical image segmentation. In this survey, we review the different types of medical images that might be utilized as data for image segmentation problems, including X-ray, CT, ultrasound, MRI, and PET images. We discuss the unique characteristics of each modality, their strengths and limitations, and how they can be utilized to enhance medical image segmentation accuracy.

A. Computed Tomography (CT)

X-rays with narrow beams are focused on the subject and rapidly spun around the body, computed tomography (CT) generates signals processed by a computer to create cross-sectional images, also known as tomographic images. Compared to conventional x-rays, these images offer more comprehensive information, allowing doctors to quickly spot important parts, cancers, or abnormalities [26, 27]. The computer creates a three-dimensional (3D) image of the patient from numerous sequential slices, which might be helpful for diagnosis and therapy planning. A CT scan may provide tomographic images of specific bodily areas, such as the brain or bladder, without requiring surgical incisions by integrating several x-ray readings collected at different angles.

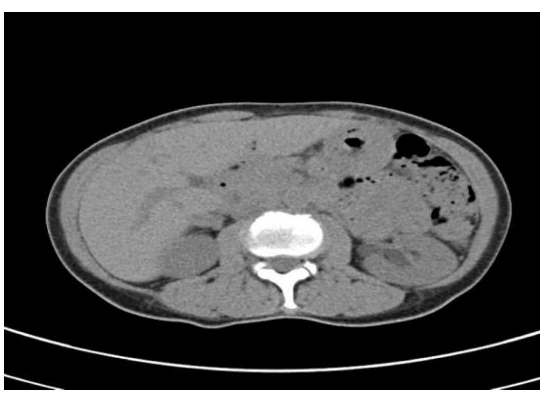

Fig. 3: Liver CT.

B. X-Ray

To aid in diagnosing damage and sickness, diagnostic imaging uses electromagnetic radiation in the form of X-rays. The body's tissues are penetrated by X-rays, which then cast a black-and-white image onto a screen. Bones, teeth and the chest may all be examined using standard X-ray techniques (lungs, heart, bones). Most people have had an X-ray taken, for example, during dental work or following a bone fracture [28, 29, 30].

Fig. 4: Chest X-Ray.

C. Magnetic Resonance Imaging (MRI)

Magnetic resonance imaging, as compared to CT scans and X-rays, is a non-radiative, three-dimensional imaging technology that generates comprehensive anatomical images. MRI utilizes potent magnetic fields and radio frequency pulses to generate finely detailed images of interior body structures, including organs, smooth tissues, and bones. A patient must remain still while positioned within a large magnet to produce an MRI image. MRIs offer images of the same body parts as CT scans but with greater depth, providing a more detailed examination of specific structures [31, 32].

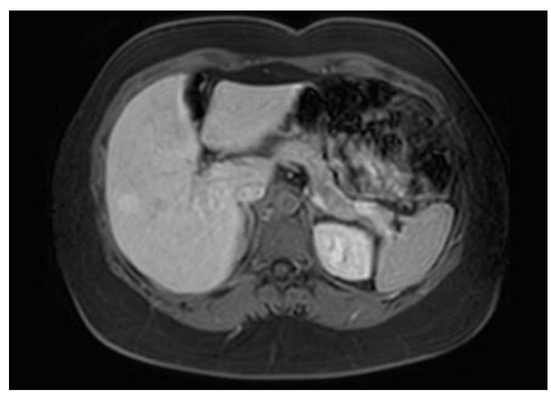

Fig. 5: Liver MRI.

D. Ultrasound Imaging (UI)

Ultrasound diagnostic imaging, also known as diagnostic sonography or ultrasonography, utilizes high-frequency sound waves to create an image of the interior organs of the body, blood vessels, joints, muscles, and tendons. This non-invasive method captures echoes from the waves penetrating the body, generating a real-time image. Unlike x-rays, ultrasound does not use ionizing radiation [33, 34].

Fig. 6: Unborn UI.

E. Positron-Emission Tomography (PET)

Positron emission tomography, a nuclear medicine technique, tracks the metabolic activity of cells in various bodily tissues by combining nuclear medicine and biochemical analysis. PET assists in visualizing the biochemical processes occurring in the body, such as the metabolism of the heart muscle, which involves the transformation of food into energy following digestion and circulation. PET is mainly used in cancer, brain disorders, or heart ailments patients [35, 36].

