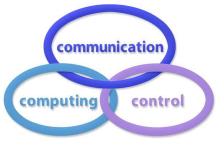
INTERNATIONAL JOURNAL OF COMPUTERS COMMUNICATIONS & CONTROL
Online ISSN 1841-9844, ISSN-L 1841-9836, Volume: 19, Issue: 2, Month: April, Year: 2024
Article Number: 6226, https://doi.org/10.15837/ijccc.2024.2.6226



CCC Publications

Dual attention U-net for liver tumor segmentation in CT images

Omar Ibrahim Alirr

Omar Ibrahim Alirr
College of Engineering and Technology
American University of the Middle East
Eqaila, 52000, Kuwait
omar.alirr@aum.edu.kw

Abstract

Segmenting liver tumors in CT scans plays a vital role in medical analysis planning. The clinicians require a detailed 3D understanding of the tumor's location and liver anatomy, to decide about the proper surgical resection approach. Manual segmentation requires a lot of efforts and time also it depends on the expertise of clinicians. An automatic U-net based method for liver tumors delineation in CT images is proposed. It relies on employing attention-based processes to enhance the performance of U-net. Hard attention and soft attention are used to orient the U-net in learning the intended features from the target CT scans. Soft attention mechanisms, spatial and channel attentions, are employed to help in extracting the long-range relationships and allow the network to successfully distinguish tumors from the surrounding parenchyma. The paper addressed the use of region based active contour implemented using Chan and Vese approach as postprocessing step to improve the predicted segmented tumors. The proposed approach is validated using a challenging big LiTS datasets. The achieved Dice score for the segmenting of liver tumors is 0.81 which shows a superior performance compared to other proposed method in the state of art. The suggested method was successful in discriminating liver tumors from surrounding tissue in heterogeneous CT scans, taking the advantage from the important preprocessing enhancement step. The method demonstrates its generalizability and reliability to be used for automatic analysis of the liver tumors in daily clinical practice. Also, the method proved its ability to achieve high accuracy in detecting stroke that proves its ability to be utilized as clinical tool for a preoperative clinical planning.

Keywords: liver tumor, soft attention, U-net, level-set, EED, segmentation, deep learning.

1 Introduction

Preoperative planning, resection risk planning, or actual surgical therapy are three areas where computer-aided technology is used in a set of procedures known as computer-aided diagnostic (CAD) for liver cancer. Liver cancer is one of the most prevalent cancers that claims a high number of lives annually. Treatment for liver cancer requires a precise diagnosis and preparation [2]. Accurate diagnosis is necessary to evaluate the size and location of the tumors and make the optimal clinical therapy selection for liver treatment options. The most popular imaging method for diagnosing liver cancer is computed tomography (CT) scanning because it provides precise anatomical details about

the human body's abdominal organs. The delineation of the liver and tumors using CT images is necessary to visualize the liver anatomy for preoperative treatment [3].

Due to the small differences between tumor and healthy tissues, especially at their borders, tumor segmentation presents an additional challenge. Additionally, the shape, size, and texture of tumors vary widely. Tumor segmentation becomes more challenging despite these obstacles; yet, the automated approach is favored because it is ideally more unbiased and does not depend on human experience [4, 5, 6]. Over the past few years, organ segmentation using CT scans has been a popular study area. Recent advancements in computer vision led to the creation of deep fully convolutional neural (DFCNs) networks, which enhanced semantic segmentation performance and enabled them to outperform peers in the field of medical imaging[7, 8, 9].With one image as the input and one label as the output, General FCN focuses on the problem of image classification. However, in medical imaging, it is necessary to identify the location of abnormality in addition to classifying it [10, 11, 12, 13].

The ISBI 2017 conference co-hosted the liver tumor segmentation (LiTS) competition challenge [14]. FCN networks were used in the competition's top-rated automated techniques. U-Net architecture was employed in many works for liver and tumors [14] [15–19]. After the FCN's success, other attempts to apply it for liver and tumor segmentation have been made; one of the best FCN architectures to be developed is the U-Net [20].U-Net has been successful in classifying images and locating specific structures. It is also able to find and distinguish borders by classifying each individual pixel. It is important to note that the preceding information is based on the most recent available information. Recently, this type of architecture has included long and short residual skip connections, or extended their length for 3D segmentation [22, 23, 24]. The feature use was enhanced in following iterations of U-Net variations including Mnet, DenseUnet, Unet++, and Unet3+ [25, 26, 27].

A number of studies put forth several strategies to raise the Unet performance, such as the attention mechanism as well as the residual mechanism. Residual convolutions enhance feature exploitation and hence boost network performance [28]. The segmentation of medical images using attention mechanisms has become more common; examples include the segmentation of the pancreas and the structure of the brain [29]. Attention-gate approaches are used to focus on local semantic information in trained systems. For example, attention-based UNets were used in recent studies to use baseline MRIs to predict ultimate ischemia lesions. However, employing 3D networks to segment medical pictures is still challenging [30, 31].

