Laplacian Salience-Gated Feature Pyramid Network for Accurate Liver Vessel Segmentation

Zhan Gao, Qiuhao Zong, Yiqi Wang[ID], Yan Yan, Yuqing Wang[ID], Ning Zhu, Jin Zhang[ID],
Yunfu Wang, and Liang Zhao[ID]

Abstract—Liver vessels generated from computed tomography are usually pretty small, which poses major challenges for satisfactory vessel segmentation, including 1) the scarcity of high-quality and large-volume vessel masks, 2) the difficulty in capturing vessel-specific features, and 3) the heavily imbalanced distribution of vessels and liver tissues. To advance, a sophisticated model and an elaborated dataset have been built. The model has a newly conceived Laplacian salience filter that highlights vessel-like regions and suppresses other liver regions to shape the vessel-specific feature learning and to balance vessels against others. It is further coupled with a pyramid deep learning architecture to capture different levels of features, thus improving the feature formulation. Experiments show that this model markedly outperforms the state-of-the-art approaches, achieving a relative improvement of Dice score by at least 1.63% compared to the existing best model on available datasets. More promisingly, the averaged Dice score produced by the existing models on the newly constructed dataset is as high as 0.734 ± 0.070, which is at least 18.3% higher than that obtained from the existing best dataset under the same settings. These observations suggest that the proposed Laplacian salience, together with the elaborated dataset, can be helpful for liver vessel segmentation.

Index Terms—Vessels, segmentation, computed tomography, neural network, pattern recognition and classification.

Zhan Gao, Qiuhao Zong, Yiqi Wang, Yan Yan, and Jin Zhang are with the Taihe Hospital, Hubei University of Medicine, Shiyan 442000, China (e-mail: gaozhan@taihehospital.com; zongqiuhao@taihehospital.com; wangyiqi@taihehospital.com; yanyan@taihehospital.com; zhjinwhu@taihehospital.com).

Yuqing Wang and Ning Zhu are with the School of Biomedical Engineering, Hubei University of Medicine, Shiyan 442000, China (e-mail: 3347460663@qq.com; 1664891351@qq.com).

Yunfu Wang is with the Taihe Hospital/Hubei Key Laboratory of Embryonic Stem Cell Research, Hubei University of Medicine, Shiyan 442000, China (e-mail: yfwang@taihehospital.com).

Liang Zhao is with the Taihe Hospital/Hubei Key Laboratory of Embryonic Stem Cell Research, Hubei University of Medicine, Shiyan 442000, China, and also with the School of Computing and Electronic Informatics, Guangxi University, Nanning 530004, China (e-mail: S080011@e.ntu.edu.sg).

Digital Object Identifier 10.1109/TMI.2023.3273528

I. Introduction

LIVER vessel segmentation of medical images generated from computed tomography (CT) provides rich information for disease diagnosis, radiotherapy positioning and surgical planning [1]. Therefore, various computational models have been developed for this task, including traditional machine learning-based and deep learning-based.

The traditional methods mainly use image filters (e.g., Frangi [2], Hessian [3], Laplacian of Gaussian [4]), clustering algorithms [5], adaptive thresholding [6], active contour models [7], [8] to pinpoint out liver vessels. While the deep learning models are more prevalent and effective due to the revolutionary design of network architectures as well as the powerful devices. The fruitful innovations in this category are mainly two-fold: new loss functions and novel architectures. The former includes active contour loss [9], boundary loss [10], Dice loss [11], Tversky loss [12], etc.; while the latter attracts even more intensive efforts, e.g., the U-Net based models [13], [14], attention-based model [15], multi-path convolutional network [16], etc.

Although dozens of methods have been proposed from every perspective, the performance is still unsatisfactory. In fact, liver vessel segmentation has never been a trivial task because vessels usually have tiny size, low contrast and heavy noise, particularly the small size, which leads to the heavily skewed distribution of vessels versus other liver tissues. Taking Fig. 1, all the ratios between vessel pixels and liver pixels contained in the raw images are less than 0.1, some even less than 0.05, presenting great challenges for shaping vessel-specific feature learning. Another consequence of the smallness is the difficulty of formulating contextual features from vessel areas. Taking Fig. 1 again, many vessels only occupy dozens of, even less than ten pixels, resulting in very limited information to be retrieved. What's more, it is a painstaking task to delineate vessels from livers due to the tiny regions they occupy, hence the scarcity of large-volume and high-quality data. These challenges collectively lead to the poor performance of existing models.

To overcome these shortcomings, we propose a novel Laplacian-based salience generator that is efficient and effective in highlighting vessel-like areas and suppressing others. For instance, by applying the filter to the liver images shown in Fig. 1, the averaged vessel-to-liver weighted pixel-level ratio is increased to 2.73 times higher than the original one. More
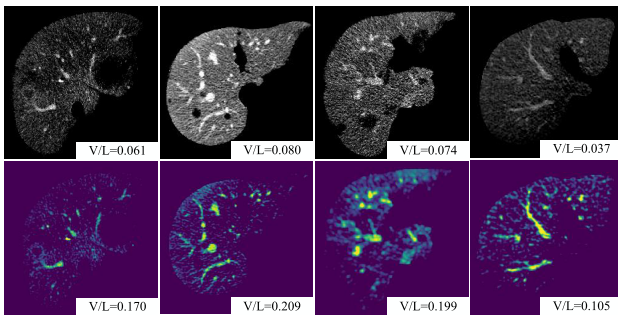
Fig. 1. Effectiveness of the proposed Laplacian salience filter. The first row contains the original images and the second is the Laplacian salience filtered ones. The value at the bottom right corner is the ratio between the weighted number of vessels and other liver tissues, where the weights are equal before filtering and are determined automatically by Laplacian salience after filtering.

importantly, the filter is insensitive to the phase of images, which can also be observed in Fig. 1. By integrating the filter into a pyramid deep learning architecture, the features can be more effectively tailored to the vessel areas, and the skewed distribution can be automatically balanced as well. In addition, the pyramid network is able to capture different levels of features and fuse them together for better vessel segmentation. The Laplacian salience automatically shapes the model focusing on vessel-like regions by attaching higher weights to these pixels and lower weights to others, thus paying more attention to vessel-like areas. For multi-scale feature fusion, it can adapt to various vessel sizes and compensate for weak features learned from small vessels by large vessels, so that the learned features can be applied to different situations.

Besides the sophisticated model, another big issue that hinders the progress of liver vessel segmentation is the small and imperfect dataset, which is also caused by the tiny size of vessels. To construct a qualified large dataset, great endeavors have been made by our team. In particular, 532 volumes of contrast-enhanced abdominal computed tomography images were collected, and the vessels were carefully delineated by three radiologists elaborately.

The big jump in Dice scores generated by our model, as well as from our data, demonstrates their goodness. The source codes and data are available at https://github.com/lzhLab/LiVS.

