Bulletin of Electrical Engineering and Informatics
Vol. 12, No. 3, June 2023, pp. 1620∼1629
ISSN: 2302-9285, DOI: 10.11591/eei.v12i3.4676

Curriculum learning based overcomplete U-Net for liver tumor segmentation from computed tomography images

Bindu Madhavi Tummala, Soubhagya Sankar Barpanda
School of Computer Science and Engineering, VIT-AP University, Andhra Pradesh, India

Article Info

Article history:

Received Aug 30, 2022
Revised Nov 15, 2022
Accepted Dec 14, 2022

Keywords:

Curriculum learning
Deep learning
Liver tumor segmentation
Overcomplete networks
U-Net architecture

ABSTRACT

In this paper, we have proposed an overcomplete U-Net to perform liver tumor segmentation jointly using a curriculum learning strategy. Liver tumor segmentation is the most prominent and primary step in treating liver cancer and can also help doctors with proper diagnosis and therapy planning. However, it is challenging because of variations in shape, position, and depth of tumors and adjacent boundaries with internal organs around the liver. We have presented a promising solution by designing a U-Net-based segmentation network with two branches: an overcomplete branch to fine grade the small structures and an undercomplete branch to fine grade the high-level structures. This combination allows the network to learn all types of tumor artifacts more accurately. We also changed the conventional learning paradigm to curriculum learning where the input images are fed to the network from easy to hard ones to achieve faster convergence. Finally, our network segments the tumors directly from the whole medical images without the need for segmented liver region of interests (ROIs). The proposed network achieved a DICE score of 75% in tumor segmentation which is a decent value when compared with some existing deep learning methods for liver tumor segmentation.

Corresponding Author:

Bindu Madhavi Tummala
School of Computer Science and Engineering, VIT-AP University
Amaravati, Andhra Pradesh, India
Email: bindumadhavi.t@vitap.ac.in

1. INTRODUCTION

According to the International Agency for Research on Cancer (IARC), the statistics in 2020 say that hepatic cancer is the sixth most commonly occurring and second death-causing cancer worldwide [1]. Liver cancer can be detected with the help of imaging modalities like ultrasound, computed tomography, and magnetic resonance imaging. Among these, computed tomography (CT) is the most used modality for visualization. It gives cross-sectional pictures of the abdomen. To detect and diagnose hepatic cancer, doctors need to perform further processing by contouring the liver and its tumors from the whole CT image. Manual liver tumor segmentation is time-consuming because in a CT scan the liver stretches over 170 slices and is error-prone due to inter-observer variabilities, low contrast ambiguities of the modality, and especially differences in shape, location, and the number of tumors present. So, a robust automatic segmentation assists the radiologists in better diagnosis and treatment planning. Early methods are based on image processing tools like edge detection filters and statistical modeling. Later comes the machine learning methods which use handcrafted features to perform the segmentation.

Nowadays, the state-of-art methods for automatic liver and liver tumor segmentation are based on deep learning. Since these methods can read and evaluate data-specific features on their own there is no hassle of extracting features manually from medical images. The majority of deep learning methods used for segmentation in computer vision and medical imaging fields are convolution neural networks (CNNs) following the encoder-decoder style of architecture. The CNNs are responsible for better feature extraction and one can easily change the number of layers to make the network dense [2] for better performance and segmentation. The first discussion was about Seg-Net which got wide recognition. The Seg-Net [3] encoding block contains a convolutional layer followed by a max pool layer and the filter size is increased as the network goes deeper to extract the high-level features like organs within the CT image. Only the primary encoding layers capture the low-level information like edges which are important for accurate delineation of the targeted organ. Considering the deep learning-based medical image segmentation methods, U-Net [4] is the major leap where skip connections are added between the encoder and decoder at each level to pass the information that helps to improve the training quality as well as the segmentation accuracy. The encoder captures the contextual information and the decoder is an expansion path that performs localization to each pixel obtained. Both are almost symmetric obtaining a U shape architecture. From then, U-Net became a baseline architecture for almost every medical image segmentation neural network.

V-Net [5] and 3D U-Net [6] were designed by converting normal convolutions of U-Net into 3D convolutions to perform segmentation on volumetric medical images. U-Net++ [7] proposed changes in skip connections to reduce the gap between the feature maps of the encoder and decoder. Simple extensions of U-Net include res-UNet [8] and dense-UNet [9] where residual connections and dense layers are introduced accordingly in the encoder-decoder blocks. Residual connections solve vanishing gradients problem and the dense networks make the present layer to reuse the features of previous layers to achieve ease in training and efficiency [10]. UNet3+ [11] proposed full-scale skip connections with deep supervision. It has the advantage of combining low-level features with high-level semantics learned from feature maps obtained from different scales. The rest of the paper is organised as follows: section 2 contains the review of different papers on liver and tumor segmentation along with their limitations, section 3 discuss the architectural details of the proposed method, section 4 shows the pre settings to be done to make the data and network ready for training and testing along with the experimental results, and section 5 contains the conclusion by discussing the scope of improvement.

