Modified Double U-Net Architecture for Medical Image Segmentation

Sagar Deep Deb [1] Rajib Kumar Jha [1]

[1]Indian Institute of Technology Patna, Bihar

Introduction

- Semantic segmentation is the process of labelling each pixel of an image with its appropriate class. In biomedical image processing, it is a critical pre-processing step [1] [2].
- The goal of image segmentation is to obtain a set of regions that collectively cover the entire image and accurately represent the structure of interest [2].
- There are various methods for medical image segmentation, including thresholding, clustering, region growing, and deep learning. These methods can be used to segment different types of medical images, such as CT, MRI, and ultrasound. The segmented images can then be used for diagnosis, treatment planning, and image-guided surgery.
- After the evolution of deep learning, specifically, the U-Net, which is based on the encoder-decoder model, researchers have extensively used several modified versions of U-Net for semantic segmentation [3].

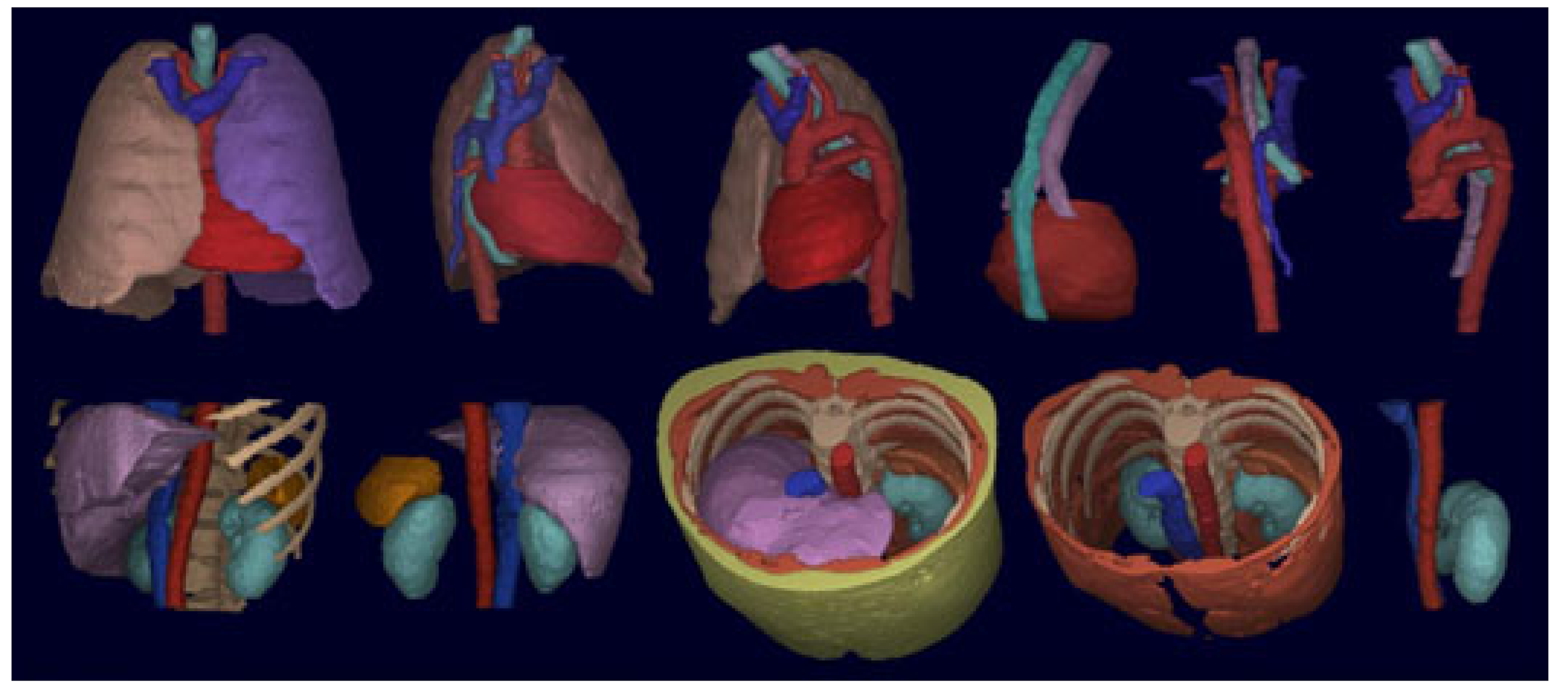

Figure 1. A typical Bio-medical Image Segmentation

- Deep learning-based biomedical image segmentation has shown to be a powerful and promising technique that can improve diagnostic accuracy and treatment outcomes for patients.

