International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 02 Issue: 11 November 2024
Page No: 2569 - 2574
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2024.0353

Medical Image Segmentation

Arpit Mohankar[1], Aishwarya Nagpure[2], Sania Shaikh[3], Khushi Singh[4], Firdous Jahan Shaikh[5]
[1,2,3,4]UG Scholar, Dept. of AI&DS, Rajiv Gandhi Institute of Tech., Mumbai, Maharashtra, India.
[5]Assistant Professor, Dept. of AI&DS, Rajiv Gandhi Institute of Tech., Mumbai, Maharashtra, India.
Emails: *arpitmohankar24@gmail.com[1], aish8403@gmail.com[2], sania.shaikh9202@gmail.com[3], khushimangal2003@gmail.com[4], firdous.khan@mctrgit.ac.in[5]*

Abstract

Medical image segmentation is a critical component in the development of computer-aided diagnosis and treatment planning systems. This paper provides a comprehensive survey of recent advances in segmentation techniques applied to various imaging modalities, including Magnetic Resonance Imaging (MRI). Traditional methods such as thresholding, region-growing, and active contours are reviewed alongside contemporary machine learning-based approaches, particularly deep learning models. The survey emphasizes the growing dominance of convolutional neural networks (CNNs) and their variants, including U-Net and Fully Convolutional Networks (FCNs), which have shown remarkable success in handling complex medical imaging challenges. Additionally, the paper discusses hybrid methods that combine classical techniques with artificial intelligence to improve accuracy and robustness in segmentation tasks. Key challenges such as class imbalance, boundary delineation, and computational efficiency are also highlighted. Future directions, including the integration of multi-modal data and advancements in self-supervised learning, are explored as potential solutions to overcome current limitations in medical image segmentation.

Keywords: *Convolutional neural networks (CNNs); Deep learning; Medical image segmentation; U-Net; Fully Convolutional Networks (FCNs)*

1. Introduction

Medical image segmentation is a crucial task in healthcare, allowing for accurate identification and outlining of anatomical structures and abnormal regions in imaging techniques like MRI, CT scans, and X-rays. Accurate segmentation is essential for various clinical applications, including disease diagnosis, treatment planning, and surgical guidance. Traditionally, this process depended on manual annotations or rule-based algorithms, which are time-consuming and prone to variability between observers. However, with the rise of deep learning, particularly convolutional neural networks (CNNs), the precision, speed, and automation of medical image segmentation have improved significantly, surpassing traditional methods with reliable pixel-wise classifications. Additionally, advancements in both computational power and model architecture have further enhanced medical image segmentation.

Models like U-Net, Fully Convolutional Networks (FCNs), and SegNet are popular due to their ability to capture both detailed and global information. These models utilize an encoder-decoder structure, where the encoder extracts high-level features and the decoder refines the segmentation output. They have demonstrated state-of-the-art performance in detecting organs, tumors, and lesions. Despite these advancements, the field still faces significant challenges. A key issue is the limited availability of large, annotated medical datasets, which are vital for training deep learning models. Furthermore, variations in imaging methods, acquisition protocols, and patient demographics make it difficult for models to generalize effectively. To address these challenges, researchers are focusing on data-efficient training strategies like semi-supervised learning, data augmentation, and transfer learning. Additionally, efforts are underway to improve model robustness and ensure consistent performance across various clinical settings, which is crucial for broader adoption in real-world healthcare environments. This paper includes a study of nine papers that showcase different methods for medical image segmentation.

The rest of the paper deals with a literature survey, results, and conclusion of the various techniques and methods used for medical image segmentation.

