International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

A Review on Methodologies and Challenges of Whole Heart Segmentation Using Deep Learning

Kotte Anusha[1], Dr. V. Kamakshi Prasad*[2]

Submitted: 25/01/2024 **Revised**: 03/03/2024 **Accepted**: 11/03/2024

Abstract: In recent years, deep-learning approaches have had an enormous impact on the analysis of medical images, particularly in whole-heart segmentation. This active research focuses on the accurate delineation of substructures within the heart, the accurate assessment of cardiac function, treatment planning, and facilitating cardiac interventional procedures. It also provides essential morphological information. Ultrasound, Magnetic Resonance Imaging (MRI), and Computed Tomography (CT) are the widely used imaging modalities to segment the substructures such as the ventricles, Atria, and Aorta within the heart. The conduct of whole-heart segmentation is challenging because of laborious manual delineation, which is tedious, subject to change, and requires meticulous analysis. This article discusses the recent advancements in whole-heart segmentation and provides a comprehensive analysis of the various deep learning approaches employed in this domain. This paper highlights different deep learning models such as U-Net, V-Net, and attention mechanisms used to achieve accurate segmentation of the whole heart. Additionally, this paper also explores the volume rendering methods that can be applied to heart structures and the advantages and limitations of these methods in obtaining accuracy and robustness in handling complex cardiac substructures. This article identifies these challenges and suggests research directions to promote reproducible and robust results.

Keywords: *Whole heart segmentation, Deep learning, neural networks, cardiac image analysis, MRI, CT*

1. Introduction

In line with World Health Organization (WHO) data, heart diseases are identified as one of the most common causes of mortality in the world [1]. Around 17.9 million people succumb every year due to various known Cardiovascular Diseases (CVDs), including congenital heart disease, coronary heart disease, cerebrovascular disease, and other heart-related complications. By the year 2030, projections suggest that a global population of 23.3 million individuals will experience the CVDs [1]. Early detection and accurate diagnosis of CVDs are essential to reducing the mortality rate. Whole-heart segmentation in medical imaging has significant clinical and research implications and is useful for the accurate delineation and identification of various cardiac substructures. This can be done using different imaging modalities such as magnetic resonance imaging (MRI) scans, computed tomography (CT), and Ultrasound. The seven heart structures that make up the whole heart segmentation are as follows: (i) left ventricle (LV), (ii) right ventricle (RV), (iii) left atrium (LA), (iv) right atrium (RA), (v) the ascending aorta, and (vi) the pulmonary artery. Whole-heart segmentation results are essential to the advancement of cardiac imaging analysis, clinical care improvement, and treatment planning in cardiac surgeries. Surgeons may plan the most effective surgical strategies using 3D renderings of the segmented heart to understand complicated anatomical variations. Figure 1 shows the whole-heart segmentation results obtained for cardiac CT images in three orthogonal views.

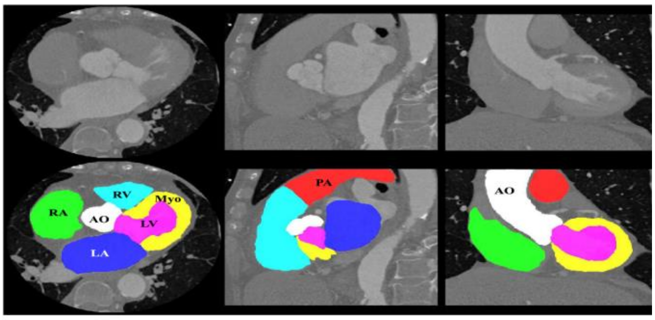


Fig 1: Cardiac CT images and its respective heart segmentation results

The substructures of the heart need to be precisely segmented to do an accurate assessment of the heart. For instance, the findings of the segmented ventricles and the segmented myocardial results are used to compute the ejection fraction and myocardial mass, respectively, which are essential markers for identifying heart illness. Automatic segmentation of the whole heart is still a challenging endeavor due to the inherent noise, the interaction that exists between the boundaries of each cardiac structure, the changing shape of each patient's heart anatomy, and the low image quality. The majority of

[1] Jawaharlal Nehru Technological University, Hyderabad – 500085, INDIA
ORCID ID: 0009-0003-5511-4398
[2] Jawaharlal Nehru Technological University, Hyderabad – 500085, INDIA
ORCID ID: 0000-0002-0430-9772
** Corresponding Author Email: anusha.jntuh@gmail.com*

research studies concentrate on the segmentation of a few structures, such as myocardial [2], [3], or ventricle segmentation [4]–[6]. Only a few studies currently specialize in whole-heart segmentation, which continues to rely on time-intensive manual segmentation (which requires 8–10 hours to manually segment substructures within the heart).

The main motivation of this paper is to review the existing techniques, challenges, and methodologies. This article is structured as follows: Some of the difficulties encountered in the process of whole-heart segmentation are presented in Section 2. The common deep learning architectures useful for heart segmentation are discussed in Section 3. Different approaches to previous works are analyzed in Section 4. The evaluation metrics useful for segmentation are discussed in Section 5 and concluded in Section 6. The relevant papers, collected from online sources including the PubMed database, the IEEE Xplore database, and the Google Scholar search engine, have been referenced and focused on papers published since 2011.