II. RELATED WORKS

A. Salience Map

A salience map highlights the focus of the human visual system, which has been demonstrated to be effective in dealing with inhomogeneous distributions as well as intra-class variations in visual tasks [17]. There have two typical branches related to salience, namely salience prediction and salience application. Salience prediction has been extensively studied from different perspectives, such as data-driven methods [18], [19], generative adversarial network-based models [20], [21], et al. Regarding salience application, it is mostly used to facilitate downstream data analysis, where the salience itself is mainly obtained by attention-based models and traditional machine learning-based methods. For instance, the

self-attention model [22] as well as its spatial variations [23], [24] and channel variations [25], [26] have been proposed to capture the salience of objects. Another work, the attention gate [27], has also been proposed to suppress irrelevant regions and highlight salient features. This study is similar to ours in using salience but different in generating salience. Although these implicit model-driven salience generators have been proven to be effective in real-world tasks, they require a large amount of high-quality training data. This restriction can be overcome by traditional approaches, such as Conditional Random Field-based [28], Bayesian inference-based [29], and clustering-based [30], but at the cost of accuracy.

B. Loss Weighting

Loss weighting is dissimilar to salience. It assigns different weights to different classes, samples, and even pixels, while salience varies in weight across different regions, which may be composed of multiple classes. Loss weighting for image analysis can be categorized into three clusters typically, i.e., class-wise [31], [32], sample-wise [11], [12], [33], and pixel-wise [9], [10]. The class-wise weighting assigns different weights to different classes to balance the uneven distributions; the sample-wise weighting assigns different weights to different samples, even if they are from the same class, so that more attention could be paid to hard samples; and the pixel-wise weighting is the most fine-grained strategy, which can also apply different weights to different classes and samples. Among them, pixel-wise weighting is more popular. For instance, Focal loss [34] is a common weighting approach for different pixels based on the predicted probability, while contour loss [9] and boundary loss [10] have been adopted to focus on ambiguous regions across different classes.

C. Tiny Object Segmentation

The cross-section of a vessel is a tiny blob since the branches of vessels are thin tube-like structures. Therefore, vessel localization in liver is similar to tiny object segmentation. Several intrinsic difficulties hinder the accurate segmentation of tiny objects, including information loss, noisy feature representation, insufficient samples and sensitive to disturbance [35]. To overcome these limitations, extensive efforts have been made from various perspectives [36], such as data augmentation [37], soft label assignment [38], scale-specific segmentation [39], feature reassembly [40], attention-based segmentation [41], similarity-aware learning [42], super-resolution-based segmentation [43], context-aware modeling [44], focus-aware segmentation [16], etc. The strategies presented here are not used separately, but in most cases, they are combined together to achieve better performance. For instance, Liu et al. [45] propose a dual-branch network with dual-sampling modulated Dice loss for hard exudate segmentation from color fundus images, in which each branch is an exudate size-specific feature learner so that more attention can be paid to small exudates. It also oversamples small exudates that are difficult to segment. In another study, Meng et al. [46] combine multi-scale feature fusion and the
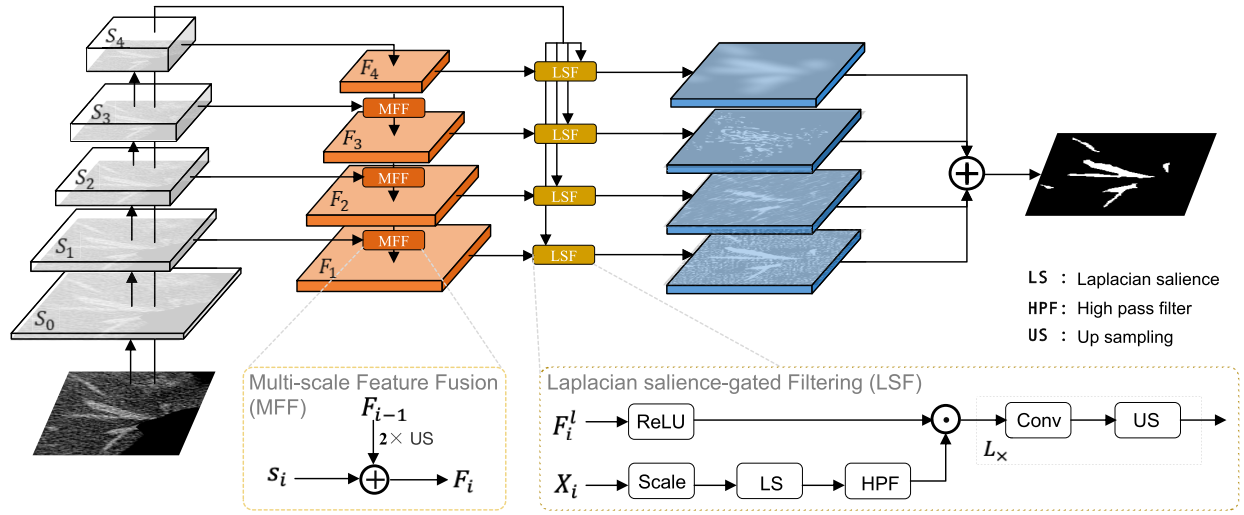
Fig. 2. Architecture of the Laplacian salience-gated feature pyramid network. It is mainly composed of two parts, including pyramid feature fusion and the Laplacian salience-gated feature filtering. The pyramid network is for multi-scale feature extraction, while the Laplacian salience is to shape the vessel-specific feature learning. Note that the Laplacian salience is applied to different scales of the input image other than the corresponding feature maps, and the final feature map is the integration of the feature maps obtained from different scales with size matched.

Transformer for fine-grained segmentation of liver tumors and vessels.

Note that object segmentation is a hot topic in computer vision, and hundreds of models have been proposed. For more details, please refer to the recent survey [36].

III. METHODS

Our model consists primarily of two parts: Laplacian salience-gated filtering and pyramid feature fusion; *cf.* Fig. 2.

A. Laplacian Salience-Gated Filtering

1) Laplacian: The Laplacian, or Laplacian operator, is the second derivative of the n-dimensional Euclidean space, defined as

$$\Delta f = \nabla \cdot \nabla f = \sum_{i=1}^{n} \frac{\partial^2 f}{x_i^2}, \tag{1}$$

where f is a function of $\mathbf{x} = \langle x_1, x_2, \cdots, x_n \rangle$. For two-dimensional space,

$$\Delta f(x, y) = \frac{\partial^2 f}{x^2} + \frac{\partial^2 f}{y^2}. \tag{2}$$

Note, the x and y are the relative distance to the interesting point (x_i, y_i), not the absolute coordinates.

Since the second derivative is very sensitive to noise, the Laplacian is further smoothed by a Gaussian, resulting in

$$\Delta f^s(x, y) = C \Delta f(x, y) e^{-\frac{x^2 + y^2}{2\sigma^2}}, \tag{3}$$

with σ the standard deviation of the Gaussian envelope and C the normalization factor.

2) Laplacian Wavelet: The Laplacian wavelet in 2D space is the Laplacian (the scaling function) modulated by a complex exponential (the wavelet function), defined as

$$g(x, y; \lambda, \phi) = \Delta f^s(x, y) \cdot e^{i(\frac{2\pi}{\lambda}(x \cos\theta + y \sin\theta) + \phi)}, \tag{4}$$

where λ is the wavelength and ϕ is the phase offset.