2. RELATED WORKS

In this section, we have briefly reviewed the deep learning architectures developed for liver tumor segmentation. We have focused more on liver tumors than the liver because of their increased complexity. According to Christ *et al.* [12] combined two U-Nets in a cascaded fashion to segment liver and liver tumors respectively. But this method used 3D conditional random field as a post processing step to refine the segmentation result. According to Chlebus *et al.* [13] proposed the same cascaded architecture with different object-based postprocessing step. According to Li *et al.* [14] combined U-Net and dense U-Net that extracts more contextual information through intra-slice features with less computational cost. Budak *et al.* [15] proposed two encoder decoder convolutional neural networks (EDCNN). The first network segments the liver and the second one segments the tumors. With liver as region of interest (ROI). Segmentation done from liver images drastically reduce the false positives. Research by Jin *et al.* [16] proposed 3D hybrid attention mechanism called RA-UNet to segment liver and tumors. There are many other papers based on medical image segmentation which can be seen in well written review papers [17], [18]. Some more observations on liver tumor segmentation are listed in Table 1 to have a quick review. Besides the improvements made by the liver tumor segmentation family of networks, there are two common drawbacks observed; i) difficult in segmenting small and even minute liver tumors because of its variations in shape, size, position and ii) to achieve liver tumor segmentation, liver segmentation is a prerequisite.

Most of the segmentation networks are U-Net variant and these fail in segmenting small liver tumors because they are designed to have more focus on high-level structures. As going deeper in the network the encoder downsamples the input image at every stage and the receptive field increases along with the number of filters and thus forces the network to have more focus on high-level features. Moreover, only the initial layers of the encoder collect the low-level features making the network prone to difficultly in segmenting small structures. For example, consider a 4-layer U-Net based network with 64, 128, 256, and 1,024. We can observe

that only first layer is extracting the low-level features (small structures, edges) and the remaining layers focus on high-level features. So, the percentage of filters capturing low-level features is $\frac{64}{64+128+256+512} = 6.67\%$ which is very much less than the filters (93.33%) learning the high-level structures.

Table 1. Overview of liver and its tumor segmentation methods

Authors	Year	Title	Dataset	Architecture	Observations
[19]	2020	Deep learning and level set approach for liver and tumor segmentation from CT scans	LiTS and IRCAD	Cascaded U-Net FCN	Used level set approach to refine the results and used liver segmentation as a preliminary
[13]	2018	Automatic liver tumor segmentation in CT with fully convolutional neural networks and object-based post processing	LiTS	Cascaded U-Net FCN	postprocessing step is used to refine results and detects bigger lession better than the smaller ones
[20]	2019	Liver tumor segmentation in CT volumes using an adversarial densely connected network	LiTS	Dense connected network with multi scale residual connections	Used liver segmentation as a preliminary
[21]	2019	A joint deep learning approach for automated liver and tumor segmentation	LiTS	U-Net variant	Segmentation of tumors is done in single step but still have misclassifications in small tumors
[14]	2017	H-DenseUNet: hybrid densely connected UNet for liver and tumor segmentation from CT volumes	LiTS and IRCAD	H-Dense UNet 2D-Dense UNet 3D-Dense UNet	Segmented liver tumor from 2D and 3D slices but suffer from high computational cost
[22]	2017	3D liver tumor segmentation in CT images using improved fuzzy C-means and graph cuts	3Dircadb1	Kernelized fuzzy C-means with spatial constraints	Undersegmentation and oversegmentation problems
[23]	2020	Liver tumor segmentation in CT scans using modified SegNet	3Dircadb1	SegNet architecture with pretrained VGG-16 as encoder	Have false positives
[24]	2020	Improving CT image tumor segmentation through deep supervision and attentional gates	IRCAD	CNN with deep supervision and attention gates	Slight computation burden but achieved good accuracy
[25]	2018	Liver lesion segmentation informed by joint liver segmentation	3Dircadb1	U-Net variant	Prone to misclassification
[26]	2017	Automatic liver segmentation using an adversarial image-to-image network	1000 annotated CT volumes	CNN with multi level supervision	Performed well in liver segmentation

The difficulty also lies in the properties of tumors. They are shape variant, and size varies from minute to big, can be present at any place of the liver, and finally they can be multiple. All these properties make segmenting liver tumor-like searching a needle-in-a-haystack. Addressing the second difficulty, all the developed networks segmented the liver very well because it is a large structure and comes under the category of high-level features. But, to segment the liver tumor firstly the liver is to be segmented and then using those predictions the liver ROI is extracted from whole CT images. Then, the extracted liver regions along with the tumor masks are fed again to the same or different network to perform tumor segmentation. This procedure includes two times training and preprocessing which takes computational cost and time.