2. Challenges

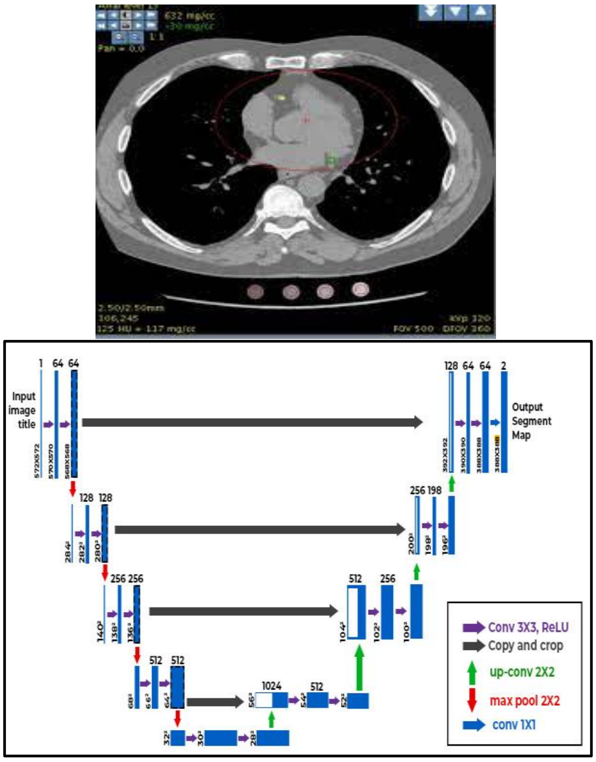
In medical imaging, the manual segmentation of cardiac substructures is a challenging and time-consuming process. Several problems are identified with manual delineation, such as subjectivity as well as observer variability, which can raise discrepancies in the correct annotations; it is not suitable for large datasets; and this process is prone to errors and lack of consistency. As a result, automatic whole-heart segmentation has become increasingly prevalent nowadays.

3. Deep Learning Architectures for Whole Heart Segmentation

A good number of deep learning architectures have evolved to address the challenges of whole-heart segmentation. This paper explains notable architectures, including U-Net, V-Net, deep lab, and others, discussing their design principles, capabilities, and significance in achieving segmentation accuracy.

3.1. U-Net

In medical imaging, for accurate delineation of cardiac structures, the most prominent architecture that has gained significant attention is U-Net, which is implemented by Ronneberger et al. [7]. It is a fully neural network (FCN) combining an encoder and decoder in a U-shaped configuration. This unique structure helps the model to obtain intricate details while maintaining contextual information. This architecture consists of two parts, namely the encoder-decoder path or contraction-expansive path. In the convolutional path, high-level features are extracted by applying convolutional and pooling layers. This path consists of two 3 x 3 convolutions, followed by ReLu and batch normalization. In this process, this path preserves the important features and reduces the spatial dimensions by applying a 2 x 2 maximum pooling operation.

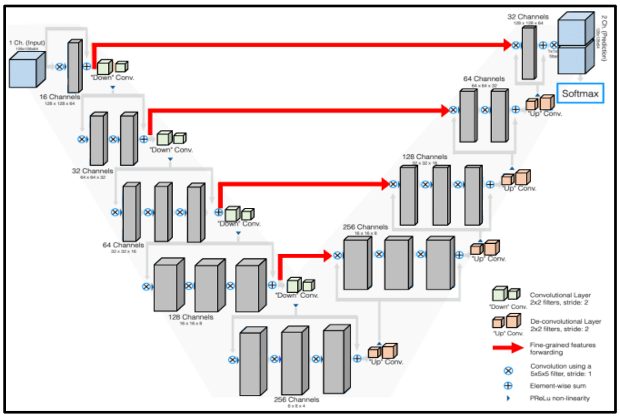


Image

On the other side, in the expansive path, 2 x 2 transposed convolutions are used to upsample the features, and the model generates a high-resolution segmentation map. Skip connections are helpful in mitigating the vanishing gradient problem and also enabling accurate localization of cardiac structures. This architecture is broadly employed in many applications like segmentation of retinal blood vessels [8–10], detection and segmentation of tumors in the brain [11], and particularly in segmenting substructures of the heart [12].

3.2 V-Net

In medical imaging, there is another powerful tool for segmenting volumetric structures. It is an extension of the U-Net and addresses the difficulties of segmenting 3D structures from MRI and CT imaging modalities. It also employs a skip connection encoder-decoder design like U-Net. In this architecture, instead of a max-pooling operation, strided convolutions are used. This is performed through convolution with 2 x 2 x 2 kernels applied with stride 2. It uses 3D convolutions to handle the volumetric data. These 3D convolutions are performed using 5 x 5 x 5 kernels in each stage with padding.



3.3 Deep Lab

Deep lab architecture is a cutting-edge solution for accurate semantic segmentation tasks. It incorporates dilated convolutions and spatial pyramid pooling for capturing both local and global context within images. It is coupled with residual connections for the enhancement of gradient flow during training.

3.4 Attention mechanisms

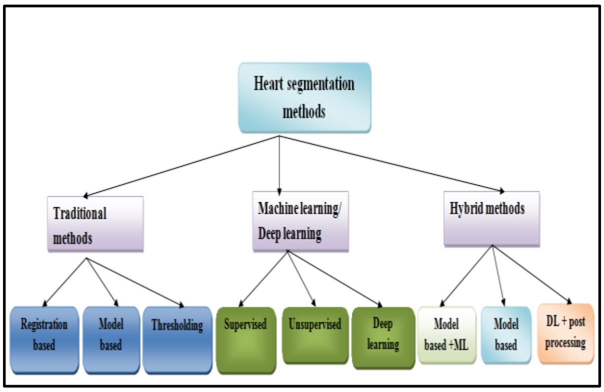
Attention mechanisms have emerged as an important tool in improving the accuracy and contextual understanding of deep learning models for whole-heart segmentation. These mechanisms emphasize the appropriate parts of the image and enhance the segmentation of cardiac structures. This can be done by selectively focusing on specific regions of an image and assigning levels of importance to different areas. The model adjusts the weights that are assigned to each pixel at the time of feature extraction. So that it leads to more contextually informed segmentation.