PET can be used with other diagnostic procedures, for example, MRI or CT, to provide more expressive information about malignant tumors and other anomalies. PET and CT technologies have been integrated into a recent technological innovation, the PET/CT scanner. PET/CT is a promising tool for diagnosing and treating lung cancer, Alzheimer's, and coronary artery disease [37, 38, 39].

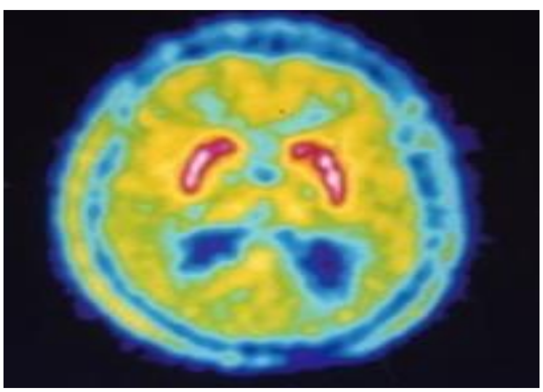

Fig. 7: Brain PET.

V. Medical Image Segmentation Datasets

Obtaining enough data to build the dataset is essential for deep-learning model segmentation. The label-standardized dataset corresponds to expert-annotated high-quality image data and the segmentation algorithm's accuracy enables meaningful comparisons among multiple systems. The benchmark datasets used in medical image segmentation will be described in this section.

Brain Tumor Segmentation (BRATS): This is a MICCAI conference-related dataset that is the result of a brain tumor segmentation competition [40, 41, 42]. The competition has been performed yearly since 2012 to evaluate various techniques for segmenting brain tumors into smaller pieces. A growing number of training sets are added each year.

Lung Image Database Consortium Image Collection (LIDC-IDRI): The dataset consists of medical images of the chest obtained from imaging techniques such as CT scans and X-rays, along with lesion labels indicating the corresponding diagnostic results. This dataset of 1018 research occurrences is used to examine the early detection of cancer in groups at increased risk [43, 44, 45].

Segmentation in Chest Radiographs (SCR): The JSRT database is a widely used collection of chest radiographs in medical imaging. The SCR database was developed to facilitate the comparison of heart, lung area, and clavicle segmentation considering traditional posterior radiographs of the chest [46, 47, 48]. The database is partitioned rigorously to create benchmarks for evaluation of the dataset; 154 images include at least one lung nodule, while the remaining 93 images do not.

Liver Tumor Segmentation Challenge: The purpose of this competition is to motivate researchers to create ways of segmenting liver lesions. The challenge data and slices are contributed from various clinical sites worldwide. The testing dataset comprises 70 computed tomography (CT) scans, while the training dataset consists of 130 CT scans [49, 50, 51].

Liver Tumor Segmentation (LiTS): Some clinical institutions worldwide provide liver segmentation and liver tumor statistics in medical imaging research. There are 130 CT scans in the training dataset and 70 scans in the test dataset. These datasets explore and compare liver segmentation techniques [52, 53, 54].

VI. Image Segmentation for different organs

The human body has diverse organs and tissues with unique characteristics. Researchers extract crucial concepts from such scenarios and develop segmentation algorithms that cater to different organs to enhance segmentation accuracy. Below, we describe the optimal method for segmenting diverse organs.

A. Brain

MRI is a paramount diagnostic tool for brain-related illnesses. Brain imaging analysis provides insights into various brain diseases, including Alzheimer's, schizophrenia, and tumors [65]. Myronenko et al. [66] developed an asymmetric FCN and a deep learning network with residual learning to segment 3D MRI brain tumors, which won first place in a 2018 competition. Nie et al. used T1, T2, and diffusion-weighted modal brain imaging to analyze the brains of eleven healthy babies. They applied a 3D FCN to segment multimodal brain MRI images by incorporating semantic context and combining features of various sizes. To achieve more accurate edge segmentation, Wang et al. [66] improved upon the FCN-8S and other key semantic segmentation networks with an accuracy of 87.31% by proposing a CRF-based edge-sensing FCN by the loss function that includes edge information. Borne et al. achieved 85% accuracy using multiple segmentations with GAN. Adversarial training was employed by Moeskops et al.

