Deep Learning Applications in Stroke Segmentation: Progress, Challenges, and Future Prospects

Xingyi Rong

Department of Computing, Harbin Institute of Technology, Harbin, 150001, China
2021111665@stu.hit.edu.cn

Abstract. Stroke is a major global health challenge, significantly contributing to disability and death worldwide. Due to the rapid progress of deep learning, the challenges in this field have the potential to be solved. This article offers a comprehensive examination of the uses of deep learning in stroke segmentation. It specifically highlights advanced models like U-Net, RCNN, and their variations. These techniques employ Convolutional Neural Networks (CNNs) to accurately and efficiently segment stroke lesions in medical pictures by classifying each pixel. Research findings suggest that there has been substantial improvement in both precision and speed, providing swift and precise outcomes for stroke detection. Nevertheless, the study also uncovers the difficulties and constraints of current techniques in managing intricate lesions, instantaneous applications, generalization capability, and model interpretability. Subsequent investigations should focus on overcoming these obstacles by investigating intricate network structures, refining optimization methods, developing streamlined models, and implementing visualization explanation approaches. This will significantly enhance the progress of stroke segmentation technology based on deep learning.

Keywords: Deep learning, stroke segmentation, artificial intelligence

1 Introduction

Stroke, being a prominent contributor to both disability and mortality on a global scale, presents substantial obstacles to public health [1-3]. A cerebrovascular accident, also known as a stroke, occurs when there is an interruption or significant reduction in blood flow to a specific area of the brain. This leads to a deprivation of vital nutrients and oxygen to the brain tissue. This can result in the death of brain cells, which can potentially result in long-term neurological harm, impairment, or even fatality. Strokes are commonly the result of either a constriction (ischemic stroke) or rupture (hemorrhagic stroke) of the blood arteries in the brain. Precise and prompt identification of stroke is essential for good treatment and prognosis. Stroke segmentation, in this particular context, is a crucial task that entails identifying and outlining the brain regions that have been impacted by the stroke. Nevertheless, conventional segmentation techniques frequently depend on manual interventions, which not only consume a significant amount of time but also are susceptible to subjectivity and

Y. Wang (ed.), *Proceedings of the 2024 2nd International Conference on Image, Algorithms and Artificial Intelligence (ICIAAI 2024)*, Advances in Computer Science Research 115,
https://doi.org/10.2991/978-94-6463-540-9_59

inconsistency. Hence, it is vital to contemplate more sophisticated methodologies in this particular scenario.

Deep learning has significantly transformed the field of medical image analysis [4, 5], particularly the process of stroke segmentation. Deep learning algorithms have shown exceptional performance in several medical imaging applications due to their capacity to learn intricate patterns from extensive datasets. Deep learning approaches have proven to be highly accurate and efficient in automatically identifying and outlining the areas of tissue damage caused by stroke in the field of stroke segmentation [6]. Xu G, Zhang X, He X, et al. have proposed a novel lightweight architecture called LeViT-UNet, which integrates multi-stage Transformer blocks in the encoder via LeViT to explore the effectiveness of the fusion between local and global features [7]. Yuqi Zhang et al. introduce CA-UNet, a model for segmenting 3D carotid computed tomography angiography (CTA) images. The purpose of this model is to extract carotid arteries in a fully automated manner [8]. Sercan Yalçın et al. introduced a convolutional deep network structure called the optimized dimensional U-Net (D-UNet). This architecture involves blocking and adaptively sequencing the convolution layers, as well as optimizing the amount of activation functions and hyperparameters.

The incorporation of deep learning into stroke segmentation is not just a technological innovation but also fulfills an essential requirement in clinical practice. With the increasing size and intricacy of medical images, there is a pressing need for precise and automated segmentation techniques to assist in diagnosis and treatment planning. Although there has been notable advancement in this area, there are still obstacles that need to be addressed, including strengthening the accuracy of segmentation, reducing the complexity of computations, and improving the models' ability to generalize.

This article aims to examine the practical uses of deep learning in the process of stroke segmentation. This study critically examined the most advanced deep learning models and analyzed their performances, constraints, and potential future developments. This paper seeks to offer a thorough and detailed examination of the topic, emphasizing the capacity of deep learning to fundamentally transform the diagnosis and prognosis of strokes. By amalgamating knowledge from recent studies, there is an optimistic expectation of facilitating progress in this crucial domain of medical imaging.

2 Method

2.1 The Introduction of Semantic Segmentation

Semantic segmentation is a sophisticated image processing approach that entails categorizing and assigning labels to each pixel in an image according to the specific item category they pertain to. Within the realm of stroke segmentation, semantic segmentation is of utmost importance as it accurately identifies and delineates the borders of stroke lesions. The process of semantic segmentation comprises multiple essential stages. At first, the system receives medical images, such as computed tomography (CT) or magnetic resonance imaging (MRI) scans. Advanced deep learning models, typically utilizing convolutional neural networks (CNNs), are subsequently employed to methodically classify each pixel of the image. The result is an accurately annotated image in which each pixel is designated as belonging to a certain tissue type or lesion, such as a region afflicted by stroke or healthy brain tissue.