Training a single network that directly segments the tumors by learning the specific features of tumors from the whole CT images is much needed to solve the small segmentation as well as the two-stage segmentation problems. Considering the above points we have proposed a curriculum learning-based overcomplete U-shaped network (CLU-Net) to perform liver tumor segmentation that segments the tumors directly (i.e. without liver ROI) and achieves an improvement in accuracy with a good learning paradigm and architectural structure. Our contributions are pointed as follows.

– Proposed a CLU-Net, where the network learns to efficiently segment the liver tumors by refining its learning progressively from easy samples to complex ones.

- The network used for training is combination of overcomplete and undercomplete branches and the architecture is U-Net variant (i.e encoder-decoder framework).
- The learning strategy is robust where we have divided the tumor masks into three categories based on size and number: big and single tumors, small and few tumors, and finally minute multiple tumors.
- The liver tumors are directly segmented without the need of liver regions.

3. METHOD

The detection and treatment of liver disorders using CT images is an essential task for segmenting the liver and its tumours. The segmentation of the liver and associated tumor is difficult due to the uneven presence, fuzzy borders, various densities, forms, and sizes of lesions. The proposed work consists of a neural network with a fusion of overcomplete and undercomplete branches and curriculum learning that send the inputs based on the degree of difficulty.In this section, we have discussed overcomplete and undercomplete representations, curriculum learning strategy and explained how these incorporations improved the overall performance of the proposed network.

3.1. Overcomplete and undercomplete representations

An overcomplete representation is first described in [27] in signal processing for creating dictionaries on the basis of more representations than the input signals. These representations show high robustness in reconstructing the signals with noise. This quality made them to be used in autoencoders [28] and recurrent neural networks [29]. Denoising autoencoders using this representation obtained better feature detection quality. Coming to the undercomplete representation, traditional encoder-decoder networks are called as undercomplete networks because the receptive field increases due to the downsampling that takes place in encoder. This increase in receptive field cause the network to concentrate more on high-level representations and make them to have less or no focus on small structures. To understand in detail, lets us consider an example. Let M be a medical image and FM1 and FM2 are the two feature maps extracted from two convolutional blocks and k is the filter size and the pooling coefficient and stride is set to the value 2 for both the networks. In an undercomplete network, the size of the receptive field is directly dependent on the variables of maxpooling layer i.e. pooling coefficient and stride. So, the receptive field size in second conv block is 2×k×2×k and 4×k×4×k for block 3 and this continues till the last block of the encoder. We can generalize this into a formula for receptive field (RF) size in conv block 'b' as (1).

$$RF_b = 2^{2*(b-1)} \times k \times k \tag{1}$$

Similarly, in an overcomplete network, the receptive field size depends upon an upsampling layer which works opposite to the maxpooling layer. Then the recptive field for conv 2 is $\frac{1}{2} \times k \times \frac{1}{2} \times k$ and for conv 3 it is $\frac{1}{4} \times k \times \frac{1}{4} \times k$ and this can be generalized to block 'b' as (2).

$$RF_b = (\tfrac{1}{2})^{2*(b-1)} \times k \times k \tag{2}$$

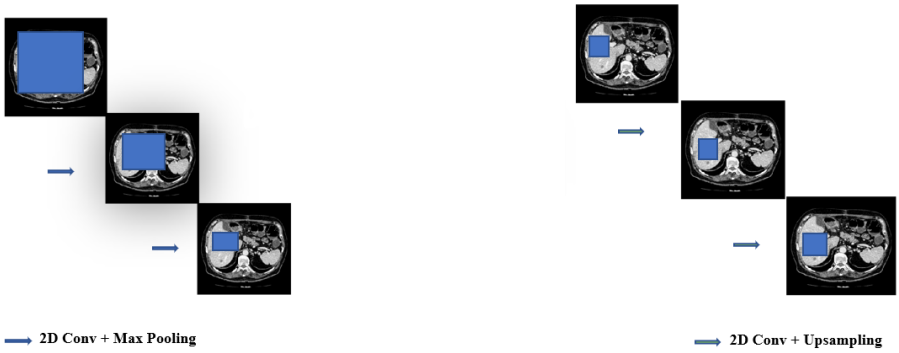
But in our architecture, we have made the receptive size neutral or the same in both the branches by considering filter size to 3x3, stride, and pooling value to 1 and is depicted in Figure 1, where Figure 1(a) denotes the receptive field increasing by each level whereas Figure 1(b) shows that the receptive field is common for all the levels. From the discussion one can conclude that the overcomplete networks are good at extracting small structures i.e., it is efficient in capturing low-level features whereas the under complete networks are good at extracting large structures i.e. efficient in capturing high-level features. Each representation has an advantage of its own. So, we have combined these two networks into a single U-Net variant network.