TABLE I: Different segmentation models with their results and datasets.

Author	Model	Dataset	Accuracy
Xue et al. [24]	SegAN	BraTS 2015	DSC: 0.85
Zhang et al. [55]	U-Net	BraTS 2017	DSC: 0.85
Zhang et al. [55]	ResU-Net	BraTS 2017	DSC: 0.86
Zhang et al. [55]	AGResU-Net	BraTS 2017	DSC: 0.87
Alqazzaz et al. [56]	SegNet	BraTS 2017	DSC: 0.85
Sun et al. [57]	3D FCN	BraTS 2018	DSC: 0.90
Zhang et al. [55]	AGResU-Net	BraTS 2019	DSC: 0.87
Aboelenein et al. [58]	MIRAU-Net	BraTS 2019	DSC: 0.88
Sheng et al. [59]	ResU-Net	BraTS 2019	DSC: 0.88
Yan et al. [60]	U-Net	BraTS 2021	DSC: 0.87
Ahmed et al. [61]	MS UNet	BraTS 2021	DSC: 0.91
Raza et al. [62]	dResU-Net	BraTS 2021	DSC: 0.86
Suji et al. [63]	U-Net	LIDC-IDRI	IoU: 0.59
Mohagheghi et al. [64]	3D U-Net	Sliver07-I	DSC: 0.97

[67] to enhance completely and dilate convolutional networks' capacity to segment brain MRIs. CNNs were trained for semantic segmentation of brain tumors by Rezaei et al. [68], leading to improved segmentation accuracy.

B. Chest

In the medical field, chest X-rays are frequently used for diagnostic purposes as they are rapid and straightforward. Chest X-rays employ low radiation doses to capture images of the chest, which can be used to segment the lung region and aid in diagnosing and tracking lung diseases, including pneumonia and cancer [69]. The SCAN technique can segment the lung regions and heart in chest X-rays. To address the problem of overfitting and the parameters number in the original U-Net model, Novikov et al. [70] introduced an all-convolutional adaptation of U-Net with stridden convolutions that reduces the parameters number by a significant margin of ten while maintaining accuracy and producing superior results.

C. Abdomen

Abdominal MRI and CT scans may be used to distinguish between the spleen, liver, kidneys, and other organs for diagnostic purposes. Improvements in GAN have also been used in the segmentation of abdominal organs. An adversarial image-to-image network, or DI2IN-AN, was suggested by Yang et al. [71] as a means of liver segmentation. During training, the generator makes segmentation predictions while the discriminator separates them from the ground truth. Due to its size and shape, the spleen presents a significant challenge when attempting to segment MRI images of the organ. The splenomegaly segmentation network (SSNet) was suggested by Huo et al. [72] to solve this problem by using the cGAN architecture. The Markovian Discriminator (PatchGAN) is employed in place of the generator to reduce false negatives and false positives, while the generator uses a Global Convolutional Network (GCN).

VII. Image Segmentation Challenges

Despite their success, deep learning-based approaches still face several challenges and open issues in medical image segmentation. The survey identified several challenges and open issues in medical image segmentation using deep learning approaches. These issues include:

- Insufficient training data: a large amount of labeled training data are necessary for deep learning-based approaches to achieve optimal performance. However, medical image datasets are often limited in size and may not contain enough labeled data for training.
- Class imbalance: Medical images often contain class imbalance, where the number of pixels belonging to the region of interest is significantly lower than the background pixels. This class imbalance can lead to biased segmentation results.
- Model generalization: Deep learning models trained on one set of data may not work well on other sets of data due to differences in how images look and how clinical protocols can change.
- Interpretability and explainability: Deep learning models are often considered ”black boxes,” making it difficult to interpret the segmentation results and provide explanations for clinical decision-making.
- Incorporating prior knowledge: Deep learning models may not incorporate prior knowledge of anatomy or pathology, which can improve segmentation accuracy and reduce false positives.