3.5 Variational auto-encoders

Variational auto-encoders (VAE) are a type of generative model consisting of auto-encoders and probabilistic modeling. It is popular for unsupervised learning tasks, especially for generating new data samples in various domains that resemble a given dataset. VAEs use stochasticity in latent space, which is represented as a probabilistic distribution. The encoder in VAEs maps the data to latent space, and the decoder generates data from sampled latent codes. It is useful in various applications like image generation, data compression, and medical imaging.

4. Related Work

In this section, we give a brief summary of existing research and literature related to automatic whole-heart segmentation. Numerous computerized methods have evolved to segment the cardiac structures. These techniques are grouped under traditional methodologies like segmentation methods using convolutional neural networks and registration-based methods.



4.1 Traditional segmentation methods

Traditional segmentation methods typically rely on techniques like registration, region growing, or thresholding to delineate anatomical structures. These methods often require manual input or predefined rules and can be less robust to variations in data quality and appearance. While they can achieve high precision under ideal conditions, their performance may degrade when applied to complex medical images with diverse shapes and intensities.

4.1.2 Registration based methods:

The method of Hortense [13], is based on multi-atlas segmentation within a non-rigid registration framework, aims to enhance the diagnosis of coronary artery disease (CAD) by providing functional information from the same CTA data, eliminating the need for additional scans. It also enables visualization of coronary arteries and serves as a region of interest for further segmentation. In validation, the method demonstrated high accuracy, with a mean segmentation error of 1.05±1.30 mm and an average Dice coefficient of 0.93 across 8 images.

In the study of Yang et al. [14], a three-step multi-atlas-based method for whole heart segmentation in CT images is proposed. The method involves initial heart region detection through alignment with low-resolution atlas images, precise heart segmentation through high-resolution atlas image registration, and refinement by minimizing dissimilarity within the heart region. The approach achieved promising results with an average Dice coefficient of 0.9051 in a leave-one-out experiment on 20 training datasets, providing accurate segmentation for cardiac chambers, aorta, and pulmonary arteries.

Zhuang X wt al. [15] developed a multi-atlas segmentation (MAS) scheme for automatic whole heart segmentation in cardiac CT images. The method employs a hierarchical registration process and introduces a novel atlas ranking criterion based on conditional entropy. Joint label fusion is used to combine multiple label estimates, resulting in a mean Dice score of 0.918 ± 0.021 and a runtime of 13.2 minutes per case. Comparative analysis shows that this MAS approach outperforms other label fusion strategies and benefits from larger atlas databases, offering improved segmentation accuracy. Oster [16] achieved a good accuracy of 86.0% by using nonlinear registration combined with nonlocal fusion techniques, which is measured with the dice coefficient. In the study of Zhuang et al. [17], an updated label fusion methodology is implemented on MRI and CT images based on multimodality atlases with a 0.899 dice score. These traditional methods for heart segmentation are not reliable and yield disappointing results when the data quality is suboptimal. Segmentation of cardiac structures was successful in the study of Galisot et al. [18], but due to missing tags in the dataset, this yields erroneous results in segmenting pulmonary arteries that impact the overall segmentation results.

4.1.2 Model based methods:

Kang [19] approach belongs to the model-based category, and begins by smoothing images to reduce noise. It proceeds to detect the volume of interest (VOI) using k-means clustering, extracting the whole heart roughly. Seed volumes are identified based on anatomical data, and the left and right heart are separated using power watershed analysis. Finally, the left and right sides of the heart are refined with an active contour model without edge, resulting in an average segmentation error of less than 5% and an average processing time of 51.66±3.35 seconds for clinical datasets. An automated approach for segmenting the four cardiac chambers in cardiac computed tomography angiography (CTA) is presented in Ho Chul Kang [20]. This involves image smoothing, k-means clustering to detect the volume of interest, and seed volume identification based on anatomical analysis. The left and right heart regions are separated using the power watershed algorithm, and the heart's sides are refined with the level-set method. The atrium and ventricle are then extracted from the left and right heart regions using the split energy function. The method was evaluated on 20 clinical datasets, demonstrating an average segmentation accuracy of less than 3.3% when compared to manual segmentations across various patients. X. Zhao et al. [21] proposed a fully automatic method for thoracic CT image whole heart segmentation. It employs a Robust Active Shape Model (Robust ASM) to mitigate outliers caused by neighboring organs with similar intensities. Additionally, a shape-constrained active contour model is employed for enhanced segmentation. Results indicate a mean point-to-surface error of 2.37mm and an averaged Dice index of 0.90 across 38 images.

4.1.3 Thresholding based methods:

Lee HY et al. [22] implemented a novel automatic left ventricle (LV) segmentation algorithm, called ITHACA, which has been developed for quantifying cardiac output and myocardial mass in clinical applications. This method combines region growth with iterative thresholding to segment the LV endocardium and utilizes active contour modeling guided by the endocardial border and myocardial signal information for epicardial extraction. Compared to both manual tracing and commercial MASS Analysis software, ITHACA significantly enhances myocardial border definition, with close agreement to manual tracing, making it a promising tool for clinical practice. The study of Katherine et al. [23] methodology involves the development and implementation of the Otsu Thresholding method, coupled with Hounsfield unit (HU) values, for heart segmentation in thorax CT scan images. The study evaluates its effectiveness by calculating balanced accuracy across 30 test datasets, with results averaging at 72.54%. Rodríguez R [24] proposes a two-stage approach involving optimal window size and scale selection.The primary objective is to accurately extract the number of blood vessels from images, with segmentation errors less than 3%. This method is a part of a broader image analysis process aimed at diagnosing and predicting malignant tumors through morphometrical analysis.