2.2 Deep Learning-based Stroke Segmentation

U-Net. Yuqi Zhang and colleagues introduced CA-UNET [9], which can be described as follows: 1) Encoder-Decoder Structure: Similar to the original U-Net, CA-UNet utilizes an encoder-decoder structure. The encoder path acquires contextual information by reducing the resolution of the input image via a sequence of convolutional layers and pooling operations. The decoder path progressively enhances the spatial resolution by upsampling the feature maps and merging them with the matching feature maps from the encoder path. 2) Revised Downsampling Scheme: The authors of the article suggest decreasing the quantity of downsampling layers in comparison to the original U-Net. The reason for avoiding excessive downsampling in carotid artery segmentation is that it can result in the loss of crucial features necessary for accurate segmentation. Through the elimination of certain downsampling layers, CA-UNet can retain a greater spatial resolution, hence safeguarding the intricate intricacies of the carotid arteries. 3) Skip Connections: CA-UNet incorporates skip connections that connect corresponding levels in the encoder and decoder routes. The connections in question merge low-level feature maps from the encoder with upsampled feature maps from the decoder. This enables the network to retain spatial information from previous levels and integrate it into the segmentation output. 4) The authors offer a multi-scale loss function to tackle the problem of imbalanced positive and negative data in carotid artery segmentation. This loss function calculates the segmentation loss at several scales, which correspond to the different layers of the decoder route. The ultimate loss is a weighted amalgamation of the losses from all scales, which aids in equalizing the learning process and enhancing the segmentation performance. 5) Optimized Model Parameters: CA-UNet achieves parameter optimization by eliminating redundant downsampling layers and integrating skip connections, resulting in a reduced parameter count compared to other 3D segmentation networks. This aids in expediting the training procedure and mitigates the potential for overfitting.

RCNN. Sangeeta Rani et al. suggested various techniques, including RCNN (region-based convolutional neural network) [10], Fast R-CNN (fast region-based convolutional neural network), Faster R-CNN (faster region-based convolutional neural network with region proposal network), YOLO (you only look once), SSD (single-shot multi-box detector), and Efficient-Det, for stroke localization and classification. The paper also includes comparisons of RCNN, Fast R-CNN, Faster R-CNN, YOLO, SSD, and Efficient-Det in terms of accuracy.

The RCNN technique was proposed as a method that utilizes convolutional neural networks to enhance the accuracy of object detection in photographs. The process depends on three primary phases: 1) Region Proposal: During this stage, the input image is initially partitioned into multiple regions. Subsequently, a discerning search technique is employed to combine comparable regions according to their resemblance, so generating clusters of regions that are probable to encompass items. This procedure is iterated until potential entities are identified and bounding boxes are defined around them. 2) Feature Extraction: The regions that have been discovered and enclosed in bounding boxes are subsequently extracted from the original image. These regions are then enlarged to a standardized dimension that is suitable for input into a Convolutional Neural Network (CNN). CNN obtains valuable characteristics from these chopped sections. To ascertain whether a certain area is part of the foreground (which includes an object) or the background, a metric called Intersection over Union (IOU) is computed by comparing the predicted bounding box with the actual ground truth box. A value of IOU that is greater than or equal to 0.5 often signifies a foreground region. 3) Classification: The characteristics extracted from the foreground regions are inputted into a classifier, such as a Support Vector Machine (SVM), to determine the class label of the identified object. This approach achieves comprehensive object detection by accurately determining the precise location and accurately recognizing the things depicted in the image.

2.3 Evaluation Metrics

Table 1. The different metrics and corresponding definitions.

Evaluation metrics	Definition
Dice Coefficient (Dice Similarity Index)	Measures the overlap between the predicted segmentation and the ground truth segmentation. It is calculated as twice the area of the intersection divided by the sum of the areas of the predicted and ground truth segmentations.
Intersection over Union (IoU)	Also known as the Jaccard Index, it measures the overlap between the predicted and ground truth segmentations. It is calculated as the area of intersection divided by the area of union between the predicted and ground truth segmentations.

Accuracy	Measures the overall proportion of correctly classified pixels (stroke vs. non-stroke). However, for imbalanced datasets (where stroke pixels are significantly fewer than non-stroke pixels), accuracy alone can be misleading.
Sensitivity (Recall)	Measures the proportion of actual stroke pixels that are correctly predicted as stroke. High sensitivity indicates that most actual stroke pixels are detected.