VIII. Conclusion

Over the past few years, approaches based on deep learning have done well at segmenting medical images because they can learn complex features from the data. Medical images often contain complex structures and require high accuracy in segmentation. Deep learning techniques for medical image segmentation have shown significant promise lately. In this survey, The studies were further categorized based on the medical imaging modalities, for example, ultrasound, CT, and MRI. The studies were also classified based on deep learning algorithms, for example, FCN, SegNet, U-Net, and SegAN.

In addition, this survey paper has provided insights into open issues that can lead to more accurate and reliable medical image segmentation, ultimately improving patient outcomes. Potential solutions to be tested are data augmentation, loss function balancing, transfer learning, and hybrid models combining deep learning and traditional methods.

References

[1] Hui Cui, Hao Wang, Ke Yan, Xiuying Wang, Wangmeng Zuo, and David Dagan Feng. Biomedical image segmentation for precision radiation oncology. In *Biomedical Information Technology*, pages 295–319. Elsevier, 2020.

[2] Manuela Lehmann. Pocus–computer-based training for improving the quality of ultrasonic findings in gallbladder changes. 2022.

[3] Mathieu Hatt, Chintan Parmar, Jinyi Qi, and Issam El Naqa. Machine (deep) learning methods for image processing and radiomics. *IEEE Transactions on Radiation and Plasma Medical Sciences*, 3(2):104–108, 2019.

[4] Dires Teferi Sileabat. *Wheat Quality Assessment Model Using Image Processing and Transfer Learning Techniques*. PhD thesis, 2022.

[5] Jie Guo, Meiting Wang, Yan Zhou, Bin Song, Yuhao Chi, Wei Fan, and Jianglong Chang. Hgan: Hierarchical graph alignment network for image-text retrieval. *arXiv preprint arXiv:2212.08281*, 2022.

[6] Nan Li, Zhe Chen, Xiaoguang Zhang, and Xiang Liu. An ultra-fast bi-phase advanced network for segmenting crop plants from dense weeds. *Biosystems Engineering*, 212:160–174, 2021.

[7] Andreas Maier, Christopher Syben, Tobias Lasser, and Christian Riess. A gentle introduction to deep learning in medical image processing. *Zeitschrift für Medizinische Physik*, 29(2):86–101, 2019.

[8] Swarnendu Ghosh, Nibaran Das, Ishita Das, and Ujjwal Maulik. Understanding deep learning techniques for image segmentation. *ACM Computing Surveys (CSUR)*, 52(4):1–35, 2019.

[9] Yu Song, Zilong Huang, Chuanyue Shen, Humphrey Shi, and David A Lange. Deep learning-based automated image segmentation for concrete petrographic analysis. *Cement and Concrete Research*, 135:106118, 2020.

[10] MO Khairandish, M Sharma, V Jain, JM Chatterjee, and NZ Jhanjhi. A hybrid cnn-svm threshold segmentation approach for tumor detection and classification of mri brain images. *IRBM*, 2021.

[11] Fabio AM Cappabianco, Pedro FO Ribeiro, Paulo AV De Miranda, and Jayaram K Udupa. A general and balanced region-based metric for evaluating medical image segmentation algorithms. In *2019 IEEE international conference on image processing (ICIP)*, pages 1525–1529. IEEE, 2019.

[12] Pramit Brata Chanda and Subir Kumar Sarkar. Study on efficient drlse-oriented edge-based medical image segmentation of cardiac images. In *Emerging Technologies in Data Mining and Information Security*, pages 823–831. Springer, 2021.

[13] Hu Cao, Yueyue Wang, Joy Chen, Dongsheng Jiang, Xiaopeng Zhang, Qi Tian, and Manning Wang. Swin-unet: Unet-like pure transformer for medical image segmentation. *arXiv preprint arXiv:2105.05537*, 2021.

[14] Jian-Hua Shu, Fu-Dong Nian, Ming-Hui Yu, and Xu Li. An improved mask r-cnn model for multiorgan segmentation. *Mathematical Problems in Engineering*, 2020, 2020.

