Alexandria Engineering Journal (2021) 60, 1231–1239

HOSTED BY

Alexandria University

Alexandria Engineering Journal

www.elsevier.com/locate/aej
www.sciencedirect.com

Efficient 3D Deep Learning Model for Medical Image Semantic Segmentation

Nasser Alalwan [a,*], Amr Abozeid [b], AbdAllah A. ElHabshy [b], Ahmed Alzahrani [a]

[a] Computer Science Department, Community College, King Saud University, Riyadh, Saudi Arabia
[b] Mathematics Department, Faculty of Science, Al-Azhar University

Received 21 June 2020; revised 27 September 2020; accepted 14 October 2020
Available online 7 November 2020

KEYWORDS

Deep learning;
Semantic segmentation;
Medical imaging;
Depthwise separable convolution;
Computed tomography

Abstract Medical image segmentation is important for disease diagnosis and support medical decision systems. The study proposes an efficient 3D semantic segmentation deep learning model "3D-DenseUNet-569" for liver and tumor segmentation. The proposed 3D-DenseUNet-569 is a fully 3D semantic segmentation model with a significantly deeper network and lower trainable parameters. The proposed model adopts Depthwise Separable Convolution (DS-Conv) as opposed to traditional convolution. The DS-Conv significantly decreases GPU memory requirements and computational cost and achieves high performance. The proposed 3D-DenseUNet-569 utilizes DensNet connections and UNet links, which preserve low-level features and produce effective results. The results of experimental study on the standard LiTS dataset demonstrate that the 3D-DenseNet-569 model is effective and efficient with respect to related studies.

1. Introduction

Nowadays, medical diagnostic scans, for example Computed Tomography (CT) and Magnetic Resonance Imaging (MRI) are important for diagnosis and assessment the treatment of many disease. In clinics, human experts including radiologists and physicians mainly perform medical image interpretation. Recent advances in deep learning techniques are aiding in segmenting, classifying, and identifying patterns or organs in medical applications [1].

Liver tumor (cancer) is the most common tumor disease worldwide, and it leads to significant fatalities on an annual basis. Precise tumor measurements (from MRI and CT), including tumor size, location, and shape, can aid doctors in making precise cancer assessment and treatment planning [2,3]. Abnormality in form and texture of liver in MRI and CT images are significant biomarkers for the diagnosis of initial disease and progression in liver tumor disease [4,5].

Manual or semi-manual (traditional) segmentation methods are used to analyze medical images for diagnosing liver cancer. Unfortunately, traditional segmentation methods are based on the operator, and thus they are subjective and very time-consuming. Therefore, automatic segmentation techniques of liver and tumor are in high demand. Previously,

* Corresponding author.
E-mail addresses: nalalwan@ksu.edu.sa (N. Alalwan), aabozeid@azhar.edu.eg (A. Abozeid), abdallah@azhar.edu.eg (AbdAllah A. ElHabshy), ahmed@ksu.edu.sa (A. Alzahrani).
Peer review under responsibility of Faculty of Engineering, Alexandria University.

https://doi.org/10.1016/j.aej.2020.10.046

many computer-aided (automatic) techniques were developed. Actually, develop effective and efficient semantic segmentation techniques are interesting research topic for many areas.

The automatic segmentation of liver and tumor faces many challenges including the contrast level between liver and tumor is relatively small, there are varying sizes and types of liver tumors [6] and anomalies in tissues (e.g., after tumour resection). In addition, there are different acquisition techniques and multi resolution scanners, which lead to unpredictable intensity variation between a liver and its tumor [3].

Recently, deep learning techniques gained significant attention to handel a lot of computer vision problems [1,7–11]. Specifically, the Convolutional Neural Networks (CNN) and the Fully CNN (FCNs) achieved significant success in medical image segmentation, classification, and recognition. The FCNs models are highly stable with different image resolution, and this encourages researchers to use them for liver tumor segmentation from CT and MRI 3D images (volumes) [1,12].

Generally, medical image segmentation using CNNs/FCNs techniques was classified into two classes. (A) 2D CNNs/FCNs, for example UNet [7], FCN based on VGG-16 model [13], and multi-channel FCN [14]. (B) 3D CNNs/FCNs, where the 2D convolution operations that applied on 2D images are generalized to 3D convolution operations to be applied on 3D images [3,9].