While deep learning FCNs have demonstrated remarkable accuracy in organ segmentation from CT images, these techniques still need a training phase on several datasets to cover all relevant characteristics of the target organ and build a trained network that can recognize that organ in the test dataset. These methods do not utilize the local characteristics in the test dataset itself to enhance the final segmentation from the target CT scan. Accordingly, new research favors combining deep learning approaches with local information-based strategies. In the field of liver and tumor segmentation, numerous intensity-based techniques have been developed to identify the liver and tumor's intensity range by statistically analyzing the intensities of CT images [6, 32, 33]. One of these more independently deployable solutions is the level set based active contour approaches. A coarse segmentation initial mask is pushed using a level set contour - based approach to more accurately match the liver/tumor border in the test CT image [34].

In this paper, a framework is proposed to segment the tumors from the target CT scans. In order to identify and segment liver tumors, an attention-based U-net in conjunction with an active contour process is considered and used in the hepatic envelope. Liver tumor segmentation is represented by the U-net network's output, which is further improved in the postprocessing stage by employing active contour. The final segmentation is generated by deforming the U-net segmented tumors to match the tumor borders in the target CT image using the level set approach as an initial mask. The postprocessing step has high impact in improve the tumors segmentation accuracy, that increase the method potential to be used in the clinical routines.

2 Materials and Methods

A deep learning-based methodology is used to learn and extract features of liver tumors from a target CT scan. The framework consists of a number of steps as depicted in figure 1, it includes the main below steps:

- Enhancement of liver tumors using edge enhancement diffusion filtering (EED),
- Liver ROI extraction.
- Attention gated ResUnet segmentation, and
- Post-processing refinement using the level-set based active contour method.

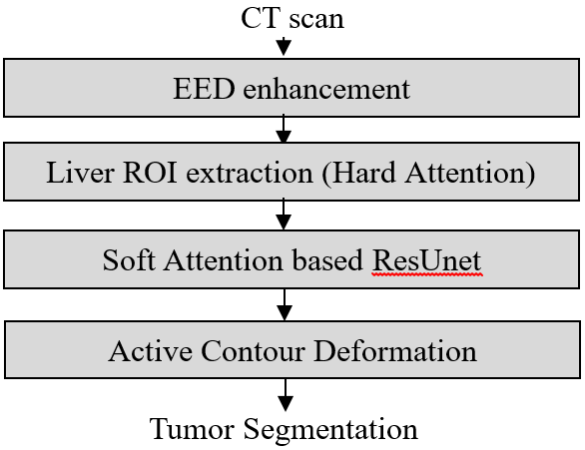


Figure 1: Flow Chart of Tumor Segmentation

In order to enhance the appearance of the liver and tumors in the target CT scans, the first step is image pre-processing. The tumors inside the liver ROI are difficult to segment due to a number of factors, such as inhomogeneous intensity, fuzzy boundaries, and low contrast. Additionally, highly variable are the liver tumor regions' texture, size, and position in the CT scans. Based on that, an enhancement step of the tumors is used, which is the tensor-based Edge Enhancing Diffusion (EED) filtering to improve the liver lesion detection[35]. Hence, to improve Unet segmentation, the research focuses on using the convolutional block attention module (CBAM)[36], which combines the channel and spatial attention modules. Lastly, during the postprocessing stage, the segmented tumors are refined using the localized level set. As preprocessing step, all CT scans undergo isotropic resampling to 1mm3 resolution.

2.1 Datasets

Experiments are conducted on publicly accessible LiTS datasets. A portion of the training LiTS datasets are used to train the U-net networks. A total of 131 contrast-enhanced CT scans from various clinical institutes are included in the LiTS training datasets. Expert manual delineation of the liver and tumors is included with the CT scans. There are 908 lesions in the LiTS dataset. The LiTS dataset is used in this work for both validation and training. Tumor segmentation training is not performed on LiTS datasets that are missing their tumor masks. For the purpose of segmenting liver tumors, 15125 2D slices were taken from LiTS datasets and used for network training. Figure 2 shows an examples of training datasets. The scans highlight the containing of a challenging tumors regions.

2.2 Liver tumors appearance enhancement

Tensor-based Edge Enhancing Diffusion (EED) filtering is used to improve the segmented liver (ROI) from the CT image in order to find the tumorous tissues. (EED) filtering modifies the diffusion

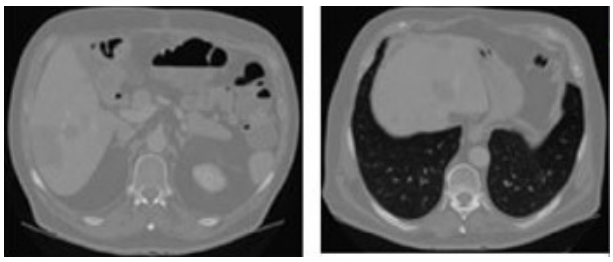


Figure 2: Dataset examples from LiTS

based on the image structure by using a diffusion tensor. The (EED) filter reduces noise while raising contrast and improving intensity uniformity, which aids in maintaining the shape's borders[36]. Anisotropic enhancing diffusion filters modify the diffusion along the image structures by means of a diffusion tensor, as opposed to a scalar diffusion. This diffusion filtering uses a structure tensor to characterize the structure of the picture. It is based on using structure characteristics or local coherence of structure as a diffusion tensor to guide the diffusion. In general the structure tensor can be written by the equation 1 as;

$$J_\rho(\nabla_{u\sigma}) = k_\rho * (\nabla_{u\sigma}\nabla_{u\sigma}^T) \tag{1}$$

Where $\nabla_{u\sigma}$ is the gradient of image u at scale σ, and k is the Gaussian kernel with standard deviation ρ. The quantity of smoothing is defined by the appropriate eigenvalues, whereas the direction of the smoothing is defined by the eigenvectors of the structure tensor using the principal axis transformation.