The Laplacian wavelet is isotropic in its original form, but various patterns should be captured, so it is further refined as

$$g(x, y; \lambda, \phi, \gamma, \theta, \sigma) = \Delta f^s(x', y') e^{i(\frac{2\pi}{\lambda} x' + \phi)} \tag{5}$$

$$= C \Delta f(x', y') e^{-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}} e^{i(\frac{2\pi}{\lambda} x' + \phi)}, \tag{6}$$

where $x' = x \cos\theta + y \sin\theta$, $y' = -x \sin\theta + y \cos\theta$, γ is the spatial aspect ratio between x and y, and θ is the rotation angle from (x, y) to (x', y'). According to the convolution theorem, the final operator applied to f can be further written as

$$g(x, y) = C(1 + \gamma - \frac{x'^2 + \gamma^2 y'^2}{\sigma^2}) e^{-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}} e^{i(\frac{2\pi}{\lambda} x' + \phi)}. \tag{7}$$

For simplicity, the parameters are ignored. In addition, the normalization factor C can be dropped out.

3) Laplacian Salience: According to the Euler's formula that $e^{ix} = \cos x + i \sin x$, the above wavelet can be split into the real and imaginary two parts, with

$$g^R(x, y) = A \cos(2\pi \frac{x}{\lambda} + \phi) \tag{8}$$

$$g^I(x, y) = A \sin(2\pi \frac{x}{\lambda} + \phi), \tag{9}$$

where

$$A = (1 + \gamma - \frac{x'^2 + \gamma^2 y'^2}{\sigma^2}) e^{-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}}. \tag{10}$$

Let

$$\Lambda^R(x, y) = \sum_{i}^{K} g^R(x, y; \lambda_i, \phi_i, \gamma_i, \theta_i, \sigma_i) \tag{11}$$

be the integral of K patterns, then by convolving $\Lambda^R(x, y)$ on an image $f(x, y)$ the specific patterns contained in $f(x, y)$ can be fished out. Analogously, the patterns defined by $\Lambda^I(x, y)$ can also be obtained from $f(x, y)$ by convolution as well.

Based on the above two integrated banks of filters, the salience map s of an image f can be obtained by

$$s(x, y) = h((f(x, y) - \bar{f}) * \Lambda^R(x, y))$$
$$+ h((f(x, y) - \bar{f}) * \Lambda^I(x, y)) \tag{12}$$

where $\bar{f}$ is the mean intensity of f, $*$ represents convolution operation and $h(\cdot)$ is a normalization function.

It is known that $\Lambda^I(x, y)$ is particularly helpful in highlighting sudden changes, such as edges, steps and corners, because of the sign change over the 0-point. To further enhance the constrastiveness, the wavelength λ is optimized to $2\sqrt{3 + \gamma}\sigma$ by solving the first-order derivative of $\Lambda^I(x, y)$ at 0. For $\Lambda^R(x, y)$, it is used to smooth the saliency map so as to withstand noise and intensity variations. Similarly, its wavelength is optimized to $4\sqrt{1 + \gamma}\sigma$.

The normalization and subtraction are critical to the generation of saliency maps. Normalization is necessary because the optimized basic wavelets are different between the odd and even waves, while subtraction is used to remove the background so that the abnormal regions can stand out.

B. Pyramid Feature Fusion

1) Model Architecture: A novel salience-gated feature pyramid architecture is proposed to distinguish vessels from other tissues; cf. Fig 2. The feature pyramid network (FPN) [47] and its variants have been intensively explored for object detection and segmentation [48] due to its superiority in multi-scale feature fusion. However, in its original form, it is unsuitable for vessel segmentation due to the erosion, or even elimination, of small vessels during downscaling. To this end, we refine the architecture in two-fold: i) bottom-up zoom-in and top-down zoom-out to enlarge tiny vessels and obtain more detailed patterns, and ii) automatic highlighting of vessels boundaries by a high-pass filter on the salience maps at different scales.

Let $X \in \mathbb{R}^{w \times h \times c}$ be an image with width w, height h, and channel c. The bottom-up zoom-in of X is carried out by passing it through a ResNet50 block [49] at each reverse pyramid level. The input of the higher level is the output of the lower level with up-sampling, which scales the output two times larger. The top-down zoom-out is conducted by adding the down-sampled output of the upper level and the feature maps of the same level in the bottom-up path. See the detailed illustration in Fig 2. The zooming operation (both in and out) is repeated three times, but not more, because of the following two main reasons: i) the size of feature maps grows exponentially with zoom-in, thus easily leading to memory exhaustion if scaled too deep; ii) three times of up-scaling ($2^3 \times 2^3 = 64$ times larger in the area) is sufficient to enlarge a tiny vessel enough for exquisite segmentation.

2) Salience-Gated Feature Generation: Unlike traditional feature generation, we modulate the feature maps with gates. Let $X^{(l)}$ be an image, either the original input or its enlarged form, at layer l, and $F^{(l)}$ be the feature maps produced at layer l. The gate $G(X^{(l)})$ to be applied to $F^{(l)}$ is calculated as

$$G(X^{(l)}) = \text{IFFT}(\text{FFT}^{\text{high-pass}}(S(X^{(l)}); \rho)), \tag{13}$$

where $S(X^{(l)})$ is the Laplacian salience of $X^{(l)}$, $\text{FFT}^{\text{high-pass}}(x)$ is the high-pass of fast Fourier transformation of x, $\text{IFFT}(x)$ is the inverse fast Fourier transformation of x, and ρ is the threshold to determine the high-pass band. Based on $G(X^{(l)})$, the modulated feature map of $F^{(l)}$ is calculated by

$$F^{(l)'} = G(X^{(l)}) \odot \sigma(F^{(l)}), \tag{14}$$

where $\odot$ is the element-wise dot product, and $\sigma(\cdot)$ is an activation function.

Intuitively, a salience map is helpful to focus on vessel regions as well as those ambiguous regions that resemble vessels. However, texture information is very limited and varies significantly due to the large number of tiny vessels presented. To this end, we further shape the importance of feature maps on the boundary regions, hence the high pass filter. There have two main reasons: i) the boundary regions are more consistent across different vessels regardless of their size, and ii) the boundary labels are imperfect, so more attention to them improves segmentation.

3) Salience-Modulated Loss: The loss function of the model is guided by the salience of the input image X, which is computed by

$$L = -\frac{1}{MN} \sum_{i}^{M} \sum_{j}^{N} [\alpha S(X_{ij})Y_{ij}\ln P_{ij}$$
$$+ (1 - \alpha)S(X_{ij})(1 - Y_{ij})\ln(1 - P_{ij})], \tag{15}$$

where Y is the label, P is the prediction, $S(X)$ is the salience of X, and α is the balancing factor within the range of 0 to 1.

Our salience-modulated loss is ostensibly similar to Focal loss [34], but de facto different. Ours is determined from the raw input, reflecting the difficulty between vessels are other regions; while Focal loss is determined from the model, unveiling the models' distinguishability, which unfortunately can vary significantly with different inputs.

Note that, the loss modulation is slightly different from the feature gating, in which the high-pass filtering is excluded in the former one. This is because the final target is the entire vessel area, while the features are primarily based on the boundary regions.

C. Implementation Details

The proposed model is implemented in PyTorch and validated on a cluster having four NVIDIA V100 GPUs. The learning rate is optimized by Adam [50] with an initial value of 10^{-2} and a reduction factor of 0.1, while the batch size is 8.