[15] Tomoyuki Shibata, Atsushi Teramoto, Hyuga Yamada, Naoki Ohmiya, Kuniaki Saito, and Hiroshi Fujita. Automated detection and segmentation of early gastric cancer from endoscopic images using mask r-cnn. *Applied Sciences*, 10(11):3842, 2020.

[16] Xin Luo, Qianwen Huang, Xiang Ji, and Jingfeng Bai. Segmentation and registration of ultrasound images of uterine fibroids for usghifu. In *2021 14th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*, pages 1–5. IEEE, 2021.

[17] Yanfei Guo and Yanjun Peng. Carnet: Cascade attentive refinenet for multi-lesion segmentation of diabetic retinopathy images. *Complex & Intelligent Systems*, 8(2):1681–1701, 2022.

[18] Suvash Sharma, John E Ball, Bo Tang, Daniel W Carruth, Matthew Doude, and Muhammad Aminul Islam. Semantic segmentation with transfer learning for off-road autonomous driving. *Sensors*, 19(11):2577, 2019.

[19] Jonathan Long, Evan Shelhamer, and Trevor Darrell. Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3431–3440, 2015.

[20] Takeshi Nakazawa and Deepak V Kulkarni. Anomaly detection and segmentation for wafer defect patterns using deep convolutional encoder–decoder neural network architectures in semiconductor manufacturing. *IEEE Transactions on Semiconductor Manufacturing*, 32(2):250–256, 2019.

[21] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.

[22] Chen Li, Yusong Tan, Wei Chen, Xin Luo, Yulin He, Yuanming Gao, and Fei Li. Anu-net: Attention-based nested u-net to exploit full resolution features for medical image segmentation. *Computers Graphics*, 90:11–20, 2020.

[23] Özgün Çiçek, Ahmed Abdulkadir, Soeren S Lienkamp, Thomas Brox, and Olaf Ronneberger. 3d u-net: learning dense volumetric segmentation from sparse annotation. In *International conference on medical image computing and computer-assisted intervention*, pages 424–432. Springer, 2016.

[24] Yuan Xue, Tao Xu, Han Zhang, L Rodney Long, and Xiaolei Huang. Segan: adversarial network with multi-scale l1 loss for medical image segmentation. *Neuroin-*

formatics, 16(3):383–392, 2018.

[25] Wei Dai, Nanqing Dong, Zeya Wang, Xiaodan Liang, Hao Zhang, and Eric P Xing. Scan: Structure correcting adversarial network for organ segmentation in chest x-rays. In *Deep learning in medical image analysis and multimodal learning for clinical decision support*, pages 263–273. Springer, 2018.

[26] Haijo Jung. Basic physical principles and clinical applications of computed tomography. *Progress in Medical Physics*, 32(1):1–17, 2021.

[27] Philip J Withers, Charles Bouman, Simone Carmignato, Veerle Cnudde, David Grimaldi, Charlotte K Hagen, Eric Maire, Marena Manley, Anton Du Plessis, and Stuart R Stock. X-ray computed tomography. *Nature Reviews Methods Primers*, 1(1):1–21, 2021.

[28] T Lurthu Pushparaj, E Fantin Irudaya Raj, and E Francy Irudaya Rani. A detailed review of contrast-enhanced fluorescence magnetic resonance imaging techniques for earlier prediction and easy detection of covid-19. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, pages 1–13, 2022.

[29] Benjamin E Rush. *Using Bioimaging Techniques as Muscle Quality Biomarkers for Sarcopenia and Cachexia Diagnosis and Treatment*. PhD thesis, The University of Wisconsin-Madison, 2022.

[30] Eric C Ledbetter and Ian R Porter. Advanced ophthalmic imaging in the horse. *Equine Ophthalmology*, pages 90–132, 2022.

[31] Pallavi Bhosle, Hujeb Pathan, Ganesh Tapadiya, and Md Irshad Alam. Case study on oropharyngeal cancer prediction and diagnosis and management based upon mri, ct scan imaging techniques. In *Disruptive Developments in Biomedical Applications*, pages 91–106. CRC Press, 2022.

