IJETIR
ISSN (Online) 2583-0554
Impact Factor: 5.731
International Journal of Emerging Technologies and Innovative Research (IJETIR)
Volume 5, Issue 12, December 2025

Advancements in Deep Learning for Medical Image Segmentation: A Comprehensive Review of Techniques across Organs and Modalities

Sarala D and Dayanand R B
Assistant Professor, Department of Computer Science and Engineering
Associate Professor, Department of Computer Science and Engineering
KLE University, Hubbali, India
MSRIT University, Bangalore, India
saralad_12@rediffmail.com and dayanandarb@msrit.edu.

Abstract: *Medical image segmentation is the process of partitioning digital images into multiple segments to simplify and change the representation of an image into something more meaningful and easier to analyse. It is a critical step in clinical diagnostic tools that helps in the visualization of the structures of interest, quantification of tissue volumes, diagnosis, treatment planning, and image-guided interventions. Despite considerable advancements in this field, the segmentation of medical images remains a complex task due to the variability of medical data and the constraints of current imaging technologies. This survey paper provides a comprehensive review of medical image segmentation methods, spanning from traditional techniques to modern deep learning approaches. We present a structured comparison of various segmentation methods, discussing their advantages, disadvantages, and shortcomings within clinical and research settings. Furthermore, this paper identifies the crucial need for advancements in this field and the importance of developing generalizable and interpretable models that can be seamlessly integrated into clinical workflows. Through this survey, we aim to delineate a clear path for future research, underscore the importance of multidisciplinary collaboration, and foster a deeper understanding of the intersection between medical image analysis and artificial intelligence*

Keywords: Medical Image Segmentation, Digital Images, Medical Data, Deep Learning, Artificial Intelligence.

I. INTRODUCTION

Prior subsequent image processing operations, such as recognizing anomalies like tumors or lesions, segmenting medical images is an essential step. Because internal organs have complicated structures, abnormalities need segmentation, and different imaging methods show distinct textures associated with different tissues and disorders, segmentation may be fairly difficult. Medical image segmentation is an important operation that frequently requires accurate findings obtained by automated or manual approaches. Extraction of the required item exactly might be difficult, particularly when working with a target organ that has a complex tissue structure. In the realm of medical imaging, medical image segmentation is an essential procedure for improving disease diagnosis, treatment planning, and monitoring. It is standard practice to segment digital medical images into portions that correspond to various organs, illnesses, tissue types, or other biologically significant properties. Since accurate segmentation has a major influence on a clinician's ability to identify and measure problems, it is an important field of research in medical informatics and image processing. A new class of segmentation algorithms has been developed as a result of recent developments in machine learning and artificial intelligence. These algorithms offer greater flexibility and precision than previous approaches. These methods, which use convolutional neural networks (CNNs) to automate the segmentation of complicated imaging data, are skilled at deep learning. Following extensive training on annotated image datasets, these algorithms learn to recognize complex patterns and changes in tissue structure, producing extremely accurate and dependable segmentation outputs [1].

These are the several categories that are comprised: Each pixel in the image is examined based on its gray values compared to the threshold values, which then categorizes the pixels appropriately. (1) Based on the gray features of the image, a segmentation method is employed to generate one or more gray thresholds. Pixels on the boundaries of discrete sections with different gray values are used in edge-based segmentation. The edges of the image correspond to the high-frequency region when it is transformed from the spatial domain to the frequency domain using the Fourier transform. This feature allows boundaries to be drawn between regions by first locating the edge pixels and then connecting them. Region-based segmentation is a technique that divides data by taking advantage of the uniform, smooth surface of an item [2]. By the splitting and merging approach or the region expansion strategy to find the area with consistent attributes. This is because the region in the image with a consistent or progressively changing intensity is represented by the image's smooth surface. By using a segmentation methodology, fuzzy theory's notion of membership degree is introduced, and pixel points are divided into regions with high membership degrees. By applying deep learning methods, scientists may use a convolutional neural network to extract the complex elements that the brain perceives, simulating the extensive visual processing that is seen in the human visual system. The deconvolution layer samples deep features and classifies each pixel independently by sampling to complete image segmentation. Medical image segmentation has shown to be quite successful with deep learning techniques. Despite the great efficacy of deep learning algorithms, there are two main problems with medical image segmentation [3]:

In medical images, labeled data is absent. The process of gathering labeled medical image data necessitates a significant commitment of professional medical expertise, time, and money since expert annotation is required. Building segmentation models for medical images can be difficult when patient privacy concerns prevent tagged medical image data from constantly being available. Medical images may also have dimensional problems and noise. Because medical images are complex, extracting features from the data can be a challenging undertaking because of their high dimension and significant noise [4]. Before the widespread adoption of deep learning, segmentation issues were usually solved with improvised methods. For every assignment, unique processing pipelines and algorithms were created to satisfy segmentation requirements and guarantee excellent outcomes. It's evident that certain techniques, intended for particular objectives, are inappropriate for use in other contexts.