3.2. Curriculum learning

Deep learning networks are state-of-the-art methods today in every field. The research improvements aremainly focusing on building deeper networks, adding residual and dense connections and new attention mechanisms that help to improve the overall performance of the network. But, less approaches on the way of training, all the networks are trained on random inputs. Since the neural networks are inspired by the human brain, it is quite achievable to inspire the way they learn. Every human brain learns a topic starting from easy tasks to complex tasks which is called a curriculum followed by all the school system and organizations. This method helps learning in a better, easy and accurate way. This same method is implied to the neural

networks first by [30]. Replacing the conventional training strategy gradient descent random batch sampling with the easy-to-hard sampling is called as curriculum learning. The intuition behind this strategy is that learning becomes more fruitful if the inputs are organized in the order of complexities and fed to the network. The benefits of this learning strategy includes; i) speed in convergence, ii) better accuracy, and iii) even small networks can obtain decent performance. The thumb rule to implement curriculum learning is the data must contain a range of incrementally easy to difficult examples. We believe a range of difficulty exists within the liver tumor segmentation tasks. It is because liver tumors progress over time and it will appear at different places within the liver, shapes and even sizes. So, we have divided our input data into three levels and followed a curriculum learning training schema to train the designed segmentation network. The samples from three levels of data are shown in Figure 2.

- Stage 1: easy images (collected images that contain single and large tumors).
- Stage 2: easy images+hard images (collected images that contain maximum two tumors of variable sizes).
- Stage 3: easy images+hard images+very hard images (collected images that contain minimum three tumors of variable sizes).



(a) (b)

Figure 1. The amount of field of view considered by the networks based on their respective receptive fields (a) receptive field simulation in undercomplete network and (b) receptive field simulation in overcomplete network

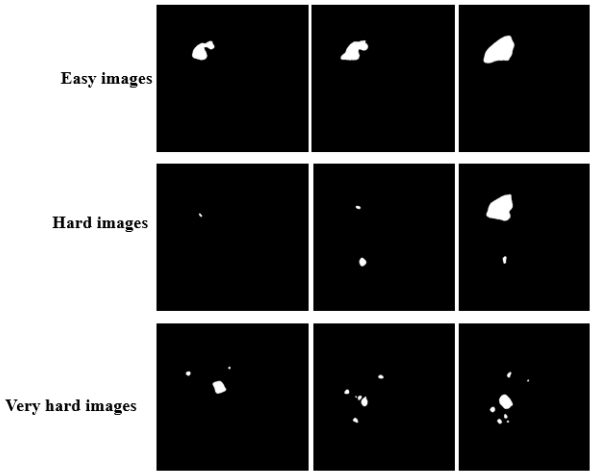


Figure 2. Example images for each difficulty level

3.3. Architecture

The curriculum learning based neural network (CLU-Net) is depicted in Figure 3. This architecture contains two branches: an overcomplete branch and an undercomplete branch. Overcomplete branch consists of Conv2D followed by a upsampling layer and ReLU. Undercomplete branch is similar to the encoder of U-Net. It contains Conv2D followed by maxpooling layer and ReLu. The decoder of the overcomplete network is the encoder of the undercomplete network and viceversa. The features maps from each block is fused (i.e. added) in the feature fusion block and sent as input to the two branches. The skip connections follow the same U-Net strategy that is the encoder and decoder share the information to have better predictions. At the final layer, the outputs from the two branches are concatenated and sent to 1×1 Conv2D output layer to generate the predictions. The overcomplete branches are good at segmenting small structures and the undercomplete branches are good at segmenting large structures. Combining the two structures yield a good performance. The learning paradigm used here is stage-wise learning where the network trains slowly from images in the order of complexity. At every stage, previous levels are added to avoid catastrophic forgetting [31].

