SJIF Impact Factor (2023): 8.574| ISI I.F. Value: 1.241| Journal DOI: 10.36713/epra2016 ISSN: 2455-7838(Online)
EPRA International Journal of Research and Development (IJRD)
Volume: 8 | Issue: 3 | March 2023 - Peer Reviewed Journal

MULTI-SCALE ATTENTION BASED U-NET MODEL FOR LIVER TUMOR SEGMENTATION

J. Jenisha[1*], A. Joel Dickson[2]

[1*] *Communication Systems, Bethlahem Institute of Engineering, Anna University.*
[2] *Electronics and Communication Engineering, Bethlahem Institute of Engineering, Anna University, Chennai.*

**Corresponding Author: J. Jenisha*

Article DOI: https://doi.org/10.36713/epra12576
DOI No: 10.36713/epra12576

ABSTRACT

It is essential to automatically evaluate the position and size of the liver tumour for radiologists, diagnosis, and the clinical process. Many U-Net-based variants have been suggested in recent years to enhance the segmentation results for medical image segmentation, but they are unable to describe the global spatial and channel relationships among lesion regions. To overcome this issues, we proposed a novel network called Multi-scale Attention UNet (MA-UNet) to address this problem by adding a self-attention mechanism into our approach to adaptively combine local features with their global dependencies. The attention mechanism of the MA-UNet allows it to capture complex contextual dependencies. Position-wise Attention Block and Multi-scale Fusion Attention Block are the two blocks that we have developed. The feature interdependencies in spatial dimensions, which represent the spatial dependencies between pixels in a global view, are modelled using the Position-wise Attention Block. A multi-scale semantic feature fusion attention block is also used to capture the channel dependencies between any feature map. On the MICCAI 2017 LiTS Competition dataset, we assess our methodology. Compared to other cutting-edge methods, the suggested way performs better. The Dice and VOE values of liver tumors segmentation are 0.749 ± 0.08 and 0.21 ± 0.06 respectively.

KEYWORDS: *Liver tumor segmentation, Attention mechanism, Deep learning.*

1. INTRODUCTION

A significant amount of people worldwide pass away from liver cancer each year [1], making it one of the most prevalent cancer diseases today. A malignant liver tumour known as liver cancer falls into two categories: main and secondary [2]. Secondary liver cancer, also known as sarcoma, is comparatively uncommon compared to primary liver cancer and develops from the epithelial or mesenchymal tissue of the liver. Primary liver cancer is a high-incidence and extremely harmful malignant tumour. The liver, the biggest solid organ in the human body, performs numerous crucial metabolic processes [3]. When malignant masses develop in the liver, the results can be severe and potentially fatal. Therefore, increasing the survival rate of liver cancer patients requires early diagnosis and treatment [4]. In order to assess liver tumours, CT-based imaging techniques are frequently used, and CT scans can plainly display the number, size, and boundaries of lesions. Segmentation of liver lesions serves as a precondition for identification and is crucial to the disease's management [5]. Liver segmentation can be split into semi-automatic and manual segments [6]. However, adding manual intervention to the semi-automatic segmentation process will introduce bias and errors because manual segmentation primarily depends on the radiologists' judgment, which takes time and is subject to error. Given the distinctive variety and spread of liver tumour shapes, it becomes extremely difficult to automatically segment liver tumour lesions.

2. LITERATURE REVIEW

Some of the papers based on the liver tumor segmentation are reviewed below,

ResNeXt50-Dilated Convolution-Transformer U-Net (RDCT-U-Net) was introduced by Li & Ma [7] for the segmentation of liver tumours. To increase the network depth, broaden the visual field, and boost the effectiveness of feature extraction without increasing the parameters, they create a backbone network with ResNeXt50 as its dominant component and dilated convolution as a supplement. Transformer is simultaneously introduced in down sampling to improve the segmentation accuracy of liver tumours and the network's overall perception and understanding of the image. High accuracy is provided by this model, but time usage is greater.