4.2 Deep learning-based methods in segmentation

Nowadays, deep learning convolutional neural networks have achieved good results in segmentation of medical images. A detailed and evaluative survey of cardiac image segmentation is presented in [25] specifically focusing on whole heart, bi-ventricles, and left atrium segmentation across various imaging modalities such as CT, MRI, and echocardiography. This paper summarizes the challenges in cardiac image segmentation, comparing existing segmentation methods, categorizing notable contributions, and critically assessing their performance and accuracy. For instance, Multi-planar deep CNNs featuring adaptive fusion were employed by Mortaz et al. [26], effectively utilizing complementary data from different 3D scan planes. In this approach, three CNNs were tailored separately for CT and MRI images across three planes, each trained from scratch for voxel-wise labeling. The model was assessed with 4-fold cross-validation on MM-WHS 2017 data, resulting in a precision and dice index of 0.93 and 0.90 for CT, and 0.87 and 0.85 for MR images, respectively. Bian et al. [27] enhances the local contrast by combining 3D FCN and deep supervision mechanisms with transfer learning to overcome the difficulties in training. Tong et al. [28] presented a combination of 3D u-Net with ROI to make the computational complexity more simple. MRI and CT data are combined for better extraction of different substructures. In the study of Smedby et al. [29], CT images are segmented at their first resolution using a two-stage U-Net architecture. Shi, Z. et al. [30] proposed a probabilistic deep voxelwise dilated residual network designed for automatic whole heart segmentation called Bayesian VoxDRN for 3D MR images. It incorporates variational dropouts to model uncertainty through approximate Bayesian inference. It effectively handles imbalanced datasets using a combination of focal loss and Dice loss. In the study of Xu et al. [31] presented a novel whole heart segmentation approach i.e., combining faster R-CNN and U-Net (CFUN) results in precise localization and heart segmentation. A 3D edge loss auxiliary loss function is used to improve the performance of the segmentation. It worked on MM-WHS dataset with average dice score of 0.859. An enhanced Fully Convolutional Network (FCN) is implemented by Yang et al. [32] by incorporating 3D operators, transfer learning, and deep supervision. Class imbalance issues are addressed when dealing with multiple substructures, and introduce a hybrid loss that balances training across classes while preserving detailed boundary information. The framework proposed by M. Habijan et al. [33] employs two 3D U-Net neural networks: one for heart localization and the other for segmentation. Performance was assessed on five CT volumes from the MICCAI 2017 Multi-Modality Whole Heart Segmentation challenge, yielding an impressive average dice score of 89%. Mahendra Khened et al. [34] introduced an efficient, highly parameter-reduced FCN-based architecture for medical image segmentation. This design incorporates innovations such as skip connections and Inception modules for multi-scale processing. They demonstrated its success on two datasets, achieving top results in the ACDC-2017 challenge and the highest Jaccard index in LV-2011, enabling an integrated framework for cardiac segmentation, parameter extraction, and disease diagnosis with clinical potential. Christian Payer et al. [35] propose a two-step approach, where the first step involves the localization of the heart on low-resolution images. Localization is done using U-Net architectures like FCNN with heatmap regression, which indicates the intensity or probabilities of certain features or objects. To streamline the process and optimize computational resources, the authors combine the localization and segmentation of convolutional neural networks (CNNs). This integration aims to minimize memory requirements and computation time, enhancing efficiency without compromising accuracy. Notably, the segmentation technique excels in MRI data, which often exhibits greater variations in anatomical field of view, intensity ranges, and acquisition artifacts compared to CT data. C. Ye et al. [36] introduces a novel deeply supervised 3D UNET architecture with multi-depth fusion for enhanced context information extraction. It pioneers the application of focal loss to image segmentation, extending its use to multi-category tasks. The proposed hybrid loss function, combining focal loss and Dice loss, effectively addresses category imbalance. Evaluation on the MICCAI 2017 whole-heart CT dataset achieved a Dice score of 90.73%. Tao Liu et al. [37] proposed a different approach for accurate and precise whole-heart segmentation. To segment the cardiac structures, a two-stage U-net architecture based 2D image segmentation technique is applied. Prior to network training, an adaptive threshold window is applied to determine the region of interest and minimize the noise. The first stage's threshold range is [-1024, 1354]. Weight maps are introduced to improve the performance of segmentation and identify the correct boundary sections of the heart. Kening Le et al. [38] introduced an enhanced Unet-GAN model, termed R2Unet-GAN, with the generator being R2U-Net and the discriminator being FCN and adopted a cosine decay learning rate policy. This model exhibited improved accuracy, reaching 88.9% during training on 15 sets, with the highest accuracy on a single set reaching 94.0%.

Sulaiman Vesal et al. [39] proposed a three-stage encoder-decoder architecture that is used to emphasize a three-dimensional understanding of spatial information. At the first stage, the course density map for target structures is estimated and localized. A 3D Dilated Residual Unet (3D DR-Unet) framework is used for accurate heart segmentation. An innovative aspect of this proposed approach is its capacity to perform these tasks without extensive pre- or post-processing steps, enabling the utilization of the full 3D volume. An innovative U-Net based GAN approach is introduced by Z. Lou et al. [40] employing a U-Net as the generative network and an FCN as the discriminator. The model was assessed using five CT datasets from MM-WHS 2017, resulting in an impressive average dice score of 86.32% (with a peak of 93.64% on the best-performing dataset). Galea et al. [41] presented a simple and practical slice-by-slice 2D segmentation approach using region of interest (ROI) localization. This architecture uses simple bilinear interpolation instead of transposed convolutional neural networks to reduce complexity and memory restrictions. Improved performance was achieved through the ensemble of various architectures.