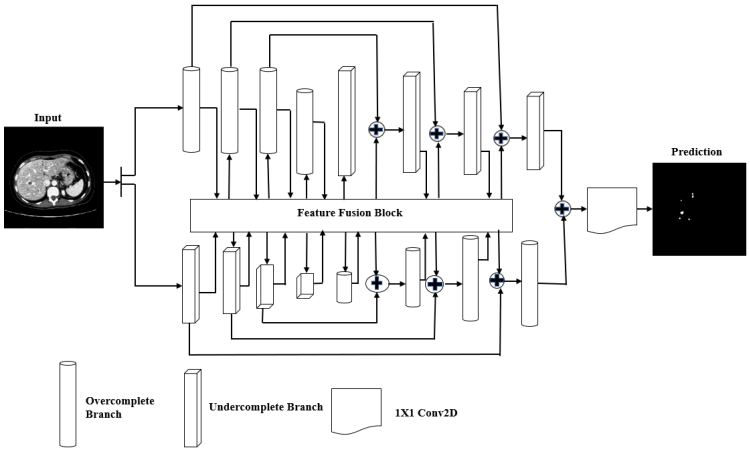


Figure 3. The CLU-Net architecture

4. RESULT AND DISCUSSION

The network architecture is based on the original UNET [4]. But due to architectural changes in the encoder and decoder branches. The network was designed from scartch and all the steps followed are explained in detail.

Dataset statistics: the 3Dircadb1 [32], a public dataset is selected. It contains the data of 20 patients with 75% of tumors. The reason behind the selection of this dataset is the number of tumors it has and the challenging complexities in terms of tumor contrasts, multiple tumors within a patient and size of some tumors which are not even visible to the naked eye. Each image is of size 512×512 with the slice range between 74 to 260. The inter-slice thickness lies within 1.25 mm to 4 mm. The medical images are raw images and should be prepared before giving input to the network. This step is known as data preprocessing. Generally, the preprocessing includes a series of substeps and these are dependent on the task the network is going to perform. So, the preprocessing steps followed in this paper are:

- Every organ present in the CT image has a set of grayscale pixels known as HU range. We can enhance the contrast or highlight our target organ by tuning the image into our desired window range. This process is called as HU windowing. Since our target is liver we have selected the window range (-100, 400) as suggested by [33]. The effect of windowing on the raw CT slice is shown in Figure 4, where Figure 4(a) contains the original CT slice where liver is not visible and Figure 4(b) contains the same slice after windowing and it is clearly evident that after windowing the liver is more visible.
- Even after windowing, the medical image has to be normalized and equalized to enhance the contrast of the liver and its tumor. It is because they both share almost same contrasts in this dataset. So to improve the visual clarity all the CT images are normalized and contrast limited adaptive histogram equalization

(CLAHE) method [34] is applied. This method enhances the local contrast and amplifies the noise.

– Every CT image has a corresponding liver mask but the tumor masks are distributed over different folders. So for ease in computation, all the tumor folders of a single patient are combined and then merged with the corresponding liver mask. The liver and the tumors are designated with grey and white color notations respectively. In Figure 5(a) contains the preprocessed CT images with more clarity than the original CT images and 5(b) contains their corresponding masks. The final preprocessed input images and its masks are shown for reference.
– All the images and corresponding masks are checked to remove the images that don't contain liver to make the network more close to the target.

(a) (b)

Figure 4. HU windowing of range (100,400) is applied to make the liver more visible (a) an example of a raw CT slice before windowing and (b) the same slice after windowing effect

(a) (b)

Figure 5. After the preprocessing steps the appearance of two sample images along with their respective masks or ground truths (a) the prerocessed medical images and (b) their corresponding masks

Network parameters: binary cross entropy (BCE) is used as a loss function and DICE and volumetric overlap error (VOE) are the evaluation metrics to verify the network. The notations along with definitions are as follows. BCE: it compares the predicted probabilities with the original ones and caluclates the penality score based on the distance to be achieved to reach the output value as (3):

$$BCE = -\frac{1}{P}\sum_{i=1}^{N} y_i.log(p(y_i)) + (1-y_i).log(1-p(y_i)) \quad (3)$$

Where P is the total number of pixels, y_i is the prediction for pixel i and $p(y_i)$ is the probability of predicted pixel in either foreground or background. DICE coefficient: it is the most popular technique used to measure the overlap between the automatic and actual predictions. Its value lies within [0,1]. It is measured as (4):

$$DICE(S,G) = \frac{2TP}{2TP+FP+FN} \quad (4)$$

VOE: it is the error metric calculated between the intersection and union of the two prediction sets. It is measured as (5):

$$VOE(S,G) = 1 - \frac{TP}{TP+FP+FN} \quad (5)$$

Where 'S' is the segmented image, 'G' is the ground-tuth image, 'TP' means true positives that denotes the pixels that are foreground and are also classified as foreground, 'FN' means false negatives that incorrectly denotes the foreground pixels as background by the classifier, and 'FP' means false positives that incorrectly denotes the background pixels as foreground by the classifier.