Modified U-net was developed by Manjunath & Kwadiki [8] for the segmentation of liver tumours. It is made up of an expansion track (right) and a retrenchment track (left). The structure is taken over using the retrenchment track (Encoder). Use of an

SJIF Impact Factor (2023): 8.574| ISI I.F. Value: 1.241| Journal DOI: 10.36713/epra2016 ISSN: 2455-7838(Online)
EPRA International Journal of Research and Development (IJRD)
Volume: 8 | Issue: 3 | March 2023 - Peer Reviewed Journal

extension track (Decoder) is required to understand the precise location. The complexity of segmentation is reduced by this technique, but different-sized tumours cannot be segmented.

For image segmentation, Deng *et al*. [9] presented a powerful and portable U-Net with deep skip connections. The higher-level features are first extracted from the image by downsampling the output of the prior encoder layer. By concatenating the feature maps from the output of the most recent encoder layer or previous decoder layer and up-sampling to maintain the scale, the output from each encoder layer is then taken into the matching decoder layer to classify the pixels. The softmax activates the output of the final encoder layer to output the segmentation result last. Although the segmentation outcome is improved, processing time is lengthy.

Residual-Attention UNet++, an expansion of the UNet++ model with a residual unit and attention mechanism for image segmentation, was proposed by Li *et al.* [10]. First, the deterioration issue is improved by the residual unit. Second, the segmentation task-unrelated background area can be suppressed while the target area can be given more weight by the attention process. This increases segmentation precision, but complicates the connections between the models.

Rahman *et al.* [11] hybrid Res-UNet model for segmenting liver tumour combined the ResNet and UNet models. The suggested technique is used to extract usable segments from images of liver tumours. To diagnose the liver and find tumours in the nearby organs, CNN is used along with data augmentation, pre-processing, and other techniques. Three distinct routes are included in Res-UNet: Route of encoding transforms the data into a precise recognition. Reverses the encoding process and classifies the image pixel by pixel in decoding route. Bridge procedure that connects the two routes. ResNet, on the other hand, employs artificial neural networks and is a condensed form of residual blocks. The skip connections concept streamlines and expedites the deep learning process in complex networks for residual blocks. The hybrid Res-UNet, in contrast, provides for full standby of convolutional blocks. This method improve the accuracy, however tumor edge is unable to segment.

By combining four effective neural networks, Popescu *et al.* [12] created an intelligent judgment system for segmenting liver and hepatic tumours (ResNet152, ResNeXt101, DenseNet201, and InceptionV3). The model operates similarly to a UNet, with the segmentation head serving as a decoder and the image classification network acting as an encoder. This reduces the dimensionality of the tensor while restoring it to its initial size. Following that, a weighted decision system will be used to combine the inputs into a singular segmentation using the segmentations that were obtained from these networks as inputs. Using only the liver as an input, this process is carried out independently, first segmenting the liver tissue, and then identifying the lesion areas. Efficiency is increased by this model, but over-fitting is an issue.

SAR-U-Net, which was developed by Wang *et al.* [13] and is based on the traditional U-Net. First, the SE block is used to selectively emphasize features important to a particular segmentation task while suppressing irrelevant regions after each convolution in the U-Net encoder; Second, the transition layer and the output layer are replaced with the atrous spatial pyramidal pooling, which also allows for the acquisition of multi-scale picture data via various receptive fields. Thirdly, the traditional convolution block is replaced with the residual structures to alleviate the gradient vanishing issue, causing the network to gain accuracy from significantly more depth. But the time commitment is greater.