[32] Toufique A Soomro, Lihong Zheng, Ahmed J Afifi, Ahmed Ali, Shafiullah Soomro, Ming Yin, and Junbin Gao. Image segmentation for mr brain tumor detection using machine learning: A review. *IEEE Reviews in Biomedical Engineering*, 2022.

[33] Jeffrey Smith, Allison N Schroeder, Alexander R Lloyd, and Kentaro Onishi. Evolution of sports ultrasound. In *Musculoskeletal Ultrasound-Guided Regenerative Medicine*, pages 437–468. Springer, 2022.

[34] Lars A Gjesteby, Joseph R Pare, and Laura J Brattain. Ultrasound for the emergency department and prehospital care. In *Engineering and Medicine in Extreme Environments*, pages 209–234. Springer, 2022.

[35] Krishna Kanta Ghosh, Parasuraman Padmanabhan, Chang-Tong Yang, David Chee Eng Ng, Mathangi Palanivel, Sachin Mishra, Christer Halldin, and Balazs Gulyas. Positron emission tomographic imaging in drug discovery. *Drug Discovery Today*, 27(1):280–291, 2022.

[36] Mai Lin, Ryan P Coll, Allison S Cohen, Dimitra K Georgiou, and Henry Charles Manning. Pet oncological radiopharmaceuticals: Current status and perspectives. *Molecules*, 27(20):6790, 2022.

[37] Domenico Albano, Francesco Dondi, Francesco Bertagna, and Giorgio Treglia. The role of [68ga] gapentixafor pet/ct or pet/mri in lymphoma: A systematic review. *Cancers*, 14(15):3814, 2022.

[38] Yuhan Yang, Bo Zheng, Yueyi Li, Yuan Li, and Xuelei Ma. Computer-aided diagnostic models to classify lymph node metastasis and lymphoma involvement in enlarged cervical lymph nodes using pet/ct. *Medical Physics*, 2022.

[39] Alfred O Ankrah, Ismaheel O Lawal, Rudi AJO Dierckx, Mike M Sathekge, and Andor WJM Glaudemans. Imaging of invasive fungal infections-the role of pet/ct. In *Seminars in Nuclear Medicine*. Elsevier, 2022.

[40] Ujjwal Baid, Satyam Ghodasara, Suyash Mohan, Michel Bilello, Evan Calabrese, Errol Colak, Keyvan Farahani, Jayashree Kalpathy-Cramer, Felipe C Kitamura, Sarthak Pati, et al. The rsna-asnr-miccai brats 2021 benchmark on brain tumor segmentation and radiogenomic classification. *arXiv preprint arXiv:2107.02314*, 2021.

[41] Sarthak Pati, Ujjwal Baid, Maximilian Zenk, Brandon Edwards, Micah Sheller, G Anthony Reina, Patrick Foley, Alexey Gruzdev, Jason Martin, Shadi Albarqouni, et al. The federated tumor segmentation (fets) challenge. *arXiv preprint arXiv:2105.05874*, 2021.

[42] Bjoern H Menze, Andras Jakab, Stefan Bauer, Jayashree Kalpathy-Cramer, Keyvan Farahani, Justin Kirby, Yuliya Burren, Nicole Porz, Johannes Slotboom, Roland Wiest, et al. The multimodal brain tumor image segmentation benchmark (brats). *IEEE transactions on medical imaging*, 34(10):1993–2024, 2014.

[43] Hanxiao Zhang, Xiao Gu, Minghui Zhang, Weihao Yu, Liang Chen, Zhexin Wang, Feng Yao, Yun Gu, and Guang-Zhong Yang. Re-thinking and re-labeling lidc-idri for robust pulmonary cancer prediction. In *Workshop on Medical Image Learning with Limited and Noisy Data*, pages 42–51. Springer, 2022.

[44] Yanbo Shao, Minghao Wang, Juanyun Mai, Xinliang Fu, Mei Li, Jiayin Zheng, Zhaoqi Diao, Airu Yin, Yulong Chen, Jianyu Xiao, et al. Lidp: A lung image dataset with pathological information for lung cancer screening. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 770–779. Springer, 2022.

