MSA^{2}Net: Multi-scale Adaptive Attention-guided Network for Medical Image Segmentation

Sina Ghorbani Kolahi[1]
sina.ghorbani@modares.ac.ir
Seyed Kamal Chaharsooghi[1]
skch@modares.ac.ir
Toktam Khatibi[1]
toktam.khatibi@modares.ac.ir
Afshin Bozorgpour[2]
afshin.bozorgpour@ur.de
Reza Azad [2,3]
reza.azad@rwth-aachen.de
Moein Heidari[4]
moein.heidari@ubc.ca
Iiker Hacihaliloglu[4]
iiker.hacihaliloglu@ubc.ca
Dorit Merhof [2,5]
dorit.merhof@ur.de

[1] Tarbiat Modares University
Tehran, Iran
[2] University of Regensburg
Regensburg, Germany
[3] RWTH Aachen University
Aachen, Germany
[4] University of British Columbia
British Columbia, Canada
[5] Fraunhofer Institute for Digital Medicine MEVIS
Bremen, Germany

Abstract

Medical image segmentation involves identifying and separating object instances in a medical image to delineate various tissues and structures, a task complicated by the significant variations in size, shape, and density of these features. Convolutional neural networks (CNNs) have traditionally been used for this task but have limitations in capturing long-range dependencies. Transformers, equipped with self-attention mechanisms, aim to address this problem. However, in medical image segmentation it is beneficial to merge both local and global features to effectively integrate feature maps across various scales, capturing both detailed features and broader semantic elements for dealing with variations in structures. In this paper, we introduce MSA^{2}Net, a new deep segmentation framework featuring an expedient design of skip-connections. These connections facilitate feature fusion by dynamically weighting and combining coarse-grained encoder features with fine-grained decoder feature maps. Specifically, we propose a Multi-Scale Adaptive Spatial Attention Gate (MASAG), which dynamically adjusts the receptive field (Local and Global contextual information) to ensure that spatially relevant features are selectively highlighted while minimizing background distractions. Extensive evaluations involving dermatology, and radiological datasets demonstrate that our MSA^{2}Net outperforms state-of-the-art (SOTA) works or matches their performance. The source code is publicly available at *https://github.com/xmindflow/MSA-2Net*.