A new U-Net variant using stacked dilated convolutions for medical image segmentation (SDU-Net) was put forth by Wang *et al.* [14]. SDU-Net uses the architecture of standard U-Net but modifies the procedures for the encoder and decoder (an operation indicates all the processing for feature maps of the same resolution). SDU-Net uses one standard convolution followed by numerous dilated convolutions and concatenates all of the outputs from the dilated convolutions as input to the following operation, in contrast to vanilla U-Net, which includes two standard convolutions in each encoder/decoder operation. Accuracy is increased, but the process takes longer. The overview of the segmentation of liver tumours as of today is shown in Table.1.

SJIF Impact Factor (2023): 8.574| ISI I.F. Value: 1.241| Journal DOI: 10.36713/epra2016 ISSN: 2455-7838(Online)
EPRA International Journal of Research and Development (IJRD)
Volume: 8 | Issue: 3 | March 2023 - Peer Reviewed Journal

Table: 1 Summary of existing liver tumor segmentation

Author	Methods used	Advantages	Disadvantages
Li & Ma [7]	RDCT-U-Net	-This model provides high accuracy	-time consumption is more.
Manjunath & Kwadiki [8]	Modified U-net	- This method decreases the complexity of segmentation	-different size tumor cannot be segmented.
Deng *et al.* [9]	efficient and lightweight U-Net	-This improves the segmentation result	-time for processing is high.
Li *et al.*[10]	Residual-Attention UNet++	-improves the segmentation accuracy	- the connection of model's get complicated.
Rahman *et al.*[11]	hybrid ResUNet model	-improve the accuracy.	- tumor edge is unable to segment.
Popescu *et al.* [12]	efficient neural networks	-increases efficiency	-over-fitting problem occurs.
Wang *et al.* [13]	SAR-U-Net	- alleviate the gradient vanishing issue	- the time commitment is greater
Wang et al. [14].	SDU-Net	-high accuracy	- the process takes longer

3. PROBLEM DEFINITION

Liver cancer causes a major threat to people's lives and overall health. So it is necessary to detect the diseases as soon as possible. Many approaches are introduced for Liver tumor segmentation. Still, it is difficult because of the tumor's irregular form and hazy boundaries in CT images. Additionally, in a global view, the dilated convolutions and pooling operations are unable to take advantage of the spatial and channel-wise connection between pixels. Likewise, when using pooling procedures, it is simple to lose details from feature map information. As a result, liver tumour feature depiction and segmentation are inaccurate.

4. PROPOSED MA- UNet MODEL

To segment liver tumours, we suggest a new network called Multi-scale Attention based U-Net (MA- UNet) with a dual attention mechanism. The MA-UNet model's design is depicted in Fig. 1. We go into great depth about the suggested technique, mentioning Res-block, Position-wise Attention Block (PAB), and Multi-scale Fusion Attention Block (MFAB). The Low-level features from Skip-Connection have more edge information while the High-level features have rich semantic information of the picture. Images' characteristics are recovered using low-level features. For high-level and low-level features, we use the channel-wise attention mechanism, accordingly. The goal is to reduce the weight of unimportant feature information and raise the importance of relevant information for each feature channel in the segmentation task.

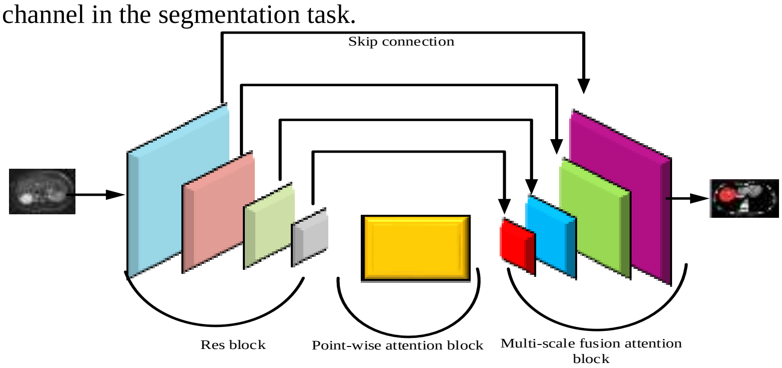

Figure 1: Architecture of MA- UNet model

The proposed MA-UNet model involves three major steps and are explained below.

