



Edge-Aware Multi-task Network for Integrating Quantification Segmentation and Uncertainty Prediction of Liver Tumor on Multi-modality Non-contrast MRI

Xiaojiao Xiao, Qinmin Vivian Hu, and Guanghui Wang^(✉)

Department of Computer Science, Toronto Metropolitan University, Toronto, Canada
wangcs@torontomu.ca

Abstract. Simultaneous multi-index quantification, segmentation, and uncertainty estimation of liver tumors on multi-modality non-contrast magnetic resonance imaging (NCMRI) are crucial for accurate diagnosis. However, existing methods lack an effective mechanism for multi-modality NCMRI fusion and accurate boundary information capture, making these tasks challenging. To address these issues, this paper proposes a unified framework, namely edge-aware multi-task network (EaMtNet), to associate multi-index quantification, segmentation, and uncertainty of liver tumors on the multi-modality NCMRI. The EaMtNet employs two parallel CNN encoders and the Sobel filters to extract local features and edge maps, respectively. The newly designed edge-aware feature aggregation module (EaFA) is used for feature fusion and selection, making the network edge-aware by capturing long-range dependency between feature and edge maps. Multi-tasking leverages prediction discrepancy to estimate uncertainty and improve segmentation and quantification performance. Extensive experiments are performed on multi-modality NCMRI with 250 clinical subjects. The proposed model outperforms the state-of-the-art by a large margin, achieving a dice similarity coefficient of 90.01 ± 1.23 and a mean absolute error of 2.72 ± 0.58 mm for MD. The results demonstrate the potential of EaMtNet as a reliable clinical-aided tool for medical image analysis.

Keywords: Segmentation · Quantification · Liver tumor · Multi-modality · Uncertainty

1 Introduction

Simultaneous multi-index quantification (i.e., max diameter (MD), center point coordinates (X_o , Y_o), and Area), segmentation, and uncertainty prediction of liver tumor have essential significance for the prognosis and treatment of patients [6, 16]. In clinical settings, segmentation and quantitation are manually performed by the clinicians through visually analyzing the contrast-enhanced MRI images (CEMRI) [9, 10, 18]. However, as shown in Fig. 1(b), Contrast-enhanced

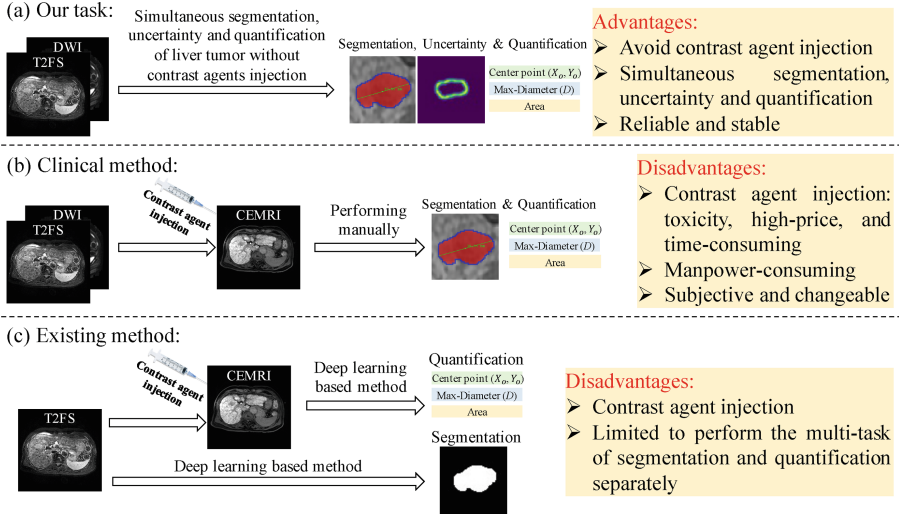


Fig. 1. Our method integrates segmentation and quantification of liver tumor using multi-modality NCMRI, which has the advantages of avoiding contrast agent injection, mutual promotion of multi-task, and reliability and stability.

MRI (CEMRI) has the drawbacks of being toxic, expensive, and time-consuming due to the need for contrast agents (CA) to be injected [2, 4]. Moreover, manually annotating medical images is a laborious and tedious process that requires human expertise, making it manpower-intensive, subjective, and prone to variation [14]. Therefore, it is desirable to provide a reliable and stable tool for simultaneous segmentation, quantification, and uncertainty analysis, without requiring the use of contrast agents, as shown in Fig. 1(a).