The salience of some liver tissues, which are very different from vessels, can be 0, so they can hardly contribute to model training. To further boost the learning of non-vessel regions, the lower bound of the salience is set to 0.05.

IV. EXPERIMENTS

A. Datasets

Two types of datasets are used to evaluate the performance of our model as well as others, including the widely used public datasets and our newly constructed dataset. The public ones are 3D-IRCADb (3D Image Reconstruction for Comparison of Algorithm Database) [51] and MSD (Medical Segmentation Decathlon) [52], while ours is LiVS (short for <u>Li</u>ver <u>V</u>essel <u>S</u>egmentation).

The 3D-IRCADb contains 20 volumes of CT images having 2,823 slices with vessel masks, and MSD consists of 303 volumes and 21,120 slices having vessel masks.

LiVS has 532 volumes and 15,984 slices, where each slice is delineated by three senior medical imaging experts and the final mask is determined by majority vote. Due to the small size of vessels, the delineation of each vessel can oscillate easily. To address this, the coincidence of each vessel among the three masks is calculated. If the majority voting over any mask is smaller than 0.5, the vessel is highlighted and sent back to the three experts for further refinement. This process is repeated until no disagreement exists. A detailed description of each volume is available at https://github.com/lzhLab/LiVS.

Although SLIVER07 [53] (having 30 volumes and 4,159 slices) and CHAOS [54] (having 50 volumes and 2,874 slices) have been frequently used in liver vessel segmentation, they actually have no *vessel* mask, and all studies using these datasets have to manually delineate vessel masks. Hence, they are excluded from this study.

B. Evaluation Metrics

The performance is measured by Dice similarity coefficient (DSC), volumetric overlap error (VOE), average symmetric surface distance (ASSD) and Hausdorff distance (HD), which are defined as:

$$DSC = 2|P \cap G|/(|P| + |G|) \tag{16}$$

$$VOE = 1 - |P \cap G|/(|P \cup G|) \tag{17}$$

$$ASSD = \frac{\sum_{x \in \partial P} d(x, \partial G) + \sum_{x \in \partial G} d(x, \partial P)}{\partial P + \partial G} \tag{18}$$

$$HD = \max\{\max_{x \in \partial P} d(x, \partial G), \max_{x \in \partial G} d(x, \partial P)\}, \tag{19}$$

where P is the prediction, G is the ground truth, ∂X is the boundary of X, $d(x, \partial Y)$ is the minimum Euclidean distance between x and ∂Y defined as $d(x, \partial Y) = \min_{y \in \partial Y} d(x, y)$.

Among them, DSC and VOE are overlap-based measurements, while ASSD and HD are boundary distance-based metrics. In practice, DSC is the most widely used one as it is more robust.

C. Performance Qualification

The performance of our newly proposed model is examined on the three datasets, and head-to-head comparisons with existing state-of-the-art models are also carried out on the three datasets with the same sets of training, validation, and testing data as well. Two types of existing models are adopted, i.e., 2D-based and 3D-based. The typical 2D models borrowed

Model	Structure	DSC ↑	VOE↓	ASSD↓	HD↓
U-Net	2D	0.673	0.491	1.671	10.912
MFSNet	2D	0.572	0.598	1.925	11.832
DCSAU-Net	2D	0.672	0.490	1.682	10.620
FPN	2D	0.637	0.530	2.035	12.900
CaraNet	2D	0.571	0.598	2.724	15.370
TransFusionNet	2D	0.639	0.529	1.816	11.255
DualNet	2D	0.604	0.567	2.119	11.588
M2MRF	2D	0.607	0.563	2.004	11.415
3D U-Net	3D	0.311	0.806	4.400	23.262
V-Net	3D	0.374	0.770	2.177	11.220
Ours	2D	0.684	0.476	1.588	10.118

in this study include U-Net [13], FPN [47], MFSNet [55], DCSAU-Net [56], CaraNet [41], DualNet [45] (with backbone of HED [57] and dual-sampling modulated Dice loss), TransFusionNet [46], and M2MRF [40] (the cascade version), while the 3D models are 3D U-Net [14] and V-Net [11]. The detailed results are as follows.

1) Performance on 3D-IRCADb: The detailed performance is shown in Table I. It can be observed that the average performance of our model on 3D-IRCADb is 0.68, 0.48, 1.59 and 10.12 for DSC, VOE, ASSD and HD, respectively. This performance is significantly better than those generated by existing models, improving DSC by 1.63% compared to the second-best. This improvement is indeed remarkable considering the difficulty of vessel segmentation, although the absolute value looks trifling. Notably, the performance of the 2D models (U-Net, FPN, MFSNet, DCSAU-Net, CaraNet, DualNet, TransFusionNet, M2MRF and ours) is markedly better than that of the 3D models (3D U-Net and V-Net), *viz.*, greater than 0.6 versus smaller than 0.4. We speculate that this is mainly attributed to the small number of training data in 3D-IRCADb, i.e., only 20 volumes. This observation inspired us to construct a large-volume dataset.

2) Performance on MSD: MSD is 15 times larger than 3D-IRCADb in terms of volume; Therefore, we expect significant improvement of all models on MSD compared to those evaluated on 3D-IRCADb. Unfortunately, except for V-Net, the DSCs do not have apparent improvement, and some show even worse results; see Table II. For example, the DSC obtained from U-Net is slightly smaller than those generated from 3D-IRCADb. Regarding most of the mentioned networks and our model, they produce better results on MSDS. Noteworthily, V-Net lifts the DSC from 0.374 to 0.612, indicating that it benefits most from the large number of training data. Although our model generates the most accurate results than others, the unstable results obtained by all these models suggest that the data quality needs to be improved. To this end, we carried out comprehensive experiments on our newly constructed large-volume and high-quality data; see next.

3) Performance on LiVS: Promisingly, the performance of all models on LiVS has experienced a significant jump, increasing the average DSC by 0.157±0.105 and 0.113±0.042 compared to 3D-IRCADb and MSD, respectively; see details in Table III. This is obviously a tremendous advancement in liver vessel segmentation.

TABLE II
LIVER VESSEL SEGMENTATION PERFORMANCE ON MSD

Model	Structure	DSC ↑	VOE↓	ASSD↓	HD↓
U-Net	2D	0.671	0.485	2.141	12.385
FPN	2D	0.648	0.511	2.023	11.348
MFSNet	2D	0.584	0.583	2.054	11.094
DCSAU-Net	2D	0.682	0.477	1.713	10.056
CaraNet	2D	0.571	0.599	1.928	9.679
TransFusionNet	2D	0.686	0.471	1.718	10.202
DualNet	2D	0.666	0.494	1.594	9.326
M2MRF	2D	0.634	0.526	2.331	13.313
3D U-Net	3D	0.369	0.762	4.192	18.774
V-Net	3D	0.612	0.552	1.230	6.600
Ours	2D	0.699	0.457	1.427	9.097