Medical image segmentation has undoubtedly been transformed by deep learning as well. First presenting, the U-Net architecture is a universal segmentation model that has been evaluated on several datasets for diverse segmentation tasks. The U-Net architecture is surprisingly basic considering its remarkable performance. Because of its simplicity, it was unclear if more intricate architectural modifications might increase segmentation accuracy without sacrificing the changed designs' universality [5]. One article cannot cover every segmentation design that has been introduced in the last several years. For instance, a PubMed search including the terms "Segmentation" and "U-Net" in abstracts and titles yields over a thousand papers during the previous four years. These publications mostly concentrate on architectures built for very specific problems, which brings to mind the era before deep learning when ad hoc methods were frequently used. Few architectural concepts have developers who assert that they are better than universal and U-Net models [6]. Offered a thorough investigation with a focus on the use of deep learning in medical image analysis[7].

Problems in Medical Image Segmentation: When segmenting medical images, several issues might arise, which can lower the process's quality. Handling noisy images might make it harder to classify them, which can lead to a problem with uncertainty. Modifications arise from variations in the pixel intensity levels caused by noise in the image. The image's intensity range is not uniformly disrupted by this change in pixel intensity values. Noise in an image can be caused by a variety of factors, including motion in the image, blurring, and lack of variation in the features. The partial volume averaging problem is the cause of the image pixels' fluctuating intensity levels. To handle this ambiguity in medical diagnostic systems, image segmentation is essential [8].

Modalities in medical image segmentation [9]:

MRI: It's clear that MRI (Magnetic resonance imaging) brain scans are essential for study when looking into segmentation applications in the medical area. This is because the images have a high signal-to-noise ratio, making it necessary to segment and enhance the image to identify the region of interest. Because of the contrast levels, working with images of different resolutions might make it more difficult to divide the image efficiently. These images pose an additional challenge. The primary uses of this framework are for brain volume extraction, brain structure outline, and the division of various problems about grey and white matter as well as cerebrospinal fluid.

CT (Computed Tomography): Segmentation is a valuable technique in the analysis of computed tomography images. Examining thoracic scans, segmenting images of various organs like the heart, stomach, brain, and liver, and outlining abdominal aortic aneurysms are the main uses of the segmentation technique discussed here. The contrast and resolution of these images are not as good as those from an MRI. Segmenting CT scans involves the use of various techniques.

Ultrasound: Segmenting the region of interest accurately is a challenging task because of the frequent imperfections in ultrasound images. Many techniques proved ineffective in segmenting ultrasound images. Despite this issue, significant progress has been made in this field. Most often, manual segmentation is performed, but these images are also used for motion estimation and pathology identification through textural classifiers.

Multimodal: In this example, many techniques are applied at the same time to identify an issue. To isolate a particular region of interest, data from many modalities are combined. This method's drawback is that it might not always be able to gather data from several sources. In this situation, another difficulty is having to concentrate mostly on an alignment method.

Digital mammography: Segmenting digital mammography images is usually necessary to identify distinct types of malignancies. Several thresholding procedures are used in common methods for segmenting mammography images.

Chest Radiography: Chest radiography involves taking a radiograph to assess or diagnose diseases in the chest and its anatomy. Key concerns about the analysis of chest infections are grouped under ABCDEF, where each letter stands for a distinct issue.

Challenges associated with medical image segmentation [10]:

Variability in Image Acquisition: Different imaging machines and protocols result in variability in image appearance. This can make it challenging to develop segmentation models that generalize well to new images.

Noise and Artifacts: Medical images can contain noise and artifacts that obscure the underlying anatomy, making accurate segmentation difficult.

Anatomical Variability: There is significant natural variation in human anatomy across individuals. This variability makes it hard to develop one-size-fits-all segmentation models.

Class Imbalance: In some medical images, certain structures of interest might be much smaller or less frequently occurring compared to the rest of the image, leading to a class imbalance problem.

Expert Annotation: Creating accurate ground truth annotations for medical image segmentation is time-consuming and requires expert knowledge. Inter-observer variability can also complicate establishing a definitive ground truth.

Computational Complexity: Medical images can be very large and complex, which can make it computationally expensive to develop and train segmentation models.

Generalization: Models trained on one dataset may not generalize well to other datasets, especially if there are differences in acquisition protocols or patient populations.

Interpretability: The output of segmentation models, particularly those based on complex architectures like deep neural networks, can be difficult to interpret. This makes it challenging to identify errors or diagnose issues with the segmentation model.

Given the complexity and variability inherent in medical images due to differing anatomies, pathologies, and imaging modalities, achieving precise and reliable segmentation is challenging. The problem is compounded by the limitations of current segmentation methods, which may struggle with noise, varying contrast levels, and the high dimensionality of data. Furthermore, the scarcity of labeled data due to the need for expert annotation, concerns about patient privacy, and the resource-intensive nature of acquiring labeled medical images stymie the development of robust segmentation models. This research aims to address these issues by developing advanced segmentation techniques that can provide accurate delineation of anatomical structures and pathologies in medical images, despite the aforementioned challenges.