Recently, an increasing number of works have been attempted on liver tumor segmentation or quantification [25, 26, 28, 30]. As shown in Fig. 1(c), the work [26] attempted to use the T2FS for liver tumor segmentation, while it ignored the complementary information between multi-modality NCMRI of T2FS and DWI. In particular, there is evidence that diffusion-weighted imaging (DWI) helps to improve the detection sensitivity of focal lesions as these lesions typically have higher cell density and microstructure heterogeneity [20]. The study in [25, 30] attempted to quantify the multi-index of liver tumor, however, the approach is limited to using multi-phase CEMRI that requires the injection of CA. In addition, all these works are limited to a single task and ignore the constraints and mutual promotion between multi-tasks. Available evidence suggests that uncertainty information regarding segmentation results is important as it guides clinical decisions and helps understand the reliability of the provided segmentation. However, current research on liver tumors tends to overlook this vital task.

To the best of our knowledge, although many works focus on the simultaneous quantization, segmentation, and uncertainty in medical images (i.e., heart

[3, 5, 11, 27], kidney [17], polyp [13]). No attempt has been made to automatically liver tumor multi-task via integrating multi-modality NCMRI due to the following challenges: (1) The lack of an effective multi-modality MRI fusion mechanism. Because the imaging characteristics between T2FS and DWI have significant differences (i.e., T2FS is good at anatomy structure information while DWI is good at location information of lesions [29]). (2) The lack of strategy for capturing the accurate boundary information of liver tumors. Due to the lack of contrast agent injection, the boundary of the lesion may appear blurred or even invisible in a single NCMRI, making it challenging to accurately capture tumor boundaries [29]. (3) The lack of an associated multi-task framework. Because segmentation and uncertainty involve pixel-level classification, whereas quantification tasks involve image-level regression [11]. This makes it challenging to integrate and optimize the complementary information between multi-tasks.

In this study, we propose an edge-aware multi-task network (EaMtNet) that integrates the multi-index quantification (i.e., center point, max-diameter (MD), and Area), segmentation, and uncertainty. Our basic assumption is that the model should capture the long-range dependency of features between multi-modality and enhance the boundary information for quantification, segmentation, and uncertainty of liver tumors. The two parallel CNN encoders first extract local feature maps of multi-modality NCMRI. Meanwhile, to enhance the weight of tumor boundary information, the Sobel filters are employed to extract edge maps that are fed into edge-aware feature aggregation (EaFA) as prior knowledge. Then, the EaFA module is designed to select and fuse the information of multi-modality, making our EaMtNet edge-aware by capturing the long-range dependency of features maps and edge maps. Lastly, the proposed method estimates segmentation, uncertainty prediction, and multi-index quantification simultaneously by combining multi-task and cross-task joint loss.

The contributions of this work mainly include: (1) For the first time, multi-index quantification, segmentation, and uncertainty of the liver tumor on multi-modality NCMRI are achieved simultaneously, providing a time-saving, reliable, and stable clinical tool. (2) The edge information extracted by the Sobel filter enhances the weight of the tumor boundary by connecting the local feature as prior knowledge. (3) The novel EaFA module makes our EaMtNet edge-aware by capturing the long-range dependency of features maps and edge maps for feature fusion. The source code will be available on the author’s website.

2 Method

The EaMtNet employs an innovative approach for simultaneous tumor multi-index quantification, segmentation, and uncertainty prediction on multi-modality NCMRI. As shown in Fig. 2, the EaMtNet inputs multi-modality NCMRI of T2FS and DWI for capturing the feature and outputs the multi-index quantification, segmentation, and uncertainty. Specifically, the proposed approach mainly consists of three steps: 1) The CNN encoders for capturing feature maps and the Sobel filters for extracting edge maps (Sect. 2.1); 2) The