2. Method

Convolutional Neural Networks (CNN)-Based Methods: CNNs are widely used in medical image segmentation due to their ability to learn spatial patterns from image data. They typically consist of layers of convolutions and pooling, which extract and downsample features. In segmentation tasks, CNNs often require additional decoders to map image classifications back to pixel-level outputs. Autoencoder-based CNNs, incorporating skip connections, enhance spatial detail retention, offering higher segmentation accuracy. These models have been effective in various medical applications, such as brain tumor classification and segmentation, where they provide robust and adaptable solutions.Fully Convolutional Networks (FCN) and U-Net Variants: Fully Convolutional Networks (FCNs) and U-Net architectures are designed specifically for pixel-wise segmentation, making them highly effective for medical image tasks. U-Net uses an encoder-decoder structure with skip connections, preserving spatial details and producing accurate segmentation maps. Enhancements like residual blocks and attention mechanisms, seen in ResUNet+, further improve performance by retaining key features. These architectures are widely used in complex segmentation tasks, such as organ delineation and tumor detection, where precise boundary detection is essential.Transformer and Hybrid Architectures: Transformer-based models have recently advanced medical image segmentation by capturing both local and global features. Using self-attention mechanisms, they excel at understanding large and complex image regions. Models like TransUNet combine Vision Transformers (ViT) with U-Net, while hybrid architectures like R50-ViT merge CNNs with transformers, offering improved segmentation performance. These methods are particularly valuable in tasks like brain tumor segmentation, where both detailed local information and broader context are crucial for accurate outcomes.

3. Related Works

This paper [1] presents three architectures for brain MRI analysis. The first model classifies MRI images into affected and unaffected regions using 75x75 patches from 256x256 images. It features 10 convolutional layers and 4 max-pooling layers, but requires a separate decoder for segmentation. The second architecture, an autoencoder with an encoder-decoder structure, uses skip connections for better spatial retention. The third model incorporates an upsampling attention module, improving feature extraction with pixel- and channel-level attention.The workflow [2] of the ResUNet+ model for brain tumor segmentation begins with the preprocessing of MRI images, where multiple modalities (T1, T2, and FLAIR) are normalized and processed to extract the Region of Interest (ROI) that highlights the tumor and edema regions. Following this, the model employs a hybrid architecture that integrates residual blocks and attention mechanisms within a U-Net framework, allowing for effective feature extraction and minimizing semantic gaps between the encoder and decoder. The model processes the ROI images separately, utilizing skip connections to retain fine details while combining features from different modalities. Finally, the ResUNet+ model is trained and evaluated on established datasets such as BraTS 2020, BraTS 2019, and BraTS 2018, demonstrating superior segmentation performance compared to existing state-of-the-art methods.The proposed system [3] preprocesses MRI brain images by converting them to 2D grayscale, applying median filtering to reduce noise, and enhancing image quality using contrast-limited adaptive histogram equalization. Data augmentation via rotation and scaling addresses varying resolutions. Segmentation isolates the skull using OTSU-based thresholding, and active contour methods outline the tumor region. Key features like asymmetry, diameter, and border irregularity are extracted for classification. A CNN, modified for brain tumor images, is used along with the Adam optimizer and binary cross-entropy loss. Post-classification, advanced segmentation methods improve tumor detection and density estimation using Gaussian kernel distribution.In [4], the preprocessing phase focuses on denoising medical images, specifically MRI brain scans, using wavelet-based thresholding. Various wavelet families, including Haar, Symlet, Morlet, and Daubechies, are applied.

These wavelets remove noise while preserving critical image features like edges. Metrics such as signal-to-noise ratio (SNR), peak signal-to-noise ratio (PSNR), and mean square error (MSE) are used to evaluate performance. Wavelets with smaller support, like Haar and Symlet, prove to be ideal for capturing closely spaced features.The digital image processing technique begins with MRI as input, which undergoes pre-processing to enhance size and shape, followed by segmentation using the Chan-Vese algorithm. In the next stage, feature extraction is performed using principal component analysis (PCA) and gray level co-occurrence matrix (GLCM) techniques to improve accuracy. The images are then classified using deep convolutional neural networks (DCNN) [5]. The authors [6] propose a methodology that includes skull stripping for brain cortex extraction, followed by the segmentation of cerebrospinal fluid (CSF), gray matter (GM), and white matter (WM) using intensity-based techniques. Tumor regions are identified through region-based algorithms assessing pixel area properties. Features such as mean, variance, entropy, and energy are extracted from the segmented images, which are then classified using a Feedforward Neural Network (FFNN). The model's performance is evaluated through training and testing metrics, enhancing brain tumor detection accuracy.The paper [7] introduces SegNet which is a deep neural network for pixel-wise semantic segmentation. It has an encoder-decoder structure where the encoder mirrors VGG16's 13 convolutional layers. The decoder upsamples feature maps using pooling indices from the encoder's max-pooling layers, eliminating the need to learn upsampling. This method produces dense feature maps for pixel-wise classification. SegNet is compared to other models like FCN and DeepLab, showcasing a trade-off between memory efficiency and segmentation accuracy.The Swin-Unet architecture for medical image segmentation is built around the Swin Transformer block and includes an encoder, bottleneck, decoder, and skip connections. The encoder splits images into patches to create sequence embeddings and uses multiple Swin Transformer blocks for feature representation, employing patch merging for downsampling. The bottleneck features two Swin Transformer blocks for deep learning. The decoder utilizes patch expanding layers for upsampling, while skip connections merge shallow and deep features to preserve spatial information. This design enhances segmentation accuracy through efficient representation learning and feature fusion [8]. This paper [9] explores two model architectures for medical image segmentation: a transformer-based encoder (ViT) with 12 layers and a hybrid encoder (R50-ViT), which combines ResNet-50 with ViT. Both are pretrained on ImageNet, using a 224x224 input resolution and 16x16 patch size. TransUNet, compared to U-Net and AttnUNet, is evaluated using metrics like Dice Similarity Coefficient and Hausdorff distance. Key findings show that TransUNet with skip connections outperforms others by capturing global context and local details, and hybrid architectures (R50-ViT) perform better than pure Transformer-based models. Factors like skip connections, resolution, and patch size are fine-tuned for optimal performance. [10-12]