TABLE III
LIVER VESSEL SEGMENTATION PERFORMANCE ON LIVS

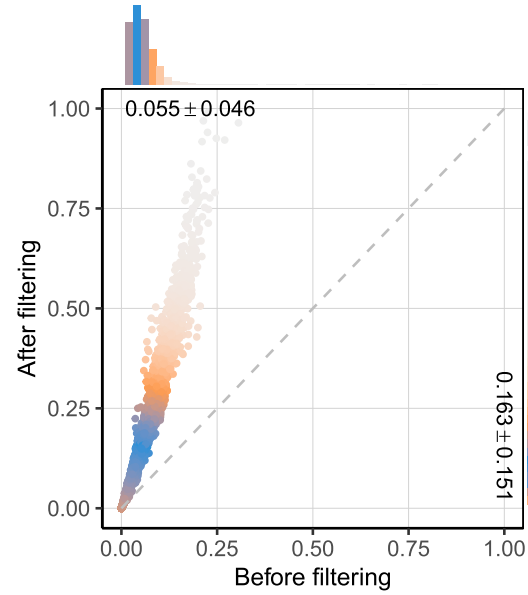| Model	Structure	DSC ↑	VOE↓	ASSD↓	HD↓
U-Net	2D	0.751	0.397	0.782	3.735
FPN	2D	0.763	0.382	0.689	3.007
MFSNet	2D	0.660	0.506	0.952	4.530
DCSAU-Net	2D	0.782	0.354	0.689	3.096
CaraNet	2D	0.654	0.512	1.105	5.923
TransFusionNet	2D	0.783	0.353	0.798	3.803
DualNet	2D	0.758	0.387	0.869	4.467
M2MRF	2D	0.755	0.392	0.740	3.393
3D U-Net	3D	0.582	0.566	5.123	14.403
V-Net	3D	0.779	0.359	0.408	1.774
Ours	2D	0.803	0.329	0.552	2.363

Fig. 3. Ratio between vessels and liver tissues before and after the proposed Laplacian salience filtering.

TABLE IV
IMPACT OF LAPLACIAN SALIENCE MODULES. EXPERIMENTS ARE
CARRIED OUT ON LIVS

Model	Laplacian Salience	High-pass Filter	DSC ↑	VOE↓	ASSD↓	HD↓
FPN	✗	✗	0.763	0.382	0.689	3.007
Ours	✗	✗	0.776	0.365	0.652	3.084
Ours	✗	✓	0.765	0.378	0.758	3.556
Ours	✓	✗	0.780	0.359	0.615	2.854
Ours	✓	✓	0.803	0.329	0.552	2.363

Interestingly, V-Net performs much better than other existing models, suggesting that it is suitable for vessel segmentation from liver images provided with large-volume and high-quality input data. In addition, 3D U-Net also generates significantly better results than that of MSD, indicating that the quality of LiVS is indeed better than MSD although they have a similar scale of data volume.

As for our model, the DSC score is as high as 0.803, which is lifted by 2.6% to 38.0% in terms of DSC compared to existing models. This observation suggests that the newly designed model effectively captures actual vascular regions and suppresses other liver tissues in feature learning, even without 3D information.

D. Salience Quantification

Performance qualification has demonstrated the effectiveness of the proposed Laplacian salience. Here we quantitatively analyze the impact of the salience on images. To this end, we calculate the ratio r of the weighted number of pixels between vessels and other liver tissues, precisely $r = \sum_{M_{x,y}=1} w_{x,y} M_{x,y} / \sum_{M_{x,y}=0} w_{x,y} (1 - M_{x,y})$, where M is the mask of an image with $M_{x,y} = 1$ for vessels at position (x, y), $M_{x,y} = 0$ for other liver tissues, and $M_{x,y} = -1$ for those outside the liver regions. The weight $w_{x,y}$ at position (x, y) is 1 for every pixel before Laplacian salience filtering, and is automatically determined by the proposed Laplacian salience after filtering.

Fig. 3 shows the distribution of r before and after Laplacian salience filtering. Before filtering, the averaged ratio between vessel and liver is 0.055 ± 0.046, while this value is markedly tripled to 0.163 ± 0.151 after filtering. It means that the proposed Laplacian salience filtering is effective in highlighting

vessel-like regions and suppressing others, which can also be observed visually in Fig. 1. Another noteworthy point is that the newly conceived filter can discriminate different classes, samples and pixels simultaneously. That being said, the imbalanced distribution of classes, the difficulty of samples, and the ambiguity of pixels can all be addressed simultaneously based on the filter.

E. Ablation Study

1) Salience Filter: The proposed Laplacian salience generator has two important modules, including a Laplacian salience generator and a high-pass filter. The former is to generate the image-wise salient map highlighting vessel-like regions, while the latter is to further shape the learning on the ambiguous regions. To investigate their influence, we toggle the two modules while leaving other parts unchanged. Results show that the combination of the two yields the best performance with DSC lifted by 4% compared to the baseline. Even with the Laplacian salience itself, the performance is still better than the original FPN model. More details can be found in Table IV.

2) Salience Loss: Obviously, salience can be incorporated into both the feature maps and the loss functions. If it is used in feature maps, the learned features will be shaped to vessels with higher attention; if it is used to modulate
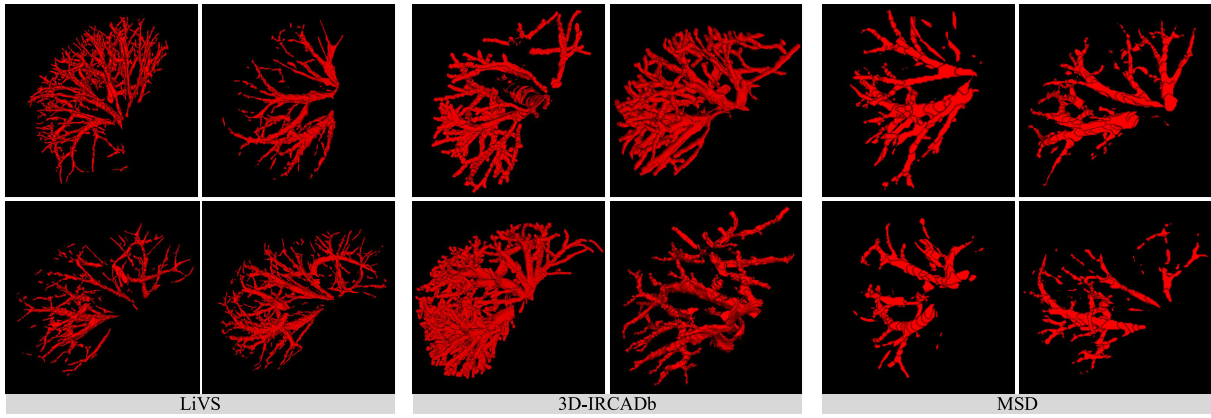
Fig. 4. Visualized 3D vessel structures built from different datasets. The density and thickness of branches produced from LiVS are evidently better than others. For 3D-IRCADb, the vessels are obviously over labeled; while MSD is under labeled on the contrary.

TABLE V
ABLATION STUDY FOR LOSS FUNCTION ON THE LiVS DATASET

Model	LS	Loss	DSC ↑	VOE↓	ASSD↓	HD↓
FPN	✗	BCE	0.763	0.382	0.689	3.007
FPN	✗	SM-BCE	0.779	0.359	0.702	3.010
FPN	✗	Focal	0.745	0.405	0.782	3.708
Ours	✗	BCE	0.776	0.365	0.652	3.084
Ours	✗	SM-BCE	0.790	0.346	0.625	2.824
Ours	✗	Focal	0.767	0.376	0.740	3.707
Ours	✓	SM-BCE	0.782	0.356	0.677	3.297
Ours	✓	Focal	0.773	0.369	0.657	3.067
Ours	✓	BCE	0.803	0.329	0.552	2.363

LS: Laplacian salience; SM-BCE: salience-modulated BCE loss.

