Image Segmentation using U-Net: A Polymorphic Architecture

Vandana Malik
Department of Computer Science,
Himachal Pradesh University
Shimla, India
panwarv525@gmail.com

A.J Singh
Department of Computer Science,
Himachal Pradesh University
Shimla, India
aj.hpucs@gmail.com

***Abstract* -** Image segmentation is an essential application of computer vision for generating meaningful segments and areas of interest for detailed image analysis. Medical service providers get assistance from AI embedded technologies for precise and accurate diagnosis of cancer cells, tumors and lesions. Medical image segmentation also plays a vital role in disease detection, prediction along with analysis and evaluation of patient's health condition. Healthcare providers can make better decision and enhance patient's outcomes by minimizing human error and improving human decision making. Electron Microscopy deals with complex biological images of cell organelles in which thin slices of images are generated for detailed analysis. In this paper electron microscopy images of mitochondria are used to implement a simple U-Net to understand its structure and workings. U-Net architecture is abundantly used for medical image segmentation. U-Net architecture is based upon Convolutional Neural Network (CNN), it is an arrangement of encoder and decoder in U shape. U-Net architecture is versatile in nature therefore it has been modified into the EM-Net, MitoNet, MLU-Net, G-Net, and LHU-Net architectures. This article signifies major drawbacks of medical image segmentation and how U-Net architecture is modified in order to overcome these drawbacks. Therefore U-Net is being called as polymorphic architecture in this article because this single architecture exists in multiple forms and used for multiple purposes.

Keywords - Image Segmentation, Electron microscopy, U-Net

1. INTRODUCTION

Image Segmentation is a sub field of Computer Vision where an image is divided into segments and regions. Image segmentation is used to get a fine grain understanding of each pixel in an image which helps in image analysis. Image segmentation helps in object detection, classification task and in identifying region of interest (ROI) (Yu, 2023). Image segmentation is used in various fields like medical imaging, satellite imaging, augmented reality, autonomous vehicles and object recognition (Manakitsa, 2024). Image segmentation uses techniques like K-Mean Clustering for clubbing pixels with same intensity together; in edge based segmentation technique boundaries of objects are detected for image segmentation. Region based and graph based techniques is also used for image segmentation (Aggarwal, 2024). Deep Learning based neural networks like Fully Convolution Network, U-Net and Mask R-CNN are mostly used for accurate segmentation (Liu, A review of deep-learning-based medical image segmentation methods, 2021). Image segmentation is of two type semantic segmentation and instance segmentation. In semantic segmentation different objects in the image is represented with different color and every pixel is highlighted and represent them as foreground and background image. Instance segmentation on the other hand tries to analyze the image and count the number of objects or instances present in the image of same class (Malhotra, Retracted] Deep Neural Networks for Medical Image Segmentation, 2022). This article aims at helping researches to know about image segmentation, its limitations and how U-Net architecture is modified in multiple ways to overcome these problems. Polymorphism means many forms and U-Net architecture changes its form as per the need therefore called as Polymorphic architecture. In this article U-Net architecture, electron microscopy, problems faced during image segmentation, research methodology for implementing U-Net architecture are highlighted. In second section literature review is included followed by section three which talks about research methodology. Section four discusses results and analysis. Fifth section highlights future scope of the study and finally in sixth section conclusion is given.

2. Review of Literature

In this article U-Net algorithm and its variations are discussed. U-Net architecture is a convolution neural network (CNN) designed in U shape with encoder and decoder network. U-Net is mostly used in medical imaging tasks because of high accuracy in segmentation and provides spatial information. U-Net is a versatile CNN architecture segmenting different medical images. Cell organelles can be segmented for nucleus, mitochondria endoplasmic reticulum, Golgi apparatus, lysosomes, plasma membrane, ribosome, chloroplast, centrosome and vacuoles. In this article variations of U-Net are discussed and segmentation of electron microscopy images is done (Conrad, 2023). Organs like brain can be segmented for white matter, gray matter, tumors cerebrospinal fluid and lesions. Heart can be segmented for left and right ventricle, myocardium. Similarly lungs can be segmented for lung lobes, tumors and nodules. Liver segmentation can be done for liver parenchyma, lesions and tumors. Kidney cysts, tumors and pancreatic tissue can also be segmented. Pancreas, prostate, breast, bones, eyes, bladder, uterus and ovaries and skin can be analyzed for tumors, lesions and tissues with the help of U-Net architecture (Liu, A review of deep-learning-based medical image segmentation methods, 2021).