Motivation

- **Improved diagnosis**: Accurate segmentation of medical images can aid in the diagnosis of diseases and conditions by highlighting relevant structures and regions in the image [1].
- **Treatment planning**: Segmentation can be used to identify and isolate specific structures or regions of interest in an image, which can be used to plan and guide surgical procedures and other treatments [3].
- **Image-guided surgery**: Segmentation can be used to create 3D models of structures and regions in the body, which can be used to guide and assist in minimally invasive and robotic surgical procedures.
- **Quantitative analysis**: Automated segmentation can help in the quantification of the region of interest and can be used for the measurement of lesion volume, perfusion and other parameters.
- **Object recognition**: Segmentation can be used for the recognition of objects in images, which can be used for image retrieval and annotation.

Materials and methods

All the five datasets used for evaluation of the proposed model is obtained from Kaggle. The interactive python code for the same is made available here.

Contributions

1. Modified Double U- Net takes full advantage of Double U-Net and ensemble learning.
2. It is the combination of two U-Net architectures stacked on top of each other. The first U-Net uses an ensemble of pre-trained Xception, DenseNet, and VGG-19 and the second U-Net is stacked at the bottom to capture more information that can be used in our semantic segmentation task.

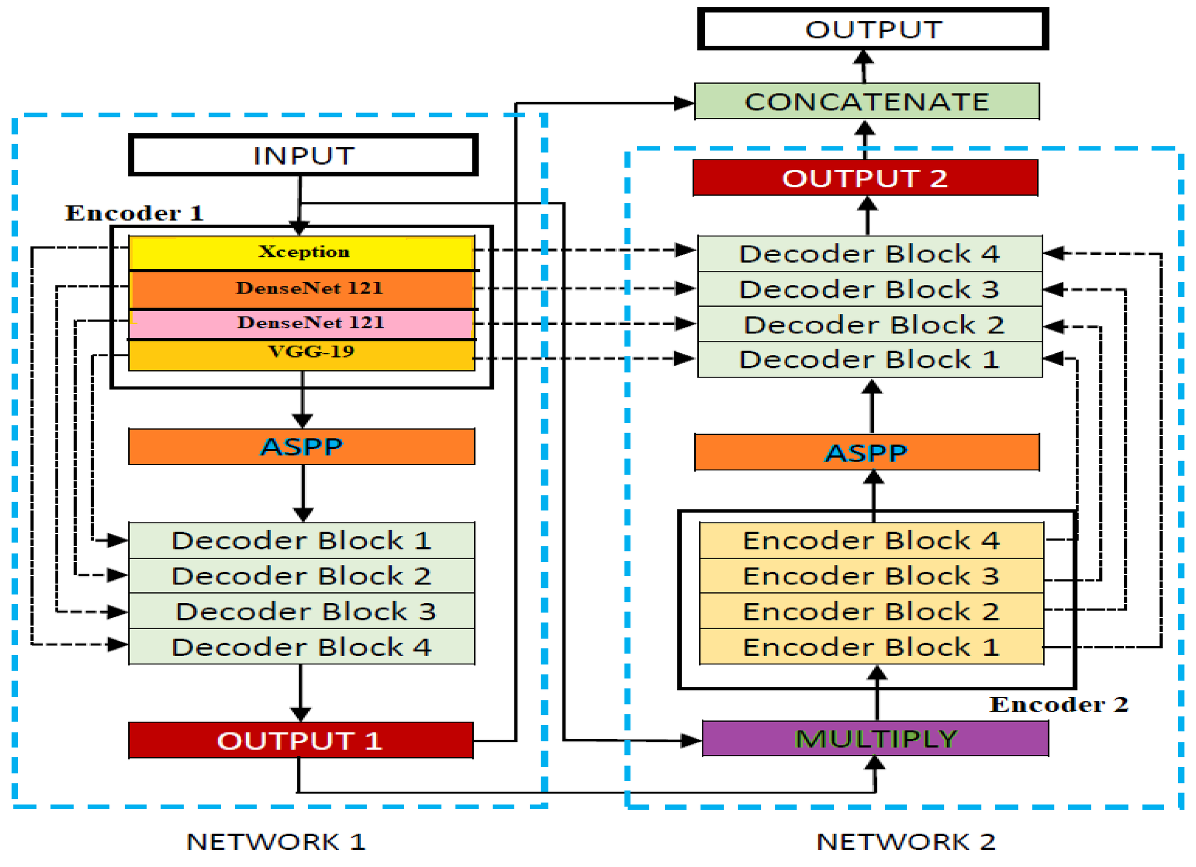

Figure 2. Block diagram of the proposed model

3. The performance of our proposed model has been evaluated on five different datasets, namely the Data Science Bowl Challenge-2018, the CVC-ClinicDB dataset, ISIC-2018 challenge dataset, the Kvasir-Instrument dataset, and the INBreast dataset.

Table 1. Hyper-parameters used for various networks and datasets

Dataset	Parameters	U-Net	BCDU-Net	DU-Net	Proposed
Data Science Bowl Challenge	Loss function	BCE	Dice Loss	Dice Loss	Dice Loss
	Optimizer	Adam	Adam	Nadam	Nadam
	Batch Size	16	16	16	16
	Learning Rate	0.00001	0.00001	0.00001	0.0001
ISIC-2018 challenge dataset	Loss function	BCE	BCE	Dice Loss	BCE
	Optimizer	Adam	Adam	Adam	Adam
	Batch Size	16	4	16	4
	Learning Rate	0.0001	0.00001	0.0001	0.000001
CVC-ClinicDB	Loss function	BCE	BCE	BCE	BCE
	Optimizer	Adam	Adam	Nadam	Nadam
	Batch Size	4	4	16	16
	Learning Rate	0.00001	0.00001	0.00001	0.00001
Kvasir-Instrument	Loss function	BCE	BCE	Dice Loss	Dice Loss
	Optimizer	Adam	Adam	Nadam	Adam
	Batch Size	8	4	4	4
	Learning Rate	0.00001	0.00001	0.00001	0.0001
INBreast	Loss function	BCE	BCE	Dice Loss	Dice Loss
	Optimizer	Adam	Adam	BCE	Adam
	Batch Size	8	4	4	4
	Learning Rate	0.00001	0.00001	0.00001	0.0001