The EED will have a diffusion in the direction of $V1$ if the eigenvalues in the direction of the eigenvectors $V1$,$V2$,$V3$ of the structure tensor are,$(\mu 1 > \mu 2 > \mu 3)$. Whether or not to do the diffusion will depend on the ratio between the second and third eigenvalues. This ratio is big for structures that resemble tubes, and small for structures that resemble blobs or plates.

Enhancement using EED filtering aims to increase the intensity uniformity inside tumor areas and retain the borders relative to the surrounding liver parenchyma, which improves the contrast of tumor areas. The areas of the tumor can now be more accurately identified and segmented. The objective of this step is to enhance U-net training by recognizing and comprehending the primary features that differentiate the tumor regions from the surrounding tissues. Using EED enhancement, the segmented liver ROI in the target CT scan is displayed in Figure 3. In contrast to the non-enhanced scans, it is evident that the tumor areas in the EED enhanced datasets (right) have strong, distinct edges and homogenous intensities (left).

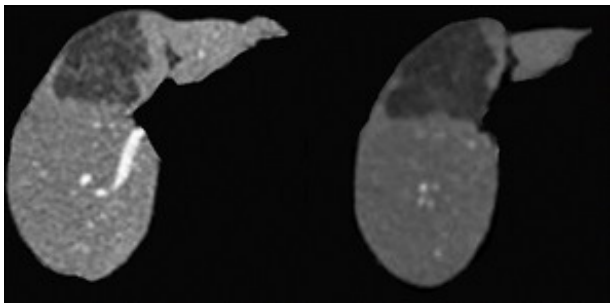


Figure 3: EED enhancement of tumors inside liver ROI

2.3 Attention-based ResU-net architecture

The proposed automatic liver tumor detection and segmentation deep neural network is shown in Figure 4. Taking its advantages from the U-Net, the proposed network has a path that contracts before expanding. A decoding path and an encoding path make up the five levels of the proposed

model. The convolution, activation function (ReLU), and batch normalization are used for each level of encoding. In each level block, these operations are deployed twice in a row, and then the block moves on to the next level after performing a max-pooling operation. Convolutional kernels use a 3x3 kernel size, while maximum pooling uses a 2x2 kernel. Each stage ends with a division of the feature's resolution in half.

The vanishing gradient problem, which is brought on by the FCN's continual rise in network depth, impedes training and degrades performance because more layers are added on top of one another before disappearing. While other deep network topologies were proposed to tackle this problem, DenseNet and ResNet are considered to be notable performance improvements. The ResDense blocks are utilized by the encoder and decoder routes of the proposed U-net network in this study. The expanding and contracting paths of the proposed network are built at each level using ResDense blocks., as shown in Figure 4.

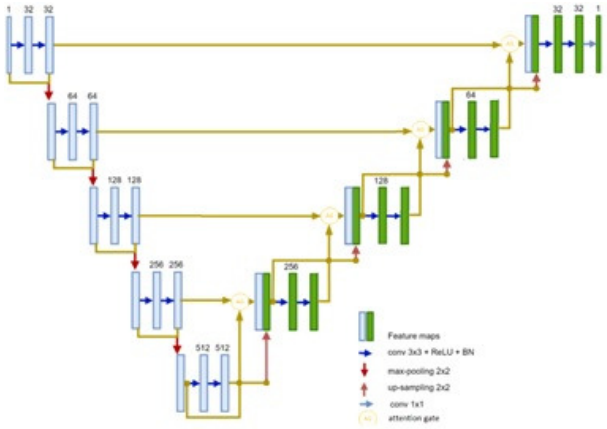


Figure 4: The suggested attention-based ResUnet-architecture.

To improve the output of the segmentation process, the suggested U-net employs the two different attention methods. Generally, the two main categories of attention mechanisms are hard and soft attentions. Cropping is used to carry out the hard attention, which focuses and highlights the significant areas in the image map and allows you to highlight key regions. Backpropagation is ineffective and hard attention is not differentiated. In this work, an embedded hard attention is employed, where the liver envelope that represent the region of interest (ROI) is segmented using our previous work [6]. Hence, the input data for the U-net is the cropped liver envelop from the target CT scans, that reduce the computation cost of U-Net significantly.

On the other hand, soft attention concentrates on weighing specific elements of the image. Higher weights are assigned to relevant image features than to irrelevant ones. It can be trained through backpropagation. Throughout training, weights are also worked out, which aids the model in concentrating on the appropriate places, based on relevance, it gives pixels weights[30, 36, 38].