The study of Habijan et al. [42] introduces a novel Feature Merge Pre-Residual Unit (FM-Pre-ResNet) connectivity structure, enabling the development of deep models without significantly increasing the parameter count compared to pre-activation residual units. Additionally, proposed a 3D encoder-decoder architecture that incorporates FM-Pre-ResNet units and a variational autoencoder (VAE). In the encoding phase, FM-Pre-ResNet units learn a low-dimensional representation, followed by VAE-based reconstruction. This approach offers robust weight regularization and mitigates overfitting. Evaluation on the MICCAI Multi-Modality Whole Heart Segmentation (MM-WHS) Challenge involving 40 test subjects yields compelling results, with an average Dice score of 90.39% for CT images and 89.50% for MRI images in whole heart segmentation.

Tianchen Wang et al. [43] performed a two-stage analysis for segmentation and classification on Congenital Heart disease (CHD) images with the ImageCHD dataset [44] using multiple U-Nets. In the first stage, segmentation-based analysis is done using 2D and 3D U-Nets. Shape-based analysis is carried out to extract the connection features between the substructures of the heart. This methodology achieved 81.9% accuracy using the dice coefficient measure. Yuhui Songb et al. [45] introduce a two-stage segmentation network architecture that consists of two modules. One is the Feature Aggregation Module (FAM), and the other is the Multi-Level Attention Mechanism (mLAM). The first module consists of four feature extraction blocks to obtain multi-scale information. As a result, high-level feature maps can be generated from low-level feature maps.

An overview of research studies focusing on whole heart segmentation using deep learning methods are presented in a table 4.1. It includes relevant references, the specific segmentation task addressed, the modalities employed (e.g., CT, MRI), the deep learning methods applied, and the reported accuracy, measured in terms of the Dice score.

Table 1. Overview of Whole Heart Segmentation Studies Using Deep Learning

S.No	Reference	Task	Imaging Modalities	Method	Accuracy
1.	Mortazi et al. [26] (2017)	Whole heart segmentation	CT & MRI	Multi planar deep convolutional network + adaptive fusion strategy	Dice index for CT-0.90 MRI- 0.85
2.	Bian et al. [27] (2018)	Whole heart segmentation	CT & MRI	Deep supervision + knowledge transfer	Dice index for CT - 80.90 MRI- 74.62
3.	Tong et al. [28] (2018)	Whole heart segmentation	CT & MRI	3D deeply supervised U-Net	CT- 0.849 MRI- 0.674
4.	Smedby et al. [29] (2018)	Whole heart segmentation	CT & MRI	Scout segmentation using Orthogonal 2D U-Net + shape context	CT- 0.807 MRI- 0.87
5.	Shi, Z. et al. [30] (2018)	Whole heart segmentation	MRI	Bayesian VOxDRN	MRI-90.83
6.	Xu et al. [31] (2018)	Whole heart segmentation	CT	Combining Faster R-CNN (localization) + U-Net	CT- 0.859
7.	Yang et al. [32] (2018)	Whole heart segmentation	CT & MRI	Deep supervision + knowledge transfer +hybrid loss	CT- 84.32 MRI-77.86
8.	M. Habijan et al. [33] (2019)	Whole heart segmentation	CT	Two 3D U-Net architectures + PCA for data augmentation	CT- 89.00
9.	Mahendra Khened et al. [34] (2018)	Whole heart segmentation	MRI	DenseNets+Multi scale processing	MRI-0.82
10.	Christian Payer et al. [35] (2019)	Whole heart segmentation	CT & MRI	U-Net + Multilable Segmentation CNN	CT- 88.9 MRI-79.0
11.	C. Ye et al. [36] (2019)	Whole heart segmentation	CT	Multi depth fusion network + focal loss	CT-90.73

12.	Liu et al. [37] (2019)	Whole heart segmentation	CT	Two stage U-Net+ Adpative threshold window	CT-0.793
13.	Kening Le et al. [38] (2020)	Whole heart segmentation	CT	Conditional generative adversarial network + R2U-Net	CT-88.9
14.	Sulaiman Vesal et al. [39] (2020)	Whole heart segmentation	MRI	Encoder-decoder network+coarse density map +3D Dilated Residual-UNet	0.928
15.	Z. Lou et al. [40] (2020)	Whole heart segmentation	CT	U-Net based GAN	CT-86.32
16.	Galea et al. [41] (2021)	Whole heart segmentation	MRI	Ensembling of U-Net and Deeplab V3+	91.87
17.	Habijan et al. [42] (2021)	Whole heart segmentation	CT & MRI	Feature merge residual unit + 3D encoder-decoder +Variational auto encoder	CT-90.39 MRI- 89.50
18.	Wang et al. [43] (2021)	Classification of Congential heart disease	CT	Multiple U-Nets used for segmentation +RoI cropping	CT-88.8
19.	Yuhui Song et al. [44] (2022)	Whole heart segmentation	CT & MRI	Feature aggregation+ multi level attention mechanism	CT- 0.94 MRI-0.934

The above studies collectively represent the state-of-the-art in automated cardiac image segmentation and contribute to advancements in medical imaging for diagnostic and clinical applications. The accuracy scores offer insights into the effectiveness of deep learning models in accurately delineating whole heart structures, a crucial aspect of cardiovascular analysis.