(a) Res-Block:

To extract high-dimensional feature information from CT images, we employ three 3 × 3 Conv blocks and one residual link. In the MA-UNet, Batch Normalization is swapped out for Group Normalization (GN). The Res-block employs group normalization. The Res-block block layout is shown in Fig. 2.

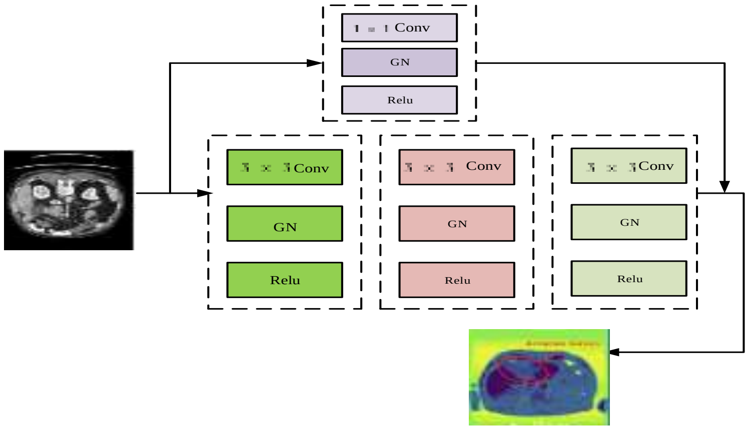

Figure 2: Block diagram of Res-Block

(b) Position-Wise Attention Block

The spatial relationships between any two position feature maps are captured using PAB. Over local feature maps, the PAB can model a broader variety of rich spatial contextual information. The block layout of a position-wise attention block is shown in Fig. 3.

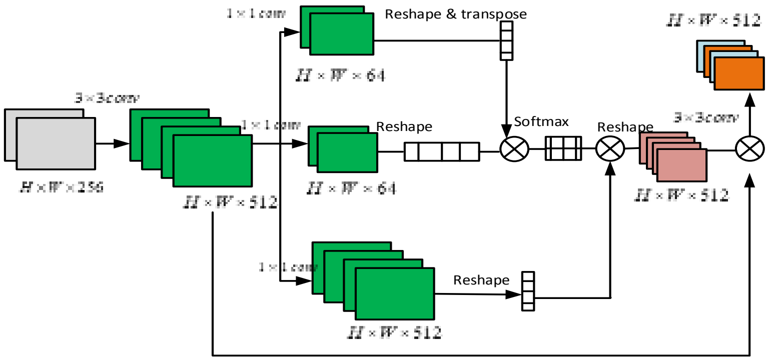

Figure 3: Block diagram of position-wise attention block

(c) Multi Scale Fusion Attention Block

We create a novel MFAB to extract the interdependence among feature channels and enhance network performance by merging the high and Low-level feature maps. The block layout for multi-scale fusion attention is shown in Fig.4. When using multi-level feature maps without additional spatial dimensions, MFAB learns the relative significance of each feature channel, enhancing the feature maps that are most helpful and suppressing the feature maps that are least helpful for the task of segmenting liver tumours.

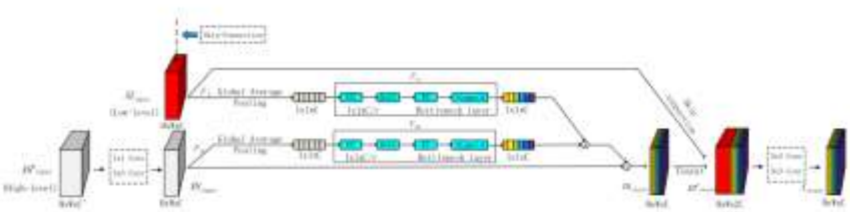

Figure 4: Block diagram of multi-scale fusion attention

5. EXPERIMENTS AND RESULTS

The proposed method is conducted on the dataset of MICCAI 2017 Liver Tumor Segmentation (LiTS) challenge. It contain 70 testing and 131 training CT scans images.