[45] Ritu Tandon, Shweta Agrawal, Rachana Raghuwanshi, Narendra Pal Singh Rathore, Lalji Prasad, and Vishal Jain. Automatic lung carcinoma identification and classification in ct images using cnn deep learning model. In *Augmented Intelligence in Healthcare: A Pragmatic and Integrated Analysis*, pages 143–166. Springer, 2022.

[46] Manawaduge Supun De Silva, Barath Narayanan Narayanan, and Russell C Hardie. A patient-specific algorithm for lung segmentation in chest radiographs. *AI*, 3(4):931–947, 2022.

[47] Rupanjali Chaudhuri, Divya Nagpal, Abhinav Azad, and Suman Pal. We-net: An ensemble deep learning model for covid-19 detection in chest x-ray images using segmentation and classification. In *International Conference on Advances in Computing and Data Sciences*, pages 112–123. Springer, 2022.

[48] Pratima Upretee and Bishesh Khanal. Fixmatchseg: Fix-

ing fixmatch for semi-supervised semantic segmentation. *arXiv preprint arXiv:2208.00400*, 2022.

[49] Devidas T Kushnure and Sanjay N Talbar. Hfrunet: High-level feature fusion and recalibration unet for automatic liver and tumor segmentation in ct images. *Computer Methods and Programs in Biomedicine*, 213:106501, 2022.

[50] Yoo Jung Kim, Hyungjoon Jang, Kyoungbun Lee, Seongkeun Park, Sung-Gyu Min, Choyeon Hong, Jeong Hwan Park, Kanggeun Lee, Jisoo Kim, Wonjae Hong, et al. Paip 2019: Liver cancer segmentation challenge. *Medical Image Analysis*, 67:101854, 2021.

[51] Jianpeng Zhang, Yutong Xie, Pingping Zhang, Hao Chen, Yong Xia, and Chunhua Shen. Light-weight hybrid convolutional network for liver tumor segmentation. In *IJCAI*, volume 19, pages 4271–4277, 2019.

[52] Patrick Bilic, Patrick Christ, Hongwei Bran Li, Eugene Vorontsov, Avi Ben-Cohen, Georgios Kaissis, Adi Szeskin, Colin Jacobs, Gabriel Efrain Humpire Mamani, Gabriel Chartrand, et al. The liver tumor segmentation benchmark (lits). *Medical Image Analysis*, 84:102680, 2023.

[53] Omar Ibrahim Alirr. Deep learning and level set approach for liver and tumor segmentation from ct scans. *Journal of Applied Clinical Medical Physics*, 21(10):200–209, 2020.

[54] Patrick Bilic, Patrick Ferdinand Christ, Eugene Vorontsov, G Chlebus, H Chen, Q Dou, CW Fu, X Han, PA Heng, J Hesser, et al. The liver tumor segmentation benchmark (lits). arxiv. *arXiv preprint arXiv:1901.04056*, 2019.

[55] Jianxin Zhang, Zongkang Jiang, Jing Dong, Yaqing Hou, and Bin Liu. Attention gate resu-net for automatic mri brain tumor segmentation. *IEEE Access*, 8:58533–58545, 2020.

[56] Salma Alqazzaz, Xianfang Sun, Xin Yang, and Len Nokes. Automated brain tumor segmentation on multimodal mr image using segnet. *Computational Visual Media*, 5:209–219, 2019.

[57] Jindong Sun, Yanjun Peng, Yanfei Guo, and Dapeng Li. Segmentation of the multimodal brain tumor image used the multi-pathway architecture method based on 3d fcn. *Neurocomputing*, 423:34–45, 2021.

[58] Nagwa M AboElenein, Songhao Piao, Alam Noor, and Pir Noman Ahmed. Mirau-net: An improved neural network based on u-net for gliomas segmentation. *Signal Processing: Image Communication*, 101:116553, 2022.

[59] Ning Sheng, Dongwei Liu, Jianxia Zhang, Chao Che, and Jianxin Zhang. Second-order resu-net for automatic mri brain tumor segmentation. *Mathematical Biosciences and Engineering*, 18(5):4943–4960, 2021.