The complete work is done on i7 computer with 16 GB RAM and NVIDIA TITAN X, 12 GB with Keras and Tensorflow backend. During training, many tuning experiments were done and the final network parameters used are. Adam is used as a optimizer with learning rate 0.0001 and the number of epochs is 75 with batch size 8 where for every 25 epochs the input images changes its stage to increase the difficulty and the α and β values are set to 0.5 and 0.8 respectively.

Training and testing: the total number of CT images in the set are 2,823 and from them the images without liver and tumor are discared and lastly the medical data with tumor (2,083 images) are saved sequentially along with their corresponding masks. The training and test set is divided into 2,000 and 83 images respectively. Again, the training data is splitted into training and validation sets in 80:20 ratio. Due to less training data available, a real time data augmentation technique with rotation 90 and tranpose methods are applied at each epoch to increase the amount of data and even the generalizability of the proposed model. No augmentation is applied on the test set. The curriculum learning categorization is done only on the training data before augmentation.

4.1. Results

The proposed CLU-network efficiently segments the liver tumor. First, the accuracy is increased by following curriculum learning strategy which increases the learning capacity of the network without the increase in network complexity. This input schema feeds the network with the images in the order of thier complexity: from easy to hard. Second, the network effectively reduced the small tumor segmentation problem by using overcomplete representation which has low receptive field and allows the network to concentrate more on minute details. Third, the network performs the segmentation of tumor directly from the CT images unlike the other networks which followed two methods to segment the tumor. The experimentation is done on U-Net [4], U-Net++ [7], CU-Net [35] a cascaded UNet, MRDU-Net [36] residual dilated encoder decoder network which segments the image by considering the inputs at different scales and our proposed network with random sampling and curriculum sampling. Both the compared architectures also segmented the tumor directly without the help of liver ROIs. The score chart is shown in Table 2 and some of the predictions are shown in Figure 6, where Figure 6(a) shows the ground-truths and Figure 6(b) contains the corresponding predictions done by the proposed network. It is evident from the figure that the predictions of the liver are predicted perfectly and the tumor predictions have some over-segmentation and under-segmentation errors.

Table 2. Comparision result of CLU-Net on liver tumor segmentation

Method	DICE (%)	VOE (%)
U-Net	60.41	37.23
U-Net++	62.28	27.24
CU-Net	64.73	26.14
MRDU-Net	65.34	26.73
CL-Net	66.78	24.38
CL-Net with curriculum	**74.58**	**19.28**

(a) (b)

Figure 6. The sample outputs of the proposed network are shown consider column-wise (a) the original ground-truths and (b) the corresponding segmentation results

5. CONCLUSION

To perform liver tumor segmentation directly from the CT images, we have presented a curriculum learning-based U-Net with two branches: overcomplete and undercomplete. The overcomplete branch refines the small structures and the undercomplete branch refines the large structures. These branch structures allow the network to accurately learn the various sizes of tumors and produce a good segmentation result. The learning schema of the proposed network is curriculum learning-based where the input images are divided into categories based on their complexity and fed to the network category-wise. The difficulty level is decided by manual examination of tumor masks by considering their size shape and quantity. This stage-wise learning introduces complexity to the network without additional computational cost. The network learned in this way rather than random sampling yields a better accuracy over the other as shown in the result section. In the future to improve the generalizability, we will implement the network along with the schema to a multimodality dataset and perform multi-organ segmentation. The input categorization will also be automized so that the division can be done to any dataset easily without any hassle on manual preprocessing.

REFERENCES

[1] J. Ferlay *et al.*, "Global cancer observatory: cancer today," *International agency for research on cancer*, vol. 3, no. 20, 2018.

[2] Yadavendra and S. Chand, "A comparative study of breast cancer tumor classification by classical machine learning methods and deep learning method," *Machine Vision and Applications*, vol. 31, no. 6, p. 46, 2020, doi: 10.1007/s00138-020-01094-1.

[3] V. Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet: a deep convolutional encoder-decoder architecture for image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 12, pp. 2481–2495, 2017, doi: 10.1109/TPAMI.2016.2644615.

[4] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*, Cham: Springer, 2015, pp. 234–241, doi: 10.1007/978-3-319-24574-4_28.

[5] F. Milletari, N. Navab, and S.-A. Ahmadi, "V-Net: fully convolutional neural networks for volumetric medical image segmentation," in *2016 Fourth International Conference on 3D Vision (3DV)*, 2016, pp. 565–571, doi: 10.1109/3DV.2016.79.