The goal is to enhance the segmentation process to support better diagnosis, treatment planning, and monitoring of diseases while minimizing the need for extensive labeled datasets and ensuring robustness against image imperfections.

Comprehensive Overview and Categorization: This paper provides a systematic and extensive survey of existing medical image segmentation methods, categorizing them into traditional image processing techniques, machine learning approaches, and advanced deep learning models. It emphasizes the evolution from manual and semi-automatic methods to the current state-of-the-art deep learning architectures, highlighting their applications across various medical imaging modalities.

Critical Evaluation of Segmentation Techniques: The paper evaluates the performance, robustness, and practical applicability of each segmentation technique, offering insights into its advantages and limitations within clinical settings. It identifies the gaps and challenges faced by current methodologies, particularly in handling data scarcity, noise, and complexity, thereby setting the stage for future research directions.

Future Directions and Methodological Advancements: In addition to summarizing the current site of medical image segmentation, the paper outlines potential research trajectories, including the integration of multimodal imaging data, the use of unsupervised and semi-supervised learning to mitigate the issue of labeled data scarcity, and the development of more generalizable models that can adapt to a wide range of tasks without extensive retraining. It also discusses the prospects of explainable AI in improving the interpretability of deep learning models for segmentation, ensuring their reliability and trustworthiness in clinical practice.

II. LITERATURE SURVEY

Image segmentation in computer vision is the division of a digital image into several segments (sets of pixels, also known as image objects). The purpose of segmentation is to reduce complexity and/or transform an image's representation into something more relevant and understandable. Since segmenting a huge image into smaller ones makes subsequent processing easier, segmentation is significant in image processing

A technique that combines ResUNet and RNN (Recurrent Neural Network) for medical image segmentation was presented in [11]. Using feature accumulation of recursive residual convolutional layers, this approach improves feature representation for image segmentation problems. The recurrent residual convolutional unit is shown in Figure 5. To improve segmentation accuracy, they modeled the temporal relationship between several brain MRI slices using LSTM (Long Sort Term Memory) and CNN (Convolutional Neural Network). To derive aortic sequence segmentation from spatiotemporal data, [12] used FCN (Fully Convolutional Network) with RNN. Taking into account the interaction between context and information, RNNs can efficiently understand both the local and global spatial characteristics of images. Remarkable temporal details and precise segmentation depend on high-quality medical images. Using the RNN design to enhance medical image segmentation performance is not prevalent. The segmentation process can be influenced by the image's many characteristics, including texture and color. A crucial component of medical image segmentation is recognizing boundaries inside a 2D (2-dimensional) or 3D (3-Dimentional) image. Medical images are also created using a wide range of modalities, including X-ray, CT, MRI, microscope, PET (positron emission tomography), SPECT (Single-photon emission computed tomography), endoscopy, OCT (Optical coherence tomography), and many more [13].

The literature has a large number of in-depth analyses of CNN-based medical image analysis. Gives a thorough rundown of medical image processing, using cutting-edge machine learning and deep learning methods. Their research demonstrates the benefits of applying particular CNN models in contrast to conventional machine learning techniques in a range of medical contexts. Similarly, [14] offered a thorough examination of the uses of deep learning in several facets of medical image processing. Furthermore, this offers a fundamental comprehension of CNN models and their use in medical image analysis. Nevertheless, the majority of these evaluations have concentrated on analyzing deep learning methods according to specific applications.

Given that CNN models are intended to operate on image data, the dimensionality of the image data has a significant influence on which model should be chosen. Typically, two-dimensional (2D) or three-dimensional (3D) imaging methods are used to create medical image data. Digital X-rays, retinal fundus images, microscopic images in pathology, mammograms, and more are examples of 2D image categories in the biomedical profession. In the field, 3D medical imaging methods including CT, MRI, and ultrasound are often employed. However, the majority of 3D medical image processing research has made extensive use of 2D CNN models [15].

In medical image processing, organ or anatomical structure segmentation is an essential step. Finding abnormalities and precisely estimating the size of organs or lesions are the main objectives. describing the major components of 3D CNN architectures and offering an overview of deep learning methods for the analysis of volumetric medical images [16]. To our knowledge, no thorough literature studies have been done specifically on 3D deep learning techniques for medical image segmentation. We performed a thorough examination of the most recent developments in deep learning for 3D medical image segmentation. There has also been a discussion of potential future developments in 3D medical image segmentation.

Numerous techniques for interactive segmentation have been put forth. Some common techniques include Graph Cuts, Random Walks, and GeoS. It is becoming more and more popular to employ machine learning to improve user experience and produce accurate results. GrabCut, for instance, segments color images using GMMs. Slic-Seg uses fetal magnetic resonance imaging (MRI) to segment the placenta using Online Random Forests (ORFs). In, 3D Computed Tomography (CT) image segmentation is achieved using active learning. They have produced better accurate segmentations with fewer user inputs than conventional interactive segmentation techniques [17].