U-net skip connection integrates spatial information from the up-sampling network with the down-sampling path in order to maintain correct spatial information. However, during this process, the inadequate feature representations from the early layers is carried over. The soft attention is used at the end of each skip connection between encoder decoder blocks in the redesigned U-net depicted in Figure 5. The skip connections that actively reduce activations at irrelevant regions are where the attention gate executes soft attention.

Figure 5 explains convolutional block attention module (CBAM)[36], that represent the attention gate's internal construction, where x and g are the gate's two inputs. The x input, which has more spatial information, reflects the skip connection from the encoder path. The bottom layer provides the g input. The g input provides greater feature representation because it originates from the deeper region of the network. These two different-shaped inputs are given to the first grey box in figure 3, where they are first subjected to 1x1 convolution operations with unique stride values for every input to produce two maps of the same dimensions that can be combined. The addition procedure strengthens the larger aligned weights. The generated feature map is then subjected to soft attention

mechanisms, channel attention operations, and spatial attention operations.

As convolution methods extract essential features by mixing cross-channel and spatial information, the soft attention mechanism in the attention gate aims to emphasize significant characteristics along the two primary dimensions, channel and spatial axes. The channel and spatial attention modules are successively used in each branch to help it acquire what to pay attention to and where in the channel and spatial axes, respectively[36]. As shown, in Figure 5, it focuses on the what and where of an incoming image to show how complimentary attention is calculated by channel and spatial attention processes. This would allow for the installation of two modules either concurrently or consecutively. A sequential architecture works better than a parallel arrangement, according to the study. The experimental findings demonstrate that the channel-first order is marginally better than the spatial-first order for the sequential process organization.

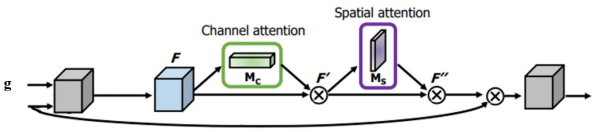


Figure 5: The architecture of gated soft attention module.

A channel attention block and a spatial attention block are combined to generate a channel-spatial attention-aware block, as seen in figure 4. The first attention block utilizes the feature correlations between channels to generate the channel attention map. Since channels on a feature map are conceived of as feature detectors, channel attention is focused on what important given an input image. To efficiently compute the channel attention, the spatial dimension of the input feature map is minimized. Average-pooling has been frequently used to combine regional data and is effective in figuring out how big the target item should be. Max-pooling captures yet another critical piece of data on particular attributes that can be used to infer better channel-wise attention. Hence, both average-pooled and max-pooled characteristics are applied at once. Compared to when they are used independently, both elements significantly increase the representation power of networks.

The spatial attention map is created by utilizing the spatial relationships between features. Spatial attention, which differs from channel attention in that it concentrates on where there is an informative section, is a supplement to channel attention. Before concatenating them to determine the spatial attention, use average-pooling and max-pooling techniques to generate an effective feature descriptor along the channel axis. The use of pooling techniques to emphasize informative regions along the channel axis is shown to be successful. The concatenated feature descriptor is then combined with a convolution layer to produce a spatial attention map, which encodes where to highlight or suppress.

2.4 Level set segmentation step

The learnt characteristics from the training datasets limit the output accuracy of any deep learning network model, therefore in some CT images, the expected segmentation might not accurately follow the boundaries of the tumors. Segmentation from the deep learning step may not catch the tumorous tissues at the surface of the tumor regions due to low contrast. This section explains the post processing step, it aims to deform the initial U-net predicted segmentation to closely match the boundaries of the tumors inside the liver ROI in the target CT scan. The U-net predicted segmentation is improved by using the level set-based active contour. The region-based active contour is a text-based segmentation technique that relies on the local information of the target dataset.

Active contour is one of the conventional techniques that have been widely used for image segmentation and boundary tracking [39]. Active contour approaches work on the premise of starting with initial boundary forms, which are represented as closed curves, or contours. Then, the contour is iteratively allowed to deform to minimize a given energy functional within the restrictions of the image, resulting in the desired segmentation. Osher and Sethian proposed a method to build active contours called level set-based active contour [40]. Actually, there are two types of level set active contours: region-based and edge-based. Although it has been found that edge-based active contour models are very image noise sensitive and depend on the initial contour position, they nevertheless

use picture gradients to detect object boundaries. Compared to edge-based level set methods, the region-based level set active contour offers advantages such resilience to initial contour placement and insensitivity to picture noise [41].

U-net predicted segmentation is expected to occur at the tumor boundary, therefore the region-based level set active contour—especially the one demonstrated by Chan and Vese seems to be more suitable than other varieties [39]. The well-known paradigm to model the interior and exterior energies in region based active contour is the one proposed by Chan-Vese, it is represented by the equation 2;

$$E_{CV}[C(s)] = \mu \int_0^{L(C)} ds + \int \int_{\Omega c} (I(x,y) - c1)^2 dxdy + \int \int_{\Omega c} (I(x,y) - c2)^2 dxdy \tag{2}$$

Where the Ωc represent the interior of the curve, c1 and c2 are the mean intensities for the interior and the exterior of the curve, respectively. The first term is the regularization term that minimize the curve length, and the second term maintain the balancing between the interior and the exterior. Assigning the value zero for μ will omit the regularization term, lead to equation represent the simple thresholding. The idea of the Chan-Vese model that it assume the image may separate into two regions, which can be logically tuned by the constant parameters c1 and c2.