5.1 Metrics for Evaluation

To accurately assess the liver tumour segmentation performance of the suggested MA-UNet method. The formulas for evaluation measures are as follows:

SJIF Impact Factor (2023): 8.574| ISI I.F. Value: 1.241| Journal DOI: 10.36713/epra2016 ISSN: 2455-7838(Online)
EPRA International Journal of Research and Development (IJRD)
Volume: 8 | Issue: 3 | March 2023 - Peer Reviewed Journal

$$Dice = \frac{2|P \cap Q|}{|A| + |B|} \tag{1}$$

$$VOE = 1 - \frac{|P \cap Q|}{|P \cup Q|} \tag{2}$$

where P and Q denote the predicted binary image and the ground true binary image respectively.

5.2 Comparative Analysis

Table.2 shows the performance of proposed MA-UNet model with existing methods such as RDCT-U-Net [7], Residual-Attention UNet++ [10], and SDU-Net [14]. The proposed MA-UNet model achieved the better segmentation performance than other methods for liver and tumors segmentation. Our method achieves dice value of 0.749 ± 0.08 and VOE value of 0.21 ± 0.06.

Table: 2 Dice and values for proposed and existing methods

Methods	Dice	VOE
RDCT-U-Net	0.633 ± 0.17	0.39 ± 0.07
Residual-Attention UNet++	0.719 ± 0.11	0.21 ± 0.04
SDU-Net	0.738 ± 0.06	0.19 ± 0.05
proposed MA-UNet	**0.749 ± 0.08**	**0.21 ± 0.06**

Input image | **Res-block** | **Position wise-attention block** | **Multi-scale fusion attention block**

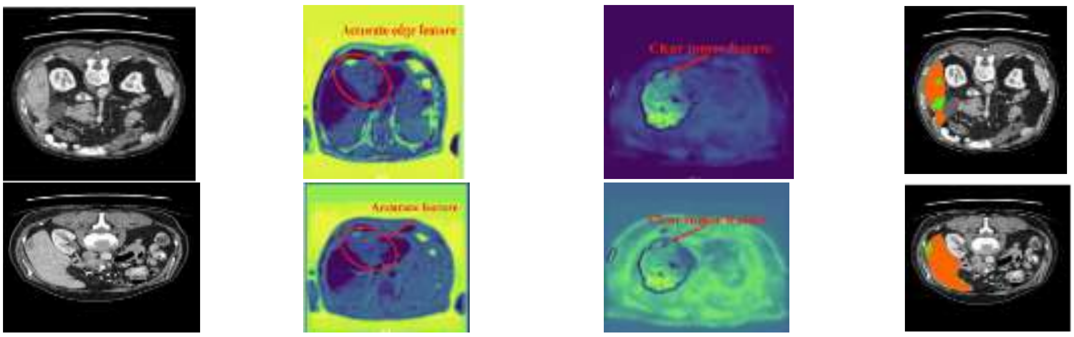

Figure 5: Results for Proposed MA- UNet model

Fig.5 shows the results of proposed MA-UNet model. The segmentation of liver tumor are considered to be the most challenging segmentation job because the shape and size of liver tumours are unpredictable. Fig. 5 shows that the suggested technique for segmenting liver tumours achieves better segmentation performance than other methods.

6. CONCLUSION

We create a novel MA-UNet model for segmenting liver tumours. In order to segment images, we add a self-attention method. In particular, we examine the Multi-scale semantic information based on the channel dependencies between any feature maps and use self-attention mechanism to capture the spatial and channel dependencies of feature maps. The tests showed that our suggested approach was superior on the 2017 LiTS dataset. The suggested approach is useful for helping the practitioner during a clinical procedure. The MA-UNet, however, also has some drawbacks and is only able to segment the liver tumour in this study. To evaluate the segmentation efficiency and robustness of MA-UNet in the future, we will investigate how MA-UNet affects other medical images.