The development of medical image segmentation and semantics was recently covered by Six categories including the deep learning-based approaches to image segmentation: sequenced models, loss function-based, deep architecture, data synthesis-based, weakly supervised, and multi-task methods. To improve a survey on medical image segmentation, they investigated a variety of deep learning architectures and conventional machine learning techniques. techniques for medical image segmentation with partial data sets examined two main data set constraints: scant and inadequate annotations. Every one of these surveys is essential to the development of medical image segmentation methods. Three factors were assessed: barriers, training techniques, and network design. The network structures section contains a thorough description of the primary network topologies used for image segmentation. The training methodologies section discusses the J Digit imaging method, which is used to train deep neural network models. The problems section lists the difficulties in applying deep learning methods for medical image segmentation. The possible uses of deep learning in radiotherapy were examined by [18]. The three categories of insufficient supervision examined in their study were incomplete, imprecise, and incorrect. Using Dice scores or Jaccard indices, concentrate on analyzing and summarizing optimization techniques for medical image segmentation. a new two-step approach to medical image segmentation was proposed, First, morphological functions are used to create a binary marker that represents the desired region of interest. Using the grab-cut approach, the binary marker image serves as a mask in the second phase of the segmentation process. Results can be further improved by using a graphical user interface (GUI) with little to no user input.

Using color images of the eye fundus, [19] studied the use of deep convolutional neural networks for the classification of diabetic retinopathy. Their suggested approach outperformed conventional feature-based categorization techniques in terms of performance. proposed deep learning models, such as Convolutional Neural Networks (CNN) and Deep Neural Networks (DNN), to classify various characteristics in retinal images, such as microaneurysms, blood vessels, exudates, fluid drops, and hemorrhages. Using particle swarm optimization, presented a method for detecting Alzheimer's illness by optimizing extracted characteristics such as variance, kurtosis, eigenvector, mean, and standard derivation. Next, the presence of Alzheimer's disease in the brain image is assessed using a decision tree classifier. Their suggested method's accuracy is merely 91%.

[19] Techniques that are frequently used to provide spatial importance maps as post-hoc classification explanations for a particular input image include class activation maps (CAM) and saliency maps. These maps can help interpret the predictions that black-box models make. Grad-CAM, for instance, is used to interpretively classify electrocardiograms, CT lung nodules, and chest X-ray images. Saliency maps were used to help classify retinal OCT images. Saliency maps from Bag Net are used to provide a comprehensible gender classification.

Because of this, in currently available commercial systems, interactive segmentation is still the preferred method for surgical planning and navigation. Even though this frequently results in more thorough segmentations, the goal of an interactive technique should be to decrease user time to lessen the load on users. Inspired by these findings, we explore how integrating CNNs with user interactions for medical image segmentation may improve segmentation robustness and accuracy while requiring fewer user interactions and time. Nevertheless, the literature on CNNs for interactive segmentation [20]. In addition to lacking image-specific adaptability, training requires a significant number of annotated images and struggles to balance the efficiency of memory space, inference time, and model complexity.

Research gap

Insufficient Labeled Data: The lack of publicly available, well-annotated medical image datasets due to privacy concerns and the requirement for expert knowledge for accurate labeling is a significant hurdle. This gap limits the training and validation of advanced segmentation models, especially deep learning approaches that are data-intensive.

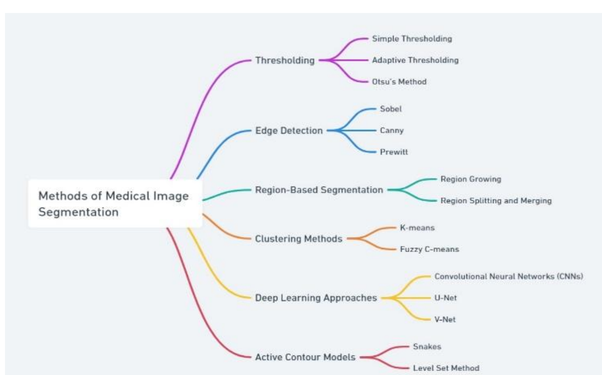
Generalizability of Models: Many segmentation models are specialized for specific imaging modalities or particular anatomical structures and lack the flexibility to be applied universally across different datasets and tasks without extensive re-tuning or retraining.

Noise and Artifact Handling: Current segmentation techniques often underperform in the presence of noise and artifacts, which are inherent in medical imaging due to patient movement, variable imaging conditions, and equipment differences. There is a need for robust algorithms that can effectively manage these issues.

Integration of Multimodal Data: There is a gap in effectively integrating information from various imaging modalities to improve segmentation accuracy. Multimodal segmentation models are not yet fully developed, and most current methods do not take full advantage of the complementary information available across different imaging techniques.

Explainability and Interpretability: As deep learning models become more complex, their "black box" nature becomes a barrier to clinical adoption. There is a need for methods that provide better interpretability of the segmentation results, allowing clinicians to understand and trust the decision-making process of AI-based segmentation tools.