4.3 Volume rendering in heart segmentation

Volume rendering is a visualization technique used to provide a dynamic and informative representation of cardiac structures. This technique transforms volumetric datasets like cardiac MRI and CT images into visualizations, which allow clinicians to understand the spatial relationships and substructures within the heart.

This technique highlights specific tissues, vessels, or chambers and thus assists in the accurate delineation of cardiac structures. This technique needs a high-dimensional transfer function to differentiate the target object from its surroundings, which is a trial-and-error process. For example, in the study of Kuanquan Wang [46] for layered medical datasets, the context-preserving volume rendering method is adopted. In their approach, the 3D cardiac gray scale is mapped to the respective optical attribute based on the anatomical structure of the heart. To improve the boundary information, local illumination ingredients are applied. F. Yang [47] conducted an investigation on the visualization of the segmented cardiac structures. An accelerated rendering method is used to improve the performance of the ray-casting rendering technique. In this approach, the shape and boundaries of cardiac structures visualization quality are improved using a transfer function. This function assigns a good opacity and color for each substructure. Overlapped cardiac substructures are visualized using carving techniques and solid rendering of surface models in the study of Tanja et al. [48]. Different visualization strategies are studied for the inner tissues of the cardiac structures in [49]. Kharche et al. [50] proposed a collaborative architecture that gives high-performance visualization for heart structures.

5. Conclusion

In this review paper, we embarked on a comprehensive journey through the realm of whole heart segmentation, shedding light on the ongoing advancements and challenges in the field of cardiac image analysis. The ability to precisely delineate cardiac structures from various imaging modalities like CT and MRI has been crucial for understanding heart diseases and assisting in clinical diagnosis and treatment.We discussed the numerous challenges associated with whole heart segmentation, emphasizing the complex and diverse anatomical variations that can be present in medical images. In Section 3, we delved into deep learning architectures, highlighting key methodologies like U-Net, V-Net, DeepLab, attention mechanisms, and variational auto-encoders. These advancements have shown remarkable promise in improving segmentation accuracy and efficiency. In section 4, we examined traditional segmentation methods, model-based strategies, registration-based techniques, and thresholding-based approaches. The comparative analysis underscores the advantages of deep learning in handling the intricate task of whole heart segmentation. In Section 5, we considered the crucial role of loss functions and evaluation metrics in assessing the performance of segmentation algorithms. We discussed the significance of balanced loss functions and how they are applied in achieving more accurate results.

Conclusively, the progress made in deep learning methods for whole heart segmentation has been instrumental in mitigating the challenges associated with complex anatomical variations and image quality. These modern architectures, such as U-Net, V-Net, and the incorporation of attention mechanisms, have significantly advanced our capabilities in accurately extracting the heart's anatomical structures. Furthermore, the adoption of balanced loss functions has contributed to reducing the impact of class imbalance in multi-class segmentation tasks.

References

[1] Geneva: World Health Organization; 2016 (WHO/DAD/2016.1). World health statistics overview 2016: Monitoring health for the SDGs (sustainable development goals) https://www.who.int/gho/ publications/world_health_statistics/2016/en/. Accessed: September 16, 2019.

[2] A. Chartsias, T. Joyce, G. Papanastasiou, S. Semple, M. Williams, D. Newby, R. Dharmakumar, and S. A. Tsaftaris, Factorized spatial representation learning: Application in semi-supervised myocardial segmentation,'' in Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent. Cham, Switzerland: Springer, 2018, pp. 490–498.

[3] C. Xu, L. Xu, G. Brahm, H. Zhang, and S. Li, ''MuTGAN: Simultaneous segmentation and quantification of myocardial infarction without contrast agents via joint adversarial learning,'' in Proc. Int. Conf. Med. Image Comput. Comput. Assist. Intervent. Cham, Switzerland: Springer, 2018, pp. 525–534.

[4] H.-Y. Lee, N. C. F. Codella, M. D. Cham, J. W. Weinsaft, and Y. Wang, ''Automatic left ventricle segmentation using iterative thresholding and an active contour model with adaptation on short-axis cardiac MRI,'' IEEE Trans. Biomed. Eng., vol. 57, no. 4, pp. 905–913, Apr. 2010.

[5] L. Zhang, M. Pereaez, S. K. Piechnik, S. Neubauer, S. E. Petersen, and A. F. Frangi, ''Multi-input and dataset-invariant adversarial learning (MDAL) for left and right-ventricular coverage estimation in cardiac MRI,'' in Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent. Cham, Switzerland: Springer, Sep. 2018, pp. 481–489.

[6] Y. Mo, F. Liu, D. McIlwraith, G. Yang, J. Zhang, T. He, and Y. Guo, ''The Deep Poincaré Map: A Novel Approach for Left Ventricle Segmentation,'' in Proc. Int. Conf. Med. Image Comput. Comput.-Assist.Intervent. Cham, Switzerland: Springer, 2018, pp. 561–568.

[7] O. Ronneberger, P. Fischer, and T. Brox, ''U-net: Convolutional networks for biomedical image segmentation,'' in Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent. Cham, Switzerland: Springer, 2015, pp. 234–241.

[8] W. Xiancheng, L. Wei, M. Bingyi, J. He, Z. Jiang, W. Xu, Z. Ji, G. Hong, and S. Zhaomeng, ''Retina blood vessel segmentation using a U-net- based Convolutional Neural Network,''inProc.Procedia Comput. Sci., Int. Conf. Data Sci., Beijing, China, Jun. 2018, pp. 8–9.

[9] C.Playout, R.Duval, and F.Cheriet,''A multitask learning architecture for simultaneous segmentation of bright and red lesions in fundus images,'' in Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent. Cham, Switzerland: Springer, 2018, pp. 101–108.