As seen in Figure 6, rather than modeling the detailed image, the curve is represented by the local areas of numerous neighborhoods. Every local region is considered separately and is further separated into local exterior and local interior. The method defines the interior and outside local regions, fits the curve at each local region to optimize energy, and then combines the Chan-Vese energy technology to prototype the local interior and exterior forces of each local area contour.

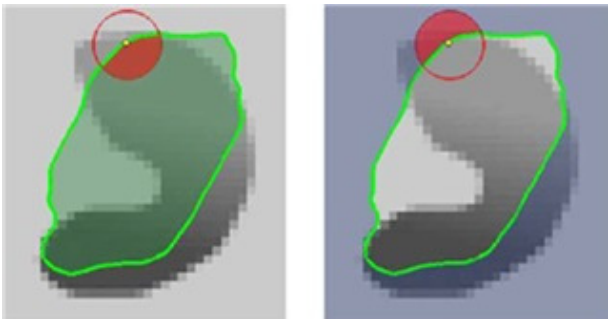


Figure 6: Local interior and local exterior regions.

2.5 Network Implementation and loss function

With patch sizes of 256 by 256, the annotated 2D slices are used to train the proposed network, demonstrating that the training images are bigger versions of the original images. All CT slices were patch-wise normalized with unit variance and zero mean normalization. The liver ROI's tumor regions are detected using the gated attention-based U-net network. The parameters are modified and the network is trained using the Adam optimizer. Starting at 0.0001, the learning rate is set. A variety of parameters are used to monitor the training process. First, the learning rate is lowered by a factor of 0.1 if the validation loss does not drop after three consecutive training epochs. Mini-batches of 32 are used to train the networks.

The trained network has undergone significant performance improvements because to the uneven class distribution of the tumor's areas. To counteract and lessen the impact of the data imbalance, dice loss and binary cross entropy (Dice-BCE) are combined. Because it assigns equal weights to each class in the dataset, this approach works well when the dataset is unbalanced. Additionally, using the loss metric, the overlap between the ground-truth patches and the area of the liver ROI that the network has determined to be infected with tumors is initially determined during the training phase. As explained in section 2.3, the proposed network is only trained inside the liver envelope ROI in order to acquire properties that only distinguish tumors from the surrounding healthy tissue. Moreover, the training method excludes any patches without an annotation mask and guarantees that the training patches used have the appropriate mask.

2.6 Performance Metrics

There are various performance metrics used to evaluate the segmentation of liver tumors from CT scans; the Dice coefficient (DSC), volume overlap error (VOE) and relative volume difference (RVD), as defined below.

$$DSC(A,B) = (2|A \cap B|)/(|A|+|B|) \tag{3}$$

$$VOE(A,B) = 1 - (|A \cap B|)/(|A|+|B|) \tag{4}$$

$$RVD(A,B) = (|B|-|A|)/|A| \tag{5}$$

Where A and B are the binary masks represent the predicted segmentation and the ground truth, respectively. the value of DSC score range is [0,1], where 1 means the perfect segmentation compared to the ground truth. The Dice score evaluates the degree of overlap between the predicted and reference segmentation masks. Similarly, the VOE assess the performance segmentation, where minimum value (0) indicates a perfect segmentation. The RVD measure compares the volume similarities between the generated segmentations and the ground truth, where the value (0) means that both have same volume. RVD directly measures the volume difference without considering the overlap between reference A and the prediction B.

3 Results and Discussion

A qualitative and quantitative evaluation is conducted on the suggested approach. In Figure 7, the liver tumor segmentation for various data sets utilizing the suggested framework is displayed. Axial slices compare a segmentation of the ground truth with the predicted output from the proposed framework. The findings of the method are displayed by superimposing the segmented tumors mask over the CT scan in the areas where the methodology was successful in recognizing the tumors. The axial slices show a comparison of three alternative masks, including the ground truth, the proposed method's segmented liver, and the plain U-Net segmentation (red). The shown datasets confirm that the proposed technique produce high performance compared to the plain U-net method.

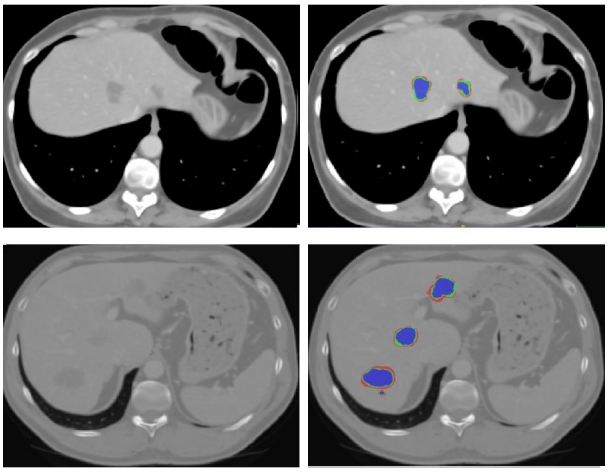


Figure 7: Method segmentation (green), plain U-Net segmentation (blue) and the ground truth (red).