4. Results and Discussion

Among the three architectures used in paper [1], an accuracy of 99.19 % is obtained, which is found to be low compared to using 5x5 filters for all 10 convolutional layers. The second architecture achieves a 98.84% segmentation accuracy and a Dice score of 91.76%, while the third achieves 99.49% segmentation accuracy and a 99.81% F-score. Figure 1 represents the flowchart used by the authors of the paper [1].

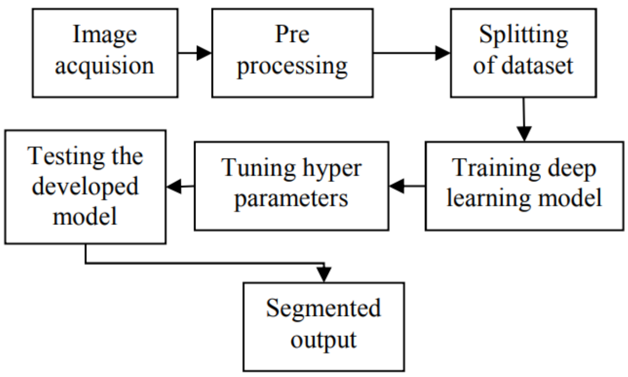

Figure 1 Flow Graph of Proposed Model

In paper [2] test results of the BraTS 2020 dataset are shared, where dice loss values were 91.90, 93.10, and

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 02 Issue: 11 November 2024
Page No: 2569 - 2574
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2024.0353

92.80 for tumor core, whole tumor, and enhancing tumor respectively.The model in paper [3] is compiled in Keras with 'Adam' optimizer and 'Binary cross entropy' loss. A default learning rate of 0.001 is used. The model is trained with a batch size of 32 for 24 epochs, and for the test images, it gives an accuracy of 95.6%. Figure 2 shows the methodology used in paper [4].