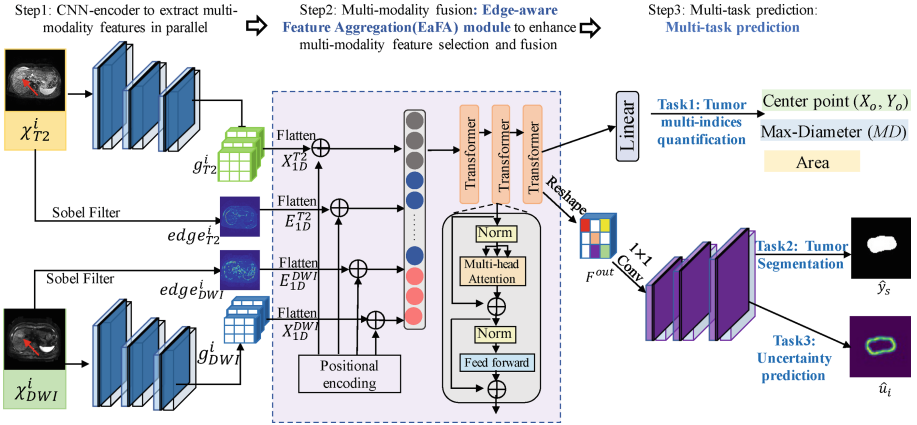


Fig. 2. Overview of the EaMtNet. It mainly consists of three steps: 1) CNN encoders for extracting local features of multi-modality NCMRI; 2) EaFA for enhancing multi-modality NCMRI feature selection and fusion; and 3) Multi-task prediction.

edge-aware feature aggregation (EaFA) for multi-modality feature selection and fusion via capturing the long-distance dependence (Sect. 2.2); and 3) Multi-task prediction module (Sect. 2.3).

2.1 CNN Encoder for Feature Extraction

In Step 1 of Fig. 2, the multi-modality NCMRI (i.e., $\chi_{T2}^i \in R^{H \times W}$, $\chi_{DWI}^i \in R^{H \times W}$) are fed into two parallel encoders and the Sobel filter to extract the feature maps (i.e., $g_{T2}^i \in R^{H \times W \times N}$, $g_{DWI}^i \in R^{H \times W \times N}$) and the corresponding edge maps (i.e., $edge_{T2}^i \in R^{H \times W}$, $edge_{DWI}^i \in R^{H \times W}$) respectively. Specifically, EaMtNet employs UNet as the backbone for segmentation because the CNN encoder has excellent capabilities in low-range semantic information extraction [15]. The two parallel CNN encoders have the same architecture where each encoder contains three shallow convolutional network blocks to capture features of adjacent slices. Each conv block consists of a convolutional layer, batch normalization, ReLU, and non-overlapping subsampling. At the same time, EaMtNet utilizes the boundary information extracted by the Sobel filter [19] as prior knowledge to enhance the weight of tumor edge information to increase the awareness of the boundary.

2.2 Edge-Aware Feature Aggregation(EaFA) for Multi-modality Feature Selection and Fusion

In Step 2 of the proposed model, the feature maps (i.e., g_{T2}^i , g_{DWI}^i) and the edge maps (i.e., $edge_{T2}^i$, $edge_{DWI}^i$) are fed into EaFA for multi-modality feature fusion with edge-aware. In particular, the EaFA makes the EaMtNet edge-aware by using the Transformer to capture the long-range dependency of feature maps

and edge maps. Specifically, the feature maps and edge maps are first flattened to the 1D sequence corresponding to $X_{1D} \in R^{N \times P^2}$ and $E_{1D} \in R^{2 \times Q^2}$, respectively. Where $N = 2 \times C$ means the channel number C of the last convolutional layer from the two parallel encoders. (P, P) and (Q, Q) represent the resolution of each feature map and each edge map, respectively. On the basis of the 1D sequence, to make the feature fusion with edge awareness, the operation of position encoding is performed not only on feature maps but also on edge maps. The yielded embeddings $Z \in R^{N \times P^2 + 2 \times Q^2}$ can serve as the input sequence length for the multi-head attention layer in Transformer. The following operations in our EaFA are similar to the traditional Transformer [22]. After the three cascade Transformer layers, the EaFA yields the fusion feature vector F for multi-task prediction. The specific computation of the self-attention matrix and multi-head attention are defined below [22]:

$$Attention(Q, K, V) = softmax(\frac{QK^T}{\sqrt{d_k}})(V) \quad (1)$$

$$MultiHead(Q, K, V) = Concat(head_1, \dots, head_h)W^O \quad (2)$$

$$head_i = Attention(QW_i^Q, KW_i^K, VW_i^V) \quad (3)$$

where query Q , key K , and value V are all vectors of the flattened 1D sequences of X_{1D} and E_{1D} . W_i^O is the projection matrix, and $\frac{1}{\sqrt{d_k}}$ is the scaling factor.