Visual inspection is done to assess how well the suggested system performs. The proposed method output is compared to the manual segmentation (gold-standard) to highlight the key differences as shown in figure 7. The computation of the soft label probability maps is carried using the attention-based U-Net. Each pixel in the target CT scan is examined, and it is then given one of the two labels—tumor or liver tissue. Based on the intensity variations around the original contour, the postprocessing step employing level set step tries to push the U-Net output (predicted segmentation) to match the border of the intended structure. The segmentation from the proposed technique, attention mechanism and postprocessing utilizing level set, outperformed the segmentation from simple U-net (blue), and it was extended to be nearly as accurate as manual segmentation, as shown in the second

Table 1: Tumor segmentation results

Task	DSC	VOE	RVD
U-net (plain)	0.65	0.52	-
U-net + level set	0.72	0.43	0.27
Proposed method	0.81	0.31	0.19

Table 2: Method comparison with state of art techniques

Method (dataset)	DSC	VOE	RVD
U-net + level set (LiTS)[6]	0.72	0.43	0.27
H-dense U-net (LiTS) [42]	0.82	0.30	-
3D MCG-FRN +ACM (3Dircadb) [43]	0.76	0.32	0.19
AHCNet (3Dircadb) [44]	0.73	-	0.13
MCG-FRN (LiTS) [43]	0.76	-	0.40
Proposed Method	0.81	0.31	0.19

column of figure 7. In comparison to the non-enhanced datasets, the EED enhancement step improves the U-net's ability to learn from clear tumors regions and also helped the level set contour extend to a clear tumor's borders.

The method was successful in separating the majority of the tumor's regions from the surrounding tissue in the target CT scans. Attention mechanisms and level-set postprocessing stages are used to solve the issue that the major variations frequently occur on the surface of tumors. However, the manual segmentation carried by clinicians or radiologists is still bigger than the predicted segmentation, even though the tumors are accurately segmented using the provided procedure. This is because they think about including marginal tissues around the tumor itself, presuming that these might be unhealthy tissues that are missed and do not appear in the screening process. Hence, the calculated accuracy shows a difference when it is compared to the manual segmentation, due to the added marginal tumorous tissue.

Quantitative study of the proposed system revealed that the performance metrics scores were fulfilled at comparable levels. Table 1 shows the results of the suggested segmentation framework, as well as the impacts of applying the attention mechanisms and the level set postprocessing phase. The table shows and contrasts the segmentation measure values obtained with plain U-net, level set following U-net, and the suggested strategy that combines the benefits of the three modules (U-net, attention, and level set).

Its clear that both types of attention have impacted the segmentation of tuomrs in a better way. The hard attention at the beginning is used to create the liver envelop as ROI to produce a more focused area of interest that is automatically identified. The soft attention mechanism provides comparable support for the identification of small tumors and the identification of tumor boundaries. The impact of using various attention strategies is obvious in the performance of the original U-net, where accuracy increases from 0.72 to 0.81.

In comparison with other methods, it's hard to make a fair comparison due to the differences from various aspects. The differences are in the source of datasets, that come from different scanning processes with variable image quantities. However, most of the proposed methods in this are U-Net based or deviated version of U-Net as outlined in table 2 [42, 43, 44]. By combining a 2-D Dense UNet for effectively extracting intra-slice characteristics with a 3-D version for hierarchically aggregating volumetric contexts, some authors proposed to use 3D U-net to make up for the shortcomings in 2D U-net, they tested the model 3D IRCAD datasets [42]. While others employed attention by combining soft and hard attention mechanism and long and short skip connections in the U-net network structure [44]. The authors of 3D MCG-FRN +ACM, for example, segmented tumors using 3D U-Net, followed by 3D fractal residual network, and finally improved the predicted tumors using active contour model (ACM) as tumor segmentation refinement. Other works combined 3U-net with conventional techniques like active contour. The new emerging research in this field is to use transformers in computer vision applications. The use of transformers in the field of medical image segmentation adds a high research potential and opens for more comparisons with current state of arts methods.

Table 2 shows a rough comparability with some other state of art methods, that used different datasets, like 3Dircadb, or the more challenging LiTS datasets. The table displays the associated work

performance for various datasets according to different measures; DSC, VOE and RVD scores. From the state of art, it has been shown that LiTS dataset is more challenging than IRCAD3 dataset, in term of the number of scans and their quality. Therefore, the comparison between methods that use different source of datasets is not fair. In addition, other important aspect that should be taken into consideration through the comparison is the level user interaction. Some methods are fully automatic, and others needs used interaction, which affect the achieved accuracy.

Despite that the LiTS is a big dataset, one of the limitation in many proposed methods including this work, is the lack of validation with the LiTS test dataset which is not publicly available. Future improvement could be done by incorporating divers datasets from many sources, this would improve the method generalizability.