Various Medical image segmentation is a critical task in image processing, aiding significantly in clinical diagnosis and treatment planning. Various techniques have been developed, each with their distinct trade-offs. Simple and Adaptive Thresholding offers ease of use and adaptability but often falls short in complex lighting conditions. Edge detection methods like Sobel, Canny, and Prewitt are fundamental for capturing structural details, yet they may suffer from noise sensitivity and accuracy issues. Region-based methods such as Region Growing and Region Splitting/Merging cater to precise segmentation tasks but can be computationally intensive. Clustering methods like K-means and Fuzzy C-means provide a balance between simplicity and flexibility, but they are not without limitations such as sensitivity to initialization and scalability issues. Deep learning approaches, particularly CNNs, U-Net, and V-Net, stand out for their accuracy and adaptability but require extensive computational resources and large labeled datasets. Lastly, active contour models, including Snakes and Level Set Methods, offer sophisticated means of capturing the intricate structures of medical images, despite their challenges with initialization and computational demands. The choice of technique is thus dictated by the specific requirements of the application, the available computational resources, and the desired level of accuracy in the segmentation outcome. Figure 1 shows the Methods of medical image segmentation. Table 1 shows the survey table.



Methods of medical image segmentation

Thresholding

Simple Thresholding: It's a basic form of image segmentation where we binarize an image by choosing a threshold value. Pixels above this value are set to one color (often white), and those below to another (often black).

Adaptive Thresholding: Instead of a single threshold value, adaptive thresholding changes the threshold dynamically over the image, which is useful for images with varying lighting conditions.

Otsu's Method: An automatic thresholding technique that chooses the threshold to minimize intra-class variance of the black and white pixels.

Edge Detection

Sobel: A gradient-based edge detection method that uses a pair of 3x3 convolutional kernels, one estimating the gradient in the x-direction and the other in the y-direction.

Canny: A multi-stage algorithm that involves blurring the image to filter out noise, finding intensity gradients, applying non-maximum suppression, and using edge tracking by hysteresis.

Prewitt: Similar to the Sobel operator, it uses a different pair of convolutional kernels to detect horizontal and vertical edges.

Region-Based Segmentation

Region Growing: This method starts with a seed point and grows a region by appending to each seed those neighboring pixels that have similar properties.

Region Splitting and merging this process involves dividing an image into a set of disjoint regions and then merging and/or splitting them according to a set of predefined rules.

Clustering Methods

K-means: A popular clustering algorithm that partitions the image into K clusters by minimizing the variance within each cluster.

Fuzzy C-means: A form of clustering in which each data point belongs to a cluster to a degree specified by a membership level.

Deep Learning Approaches

Convolutional Neural Networks (CNNs): A class of deep neural networks, commonly applied to analyzing visual imagery, which uses a variation of multilayer perceptrons designed to require minimal preprocessing.

U-Net: A convolutional network architecture for fast and precise segmentation of images.

V-Net: An extension of U-Net designed for volumetric (3D) image segmentation.

Active Contour Models

Snakes: These are curves defined within an image domain that can move under the influence of internal forces coming from within the curve itself and external forces computed from the image data.

Level Set Method: A numerical technique for tracking interfaces and shapes. The advantage of the level set model is that it is implicit, non-parametric, and can handle topological changes naturally.

Survey Table

Method	Advantages	Disadvantages	Shortcomings
Simple Thresholding	Easy to implement; fast	Not adaptable to variable lighting; simplistic	Poor results on complex images
Adaptive Thresholding	Better for varying lighting; more flexible	More complex than simple thresholding; slower	May still fail with very complex lighting conditions
Otsu's Method	Automatically determines	May not be optimal for non-	Can struggle with noise and

	the threshold	bimodal histograms	texture
Sobel	Simple; detects edges in specific directions	Sensitive to noise; only approximates edge magnitude	Not as accurate as more advanced methods
Canny	Accurate; good at detecting weak edges	Complex; requires fine-tuning of parameters	May produce false edges
Prewitt	Simple; emphasizes horizontal or vertical edges	Less accurate than other methods; sensitive to noise	Similar shortcomings to Sobel
Region Growing	Intuitive; can produce precise regions	Sensitive to noise; dependent on seed selection	May not work well on non-uniform intensity regions
Region Splitting/Merging	Can handle variable textures and intensities	Can be computationally expensive; complex rules needed	This can result in over-segmentation
K-means	Simple; relatively efficient	Requires specification of cluster number; sensitive to noise	Can converge to local minima
Fuzzy C-means	Allows for soft clustering; more flexible than K-means	Computationally intensive; sensitive to initialization	Not suitable for very large datasets
CNNs	Highly accurate; adaptable	Requires large labeled datasets; computationally intensive	May overfit to training data
U-Net	Designed for small datasets; precise segmentation	Training still requires sufficient data; computationally demanding	Complex architecture
V-Net	Extends U-Net to 3D images; very precise	Even more computationally intensive than U-Net	Requires a lot of data and resources
Snakes	Flexible to represent shape; intuitive to control	Sensitive to initialization; can get trapped in local optima	Struggle with concavities
Level Set Method	Can naturally handle topological changes; implicit surface	Computationally expensive; parameter tuning required	Difficult to implement and optimize

Workflow of steps involved in medical image segmentation

Acquire Medical Image: A medical image is obtained. This could be any type of medical imaging, such as an X-ray, MRI, CT scan, etc.