Biomedical images are segmented with the help of the U-NET architecture, a convolutional neural network. It has a unique U-shaped structure representing the encoder and decoder networks. Input images at the encoder end are compressed into lower dimensions for extracting essential features. The decoder uncompressed the segmented image in order to maintain the spatial information. High accuracy is obtained by utilizing skip connections to limit the training data. Lower-level information is combined with higher-level information by connecting the encoder layer to the decoder layer with the help of a skip connection, which results in precise segmentation as output. U-Net has become the epitome of medical image segmentation as it generates excellent results (Siddique, 2021). Researchers and practitioners have enhanced the U-Net architecture by introducing various modifications and applying them in different ways, thereby improving its performance and achieving more precise and accurate image segmentation. This article explores the different types of U-Nets and the specific problems they are addressing.

Electron microscopy (EM) is a highly detailed biological image produced with the help of rays of electrons at a nanometer scale. Electron microscopy is of two types: Transmission electron microscopy (TEM): thin slices of high-resolution images of the sample are generated. In scanning electron microscopy (SEM), a forced electron beam is used to scan a sample for the detailed surface of an image. Mitochondrial segmentation of electron microscopy images is performed by using a fully convolutional neural network, which results in a rough segmentation that is not so smooth with discontinuities and false positives. Large memory is required for training a model and annotating a large volume of 3D image data. To overcome this limitation, a multi-task network, EM-Net, was introduced to identify the centerline and limit the annotation data to improve segmentation accuracy and robustness. A novel hierarchical view ensemble convolutional network reduced the number of parameters used for training and led to a reduction in computational cost (Conrad, 2023; Yuan, 2020).

Manual annotation of medical images is a laborious task in image segmentation. There is a need for deep learning models for accurate and precise identification and annotation of mitochondria in electron microscopy images. As mitochondria have inconsistent shape, size, context, and quality, it's difficult to develop a robust model for accurate segmentation. MitoNet is a deep learning model trained on multifarious 1.5 million electron microscopic images, of which 135,000 are of segmented mitochondria. Python libraries like Empanda and the Napari plugin help MitoNet inference, clean up, and refine the segmentation process. Zooniverse was used for crowd annotation, which improved the quality of the annotation, and consensuses were made on the instance mask. The CEM1.5M and CEM-MitoLab datasets were created for robust model training (Conrad, 2023).