4 Conclusion

The suggested framework utilized the different attention processes in order to improve a ResDense Unet network to segment the tumors from the liver envelope in the target CT scans. Using EED, a preprocessing step is used to increase and enhance the tumor's appearance in the target CT images. The method showed how the addition of hard and soft attention processes improved the ability to capture the long-range connections between infections. In addition, the work highlighted the use of the region-based level set as postprocessing step to enhance the predicted segmentation. It is demonstrated that the suggested method outperformed existing models for the tumor segmentation in a challenging LiTS datasets that present it as a promising clinical tool that can be used to do a proper preoperative planning of liver treatments. The quantitative measurements showed how effective the suggested strategy in segmenting the liver tumor with acceptable results. However, future research continues to improve the segmentation process by utilizing newly emerging transformer-based method. In addition, more validation with new datasets from different sources will add more improvement to the proposed method.

Conflict of interest

The authors declare no conflict of interest.

References

[1] Chen, J.; Konstan, J.A. (2010). Conference paper selectivity and impact, *Communications of the ACM*, 53(6), 79–83, 2010.

[2] Mohammed, FA; Viriri, S. (2017). Liver segmentation: A survey of the state-of-the-art. *In: 2017 Sudan Conference on Computer Science and Information Technology (SCCSIT).* IEEE pp 1–6

[3] Alirr, OI; Rahni, AAAbd. (2019). Survey on Liver Tumour Resection Planning System: Steps, Techniques, and Parameters. *J* Digit Imaging 1–20.

[4] Anter, AM; Elsoud, MA; Hassanien, AE. (2013). Automatic Liver Parenchyma Segmentation from Abdominal CT Images. 32–36

[5] Alirr,O; Alshatti, R; Altemeemi, S; et al. (2023). Automatic Brain Tumor Segmentation from MRI Scans using U-net Deep Learning. *BioSMART 2023 - Proceedings: 5th International Conference on Bio-Engineering for Smart Technologies.*

[6] Alirr, OI. (2020). Deep learning and level set approach for liver and tumor segmentation from CT scans. *J Appl Clin Med Phys*, 21:200–209.

[7] Hesamian, MH; Jia, W; He, X; Kennedy P. (2019). Deep Learning Techniques for Medical Image Segmentation: Achievements and Challenges. *J Digit Imaging*, 32:582–596.

[8] Gong, M; Zhao, B; Soraghan, J; et al. (2022). Hybrid attention mechanism for liver tumor segmentation in CT images. *2022 10th European Workshop on Visual Information Processing (EUVIP)*, 1–6.

[9] Gruber, N; Antholzer, S; Jaschke, W; et al.(2020). A Joint Deep Learning Approach for Automated Liver and Tumor Segmentation. *Institute of Electrical and Electronics Engineers (IEEE),* pp 1–5

[10] Li, W; Jia, F; Hu, Q (2015). Automatic Segmentation of Liver Tumor in CT Images with Deep Convolutional Neural Networks. *Journal of Computer and Communications* 03:146–151.

[11] Alirr, OI; Rahni, AAAbd; Golkar, E (2018). An automated liver tumour segmentation from abdominal CT scans for hepatic surgical planning. *Int J Comput Assist Radiol Surg* 13:1169–1176.

[12] Alirr, OI; Rahni, AAAbd (2019). Survey on Liver Tumour Resection Planning System: Steps, Techniques, and Parameters. *J Digit Imaging.*

[13] Alirr, OI; Ashrani, AA (2020). Automatic atlas-based liver segmental anatomy identification for hepatic surgical planning. *Int J Comput Assist Radiol Surg* 15:239–248.

[14] Bilic, P; Christ, PF; Vorontsov, E; et al (2019). The Liver Tumor Segmentation Benchmark (LiTS)

[15] Chlebus, G; Schenk, A; Moltz, JH; et al (2018). Automatic liver tumor segmentation in CT with fully convolutional neural networks and object-based postprocessing. *Sci Rep* 8:1–7.

[16] Han, X (2017). MR-based synthetic CT generation using a deep convolutional neural network method. *Med Phys* 44:.

[17] Vorontsov, E; Tang, A; Pal, C; Kadoury, S (2018). Liver lesion segmentation informed by joint liver segmentation. *In: Proceedings - International Symposium on Biomedical Imaging. IEEE Computer Society,* pp 1332–1335

[18] Alirr OI (2022) Automatic deep learning system for COVID-19 infection quantification in chest CT. *Multimed Tools Appl* 81:527–541.

[19] Ronneberger, O; Fischer, P; Brox, T (2015). U-net: Convolutional networks for biomedical image segmentation. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 9351:234–241.