[6] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger, "3D U-Net: learning dense volumetric segmentation from sparse annotation," in *International conference on medical image computing and computer-assisted intervention*, Cham: Springer, 2016, pp. 424–432, doi: 10.1007/978-3-319-46723-8_49.

[7] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh, and J. Liang, "UNet++: a nested U-Net architecture for medical image segmentation," in *Deep learning in medical image analysis and multimodal learning for clinical decision support*, Cham: Springer, 2018, pp. 3–11, doi: 10.1007/978-3-030-00889-5_1.

[8] X. Xiao, S. Lian, Z. Luo, and S. Li, "Weighted res-UNet for high-quality retina vessel segmentation," in *2018 9th International Conference on Information Technology in Medicine and Education (ITME)*, 2018, pp. 327–331, doi: 10.1109/ITME.2018.00080.

[9] G. Huang, Z. Liu, L. V. D. Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 2261–2269, doi: 10.1109/CVPR.2017.243.

[10] X. Yu, N. Zeng, S. Liu, and Y.-D. Zhang, "Utilization of DenseNet201 for diagnosis of breast abnormality," *Machine Vision and Applications*, vol. 30, no. 7–8, pp. 1135–1144, 2019, doi: 10.1007/s00138-019-01042-8.

[11] H. Huang *et al.*, "UNet 3+: a full-scale connected unet for medical image segmentation," in *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing*, 2020, pp. 1055–1059, doi: 10.1109/ICASSP40776.2020.9053405.

[12] P. F. Christ *et al.*, "Automatic liver and lesion segmentation in ct using cascaded fully convolutional neural networks and 3D conditional random fields," in *International conference on medical image computing and computer-assisted intervention*, Cham: Springer, 2016, pp. 415–423, doi: 10.1007/978-3-319-46723-8_48.

[13] G. Chlebus, A. Schenk, J. H. Moltz, B. v. Ginneken, H. K. Hahn, and H. Meine, "Automatic liver tumor segmentation in CT with fully convolutional neural networks and object-based postprocessing," *Scientific Reports*, vol. 8, no. 1, pp. 1–7, 2018, doi: 10.1038/s41598-018-33860-7.

[14] X. Li, H. Chen, X. Qi, Q. Dou, C.-W. Fu, and P.-A. Heng, "H-denseunet: hybrid densely connected UNet for liver and tumor segmentation from CT volumes," *IEEE Transactions on Medical Imaging*, vol. 37, no. 12, pp. 2663–2674, 2018, doi: 10.1109/TMI.2018.2845918.

[15] Ü. Budak, Y. Guo, E. Tanyildizi, and A. Şengür, "Cascaded deep convolutional encoder-decoder neural networks for efficient liver tumor segmentation," *Medical Hypotheses*, vol. 134, p. 109431, 2020, doi: 10.1016/j.mehy.2019.109431.

[16] Q. Jin, Z. Meng, C. Sun, H. Cui, and R. Su, "RA-UNet: a hybrid deep attention-aware network to extract liver and tumor in CT scans," *Frontiers in Bioengineering and Biotechnology*, vol. 8, pp. 1–15, 2020, doi: 10.3389/fbioe.2020.605132.

[17] M. H. Hesamian, W. Jia, X. He, and P. Kennedy, "Deep learning techniques for medical image segmentation: achievements and challenges," *Journal of Digital Imaging*, vol. 32, no. 4, pp. 582–596, 2019, doi: 10.1007/s10278-019-00227-x.

[18] G. Litjens *et al.*, "A survey on deep learning in medical image analysis," *Medical Image Analysis*, vol. 42, pp. 60–88, 2017, doi: 10.1016/j.media.2017.07.005.

[19] O. I. Alirr, "Deep learning and level set approach for liver and tumor segmentation from CT scans," *Journal of Applied Clinical Medical Physics*, vol. 21, no. 10, pp. 200–209, Oct. 2020, doi: 10.1002/acm2.13003.

[20] L. Chen *et al.*, "Liver tumor segmentation in CT volumes using an adversarial densely connected network," *BMC Bioinformatics*, vol. 20, no. S16, pp. 1–13, 2019, doi: 10.1186/s12859-019-3069-x.