[10] Y. Zhang and A. C. S. Chung, ''Deep supervision with additional labels for retinal vessel segmentation task,'' in Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent. Cham, Switzerland: Springer, 2018, pp. 83–91.

[11] H. Dong, G. Yang, F. Liu, Y. Mo, and Y. Guo, ''Automatic brain tumor detection and segmentation using U-Net based fully convolutional networks,'' in Proc. Annu. Conf. Med. Image Understand. Anal. Cham, Switzerland: Springer, 2017, pp. 506–517.

[12] Wang, C., Smedby, O. 2017. Automatic whole heart segmentation using deep learning and shape context. In: International Workshop on Statistical Atlases and

Computational Models of the Heart, pp. 242–249.

[13] Hortense A. Kirisli, Michiel Schaap, Stefan Klein, Lisan A. Neefjes, Annick C. Weustink, Theo Van Walsum, and Wiro J. Niessen "Fully automatic cardiac segmentation from 3D CTA data: a multi-atlas based approach", Proc. SPIE 7623, Medical Imaging 2010: Image Processing, 762305 (12 March 2010); https://doi.org/10.1117/12.838370

[14] Yang, Guanyu et al. “Automatic Whole Heart Segmentation in CT Images Based on Multi-atlas Image Registration.” STACOM@MICCAI (2017).

[15] Zhuang X., Bai W., Song J., Zhan S., Qian X., Shi W., Lian Y., Rueckert D. Multiatlas whole heart segmentation of CT data using conditional entropy for atlas ranking and selection Med. Phys., 42 (7) (2015), pp. 3822-3833

[16] M.P.Heinrich and J. Oster, MRII whole heart segmentation using discrete nonlinear registration and fast non-local fusion,’’ in Proc. Int. Workshop Stat. Atlases Comput. Models Heart. Cham, Switzerland: Springer, 2017, pp. 233–241.

[17] X. Zhuang and J. Shen, ‘‘Multi-scale patch and multi-modality atlases for whole heart segmentation of MRI,’’ Med. Image Anal., vol. 31, pp. 77–87, Jul. 2016.

[18] G. Galisot, T. Brouard, and J.-Y. Ramel, ‘‘Local probabilistic atlases and a posteriori correction for the segmentation of heart images,’’ in Proc. Int. Workshop Stat. Atlases Comput. Models Heart. Cham, Switzerland: Springer, 2017, pp. 207–214.

[19] Kang, H.C., Kim, B., Lee, J. et al. Automatic left and right heart segmentation using power watershed and active contour model without edge. Biomed. Eng. Lett. 4, 355–361 (2014). https://doi.org/10.1007/s13534-014-0164-9.

[20] Ho Chul Kang, Jeongjin Lee, Juneseuk Shin, Automatic Four-Chamber Segmentation Using Level-Set Method and Split Energy Function, Healthc Inform Res. 2016;22(4):285-292.

[21] X. Zhao, Y. Wang and G. Jozsef, "Robust shape-constrained active contour for whole heart segmentation in 3-D CT images for radiotherapy planning," 2014 IEEE International Conference on Image Processing (ICIP), Paris, France, 2014, pp. 1-5, doi: 10.1109/ICIP.2014.7024999.

[22] Lee HY, Codella NC, Cham MD, Weinsaft JW, Wang Y. Automatic left ventricle segmentation using iterative thresholding and an active contour model with adaptation on short-axis cardiac MRI. IEEE Trans Biomed Eng. 2010 Apr;57(4):905-13. doi: 10.1109/TBME.2009.2014545. Epub 2009 Feb 6. PMID: 19203875.

[23] Katherine et al 2021 J. Phys.: Conf. Ser. 1816 012080. DOI 10.1088/1742-6596/1816/1/012080

[24] Rodríguez R. A strategy for blood vessels segmentation based on the threshold which combines statistical and scale space filter. Application to the study of angiogenesis. Comput Methods Programs Biomed. 2006 Apr;82(1):1-9. doi: 10.1016/j.cmpb.2005.10.008. Epub 2006 Mar 23. PMID: 16556472.

[25] Marija Habijan, Danilo Babin, Irena Gali´c, Hrvoje Leventi´c, Kreˇsimir Romi´c, Lazar Velicki, Aleksandra Piˇzurica, Overview of the whole heart and heart chamber segmentation methods, Cardiovasc. Eng. Technol. (2020) 1–23.

[26] Mortazi,A., Burt, J., Bagci, U., 2017. Multi-planar deep segmentation networks for cardiac substructures from MRI and CT. In: International Workshop on Statistical Atlases and Computational Models of the Heart, pp. 199–206.

[27] Yang, X., Bian, C., Yu, L., Ni, D.,Heng, P.-A., 2017. 3D convolutional networks for fully automatic fine-grained whole heart partition. In: International Workshop on Statistical Atlases and Computational Models of the Heart, pp. 181–189.

[28] Tong, Q. , Ning, M. , Si, W. , Liao, X. , Qin, J. , 2017. 3D deeply-supervised U-Net based whole heart segmentation. In: International Workshop on Statistical Atlases and Computational Models of the Heart, pp. 224–232.

[29] Yang, G. , Sun, C. , Chen, Y. , Tang, L. , Shu, H. , Dillenseger, J.-l. , 2017. Automatic whole heart segmentation in CT images based on multi-atlas mage registration. In: International Workshop on Statistical Atlases and Computational Models of the Heart, pp. 250–257.

[30] Wang, C., Smedby, O. 2017. Automatic whole heart segmentation using deep learning and shape context. In: International Workshop on Statistical Atlases and Computational Models of the Heart, pp. 242–249.