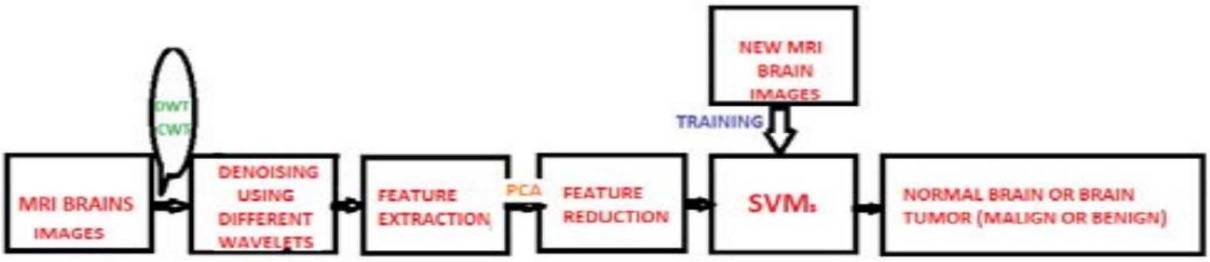

Figure 2 Detection & Classification Algorithm

Higher-level wavelets, such as Daubechies level 3 and Haar wavelets, produce better results, with Haar wavelets showing the best SNR and minimal MSE. Binary SVM yields an accuracy of 92%, binary linear classification of 91%, and binary kernel classification of 99%. Figure 3 shows the preprocessing done by the authors in paper [5] using the Chan-Vese algorithm followed by methods like GLCM, PCA, and DCNN. with 12 layers and a hybrid encoder (R50-ViT), which combines ResNet-50 with ViT. Both are pretrained on ImageNet, using a 224x224 input resolution and 16x16 patch size. TransUNet, compared to U-Net and AttnUNet, is evaluated using metrics like Dice Similarity Coefficient.

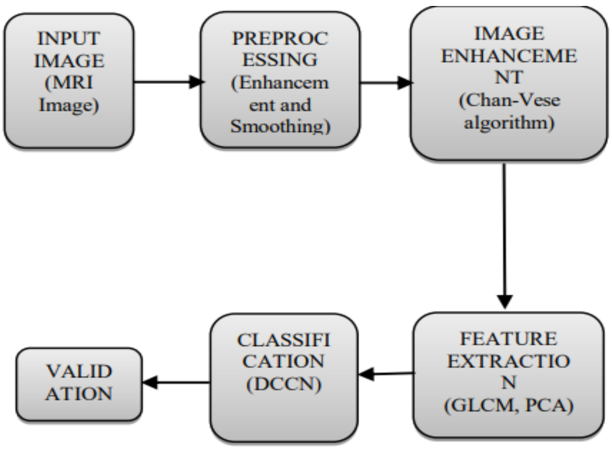

Figure 3 Pre-Processing

The steps are repeated for 50 images, giving efficient results. Figure 4 represents the techniques used in paper [6].

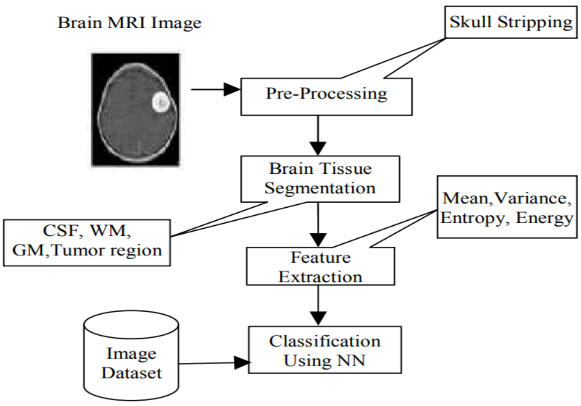

Figure 4 Block Diagram of the Proposed Technique

The authors in this paper [6] achieve an accuracy of 67% using KNN classification, 83% using neural networks, and 67% using Bayesian classification. SegNet outperforms all the other methods in paper [7], including those using depth, video, and/or CRF's on the majority of classes. In comparison with the CRF based methods, SegNet predictions are more accurate in 8 out of the 11 classes. It also shows a good ~10 percent improvement in class average accuracy when trained on a large dataset of 3.5 K images. Particularly noteworthy are the significant improvements in accuracy for the smaller/thinner classes. It achieves a Mean Intersection Over Union (mIOU) score of 60.10. The experimental results [8] show that the Unet-like pure transformer model achieves a segmentation accuracy of 79.13% in terms of Dice Similarity Coefficient (DSC↑) and 21.55% for the Hausdorff Distance (HD↓). Compared to Att-

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 02 Issue: 11 November 2024
Page No: 2569 - 2574
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2024.0353

Unet and the recent TransUnet method, while there is only a slight improvement in the DSC metric, the model delivers a notable accuracy improvement of approximately 4% and 10% on the HD metric. Figure 5 shows the schematic for the proposed TransUNet model.

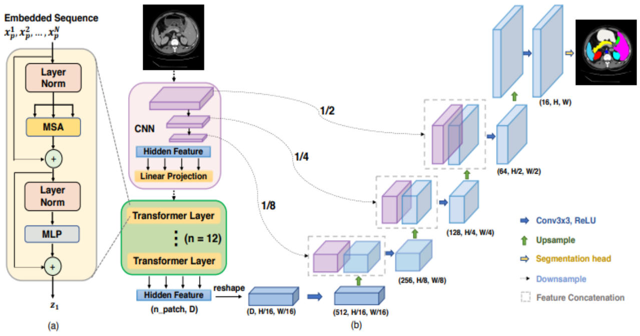

Figure 5 Architecture of the Proposed TransUNet.