2.3 Multi-task Prediction

In Step 3 of Fig. 2, the EaMtNet outputs the multi-modality quantification \hat{y}_Q (i.e., MD, X_o , Y_o and Area), segmentation result \hat{y}_s and uncertainty map \hat{u}_i . Specifically, for the quantification path, \hat{y}_Q is directly obtained by performing a linear layer to the feature F from EaFA. For the segmentation and uncertainty path, the output feature F from EaFA is first reshaped into a 2D feature map F^{out} . Then, to scale up to higher-resolution images, a 1×1 convolution layer is employed to change the channel number of F^{out} for feeding into the decoder. After upsampling by the CNN decoder, EaMtNet predicts the segmentation result \hat{y}_s with $H \times W$ and uncertainty map \hat{u}_i with $H \times W$. The CNN decoder contains three shallow deconv blocks, which consist of deconv layer, batch normalization, and ReLU. Inspired by [24], we select the entropy map as our uncertainty measure. Given the prediction probability after softmax, the entropy map is computed as follows:

$$H[x] = - \sum_{i=1}^K z_i(x) * \log_2(z_i(x)) \quad (4)$$

where z_i is the probability of pixel x belonging to category i . When a pixel has high entropy, it means that the network is uncertain about its classification. Therefore, pixels with high entropy are more likely to be misclassified. In other words, its entropy will decrease when the network is confident in a pixel's label.

Under the constraints of uncertainty, the EaMtNet can effectively rectify the errors in tumor segmentation because the uncertainty estimation can avoid overconfidence and erroneous quantification [23]. Moreover, the EaMtNet novelly make represent different tasks in a unified framework, leading to beneficial interactions. Thus, the quantification performance is improved through back-propagation by the joint loss function $L_{multi-task}$. The function comprises segmentation loss L_{seg} and quantification loss L_{qua} , where the loss function L_{seg} is utilized for optimizing tumor segmentation, and L_{qua} is utilized for optimization of multi-index quantification. It can be defined as:

$$L_{Dice} = \frac{2 \sum_i^N y_i \hat{y}_i}{\sum_i^S y_i^2 + \sum_i^N \hat{y}_i^2} \quad (5)$$

$$L_{qua}(\hat{y}_{task}^i, y_{task}^i) = \sum_{i=1} |y_{task}^i - \hat{y}_{task}^i| \quad (6)$$

where \hat{y}_s represents the prediction, and y_i represents the ground truth label. The sum is performed on S pixels, \hat{y}_{task}^i represents the predicted multi-index value, and y_{task}^i represents the ground truth of multi-index value, $task \in \{MD, X, Y, Area\}$.

3 Experimental Results and Discussion

For the first time, EaMtNet has achieved high performance with the dice similarity coefficient (DSC) up to $90.01 \pm 1.23\%$, and the mean absolute error (MAE) of the MD, X_o , Y_o and Area are down to 2.72 ± 0.58 mm, 1.87 ± 0.76 mm, 2.14 ± 0.93 mm and 15.76 ± 8.02 cm^2 , respectively.

Dataset and Configuration. An axial dataset includes 250 distinct subjects, each underwent initial standard clinical liver MRI protocol examinations with corresponding pre-contrast images (T2FS [4mm]) and DWI [4mm]) was collected. The ground truth was reviewed by two abdominal radiologists with 10 and 22 years of experience in liver imaging, respectively. If any interpretations demonstrated discrepancies between the reviewers, they would re-evaluate the examinations together and reach a consensus. To align the paired images of T2 and DWI produced at different times. We set the T2 as the target image and the DWI as the source image to perform the pre-processing of non-rigid registration between T2 and DWI by using the Demons non-rigid registration method. It has been widely used in the field of medical image registration since it was proposed by Thirion [21]. We perform the Demons non-rigid registration on an open-source toolbox DIRART using Matlab 2017b.