Results

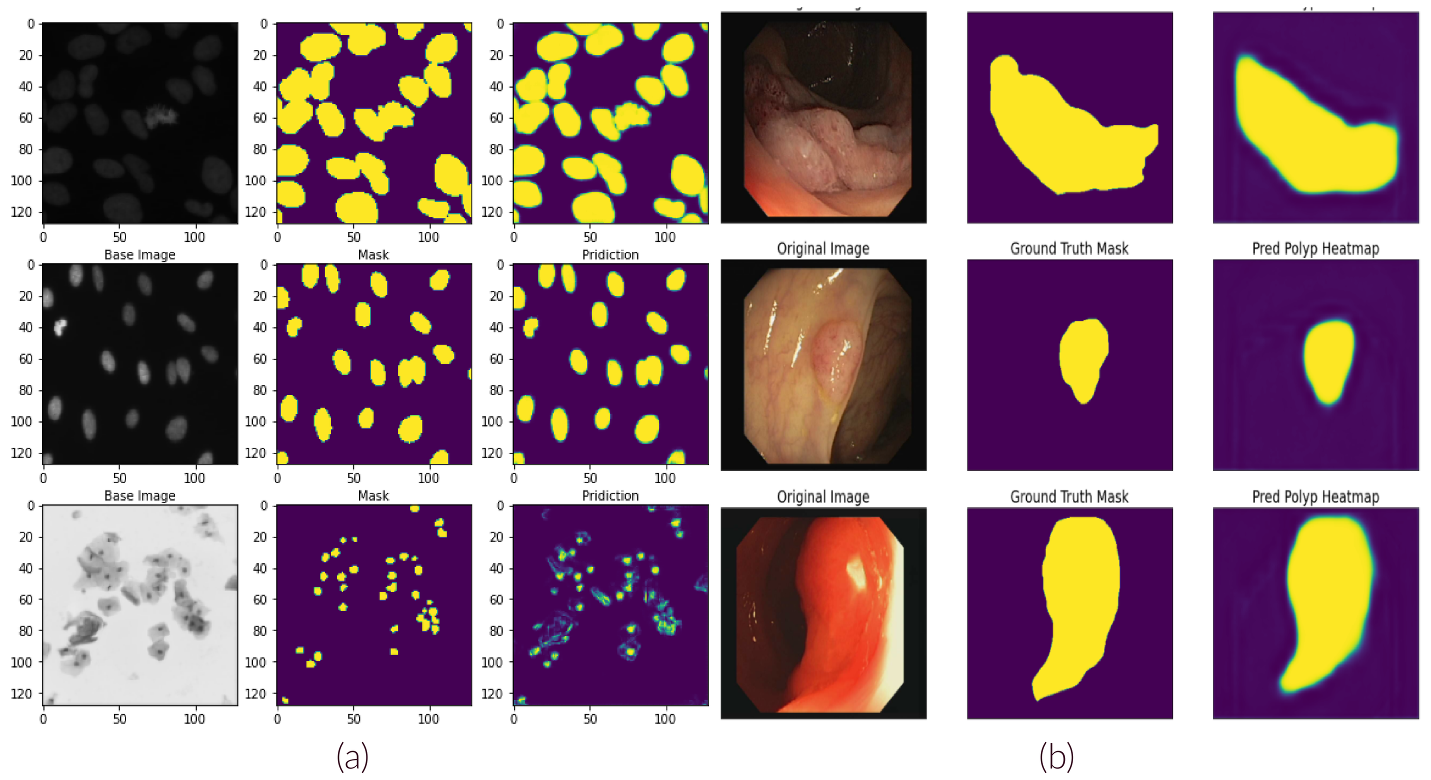

Figure 3. Few examples of the segmentation performance of our model on Data Science Bowl Challenge dataset and CVC-ClinicDB dataset.

Table 2. Comparison of the proposed model on Data-science-bowl Dataset with other models

Architectures	F1 Score (Dice)	Sensitivity	Specificity	Accuracy	AUC	Jaccard (IoU)	Parameters	FLOPs (GMac)	FPS
U-Net	92.25	90.81	98.25	96.59	94.53	85.61	9,091,461	40.7	164.93
BCDU-Net	92.28	91.33	98.10	96.59	94.72	88.66	20,660,869	45.6	105.97
SegNet	82.66	80.05	88.76	86.82	94.41	70.44	9,004,381	40.7	164.93
DU-Net	92.96	91.29	98.53	96.92	94.91	86.85	24,068,210	28.3	142.22
Proposed	**94.67**	**91.37**	**99.69**	**96.81**	**94.48**	**89.87**	23,599,122	27.7	128.25

Table 3. Comparison of the proposed model on skin lesion Dataset with other models

Architectures	F1 Score (Dice)	Sensitivity	Specificity	Accuracy	AUC	Jaccard (IoU)	Parameters	FLOPs (GMac)	FPS
U-Net	88.72	84.27	97.86	93.88	90.62	79.72	9,091,461	40.7	158.89
BCDU-Net	88.64	84.12	97.78	93.86	90.88	79.60	20,660,869	45.6	98.62
SegNet	86.84	82.23	95.10	91.22	89.80	76.74	9,004,381	42.6	168.62
DU-Net	90.22	84.56	98.77	94.54	90.23	82.19	24,068,210	28.3	138.22
Proposed	**90.99**	**85.81**	**98.81**	**94.94**	**91.67**	**83.47**	23,599,122	27.7	122.52

References

[1] Debesh Jha, Michael A Riegler, Dag Johansen, Pål Halvorsen, and Håvard D Johansen. Doubleu-net: A deep convolutional neural network for medical image segmentation. In *2020 IEEE 33rd International symposium on computer-based medical systems (CBMS)*, pages 558–564. IEEE, 2020.

[2] Matthieu Lê, Jan Unkelbach, Nicholas Ayache, and Hervé Delingette. Gpssi: Gaussian process for sampling segmentations of images. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 38–46. Springer, 2015.

[3] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.

sagardeepdeb@gmail.com