REFERENCES

1. *Calderaro J, Ziol M, Paradis V, Zucman-Rossi J (2019) Molecular and histological correlations in liver cancer. Journal of hepatology. 71(3):616-30.*
2. *Wang C, Vegna S, Jin H, Benedict B, Lieftink C, Ramirez C, de Oliveira RL, Morris B, Gadiot J, Wang W, du Chatinier A (2019) Inducing and exploiting vulnerabilities for the treatment of liver cancer. Nature. 574(7777):268-72.*
3. *Liu Z, Jiang Y, Yuan H, Fang Q, Cai N, Suo C, Jin L, Zhang T, Chen X (2019) The trends in incidence of primary liver cancer caused by specific etiologies: results from the Global Burden of Disease Study 2016 and implications for liver cancer prevention. Journal of hepatology. 70(4):674-83*

SJIF Impact Factor (2023): 8.574| ISI I.F. Value: 1.241| Journal DOI: 10.36713/epra2016 ISSN: 2455-7838(Online)
EPRA International Journal of Research and Development (IJRD)
Volume: 8 | Issue: 3 | March 2023 - Peer Reviewed Journal

4. *Satriano L, Lewinska M, Rodrigues PM, Banales JM, Andersen JB (2019) Metabolic rearrangements in primary liver cancers: cause and consequences. Nature reviews Gastroenterology & hepatology. 16(12):748-66.*
5. *Li X, Ramadori P, Pfister D, Seehawer M, Zender L, Heikenwalder M (2021) The immunological and metabolic landscape in primary and metastatic liver cancer. Nature Reviews Cancer. 21(9):541-57.*
6. *Masthoff M, Schindler P, Harders F, Heindel W, Wilms C, Schmidt HH, Pascher A, Stegger L, Rahbar K, Köhler M, Wildgruber M (2021) Analysis of failed therapy evaluations in radioembolization of primary and secondary liver cancers. Journal of Cancer Research and Clinical Oncology. 147(5):1537-45.*
7. *Li L, Ma H (2022) RDCTrans U-Net: A Hybrid Variable Architecture for Liver CT Image Segmentation. Sensors. 22(7):2452.*
8. *Manjunath RV, Kwadiki K (2022) Modified U-NET On CT images For Automatic segmentation of liver and its tumor. Biomedical Engineering Advances. 100043.*
9. *Deng Y, Hou Y, Yan J, Zeng D (2022) ELU-Net: An Efficient and Lightweight U-Net for Medical Image Segmentation. IEEE Access. 10:35932-41.*
10. *Li Z, Zhang H, Li Z, Ren Z (2022) Residual-Attention UNet++: A Nested Residual-Attention U-Net for Medical Image Segmentation. Applied Sciences. 12(14):7149.*
11. *Benedetti P, Femminella M, Reali G (2023) Mixed-Sized Biomedical Image Segmentation Based on U-Net Architectures. Applied Sciences. 13(1):329.*
12. *Popescu D, Stanciulescu A, Pomohaci MD, Ichim L (2022) Decision Support System for Liver Lesion Segmentation Based on Advanced Convolutional Neural Network Architectures. Bioengineering. 9(9):467.*
13. *Wang J, Lv P, Wang H, Shi C (2021) SAR-U-Net: squeeze-and-excitation block and atrous spatial pyramid pooling based residual U-Net for automatic liver segmentation in Computed Tomography. Computer Methods and Programs in Biomedicine. 208:106268.*
14. *Wang S, Hu SY, Cheah E, Wang X, Wang J, Chen L, Baikpour M, Ozturk A, Li Q, Chou SH, Lehman CD (2020) U-net using stacked dilated convolutions for medical image segmentation. arXiv preprint arXiv:2004.03466.*