3 Discussion

The introduction of deep learning algorithms has greatly progressed the field of stroke segmentation. Section 2 has provided evidence that deep learning models, including U-Net, have achieved impressive outcomes in accurately and efficiently segmenting stroke lesions. Out of these models, the RCNN method stood out as a significant approach, representing a big advancement in object detection. The integration of region suggestions with CNNs was effectively demonstrated, highlighting the power of this combination. However, this approach has significant drawbacks, particularly the computational inefficiencies caused by individually processing each area suggestion through the CNN. This limitation led to the emergence of more advanced iterations, including Fast R-CNN and Faster R-CNN. These advances aimed to address the inefficiencies by implementing the sharing of computational tasks and the introduction of region proposal networks (RPNs). Notwithstanding these improvements, there are still various obstacles and restrictions that must be resolved to achieve further advancement.

Although the evaluated models have attained excellent performance, there is still potential for enhancing the accuracy of segmentation, particularly for complicated and diverse stroke lesions. Managing fluctuations in stroke dimensions, form, and placement continues to be a formidable task. Further investigation could focus on developing more advanced network topologies and optimization strategies to improve the model's ability to distinguish between different classes.

In addition, the computational intricacy of deep learning models may impede their use in real-time applications. The evaluated models, particularly those based on 3D convolutional neural networks (CNNs), frequently necessitate substantial processing resources, which may be impractical in contexts with limited resources. Hence, it is imperative to create models that are both lightweight and economical, capable of achieving high segmentation accuracy while minimizing computing demands.

Furthermore, the capacity of deep learning models to generalize to unfamiliar data continues to be a challenge. The efficacy of these models is highly dependent on the caliber and variety of the training datasets. Nevertheless, the task of gathering extensive and varied stroke datasets with precise annotations can present difficulties and consume a significant amount of time. Subsequent investigations may go into methodologies such as data augmentation, transfer learning, and domain adaptation to enhance the resilience and generalization capacities of the models.

Plus, the comprehensibility of deep learning models is an additional domain that requires additional investigation. Although these models can attain impressive performance, their decision-making process is frequently obscure and challenging to explain. The absence of interpretability can impede their acceptance in clinical treatment, where trust and transparency are essential. Hence, it would be advantageous to devise methods for visualizing and elucidating the predictions made by the models.

To summarize, although deep learning techniques like RCNN and their variations have achieved substantial gains in stroke segmentation, there are still areas that can be enhanced, including segmentation accuracy, computational efficiency, generalization ability, and interpretability. Subsequent investigations should strive to tackle these obstacles and expand the limits of stroke segmentation using deep learning.

4 Conclusion

The models examined in this paper, such as U-Net, and RCNN, and their variations, have shown exceptional precision and effectiveness in autonomously segmenting stroke lesions from medical images. These procedures have the capacity to transform stroke diagnosis by offering quicker and more precise outcomes in contrast to conventional manual segmentation methods. Nevertheless, it is imperative to recognize the difficulties and constraints that come with using deep learning for stroke segmentation. The objectives encompass enhancing the accuracy of segmenting intricate lesions, minimizing the computational cost for applications requiring real-time processing, augmenting the ability to generalize to unfamiliar data, and refining the interpretability of the model. Subsequent investigations should focus on tackling these obstacles and pushing the boundaries of stroke segmentation.

References

1. Williams, L.S., Bruno, A., Rouch, D., Marriott, D.J., MAS: Stroke patients' knowledge of stroke: influence on time to presentation. Stroke 28(5), 912-915 (1997).
2. BSc, T.B., Grypdonck, M.: Being a stroke patient: a review of the literature. Journal of Advanced Nursing 26(3), 580-588 (1997).
3. Jaracz, K., Kozubski, W.: Quality of life in stroke patients. Acta Neurologica Scandinavica 107(5), 324-329 (2003).

4. Qiu, Y., Wang, J., Jin, Z., Chen, H., Zhang, M., Guo, L.: Pose-guided matching based on deep learning for assessing quality of action on rehabilitation training. Biomedical Signal Processing and Control 72, 103323 (2022).
5. Kayalibay, B., Jensen, G., van der Smagt, P.: CNN-based segmentation of medical imaging data. arXiv preprint arXiv:1701.03056 (2017).
6. Xu, G., Zhang, X., He, X., et al.: Levit-unet: Make faster encoders with transformer for medical image segmentation. In: Chinese Conference on Pattern Recognition and Computer Vision (PRCV), pp. 42-53. Springer Nature Singapore, Singapore (2023).
7. Zhang, Y., Yu, M., Tong, C., et al.: CA-UNet Segmentation Makes a Good Ischemic Stroke Risk Prediction. Interdisciplinary Sciences: Computational Life Sciences 16(1), 58-72 (2024).
8. Yalçın, S., Vural, H.: Brain stroke classification and segmentation using encoder-decoder based deep convolutional neural networks. Computers in Biology and Medicine 149, 105941 (2022).
9. Nazari-Farsani, S., Yu, Y., Armindo, R.D., et al.: Predicting final ischemic stroke lesions from initial diffusion-weighted images using a deep neural network. NeuroImage: Clinical 37, 103278 (2023).
10. Rani, S., Singh, B.K., Koundal, D., et al.: Localization of stroke lesion in MRI images using object detection techniques: a comprehensive review. Neuroscience Informatics 2(3), 100070 (2022).