The TransUNet model [9] achieves a Dice similarity coefficient of 77.48 and Hausdorff Distance of 31.69.

Conclusion

Medical image segmentation has evolved significantly, transitioning from classical methods like thresholding and region-based algorithms to advanced machine learning techniques. Deep learning models, particularly Convolutional Neural Networks (CNNs) and U-Net variants, now dominate, incorporating features like skip connections, attention mechanisms, and residual blocks for better performance. Recent innovations include transformer-based architectures and hybrid models that improve accuracy by integrating local and global features. However, challenges such as class imbalance and precise boundary detection persist. Future developments may focus on self-supervised learning and multi-modal data fusion, enhancing clinical decision-making and patient outcomes.

References

[1]. S. Subramaniam, K. B. Jayanthi, C. Rajasekaran, and R. Kuchelar. (2020). Deep Learning Architectures for Medical Image Segmentation, IEEE 33rd Int. Symp. Computer Based Medical Systems (CBMS), Rochester, MN, USA, pp. 579-584.

[2]. Metlek, S., & Çetıner, H. (2023). ResUNet+: A New Convolutional and Attention Block-Based Approach for Brain Tumor Segmentation. IEEE Access, 11, 69884-69902.

[3]. Sinha, A. R. P., Suresh, M., Mohan, N. R., D., A., & Singerji, A. G. (2021). Brain Tumour Detection Using Deep Learning. In Seventh International Conference on Bio Signals, Images, and Instrumentation (ICBSII), Chennai, India, pp. 15-20.

[4]. Gurbină, M., Lascu, M., & Lascu, D. (2019). Tumor Detection and Classification of MRI Brain Image Using Different Wavelet Transforms and Support Vector Machines. In 42nd International Conference on Telecommunications and Signal Processing (TSP), Budapest, Hungary, pp. 505-508.

[5]. Somasundaram, S., & Gobinath, R. (2019). Early Brain Tumour Prediction Using an Enhancement Feature Extraction Technique and Deep Neural Networks. International Journal of Innovative Technology and Exploring Engineering (IJITEE), 8(10S), 160-164.

[6]. Damodharan, S., & Raghavan, D. (2015). Combining Tissue Segmentation and Neural Network for Brain Tumor Detection. International Arab Journal of Information Technology, 12(1), 37-45.

[7]. Badrinarayanan, V., Kendall, A., & Cipolla, R. (2017). SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 39(12), 2481-249.

[8]. Chen, J., Lu, Y., Yu, Q., Luo, X., Adeli, E., Wang, Y., Lu, L., Yuille, A. L., & Zhou, Y. (2021). TransUNet: Transformers Make Strong Encoders for Medical Image Segmentation. arXiv:2102.04306

[9]. Cao, H., Wang, Y., Chen, J., Jiang, D., Zhang, X., Tian, Q., & Wang, M. (2022). SwinUNet: U-Net-like Pure Transformer for Medical Image Segmentation. In European Conference on Computer Vision, pp. 205-218.

[10]. Sivaguru, R., Saranya, K., Murugasan, J., & Kiruthika, E. (2024). Applications of Artificial Intelligence in the Medical Field: A Survey. International Research Journal on Advanced Engineering Hub (IRJAEH), 2(03), 407-415. https://doi.org/10.47392 /IRJAEH.2024.0060

[11]. Bokade, R., & Gharat, N. (2024). Implementation of FFT-Based Convolutional Neural Networks on Medical Imagery for Studying Performance Boost. International Research Journal on Advanced Engineering Hub (IRJAEH), 2(03), 508-513. https://doi.org/10.47392/IRJAEH.2024.0073

[12]. Shinde, S., Chakurkar, P., & Rane, R. (2024). Deep Learning-Based Small Face Detection from Hard Image. International Research Journal on Advanced Engineering Hub (IRJAEH), 2(03), 579-588. https://doi.org/10.47392/IRJAEH.2024.0084