TABLE VI
PERFORMANCE OF OUR MODEL ON LiVS WITH DIFFERENT PYRAMID LEVELS

Model	Minimum sidelength	DSC ↑	VOE↓	ASSD↓	HD↓
FPN	1/32	0.763	0.382	0.689	3.007
Ours	1/4	0.725	0.427	1.241	7.353
Ours	1/8	0.803	0.329	0.552	2.363
Ours	1/16	0.783	0.355	0.652	2.979
Ours	1/32	0.760	0.386	0.718	3.519

the loss functions, regions with higher salience values will be more sensitive to variations. To examine which is more influential and effective, we switch on/off the modules and conduct experiments on LiVS.

Results in Table V show that the salience-gated feature together with the BCE loss achieves the best performance, while the salience-modulated loss training without the salience-gated feature gains the second best. If the model includes both the salience-gated feature and the salience-modulated loss, it unexpectedly generates slightly worse results. We speculate that the combination of salience gate and salience loss hinders the efficient learning of non-vessel features due to the very little attention paid to these regions.

Another interesting point is that the original FPN can also benefit from the use of our salience-modulated loss. This suggests that our salience is helpful in capturing vessels even if they are very small.

3) Pyramid Level: The feature pyramid network has proved to be effective in capturing various abstract features. However, the size of vessels is usually very small. Hence, very deep down-sampling will reduce or even eliminate vessels from the input images. To examine which level generates the best performance, we have varied the down-sampling times and conducted experiments on LiVS.

Results show that the sidelength reduced to 1/8 of the original produces the best performance. That is, the number of down-sampling pyramid levels is 3. Note, however, that the original input is first enlarged by a super-resolution layer, thus the level is increased to 5, *cf.* Fig. 2. Not surprisingly,

either increasing or decreasing the number of down-samplings produces worse results. For more details, please refer to Table VI.

As a comparison, the performance of the original FPN is shown in Table VI as well. It can be seen that our pyramid architecture, coupled with salience, generates much better results than FPN, even with a shallower network.

F. Mask Examination

It is obvious that obtaining elaborated masks of vessels from liver images is very challenging because of their tiny size, heavy noise and low contrast. From the performance analysis (Section IV-C), we can see that both the volume and the quality of the datasets are critical to achieve satisfactory vessel segmentation. Precisely, LiVS has the largest number of volumes as well as the best quality, resulting in a big jump in DSC scores compared to MSD and 3D-IRCADb. Regarding the other two, they all have their own merits, with MSD having a larger volume and 3D-IRCADb having a better quality. Hence the performance of the two is comparable.

For ease of understanding, we constructed several vessel structures generated from the masks of the three datasets; see Fig. 4. Clearly, the 3D vessel structures constructed from LiVS show that the thickness of the branches is proportionate and the density of the vessels looks rational. Particularly, even skinny and tiny vessels can be visualized. Regarding 3D-IRCADb, many vessel branches are intertwined and mixed together, which is evidently impractical. For MSD, the vessel masks are heavily under-labeled, resulting in very sparse vessel structures, which is also unrealistic.
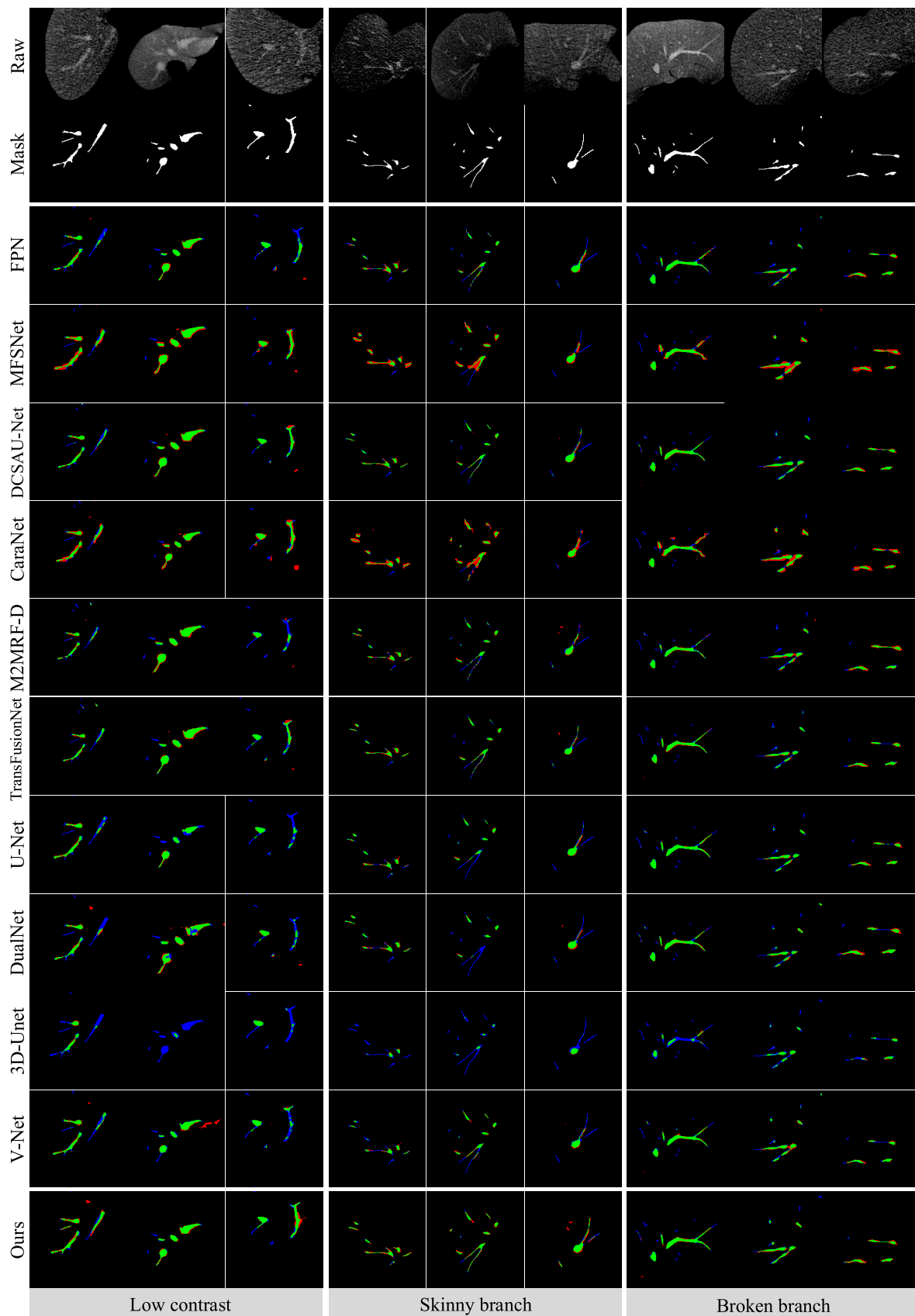
Fig. 5. Prediction visualization on LiVS. Three typical hard scenarios of vessel segmentation are presented, including vessels with low contrast, skinny branches and broken branches. The predictions are in red, while the original labels are in blue, and their superimposition is in green.

