CSIAM Trans. Appl. Math.
doi: 10.4208/csiam-am.SO-2024-0043

Vol. **7**, No. 1, pp. 178-206
March 2026

Double-Domain Driven Unet for Selective Segmentation of Medical Image

Zhi-Feng Pang[1], Lin Yang[1], Mingxiang Wu[1], Ziyu Niu[1], Raymond Chan[2] and Xue-Cheng Tai[3,*]

[1] *College of Mathematics and Statistics, Henan University, Kaifeng 475004, China.*
[2] *School of Data Science, Lingnan University, Tuen Mun 999077, Hong Kong, SAR, China.*
[3] *Norwegian Research Centre, Nygardstangen, Bergen NO-5838, Norway.*

Received 9 September 2024; Accepted 13 June 2025

Abstract. Selective segmentation has been a notable increase in interest in the use of interactive deep learning-based methods. However, developing an efficient and accurate segmentation method for medical applications remains a formidable challenge, primarily due to the inherent complexity and diversity of medical image structures. To address these challenges, we introduce a novel deep learning approach, termed the double-domain driven Unet method (DDUM), designed for the selective segmentation of medical images. Our approach utilizes a threshold geodesic distance in conjunction with the original images as input to construct a parallel Unet architecture that captures information from both the image domain and the geodesic distance domain. To further enhance the accuracy and efficacy of the segmentation process, we employ soft threshold dynamics as a replacement for the sigmoid activation function in the final layer. The efficacy of our proposed DDUM is substantiated through extensive experiments conducted on multiple medical image datasets. In particular, the DDUM method exhibits exceptional performance in terms of both segmentation accuracy and robustness.

AMS subject classifications: 68U10, 65K05

Key words: Image segmentation, regions of interest, threshold geodesic distance, double-domain driven Unet method, soft threshold dynamics.

1 Introduction

Image segmentation plays a critical component in medical imaging, enabling precise diagnosis and treatment planning through the delineation of anatomical structures or

*Corresponding author. *Email addresses:* zhifengpang@henu.edu.cn (Z.-F. Pang), yanglin_henu@163.com (L. Yang), mingxiangwu2022@163.com (M. Wu), niuziyu2892@163.com (Z. Niu), raymond.chan@ln.edu.hk (R. Chan), xtai@norceresearch.no (X.-C. Tai)

pathological regions. Current segmentation methodologies can be broadly categorized into two principal approaches: semantic segmentation [10,25,35,37,56] and selective segmentation [13, 38, 47]. Semantic segmentation involves the comprehensive separation of all foreground objects from the background within an image, while selective segmentation focuses exclusively on isolating specific subsets of foreground objects or regions of interest (ROI). Despite substantial progress in selective segmentation techniques in recent decades, medical image analysis continues to face significant challenges. These include the presence of noise, imaging artifacts, low contrast variability, and substantial inter-patient anatomical differences, all of which contribute to inconsistent segmentation accuracy, particularly in cases involving complex anatomical variations and intricate pathological presentations [44, 61]. Consequently, we focus here emphasizes the advancement of selective segmentation techniques for medical imaging, aiming to develop methods that deliver clinically applicable levels of accuracy and robustness.

In general, the selective segmentation method uses user annotations to guide the segmentation process, resulting in a faster and more accurate segmentation result for the ROI [4]. Traditional selective segmentation methods, such as graph cut-based methods [2,5,42,60], grab cut-based methods [20,24], random walk-based methods [17], geodesic-based methods [3, 9], and grow cut-based methods [51], typically rely on shallow image features and leverage annotation information to optimize an energy functional. However, such shallow features often fail to adequately capture the rich contextual information of the image, particularly in scenarios involving low contrast or significant noise. To address these limitations, recent advancements in active contour models [14] have emphasized the integration of prior annotation information, such as convexity priors [6, 29] or distance-based constraints [16, 38], to enhance segmentation accuracy and robustness. For example, Nguyen *et al.* [34] integrated the convex active contour [6] with the annotation information, computed the geodesic map to obtain the initial contour and achieved fast segmentation. Gout *et al.* [16] used the distance of some marked points as weights for the edge function and constructed an energy functional based on the level set method. Spencer and Chen [47] incorporated a regularization term based on Euclidean distance into the original active contour model for selective segmentation. Roberts *et al.* [38] proposed replacing the Euclidean distance with an edge-weighted geodesic distance, which adaptively increases near image edges, thereby offering enhanced suitability for selective segmentation tasks. Despite these advancements, achieving accurate segmentation remains heavily dependent on extensive user input and manually engineered features, limiting their scalability and practical applicability in complex scenarios.

In recent years, convolutional neural networks (CNNs) have demonstrated the capacity to achieve state-of-the-art performance in image segmentation, largely due to their ability to automatically learn advanced semantic features (for further details, please see references [1, 11, 28, 45]). The Unet architecture [39], with its distinctive encoder-decoder framework, has gained widespread adoption in medical image segmentation tasks. Nevertheless, a significant challenge persists in this domain: the extensive variability in the shapes and sizes of segmented objects inherently limits the network's ability to effectively