[20] Çiçek, Ö; Abdulkadir, A; Lienkamp, SS, et al (2016). 3D U-net: Learning dense volumetric segmentation from sparse annotation. *In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*

[21] Ronneberger, O; Fischer, P; Brox, T (2015). U-net: Convolutional networks for biomedical image segmentation. *In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics).* Springer Verlag, pp 234–241

[22] Milletari, F; Navab, N; Ahmadi SA (2016). V-Net: Fully convolutional neural networks for volumetric medical image segmentation. *In: Proceedings - 2016 4th International Conference on 3D Vision, 3DV 2016. Institute of Electrical and Electronics Engineers Inc.,* pp 565–571

[23] Zhou, Z; Rahman, Siddiquee MM; Tajbakhsh, N; Liang J (2018). UNet++: A Nested U-Net Architecture for Medical Image Segmentation. Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. *4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, held in conjunction with MICCAI 2018, Granada, Spain, S.* 11045:3–11.

[24] Mehta R, Sivaswamy J (2017). M-net: A Convolutional Neural Network for deep brain structure segmentation. *Proceedings - International Symposium on Biomedical Imaging* 437–440.

[25] Alirr O (2023). Severity Quantification of COVID-19 Infection using ResDense U-net in Chest X-ray. *BioSMART 2023 - Proceedings: 5th International Conference on Bio-Engineering for Smart Technologies.*

[26] Alirr OI, Aizzuddin A, Rahni A (2023). Hepatic vessels segmentation using deep learning and preprocessing enhancement. *J Appl Clin Med Phys* e13966.

[27] Alirr OI, Rahni AAA (2021). An Automated Liver Vasculature Segmentation from CT Scans for Hepatic Surgical Planning. *International Journal of Integrated Engineering* 13:188–200.

[28] Liu X, Song L, Liu S, Zhang Y (2021). A Review of Deep-Learning-Based Medical Image Segmentation Methods. *Sustainability 2021,* Vol 13, Page 1224 13:1224.

[29] Oktay O, Schlemper J, Folgoc L Le, et al (2022). Attention U-Net: Learning Where to Look for the Pancreas

[30] Khanh TLB, Dao DP, Ho NH, et al (2020). Enhancing U-Net with Spatial-Channel Attention Gate for Abnormal Tissue Segmentation in Medical Imaging. *Applied Sciences 2020,* Vol 10, Page 5729 10:5729.

[31] Chen J, Lu Y, Yu Q, et al (2021). TransUNet: Transformers Make Strong Encoders for Medical Image Segmentation.

[32] Alirr OI, Rahni AAA, Golkar E (2018). An automated liver tumour segmentation from abdominal CT scans for hepatic surgical planning. *Int J Comput Assist Radiol Surg* 13:1169–1176.

[33] Altarawneh NM, Luo S, Regan B, Sun C (2015). A MODIFIED DISTANCE REGULARIZED LEVEL SET MODEL FOR LIVER SEGMENTATION. 6:1–11

[34] Alirr OI, Rahni AAA (2018) Automatic liver segmentation from ct scans using intensity analysis and level-set active contours. *Journal of Engineering Science and Technology* 13:

[35] Jumakulyyev I, Schultz T (2021). Fourth-Order Anisotropic Diffusion for Inpainting and Image Compression. *Math Vis* 99–124.

[36] Woo S, Park J, Lee JY, Kweon IS (2018). CBAM: Convolutional block attention module. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 11211 LNCS:3–19.

[37] Mendrik AM, Vonken EJ, Rutten A, et al (2009). Noise reduction in computed tomography scans using 3-D anisotropic hybrid diffusion with continuous switch. *IEEE Trans Med Imaging* 28:1585–1594.

[38] Schlemper J, Oktay O, Schaap M, et al (2019). Attention gated networks: Learning to leverage salient regions in medical images. *Med Image Anal* 53:197–207.

[39] Chan T, Vese L (2001) Active contours without edges. *Image processing, IEEE transactions on image processing*, 10(2), 266-277.

[40] Osher S, Fedkiw R (2001) Level set methods: an overview and some recent results. *Journal of Computational physics,* 169(2), 463-502.

[41] Lankton S, Tannenbaum A (2008) Localizing Region-Based Active Contours. *IEEE Transactions on Image Processing* 17:2029–2039.

[42] Liu X, Guo S, Yang B, et al (2018) Automatic Organ Segmentation for CT Scans Based on Super-Pixel and Convolutional Neural Networks. *Journal of digital imaging,* 31, 748-760.

[43] Bai Z, Jiang H, Li S, Yao YD (2019) Liver Tumor Segmentation Based on Multi-Scale Candidate Generation and Fractal Residual Network. *IEEE Access* 7:82122–82133.

[44] Jiang H, Shi T, Bai Z, Huang L (2019) AHCNet: An Application of Attention Mechanism and Hybrid Connection for Liver Tumor Segmentation in CT Volumes. *IEEE Access* 7:24898–24909.

Journal's webpage: http://univagora.ro/jour/index.php/ijccc/

This journal is a member of, and subscribes to the principles of,
the Committee on Publication Ethics (COPE).
https://publicationethics.org/members/international-journal-computers-communications-and-control

Cite this paper as:
Alirr, O. I. (2024). Dual attention U-net for liver tumor segmentation in CT images, *International Journal of Computers Communications & Control*, 19(2), 6226, 2024.
https://doi.org/10.15837/ijccc.2024.2.6226