Inspired by the work [22], we set the scaling factor d_k to 64 in equation (1). All experiments were assessed with a 5-fold cross-validation test. To quantitatively evaluate the segmentation results, we calculated the dice coefficient scores (DSC) metric that measures the overlapping between the segmentation prediction and ground truth [12]. To quantitatively evaluate the quantification results,

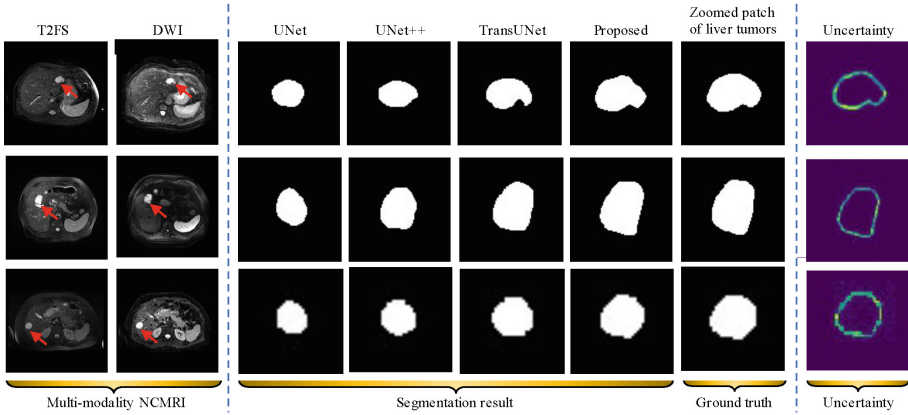


Fig. 3. The comparison of segmentation results between the proposed EaMtNet and three SOTA methods. The results show that our network yields high performance.

Table 1. Segmentation and quantification performance of the EaMtNet under different configurations. DSC is used to evaluate the segmentation performance. MAE is used to evaluate the quantification performance.

	DSC(%)	MD(mm)	$X_o(mm)$	$Y_o(mm)$	Area(cm^2)
No-EaFA	85.82 ± 1.92	3.49 ± 0.94	2.51 ± 1.43	3.12 ± 1.84	28.55 ± 9.75
No-Uncertainty	88.37 ± 2.71	3.25 ± 0.77	2.36 ± 0.92	2.78 ± 1.18	24.15 ± 9.19
Our method	90.01 ± 1.23	2.72 ± 0.58	1.87 ± 0.76	2.14 ± 0.93	15.76 ± 8.02

we calculated the mean absolute error (MAE). Our EaMtNet was implemented using Ubuntu 18.04 platform, Python v3.6, PyTorch v0.4.0, and running on two NVIDIA GTX 3090Ti GPUs.

Accurate Segmentation. The segmentation performance of EaMtNet has been validated and compared with three state-of-the-art (SOTA) segmentation methods (TransUNet [1], UNet [15], and UNet++ [31]). Furthermore, to ensure consistency in input modality, the channel number of the first convolution layer in the three comparison methods is set to 2. The visual examples of liver tumors are shown in Fig. 3, it is evident that our proposed EaMtNet outperforms the three SOTA methods. Some quantitative analysis results are shown in Table 1 and Table 2, our network achieves high performance with the DSC of $90.01 \pm 1.23\%$ (5.39% higher than the second-best). The results demonstrate that edge-aware, multi-modality fusion, and uncertainty prediction are essential for segmentation.

Ablation Study. To verify the contributions of edge-aware feature aggregation (EaFA) and uncertainty, we performed ablation study and compared and performance of different networks. First, we removed the EaFA and used concatenate,

Table 2. The quantitative evaluation of segmentation. DSC is used to evaluate the performance of our EaMtNet and three SOTA methods.

	UNet	UNet++	TransUNet	Our method
DSC(%)	76.59 ± 1.86	80.97 ± 2.37	84.62 ± 1.45	90.01 ± 1.23

Table 3. The quantitative evaluation of the multi-index quantification. The criteria of MAE is used to evaluate the performance of our EaMtNet and two SOTA methods.