Pre-process Image: The acquired image is then pre-processed. This step may involve noise reduction, contrast enhancement, or any other technique to prepare the image for further processing.

Is Image Quality Sufficient? This decision step involves evaluating whether the pre-processed image is of high enough quality to proceed. If yes, the process moves to the next step. If not, the flowchart is directed to enhance the image quality.

Enhance Image Quality: If the image quality is deemed insufficient, this step involves using various techniques to improve the clarity, contrast, and overall quality of the image.

Select Segmentation Method: Once the image quality is sufficient, a segmentation method is selected. This could involve choosing between different algorithms or techniques that best suit the image characteristics and the desired outcome.

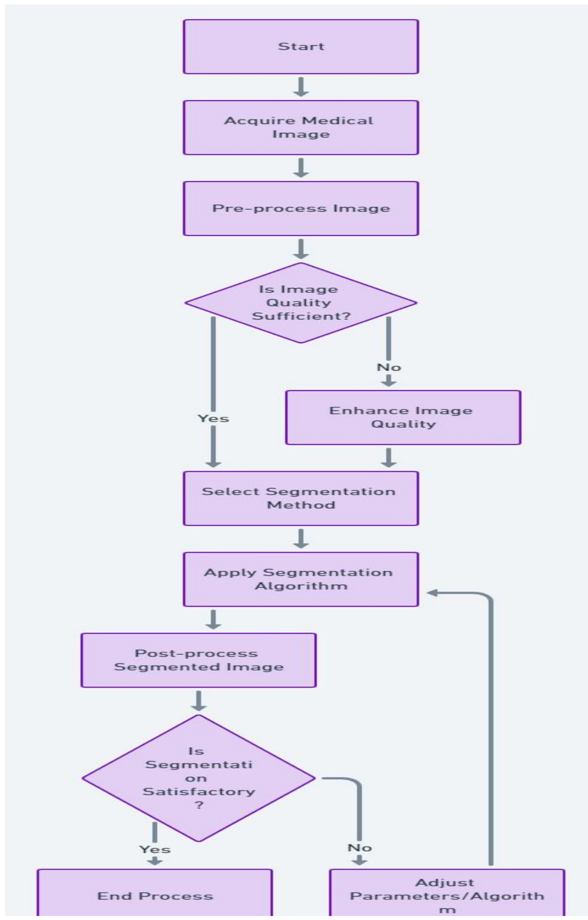
Apply Segmentation Algorithm: The chosen segmentation algorithm is applied to the image. Segmentation involves dividing the image into parts to isolate and analyze specific regions of interest, like separating a tumor from healthy tissue.

Post-process Segmented Image: After segmentation, the image may need further post-processing. This could include refining the edges of the segmented regions, removing small artifacts, or preparing the segmented image for analysis.

Is Segmentation Satisfactory? Another decision step is to evaluate the quality of the segmentation. If the segmented image is satisfactory, the process can end. If not, adjustments need to be made.

Adjust Parameters/Algorithm: If the segmentation is not satisfactory, this step requires making adjustments to the parameters of the segmentation algorithm or even choosing a different algorithm to improve the results.

End Process: This final step concludes the workflow. If the segmentation is satisfactory, the processed image is likely ready for further analysis or interpretation by medical professionals. Figure 1 shows the workflow of medical image segmentation.



Workflow of medical image segmentation

III. CONCLUSION

In conclusion, our comprehensive survey of medical image segmentation has elucidated the dynamic landscape of this field, showcasing a transition from manual delineation to semi-automatic methods, and evolving into sophisticated automated techniques powered by deep learning. The critical examination of various segmentation methods has unveiled persistent challenges, such as the handling of noise, the requirement for large annotated datasets, and the generalization across diverse pathologies and imaging modalities. Although deep learning models, particularly convolutional neural networks, have set new benchmarks in segmentation accuracy, their performance is often bottlenecked by the availability of quality data and the need for extensive computational resources. Future work in medical image segmentation should pivot towards the development of robust algorithms that can operate with minimal supervision and adapt to new tasks with limited retraining. There is a promising avenue in the exploitation of unsupervised and semi-supervised learning paradigms to alleviate the dependency on labeled datasets. Furthermore, the integration of multimodal imaging data presents an untapped potential for improving the accuracy and reliability of segmentation results. The pursuit of explainable AI in segmentation will enhance the trust and interpretability of AI-assisted diagnosis systems, ensuring alignment with clinical rationale. Addressing these gaps will not only improve the efficacy of clinical diagnostics but also pave the way for personalized medicine, where tailored treatments are informed by precise, AI-enhanced imaging analyses. Lastly, fostering collaboration across disciplines—merging insights from radiology, computer science, and bioinformatics—is essential to harness the full potential of AI in improving patient care through advanced medical image segmentation.