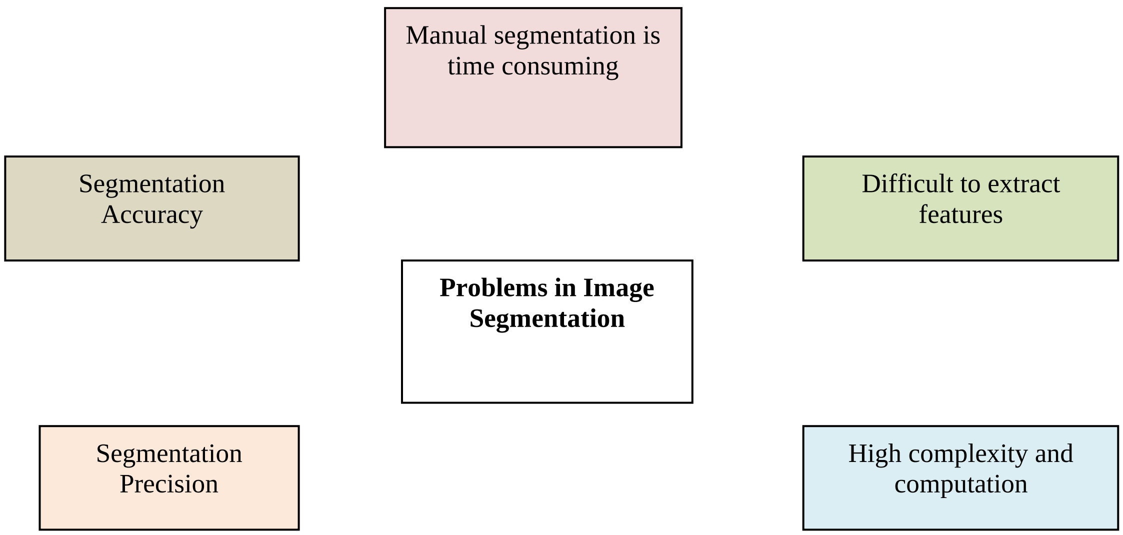

Fig.1. Problems in Image Segmentation

As manual segmentation is a time-consuming and tedious task, using supervised, unsupervised, and self-supervised learning can help in segmenting cellular and sub-cellular electron microscopic images and overcoming heterogeneity and spatial complexity (Aswath, 2023). It is difficult to extract feature information during intensive structure segmentation. To prevent loss of features, a frequency representation approach is incorporated during the encoding and decoding phases. Tokenized multilayer perceptron (MLP) gathers space information about a particular thing from deep core intensive structure segmentation like brain tumor segmentation. For this purpose, MLU-Net was proposed. It is a type of U-Net architecture integrated with a frequency representation approach and a tokenized multiplayer perceptron (Feng, 2024). Preprocessing and postprocessing of data are crucial for developing an accurate and precise image segmentation model. Generalized U-Net (G-Net) is an integration of deep learning architectures like CNN, ResNet, and DenseNet created for liver tumor segmentation. Preprocessing techniques like Hounsfield unit, windowing, histogram equalization, and postprocessing methods like conditional random field help improve segmentation accuracy by 3.35% (DJ, 2024). Transformer architectures are incorporated into medical image analysis, which leads to an increase in the complexity and computational requirements of hybrid models. Segmentation is negatively impacted by the failure of models to account for the interaction between spatial and channel features. Existing models are unable to effectively represent organs in medical image segmentation, which ultimately results in fragmented and scattered segmentations. LHU-Net is a Light Hybrid U-Net architecture that is modified to give priority to spatial feature analysis in the initial layers and channel-based features in the deeper layers. Spatial attention is integrated for extracting more features, and channel attention is integrated for feature refinement. LHU-Net performance is highly accurate with a limited number of computational parameters. A dice score of 92.66 is achieved on the ACDC dataset, reducing parameters by 85%. (Sadegheih, 2024).

Fig. 2. Variations in a U-Net Architecture

Electron microscopy images are available in datasets like EPFL and it is utilized in making EM-Net it is beneficial for testing and validating the ensemble convolutional network (Yuan, 2020). CEM1.5 and CEM-Mito-Lab dataset are created by using 1.5 million unlabeled and diverse EM images this dataset was used by Mito-Net (Conrad, 2023). While conducting study about cellular electron microscopy SNEM13D dataset, Drosophila larvae dataset and CREMI3D dataset were used with deep learning (Aswath, 2023). MICCAI Multi-Atlas Abdomen Labeling Challenge (Synapse), ACDC Automated Cardiac Diagnosis Challenge (Synapse), Left Atrial Dataset and BraTS-2018 dataset contains MRI and CT scans of brain and abdominal organs (Sadegheih, 2024).

3. METHODOLOGY

Image segmentation of medical images is a beneficial task for medical practitioners as it helps in the diagnosis of abnormal growth of cells, tumors, and lesions. Cell organelles can also be segmented, like mitochondria, which will be carried out in this article. There are multiple preprocessing steps to be performed before image segmentation. In this experiment, a dataset was downloaded from a website containing a dataset of electron microscopic images: https://www.epfl.ch/labs/cvlab/data/data-em/. Images in this dataset were in .tiff format, so firstly, they were converted into .jpg format. The original dimensions of the images were 1024 x 768, where the width was 1024 pixels and the height was 768 pixels. Image size is quite high, which increases computation time, so images were resized into 256*256 dimensions and made smaller. The U-Net architecture is a CNN-based model with an encoder and decoder at its core. Electron microscopic images of mitochondria are provided as training, testing, and ground truth images on which experiments are conducted. Experimentation is conducted on Google Collab platform. (https://colab.research.google.com/drive/1anFup9RRJXtULkjTeTpdhNoe6juHem_C).

4. RESULTS & ANALYSIS

After training the model, we get total parameters of 1940817, trainable parameters of 1940817, and zero non-trainable parameters. The model was trained for 3 epochs with a validation accuracy of 0.9676. Results obtained clearly marks the mitochondria with white color and background as black color. The following image is from the dataset displaying segmented mitochondria.