	MD(mm)	X(mm)	Y(mm)	Area(cm^2)
ResNet-50	6.24 ± 2.81	3.15 ± 1.25	3.38 ± 1.27	31.32 ± 8.47
DenseNet	4.85 ± 1.67	2.73 ± 0.89	2.95 ± 1.15	25.37 ± 7.63
Our method	2.72 ± 0.58	1.87 ± 0.76	2.14 ± 0.93	15.76 ± 8.02

meaning we removed fusion multi-modality (No-EaFA). Then, we removed the uncertainty task (No-Uncertainty). The quantitative analysis results of these ablation studies are shown in Table 1. Our method exhibits high performance in both segmentation and quantification, indicating that each component of the EaMtNet plays a vital role in liver tumor segmentation and quantification.

Performance Comparison with State-of-the-Art. The EaMtNet has been validated and compared with three SOTA segmentation methods and two SOTA quantification methods (i.e., ResNet-50 [7] and DenseNet [8]). Furthermore, the channel number of the first convolution layer in the two quantification comparison methods is set to 2 to ensure the consistency of input modalities. The visual segmentation results are shown in Fig. 3. Moreover, the quantitative results (as shown in Table 2) corresponding to the visualization results (i.e., Fig. 3) obtained from the existing experiments further demonstrate that our method outperforms the three SOTA methods. Specifically, compared with the second-best approach, the DSC is boosted from $84.62 \pm 1.45\%$ to $90.01 \pm 1.23\%$. The quantitative analysis results are shown in Table 3. It is evident that our method outperforms the two SOTA methods with a large margin in all metrics, owing to the proposed multi-modality fusing and multi-task association.

4 Conclusion

In this paper, we have proposed an EaMtNet for the simultaneous segmentation and multi-index quantification of liver tumors on multi-modality NCMRI. The new EaFA enhances edge awareness by utilizing boundary information as prior knowledge while capturing the long-range dependency of features to improve feature selection and fusion. Additionally, multi-task leverages the prediction discrepancy to estimate uncertainty, thereby improving segmentation and quantification performance. Extensive experiments have demonstrated the proposed

model outperforms the SOTA methods in terms of DSC and MAE, with great potential to be a diagnostic tool for doctors.

Acknowledgements. This work is partly supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) and TMU FOS Postdoctoral Fellowship.

References

1. Chen, J., et al.: Transunet: transformers make strong encoders for medical image segmentation. arXiv preprint [arXiv:2102.04306](https://arxiv.org/abs/2102.04306) (2021)
2. Danet, I.M., Semelka, R.C., Braga, L.: Mr imaging of diffuse liver disease. *Radiologic Clinics* **41**(1), 67–87 (2003)
3. Du, X., Tang, R., Yin, S., Zhang, Y., Li, S.: Direct segmentation-based full quantification for left ventricle via deep multi-task regression learning network. *IEEE J. Biomed. Health Inform.* **23**(3), 942–948 (2018)
4. Fishbein, M., et al.: Hepatic MRI for fat quantitation: its relationship to fat morphology, diagnosis, and ultrasound. *J. Clin. Gastroenterol.* **39**(7), 619–625 (2005)
5. Ge, R., et al.: K-net: integrate left ventricle segmentation and direct quantification of paired echo sequence. *IEEE Trans. Med. Imaging* **39**(5), 1690–1702 (2019)
6. Gonzalez-Guindalini, F.D., et al.: Assessment of liver tumor response to therapy: role of quantitative imaging. *Radiographics* **33**(6), 1781–1800 (2013)
7. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778 (2016)
8. Huang, G., Liu, Z., Van Der Maaten, L., Weinberger, K.Q.: Densely connected convolutional networks. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4700–4708 (2017)
9. Huo, J., Wu, J., Cao, J., Wang, G.: Supervoxel based method for multi-atlas segmentation of brain MR images. *Neuroimage* **175**, 201–214 (2018)
10. Lee, Y.J., et al.: Hepatocellular carcinoma: diagnostic performance of multidetector CT and MR imaging—a systematic review and meta-analysis. *Radiology* **275**(1), 97–109 (2015)
11. Luo, G., et al.: Commensal correlation network between segmentation and direct area estimation for bi-ventricle quantification. *Med. Image Anal.* **59**, 101591 (2020)
12. Milletari, F., Navab, N., Ahmadi, S.A.: V-net: fully convolutional neural networks for volumetric medical image segmentation. In: *2016 Fourth International Conference on 3D Vision (3DV)*, pp. 565–571. IEEE (2016)
13. Patel, K.B., Li, F., Wang, G.: Fuzzynet: a fuzzy attention module for polyp segmentation. In: *NeurIPS’22 Workshop*
14. Petittlerc, L., Sebastiani, G., Gilbert, G., Cloutier, G., Tang, A.: Liver fibrosis: review of current imaging and MRI quantification techniques. *J. Magn. Reson. Imaging* **45**(5), 1276–1295 (2017)
15. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) *MICCAI 2015. LNCS*, vol. 9351, pp. 234–241. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24574-4_28
16. Rovira, À., León, A.: MR in the diagnosis and monitoring of multiple sclerosis: an overview. *Eur. J. Radiol.* **67**(3), 409–414 (2008)