ACKNOWLEDGEMENT

I would like to express our sincere gratitude to all those who have supported and contributed to this research project. Primarily, I extend our heartfelt thanks to our guide for his unwavering guidance, invaluable insights, and encouragement throughout the research process. No funding is raised for this research.

REFERENCES

[1]. S. Pang *et al.*, "SpineParseNet: Spine Parsing for Volumetric MR Image by a Two-Stage Segmentation Framework With Semantic Image Representation," in *IEEE Transactions on Medical Imaging*, vol. 40, no. 1, pp. 262-273, Jan. 2021, doi: 10.1109/TMI.2020.3025087.

[2]. T. Hassanzadeh, D. Essam and R. Sarker, "2D to 3D Evolutionary Deep Convolutional Neural Networks for Medical Image Segmentation," in *IEEE Transactions on Medical Imaging*, vol. 40, no. 2, pp. 712-721, Feb. 2021, doi: 10.1109/TMI.2020.3035555.

[3]. L. Chen, P. Bentley, K. Mori, K. Misawa, M. Fujiwara and D. Rueckert, "DRINet for Medical Image Segmentation," in *IEEE Transactions on Medical Imaging*, vol. 37, no. 11, pp. 2453-2462, Nov. 2018, doi: 10.1109/TMI.2018.2835303.

[4]. Y. Xie, J. Zhang, H. Lu, C. Shen and Y. Xia, "SESV: Accurate Medical Image Segmentation by Predicting and Correcting Errors," in *IEEE Transactions on Medical Imaging*, vol. 40, no. 1, pp. 286-296, Jan. 2021, doi: 10.1109/TMI.2020.3025308.

[5]. G. Wang *et al.*, "Interactive Medical Image Segmentation Using Deep Learning With Image-Specific Fine Tuning," in *IEEE Transactions on Medical Imaging*, vol. 37, no. 7, pp. 1562-1573, July 2018, doi: 10.1109/TMI.2018.2791721.

[6]. Z. Gu *et al.*, "CE-Net: Context Encoder Network for 2D Medical Image Segmentation," in *IEEE Transactions on Medical Imaging*, vol. 38, no. 10, pp. 2281-2292, Oct. 2019, doi: 10.1109/TMI.2019.2903562.

[7]. J. Du, X. Zhang, P. Liu and T. Wang, "Coarse-Refined Consistency Learning Using Pixel-Level Features for Semi-Supervised Medical Image Segmentation," in *IEEE Journal of Biomedical and Health Informatics*, vol. 27, no. 8, pp. 3970-3981, Aug. 2023, doi: 10.1109/JBHI.2023.3278741.

[8]. Z. Li *et al.*, "LViT: Language Meets Vision Transformer in Medical Image Segmentation," in *IEEE Transactions on Medical Imaging*, vol. 43, no. 1, pp. 96-107, Jan. 2024, doi: 10.1109/TMI.2023.3291719.

[9]. D. Gut, Z. Tabor, M. Szymkowski, M. Rozynek, I. Kucybała and W. Wojciechowski, "Benchmarking of Deep Architectures for Segmentation of Medical Images," in *IEEE Transactions on Medical Imaging*, vol. 41, no. 11, pp. 3231-3241, Nov. 2022, doi: 10.1109/TMI.2022.3180435.

[10]. Y. Sun, P. Yuan, Y. Sun and Z. Zhai, "Hybrid Segmentation Algorithm for Medical Image Segmentation Based on Generating Adversarial Networks, Mutual Information and Multi-Scale Information," in *IEEE Access*, vol. 8, pp. 118957-118968, 2020, doi: 10.1109/ACCESS.2020.3005384.

[11]. X. Zhao *et al.*, "Prior Attention Network for Multi-Lesion Segmentation in Medical Images," in *IEEE Transactions on Medical Imaging*, vol. 41, no. 12, pp. 3812-3823, Dec. 2022, doi: 10.1109/TMI.2022.3197180.

[12]. S. Mo et al., "Mutual Information-Based Graph Co-Attention Networks for Multimodal Prior-Guided Magnetic Resonance Imaging Segmentation," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 32, no. 5, pp. 2512-2526, May 2022, doi: 10.1109/TCSVT.2021.3112551.

[13]. W. Lei *et al.*, "One-Shot Weakly-Supervised Segmentation in 3D Medical Images," in *IEEE Transactions on Medical Imaging*, vol. 43, no. 1, pp. 175-189, Jan. 2024, doi: 10.1109/TMI.2023.3294975.

[14]. Premakumari Pujar, Ashutosh Kumar, Vineet Kumar, "Efficient plant leaf detection through machine learning approach based on corn leaf image classification" IAES International Journal of Artificial Intelligence (IJ-AI), Vol. 13, No. 1, March 2024, pp. 1139~1148, ISSN: 2252-8938, DOI: 10.11591/ijai.v13.i1.pp1139-1148.