[31] Shi, Z. et al. (2018). Bayesian VoxDRN: A Probabilistic Deep Voxelwise Dilated Residual Network for Whole Heart Segmentation from 3D MR Images. In: Frangi, A., Schnabel, J., Davatzikos, C., Alberola-López, C., Fichtinger, G. (eds) Medical Image Computing and Computer Assisted Intervention – MICCAI 2018. MICCAI 2018. Lecture Notes in Computer Science(), vol 11073. Springer, Cham. https://doi.org/10.1007/978-3-030-00937-3_65

[32] Xu, Zhanwei, Ziyi Wu, and Jianjiang Feng. "CFUN: Combining faster R-CNN and U-net network for efficient whole heart segmentation." arXiv preprint arXiv:1812.04914 (2018).

[33] Yang, X. , Bian, C. , Yu, L. , Ni, D. , Heng, P.-A. , 2017. Hybrid loss guided convolutional networks for whole heart parsing. In: International Workshop on Statistical Atlases and Computational Models of the Heart, pp. 215–223.

[34] M. Habijan, H. Leventić, I. Galić and D. Babin, "Whole Heart Segmentation from CT images Using 3D U-Net architecture," 2019 International Conference on Systems, Signals and Image Processing (IWSSIP), Osijek, Croatia, 2019, pp. 121-126, doi: 10.1109/IWSSIP.2019.8787253.

[35] Mahendra Khened, Alex Varghese, Ganapathy Krishnamurthi: Fully convolutional multi-scale residual DenseNets for cardiac segmentation and automated cardiac diagnosis using ensemble of classifiers. Medical Image Anal. 51: 21-45 (2019)

[36] Christian Payer, Darko Stern, Horst Bischof, and Martin Urschler "Multi-Label Whole Heart Segmentation Using CNNs and Anatomical Label Configurations", Statistical Atlases and Computational Models of the Heart. ACDC and MMWHS Challenges, Springer 2019.

[37] C. Ye, W. Wang, S. Zhang and K. Wang, "Multi-Depth Fusion Network for Whole-Heart CT Image Segmentation," in IEEE Access, vol. 7, pp. 23421-23429, 2019, doi: 10.1109/ACCESS.2019.2899635.

[38] Tao Liu, Yun Tian, Shifeng Zhao, And Qingjun Wang "Automatic Whole Heart Segmentation Using a Two-Stage U-Net Framework and an Adaptive Threshold Window", IEEE VOLUME 7, 2020.

[39] Kening Le, Zeyu Lou, Weiliang Huo and Xiaolin Tian, "Auto Whole Heart Segmentation from CT images Using an Improved Unet-GAN", Journal of Physics: Conference Series 1769 (2021) 012016,doi:10.1088/1742-6596/1769/1/012016

[40] Sulaiman Vesal, Andreas Maier, "Fully Automated 3D Cardiac MRI Localisation and Segmentation Using Deep Neural Networks", Journal of Imaging, 6(7), 65, 2020.

[41] Z. Lou, W. Huo, K. Le and X. Tian, "Whole Heart Auto Segmentation of Cardiac CT Images Using U-Net Based GAN," 2020 13th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), Chengdu, China, 2020, pp. 192-196, doi: 10.1109/CISP-BMEI51763.2020.9263532.

[42] Galea, R.; Diosan, L.; Andreica, A.; Popa, L.; Manole, S.; Bálint, Z. Region-of-Interest Based Cardiac Image Segmentation with Deep Learning. Appl. Sci. 2021, 11, 1965. https://doi.org/10.3390/app11041965

[43] Habijan, M.; Galic´, I.; Leventic´, L.; Romic´, K. Whole Heart Segmentation Using 3D FM-Pre-ResNet Encoder–Decoder Based Architecture with Variational Autoencoder Regularization. Appl. Sci. 2021, 11, 3912. https://doi.org/ 10.3390/app11093912

[44] Xu, X. et al. (2020). ImageCHD: A 3D Computed Tomography Image Dataset for Classification of Congenital Heart Disease. In: , et al. Medical Image Computing and Computer Assisted Intervention – MICCAI 2020. MICCAI 2020. Lecture Notes in Computer Science(), vol 12264. Springer, Cham. https://doi.org/10.1007/978-3-030-59719-1_8

[45] https://github.com/XiaoweiXu/ImageCHD-A-3D-Computed-Tomography-ImageDataset-for-Classification-of-Congenital-Heart-Disease

[46] Yuhui Songb, Xiuquan Dua,Yanping Zhanga, Shuo Li Two-stage segmentation network with feature aggregation and multi-level attention mechanism for multi-modality heart images, Elsevier (2022)

[47] Wang, K., Zhang, L., Gai, C., and Zuo, W., Illustrative visualization of segmented human cardiac anatomy based on context-preserving model. In: Computing in Cardiology, pp. 485– 488, 2011.

[48] Yang, F., Zuo, W., Wang, K., and Zhang, H., Visualization of segmented cardiac anatomy with accelerated rendering method. In: Computing in Cardiology, IEEE, pp. 789–792, 2009.

[49] Kurzendorfer, T., Girard, E., Gralewski, K., Kleinoeder, A., Kiraly, A. P., Strobel, N., and Dori, Y., New biplane x-ray magnetic resonance image fusion prototype for 3d enhanced cardiac catheterization in congenital heart diseases. J. Cardiovasc. Magn. Reson. 16(Suppl1):O103, 2014

[50] Gai, C., Wang, K., Zhang, L., and Zuo, W., Strategy of statistics based visualization for segmented 3d cardiac volume dataset. Adv. Int. Comp. 6838:250–256, 2012.