Note, this investigation is not to point out the imperfect quality of the existing or our own data. Instead, we attempt to draw attention to vessel mask purification.

G. Case Visualization

Three main hard situations in vessel segmentation are low contrast, skinny branches and broken branches. To illustrate the performance, prediction results generated from LiVS by existing state-of-the-art models as well as ours are exemplified in Fig. 5. For vessels with low contrast, CaraNet, TransFusionNet and ours produce markedly better results than others. For those with skinny branches, only TransFusionNet and our model can generate satisfactory results, and other models miss the very skinny branches. For the broken branches, V-Net and ours perform better than the others. In short, only our newly proposed model can handle the three thorny scenarios simultaneously. This further demonstrates that the proposed Laplacian salience is sensitive to vessels and effective in vessel segmentation.

V. CONCLUDING REMARKS

Accurate liver vessel segmentation faces two difficulties, namely the scarcity of high-quality and large-volume data as well as the sophisticated model adhering to small vessel-like regions, hence the poor performance of existing models. To this end, we have made great efforts to prepare a qualified large liver vessel dataset having 532 volumes and 15,984 slices. To investigate its goodness, a Laplacian salience-gated feature pyramid network is proposed, which can automatically focus on vessel-like regions and compensate for the skewed distribution of vessels in liver tissues. In addition, it can also put more effort into poorly discriminable regions, so that low-contrast vessels, skinny branches and broken branches can be effectively fished out. Experimental results show that the newly proposed model significantly outperforms other state-of-the-art models on both our newly proposed data and existing data. More promisingly, the Dice scores generated from our data are improved by at least 18.3% compared to those produced from other datasets. These observations suggest that our newly conceived model, as well as the data, can be of great help in advancing liver vessel segmentation.

REFERENCES

[1] M. Ciecholewski and M. Kassjański, "Computational methods for liver vessel segmentation in medical imaging: A review," *Sensors*, vol. 21, no. 6, p. 2027, Mar. 2021.

[2] Y. Z. Zeng, Y. Q. Zhao, M. Liao, B. J. Zou, X. F. Wang, and W. Wang, "Liver vessel segmentation based on extreme learning machine," *Phys. Medica*, vol. 32, no. 5, pp. 709–716, May 2016.

[3] X. Zhao, "Liver vessel segmentation algorithm based on fuzzy enhance and Hessian matrix," *J. Zhejiang Ind. Trade Vocational College*, vol. 18, no. 2, pp. 48–51, 2018.

[4] M. M. Fraz, A. R. Rudnicka, C. G. Owen, and S. A. Barman, "Delineation of blood vessels in pediatric retinal images using decision trees-based ensemble classification," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 9, no. 5, pp. 795–811, Sep. 2014.

[5] K. Ahmadi, A. Karimi, and B. F. Nia, "New technique for automatic segmentation of blood vessels in CT scan images of liver based on optimized fuzzy-means method," *Comput. Math. Methods Med.*, vol. 2016, Dec. 2016, Art. no. 5237191.

[6] X. Guo, S. Huang, X. Fu, B. Wang, and X. Huang, "Vascular segmentation in hepatic CT images using adaptive threshold fuzzy connectedness method," *Biomed. Eng. OnLine*, vol. 14, no. 1, p. 57, Dec. 2015.

[7] S.-H. Lee and S. Lee, "Adaptive Kalman snake for semi-autonomous 3D vessel tracking," *Comput. Methods Programs Biomed.*, vol. 122, no. 1, pp. 56–75, 2015.

[8] Y. Cheng, X. Hu, J. Wang, Y. Wang, and S. Tamura, "Accurate vessel segmentation with constrained B-snake," *IEEE Trans. Image Process.*, vol. 24, no. 8, pp. 2440–2455, Aug. 2015.

[9] X. Chen, B. M. Williams, S. R. Vallabhaneni, G. Czanner, R. Williams, and Y. Zheng, "Learning active contour models for medical image segmentation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 11624–11632.

[10] H. Kervadec, J. Bouchtiba, C. Desrosiers, E. Granger, J. Dolz, and I. B. Ayed, "Boundary loss for highly unbalanced segmentation," *Med. Image Anal.*, vol. 67, Jan. 2021, Art. no. 101851.

[11] F. Milletari, N. Navab, and S.-A. Ahmadi, "V-Net: Fully convolutional neural networks for volumetric medical image segmentation," in *Proc. 4th Int. Conf. 3D Vis. (3DV)*, Oct. 2016, pp. 565–571.

[12] S. S. M. Salehi, D. Erdogmus, and A. Gholipour, "Tversky loss function for image segmentation using 3D fully convolutional deep networks," in *Machine Learning in Medical Imaging*, vol. 10541, Q. Wang, Y. Shi, H. Suk, and K. Suzuki, Eds. Cham, Switzerland: Springer, 2017, pp. 379–387.

[13] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI*, vol. 9351, N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi, Eds. Cham, Switzerland: Springer, 2015, pp. 234–241.

[14] O. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger, "3D U-Net: Learning dense volumetric segmentation from sparse annotation," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI*, S. Ourselin, L. Joskowicz, M. R. Sabuncu, G. Unal, and W. Wells, Eds. Cham, Switzerland: Springer, 2016, pp. 424–432.

[15] Q. Yan et al., "Attention-guided deep neural network with multi-scale feature fusion for liver vessel segmentation," *IEEE J. Biomed. Health Informat.*, vol. 25, no. 7, pp. 2629–2642, Jul. 2021.

[16] T. Kitrungrotsakul et al., "VesselNet: A deep convolutional neural network with multi pathways for robust hepatic vessel segmentation," *Comput. Med. Imag. Graph.*, vol. 75, pp. 74–83, Jul. 2019.

[17] G. Sharma, F. Jurie, and C. Schmid, "Discriminative spatial saliency for image classification," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2012, pp. 3506–3513.

[18] S. Jetley, N. Murray, and E. Vig, "End-to-end saliency mapping via probability distribution prediction," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 5753–5761.

[19] J. Pan, E. Sayrol, X. Giro-I-Nieto, K. McGuinness, and N. E. O'Connor, "Shallow and deep convolutional networks for saliency prediction," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 598–606.

[20] J. Pan et al., "SalGAN: Visual saliency prediction with generative adversarial networks," 2017, *arXiv:1701.01081*.

[21] C. Wang, S. Dong, X. Zhao, G. Papanastasiou, H. Zhang, and G. Yang, "SaliencyGAN: Deep learning semisupervised salient object detection in the fog of IoT," *IEEE Trans. Ind. Informat.*, vol. 16, no. 4, pp. 2667–2676, Apr. 2020.

[22] A. Vaswani et al., "Attention is all you need," in *Proc. 31st Int. Conf. Neural Inf. Process. Syst.*, I. Guyon et al., Eds., vol. 30, 2017, pp. 6000–6010.

[23] D. Karimi, S. D. Vasylechko, and A. Gholipour, "Convolution-free medical image segmentation using transformers," in *Medical Image Computing and Computer Assisted Intervention—MICCAI*, vol. 12901, M. de Bruijne et al., Eds. Cham, Switzerland: Springer, 2021, pp. 78–88.