[21] N. Gruber, S. Antholzer, W. Jaschke, C. Kremser, and M. Haltmeier, "A joint deep learning approach for automated liver and tumor segmentation," in *2019 13th International conference on Sampling Theory and Applications (SampTA)*, 2019, pp. 1–5, doi:

10.1109/SampTA45681.2019.9030909.
[22] W. Wu, S. Wu, Z. Zhou, R. Zhang, and Y. Zhang, "3D liver tumor segmentation in CT images using improved fuzzy C -means and graph cuts," *BioMed Research International*, vol. 2017, pp. 1–11, 2017, doi: 10.1155/2017/5207685.
[23] S. Almotairi, G. Kareem, M. Aouf, B. Almutairi, and M. A.-M. Salem, "Liver tumor segmentation in CT scans using modified SegNet," *Sensors*, vol. 20, no. 5, p. 1516, 2020, doi: 10.3390/s20051516.
[24] A. Turečková, T. Tureček, Z. K. Oplatková, and A. R. -Sánchez, "Improving CT image tumor segmentation through deep supervision and attentional gates," *Frontiers in Robotics and AI*, vol. 7, pp. 1–21, 2020, doi: 10.3389/frobt.2020.00106.
[25] E. Vorontsov, A. Tang, C. Pal, and S. Kadoury, "Liver lesion segmentation informed by joint liver segmentation," in *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, 2018, pp. 1332–1335, doi: 10.1109/ISBI.2018.8363817.
[26] D. Yang *et al.*, "Automatic liver segmentation using an adversarial image-to-image network," in *International conference on medical image computing and computer-assisted intervention*, Cham: Springer, 2017, pp. 507–515, doi: 10.1007/978-3-319-66179-7_58.
[27] M. S. Lewicki and T. J. Sejnowski, "Learning overcomplete representations," *Neural Computation*, vol. 12, no. 2, pp. 337–365, 2000, doi: 10.1162/089976600300015826.
[28] P. Vincent, H. Larochelle, Y. Bengio, and P.-A. Manzagol, "Extracting and composing robust features with denoising autoencoders," in *Proceedings of the 25th international conference on Machine learning*, 2008, pp. 1096–1103, doi: 10.1145/1390156.1390294.
[29] S. Druckmann and D. Chklovskii, "Over-complete representations on recurrent neural networks can support persistent percepts," in *Advances in Neural Information Processing Systems*, 2010, vol. 23.
[30] Y. Bengio, J. Louradour, R. Collobert, and J. Weston, "Curriculum learning," in *Proceedings of the 26th Annual International Conference on Machine Learning*, 2009, pp. 1–8, doi: 10.1145/1553374.1553380.
[31] R. French, "Catastrophic forgetting in connectionist networks," *Trends in Cognitive Sciences*, vol. 3, no. 4, pp. 128–135, 1999, doi: 10.1016/S1364-6613(99)01294-2.
[32] 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*, vol. 1, no. 1, 2010.
[33] P. F. Christ *et al.*, "Automatic liver and tumor segmentation of CT and MRI volumes using cascaded fully convolutional neural networks," *Arxiv-Computer Science*, pp. 1–20, 2017.
[34] A. M. Reza, "Realization of the contrast limited adaptive histogram equalization (clahe) for real-time image enhancement," *The Journal of VLSI Signal Processing-Systems for Signal, Image, and Video Technology*, vol. 38, no. 1, pp. 35–44, 2004, doi: 10.1023/B:VLSI.0000028532.53893.82.
[35] A. A. Albishri, S. J. H. Shah, and Y. Lee, "CU-Net: cascaded U-Net model for automated liver and lesion segmentation and summarization," in *2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 2019, pp. 1416–1423, doi: 10.1109/BIBM47256.2019.8983266.
[36] B. M. Tummala and S. S. Barpanda, "Liver tumor segmentation from computed tomography images using multiscale residual dilated encoder-decoder network," *International Journal of Imaging Systems and Technology*, vol. 32, no. 2, pp. 600–613, 2022, doi: 10.1002/ima.22640.

BIOGRAPHIES OF AUTHORS

Bindu Madhavi Tummala received her B.Tech and M.Tech degrees in computer science and engineering from DVR & Dr. HS MIC College of Technology, Vijayawada, Andhra Pradesh. She is currently pursuing her Ph.D. degree in VIT-AP University, Andhra Pradesh, India. Her research areas include image processing, medical image analysis, and deep learning. She can be contacted at email: bindumahavi.t@vitap.ac.in.

Soubhagya Sankar Barpanda received his B.Tech degree in computer science and engineering from the Biju Patnaik University of Technology, Rourkela, M.Tech degree in software engineering from the NIT Rourkela, and Ph.D. degree from the NIT Rourkela. He is currently working asassociate professor in School of Computer Science and Engineering, VIT-AP University, Andhra Pradesh, India. His research area is image processing, software engineering, and biometrics. He can be contacted at email:soubhagya.barpanda@vitap.ac.in.