[60] Benjamin B Yan, Yujia Wei, Jaidip Manikrao M Jagtap, Mana Moassefi, Diana V Vera Garcia, Yashbir Singh, Sanaz Vahdati, Shahriar Faghani, Bradley J Erickson, and Gian Marco Conte. Mri brain tumor segmentation using deep encoder-decoder convolutional neural networks. In *International MICCAI Brainlesion Workshop*, pages 80–89. Springer, 2022.

[61] Parvez Ahmad, Saqib Qamar, Linlin Shen, Syed Qasim Afser Rizvi, Aamir Ali, and Girija Chetty. Ms unet: Multi-scale 3d unet for brain tumor segmentation. In *International MICCAI Brainlesion Workshop*, pages 30–41. Springer, 2022.

[62] Rehan Raza, Usama Ijaz Bajwa, Yasar Mehmood, Muhammad Waqas Anwar, and M Hassan Jamal. dresu-net: 3d deep residual u-net based brain tumor segmentation from multimodal mri. *Biomedical Signal Processing and Control*, 79:103861, 2023.

[63] R Jenkin Suji, W Wilfred Godfrey, and Joydip Dhar. Exploring pretrained encoders for lung nodule segmentation task using lidc-idri dataset. *Multimedia Tools and Applications*, pages 1–24, 2023.

[64] Saeed Mohagheghi and Amir Hossein Foruzan. Incorporating prior shape knowledge via data-driven loss model to improve 3d liver segmentation in deep cnns. *International journal of computer assisted radiology and surgery*, 15:249–257, 2020.

[65] Tran Anh Tuan, The Bao Pham, Jin Young Kim, and Joao Manuel RS Tavares. Alzheimer's diagnosis using deep learning in segmenting and classifying 3d brain mr images. *International Journal of Neuroscience*, 132(7):689–698, 2022.

[66] Andriy Myronenko. 3d mri brain tumor segmentation using autoencoder regularization. In *International MICCAI brainlesion workshop*, pages 311–320. Springer, 2019.

[67] Pim Moeskops, Mitko Veta, Maxime W Lafarge, Koen AJ Eppenhof, and Josien PW Pluim. Adversarial training and dilated convolutions for brain mri segmentation. In *Deep learning in medical image analysis and multimodal learning for clinical decision support*, pages 56–64. Springer, 2017.

[68] Mina Rezaei, Konstantin Harmuth, Willi Gierke, Thomas Kellermeier, Martin Fischer, Haojin Yang, and Christoph Meinel. A conditional adversarial network for semantic segmentation of brain tumor. In *International MICCAI Brainlesion Workshop*, pages 241–252. Springer, 2017.

[69] Abhir Bhandary, G Ananth Prabhu, Venkatesan Rajinikanth, K Palani Thanaraj, Suresh Chandra Satapathy, David E Robbins, Charles Shasky, Yu-Dong Zhang, João Manuel RS Tavares, and N Sri Madhava Raja. Deep-learning framework to detect lung abnormality–a study with chest x-ray and lung ct scan images. *Pattern Recognition Letters*, 129:271–278, 2020.

[70] Alexey A Novikov, Dimitrios Lenis, David Major, Jiří Hladvka, Maria Wimmer, and Katja Bühler. Fully convolutional architectures for multiclass segmentation in chest radiographs. *IEEE transactions on medical imaging*, 37(8):1865–1876, 2018.

[71] Dong Yang, Daguang Xu, S Kevin Zhou, Bogdan Georgescu, Mingqing Chen, Sasa Grbic, Dimitris Metaxas, and Dorin Comaniciu. Automatic liver segmentation using an adversarial image-to-image network. In *International conference on medical image computing and computer-assisted intervention*, pages 507–515. Springer, 2017.

[72] Yuankai Huo, Zhoubing Xu, Shunxing Bao, Camilo

Bermudez, Andrew J Plassard, Jiaqi Liu, Yuang Yao, Albert Assad, Richard G Abramson, and Bennett A Landman. Splenomegaly segmentation using global convolutional kernels and conditional generative adversarial networks. In *Medical Imaging 2018: Image Processing*, volume 10574, pages 45–51. SPIE, 2018.