17. Ruan, Y., et al.: MB-FSGAN: joint segmentation and quantification of kidney tumor on CT by the multi-branch feature sharing generative adversarial network. *Med. Image Anal.* **64**, 101721 (2020)
18. Sirlin, C.B., et al.: Consensus report from the 6th international forum for liver MRI using gadoteric acid. *J. Magn. Reson. Imaging* **40**(3), 516–529 (2014)
19. Sobel, I., Feldman, G.: A 3x3 isotropic gradient operator for image processing. a talk at the Stanford Artificial Project in pp. 271–272 (1968)
20. Tang, L., Zhou, X.J.: Diffusion MRI of cancer: from low to high b-values. *J. Magn. Reson. Imaging* **49**(1), 23–40 (2019)
21. Thirion, J.P.: Non-rigid matching using demons. In: *Proceedings CVPR IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 245–251. IEEE (1996)
22. Vaswani, A., et al.: Attention is all you need. In: *Advances in Neural Information Processing Systems* 30 (2017)
23. Wang, G., Li, W., Aertsen, M., Deprest, J., Ourselin, S., Vercauteren, T.: Aleatoric uncertainty estimation with test-time augmentation for medical image segmentation with convolutional neural networks. *Neurocomputing* **338**, 34–45 (2019)
24. Wang, H., Wang, Y., Zhang, Q., Xiang, S., Pan, C.: Gated convolutional neural network for semantic segmentation in high-resolution images. *Remote Sens.* **9**(5), 446 (2017)
25. Xiao, X., Zhao, J., Li, S.: Task relevance driven adversarial learning for simultaneous detection, size grading, and quantification of hepatocellular carcinoma via integrating multi-modality mri. *Med. Image Anal.* **81**, 102554 (2022)
26. Xiao, X., Zhao, J., Qiang, Y., Chong, J., Yang, X.T., Kazihise, N.G.-F., Chen, B., Li, S.: Radiomics-guided GAN for segmentation of liver tumor without contrast agents. In: Shen, D., Liu, T., Peters, T.M., Staib, L.H., Essert, C., Zhou, S., Yap, P.-T., Khan, A. (eds.) *MICCAI 2019. LNCS*, vol. 11765, pp. 237–245. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-32245-8_27
27. Xu, C., Howey, J., Ohorodnyk, P., Roth, M., Zhang, H., Li, S.: Segmentation and quantification of infarction without contrast agents via spatiotemporal generative adversarial learning. *Med. Image Anal.* **59**, 101568 (2020)
28. Zhang, D., Chen, B., Chong, J., Li, S.: Weakly-supervised teacher-student network for liver tumor segmentation from non-enhanced images. *Med. Image Anal.* **70**, 102005 (2021)
29. Zhao, J., et al.: United adversarial learning for liver tumor segmentation and detection of multi-modality non-contrast MRI. *Med. Image Anal.* **73**, 102154 (2021)
30. Zhao, J., et al.: mfTrans-Net: quantitative measurement of hepatocellular carcinoma via multi-function transformer regression network. In: de Bruijne, M., Cat-tin, P.C., Cotin, S., Padoy, N., Speidel, S., Zheng, Y., Essert, C. (eds.) *MICCAI 2021. LNCS*, vol. 12905, pp. 75–84. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-87240-3_8
31. Zhou, Z., Siddiquee, M.M.R., Tajbakhsh, N., Liang, J.: Unet++: redesigning skip connections to exploit multiscale features in image segmentation. *IEEE Trans. Med. Imaging* **39**(6), 1856–1867 (2019)