[24] A. Dosovitskiy et al., "An image is worth 16×16 words: Transformers for image recognition at scale," in *Proc. Int. Conf. Learn. Represent.*, 2021.

[25] J. Chen et al., "TransUNet: Transformers make strong encoders for medical image segmentation," 2021, *arXiv:2102.04306*.

[26] H. Wang, P. Cao, J. Wang, and O. R. Zaiane, "UCTransNet: Rethinking the skip connections in U-Net from a channel-wise perspective with transformer," in *Proc. AAAI Conf. Artif. Intell.*, 2022, pp. 2441–2449.

[27] J. Schlemper et al., "Attention gated networks: Learning to leverage salient regions in medical images," *Med. Image Anal.*, vol. 53, pp. 197–207, Apr. 2019.

[28] P. Guo et al., "Robust vessel detection and segmentation in ultrasound images by a data-driven approach," *Proc. SPIE*, vol. 9034, Mar. 2014, Art. no. 903435.

[29] X.-H. Han, J. Wang, Y. Konno, and Y.-W. Chen, "Bayesian saliency model for focal liver lesion enhancement and detection," in *Proc. Asian Conf. Comput. Vis.*, in Lecture Notes in Computer Science, vol. 10118, C.-S. Chen, J. Lu, and K.-K. Ma, Eds. Cham, Switzerland: Springer, 2016, pp. 32–45.

[30] Y. Yuan, Q. H. Meng, W. Qin, and L. Xing, "Liver lesion detection based on two-stage saliency model with modified sparse autoencoder," in *Medical Image Computing and Computer Assisted Intervention—MICCAI*, in Lecture Notes in Computer Science, vol. 10435, M. Descoteaux, L. Maier-Hein, A. Franz, P. Jannin, D. L. Collins, and S. Duchesne, Eds. Cham, Switzerland: Springer, 2017, pp. 577–585.

[31] S. Sinha, H. Ohashi, and K. Nakamura, "Class-wise difficulty-balanced loss for solving class-imbalance," in *Proc. Asian Conf. Comput. Vis.*, in Lecture Notes in Computer Science, vol. 12627, H. Ishikawa, C.-L. Liu, T. Pajdla, and J. Shi, Eds. Cham, Switzerland: Springer, 2020, pp. 549–565.

[32] H. Zheng et al., "Alleviating class-wise gradient imbalance for pulmonary airway segmentation," *IEEE Trans. Med. Imag.*, vol. 40, no. 9, pp. 2452–2462, Sep. 2021.

[33] D. Zhou et al., "IoU loss for 2D/3D object detection," in *Proc. Int. Conf. 3D Vis. (3DV)*, Sep. 2019, pp. 85–94.

[34] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 42, no. 2, pp. 318–327, Feb. 2020.

[35] G. Cheng et al., "Towards large-scale small object detection: Survey and benchmarks," 2022, *arXiv:2207.14096*.

[36] S. Minaee, Y. Y. Boykov, F. Porikli, A. J. Plaza, N. Kehtarnavaz, and D. Terzopoulos, "Image segmentation using deep learning: A survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 7, pp. 3523–3542, Jul. 2022.

[37] J. Zhang, Y. Zhang, and X. Xu, "ObjectAug: Object-level data augmentation for semantic image segmentation," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2021, pp. 1–8.

[38] J. Lourenço-Silva and A. L. Oliveira, "Using soft labels to model uncertainty in medical image segmentation," in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, vol. 12963, A. Crimi and S. Bakas, Eds. Cham, Switzerland: Springer, 2022, pp. 585–596.

[39] T. V. Quyen and M. Y. Kim, "Feature pyramid network with multi-scale prediction fusion for real-time semantic segmentation," *Neurocomputing*, vol. 519, pp. 104–113, Jan. 2023.

[40] Q. Liu, H. Liu, W. Ke, and Y. Liang, "M2MRF: Many-to-many reassembly of features for tiny lesion segmentation in fundus images," 2021, *arXiv:2111.00193*.

[41] A. Lou, S. Guan, H. Ko, and M. H. Loew, "CaraNet: Context axial reverse attention network for segmentation of small medical objects," *Proc. SPIE*, vol. 12032, pp. 81–92, Apr. 2022.

[42] H. Bhangale, R. Bansal, S. Jain, and J. Sarvaiya, "Multi-feature similarity based deep learning framework for semantic segmentation," in *Proc. Int. Conf. Control, Autom., Power Signal Process. (CAPS)*, Dec. 2021, pp. 1–4.

[43] L. Wang, D. Li, Y. Zhu, L. Tian, and Y. Shan, "Dual super-resolution learning for semantic segmentation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 3773–3782.

[44] Y. Zhou, X. Sun, Z.-J. Zha, and W. Zeng, "Context-reinforced semantic segmentation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 4041–4050.

[45] Q. Liu, H. Liu, Y. Zhao, and Y. Liang, "Dual-branch network with dual-sampling modulated dice loss for hard exudate segmentation in color fundus images," *IEEE J. Biomed. Health Informat.*, vol. 26, no. 3, pp. 1091–1102, Mar. 2022.

[46] X. Meng et al., "Exploiting full resolution feature context for liver tumor and vessel segmentation via integrate framework: Application to liver tumor and vessel 3D reconstruction under embedded microprocessor," 2021, *arXiv:2111.13299*.

[47] T.-Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, and S. Belongie, "Feature pyramid networks for object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 936–944.

[48] A. Kirillov, R. Girshick, K. He, and P. Dollár, "Panoptic feature pyramid networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 6399–6408.

[49] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.

[50] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. Int. Conf. Learn. Represent. (ICLR)*, Y. Bengio and Y. LeCun, Eds., 2015.

[51] L. Soler et al., "3D image reconstruction for comparison of algorithm database: A patient specific anatomical and medical image database," IRCAD, Strasbourg, France, Tech. Rep., 2010. [Online]. Available: https://www.ircad.fr/research/data-sets/liver-segmentation-3d-ircadb-01/

[52] A. L. Simpson et al., "A large annotated medical image dataset for the development and evaluation of segmentation algorithms," 2019, *arXiv:1902.09063*.

[53] T. Heimann, B. van Ginneken, and M. Styner, "Comparison and evaluation of methods for liver segmentation from CT datasets," *IEEE Trans. Med. Imag.*, vol. 28, no. 8, pp. 1251–1265, Feb. 2009.

[54] A. E. Kavur et al., "CHAOS challenge-combined (CT-MR) healthy abdominal organ segmentation," *Med. Image Anal.*, vol. 69, Apr. 2021, Art. no. 101950.

[55] H. Basak, R. Kundu, and R. Sarkar, "MFSNet: A multi focus segmentation network for skin lesion segmentation," *Pattern Recognit.*, vol. 128, Aug. 2022, Art. no. 108673.

[56] Q. Xu, Z. Ma, N. He, and W. Duan, "DCSAU-Net: A deeper and more compact split-attention U-Net for medical image segmentation," *Comput. Biol. Med.*, vol. 154, Mar. 2023, Art. no. 106626.

[57] S. Xie and Z. Tu, "Holistically-nested edge detection," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2015, pp. 1395–1403.