[15]. Sreedhara, S.H., Kumar, V., Salma, S. (2023). Efficient Big Data Clustering Using Adhoc Fuzzy C Means and Auto-Encoder CNN. In: Smys, S., Kamel, K.A., Palanisamy, R. (eds) Inventive Computation and Information Technologies. Lecture Notes in Networks and Systems, vol 563. Springer, Singapore. https://doi.org/10.1007/978-981-19-7402-1_25

[16]. Y. Ban *et al.*, "Concept Graph Neural Networks for Surgical Video Understanding," in *IEEE Transactions on Medical Imaging*, vol. 43, no. 1, pp. 264-274, Jan. 2024, doi: 10.1109/TMI.2023.3299518.

[17]. R. Cheke, C. Eising, P. Denny and P. v. d. Ven, "PatchLoc: Embedded Patch Localisation Pretext Task for Tumor Segmentation in Medical Images," in IEEE Access, doi: 10.1109/ACCESS.2024.3399628.

[18]. X. Huang, H. Gong and J. Zhang, "HST-MRF: Heterogeneous Swin Transformer With Multi-Receptive Field for Medical Image Segmentation," in IEEE Journal of Biomedical and Health Informatics, doi: 10.1109/JBHI.2024.3397047.

[19]. L. Teng et al., "Unified Model for Children's Brain Image Segmentation with Co-Registration Framework Guided by Longitudinal MRI," in IEEE Journal of Biomedical and Health Informatics, doi: 10.1109/JBHI.2024.3393974.

[20]. Y. Zhuang *et al.*, "A 3-D Anatomy-Guided Self-Training Segmentation Framework for Unpaired Cross-Modality Medical Image Segmentation," in *IEEE Transactions on Radiation and Plasma Medical Sciences*, vol. 8, no. 1, pp. 33-52, Jan. 2024, doi: 10.1109/TRPMS.2023.3332619.

[21]. Y. Ling, Y. Wang, W. Dai, J. Yu, P. Liang and D. Kong, "MTANet: Multi-Task Attention Network for Automatic Medical Image Segmentation and Classification," in *IEEE Transactions on Medical Imaging*, vol. 43, no. 2, pp. 674-685, Feb. 2024, doi: 10.1109/TMI.2023.3317088.

[22]. J. Hou *et al.*, "Research on Tumor Image Segmentation in Medical Imaging based on Extremum Adaptive Median Filtering," *2022 IEEE International Conference on Advances in Electrical Engineering and Computer Applications (AEECA)*, Dalian, China, 2022, pp. 648-652, doi: 10.1109/AEECA55500.2022.9918848.

[23]. Q. Ming and X. Xiao, "Towards Accurate Medical Image Segmentation With Gradient-Optimized Dice Loss," in *IEEE Signal Processing Letters*, vol. 31, pp. 191-195, 2024, doi: 10.1109/LSP.2023.3329437.

[24]. X. Chen *et al.*, "Anatomy-Regularized Representation Learning for Cross-Modality Medical Image Segmentation," in *IEEE Transactions on Medical Imaging*, vol. 40, no. 1, pp. 274-285, Jan. 2021, doi: 10.1109/TMI.2020.3025133.

[25]. X. Zhao *et al.*, "Prior Attention Network for Multi-Lesion Segmentation in Medical Images," in *IEEE Transactions on Medical Imaging*, vol. 41, no. 12, pp. 3812-3823, Dec. 2022, doi: 10.1109/TMI.2022.3197180.0858

BIOGRAPHIES OF AUTHORS

	**Sarala D** earned her Bachelor's of Engineering BE degree in CSE from VTU, Belagavi in 2008. She obtained her master's degree in M.Tech (CSE) from VTU in 2013. Currently, she is a research scholar at VTU doing her Ph.D. in Computer Science and Engineering and also working as an Assistant Professor in KLE Technological University. She has attended many workshops and induction programs conducted by various universities. Her areas of interest are Machine Learning, Biomedical imaging. She can be contacted at email: saralad_12@rediffmail.com. https://scholar.google.com/citations?user=8fqWHvsAAAAJ&hl=en https://orcid.org/0009-0004-1494-785X
	**Dr. Dayananda R. B. ,** Associate Professor in the Department of Computer Science and Engineering, MSRIT, Bengaluru. He has academic experience of 15+ years and 4 years of Industry experience. He worked as Professor, HOD, IQAC Director, Research Center Head, and Vice Principal at Prestigious Engineering Institutions. Under his guidance Ph.D. degree is awarded to a research scholar from VTU on machine learning in the year 2021. His research mainly focuses on the Design and Development of a Cloud Computing Architecture for data security and machine learning. He can be contacted at this email: dayanandarb@msrit.edu. DAYANANDA R B - Google Scholar Orcid ID - 0000-0002-4780-0932 Scopus ID - 57410751300