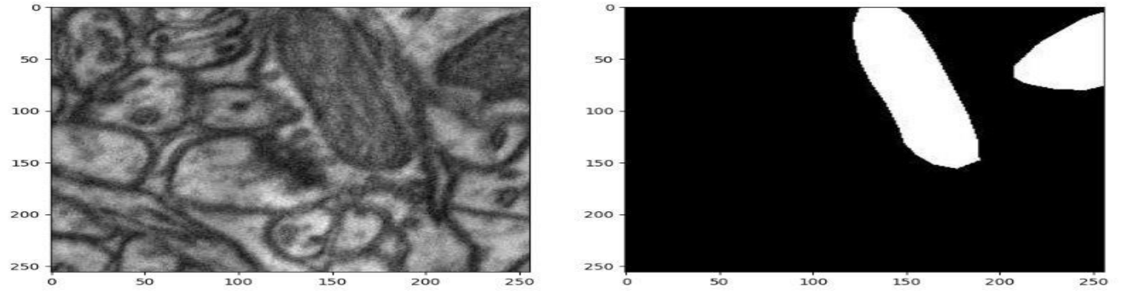

Fig.3. Mitochondria segmentation (https://colab.research.google.com/drive/1anFup9RRJXtULkjTeTpdhNoe6juHem_C)

5. FUTURE SCOPE OF THE STUDY

Image segmentation on medical images using different architectures is being applied for analyzing images to help in the early diagnosis of diseases. There is a requirement of robust models performing accurate and precise image segmentation with less complexity and computation time. Researchers can develop preprocessing and postprocessing techniques for improving image segmentation tasks. (Aswath, 2023) (Feng, 2024)

6. REFERENCES

Aggarwal, A. a. (2024). Unsupervised learning-based approach for detecting 3D edges in depth maps. Scientific Reports , 14, 796. https://doi.org/10.1038/s41598-023-50899-3

Aswath, A. a. (2023). Segmentation in large-scale cellular electron microscopy with deep learning: A literature survey. {Medical image analysis , 102920. https://doi.org/10.1016/j.media.2023.102920

Conrad, R. a. (2023). Instance segmentation of mitochondria in electron microscopy images with a generalist deep learning model trained on a diverse dataset. Cell Systems , 14, 58--71. DOI:https://doi.org/10.1016/j.cels.2022.12.006

DJ, D. a. (2024). Liver tumor segmentation using G-Unet and the impact of preprocessing and postprocessing methods. Multimedia Tools and Applications , 1--29. https://doi.org/10.1007/s11042-024-18759-y

Feng, L. a. (2024). MLU-Net: A Multi-Level Lightweight U-Net for Medical Image Segmentation Integrating Frequency Representation and MLP-Based Methods. IEEE Access . doi: 10.1109/ACCESS.2024.3360889.

Liu, X. a. (2021). A review of deep-learning-based medical image segmentation methods. Sustainability , 13, 1224. https://doi.org/10.3390/su13031224

Malhotra, P. a. (2022). Retracted] Deep Neural Networks for Medical Image Segmentation. Journal of Healthcare Engineering , 9580991. https://doi.org/10.1155/2023/9781975

Manakitsa, N. a. (2024). A review of machine learning and deep learning for object detection, semantic segmentation, and human action recognition in machine and robotic vision. Technologies , 12, 15. https://doi.org/10.3390/technologies12020015

Sadegheih, Y. a. (2024). LHU-Net: A Light Hybrid U-Net for Cost-Efficient, High-Performance Volumetric Medical Image Segmentation. arXiv preprint arXiv:2404.05102 . https://doi.org/10.48550/arXiv.2404.05102

Siddique, N. a. (2021). U-net and its variants for medical image segmentation: A review of theory and applications. Ieee Access , 9, 82031--82057. doi: 10.1109/ACCESS.2021.3086020

Yu, Y. a. (2023). Techniques and challenges of image segmentation: A review. Electronics , 12, 1199. https://doi.org/10.3390/electronics12051199

Yuan, Z. a. (2020). Em-net: Centerline-aware mitochondria segmentation in em images via hierarchical view-ensemble convolutional network. In 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI) (pp. 1219--1222). IEEE. doi: 10.1109/ISBI45749.2020.9098328.