As opposed to analyzing slice by slice, the 3D FCNs models analyze the input slices as volumetric and utilize global features between the input slices. Thus, accuracy of 3D FCNs models are better than accuracy of 2D FCNs models. Unfortunately, memory footprint and execution time of 3D FCNs are relatively high. This limits the input size, network depth, and filter size, which are important factors for achieving high performance [15]. Additionally, the high computational cost of the 3D FCNs impedes training on a large dataset.

The study proposes 3D-DenseUNet-569, which is an efficient 3D deep learning model for semantic liver and its tumor segmentation. The 3D-DenseUNet-569 architecture exhibits the following advantages:

1. Utilizes the advantages of both UNet links and DenseNets connections, and this maintains low-level features and leads to fast training and results that are more accurate.
2. Decreases the pooling layer by adopting a standard convolution with strides, and this maintains the resolution and decreases the memory consumption.
3. Adopts the DS-Conv as opposed to traditional convolution, and this significantly decreases memory consumption and computation while improving performance [16].

The 3D-DenseUNet-569 model was trained and evaluated on a standard Liver Tumor Segmentation (LiTS-2017) dataset [17]. The experimental results show the 3D-DenseUNet-569 model achieves high accuracy (96.7% for liver segmentation and 80.7% for tumor segmentation) which outperforms related models for liver and tumor segmentation tasks.

The paper organization is: Section 2 presents a related work. Section 3 explains and discuss the proposed 3D-DenseUNet-569 methodology and architecture. Section 4 presents our experiments and analyzes their results. Finally, Section 5 concludes the study and suggests some future research.

2. Related work

Many automatic segmentation methods are proposed including texture based methods [4], statistical shape models [18], region growing and threshold methods [19], graph cut techniques [20], and sigmoid edge techniques [21]. Unfortunately, the techniques depend on handcrafted features and have poor feature representation.

Recently, CNNs and FCNs gained significant success in solving many problems in computer vision fields. Several studies adapted this direction and applied various CNNs/FCNs for liver and tumor segmentation [7,22,23].

Ben-Cohen et al. [13] proposed a FCN architecture for liver and tumor segmentation from CT abdomen images. Christ et al. [24] presented an automatic method for liver and lesions segmentation from CT images by utilizing a cascaded FCNs and 3D Conditional Random Fields (CRFs).

Sun et al. [14] proposed a multi-channel FCN to segment liver tumors from multi contrast enhanced CT. This model includes three channels which can be individually trained to extract different features from a CT scans. Subsequently, a feature fusion of CECT images is used to generate probability maps.

Most of the aforementioned methods are based on 2D FCN with different network depths that process the input images slice-by-slice. However, all the 2D FCN methods ignore the 3D context features, which restrict global feature extraction capability and decrease segmentation accuracy.

During the Liver Tumor Segmentation challenge (LiTS-2017) [25], Han [26] proposed an optimal model of the first round and presented 2.5D CFN model that takes a five adjacent slices as one input slice and constructs a segmentation relative to the center slice. The model includes 32 layers and utilizes short-range residual connections from ResNet [27] and long-range concatenation connections of U-Net [7]. The optimal model for the second round of LiTS was proposed by the Lenovo Research group, China. The model employed two cascaded neural networks for liver and its tumor segmentation. This model costists of cascaded 2D and 2.5D U-Net models.

Recently, an efficient design of 3D FCN models and advancing GPU technology facilitate appling these methods to 3D medical imaging. Çiçek et al. [28] proposed a successful 3D FCN architecture, termed as3D UNet, with skip connections to 3D segmentation from sparse annotation of biomedical images. Roth et al. [9] presented a cascaded 3D FCN and an enhanced 3D UNet for multi-organ (e.g., liver, spleen, and pancreas) segmentation.

With respect to 2D and 3D, 3D FCN exhibit high accuracy but it is more complex and required high memory [2]. The high complexity impedes the model in training a larger dataset. The high memory footprint leads to reduce the network depth and the filters size, which Which negatively affects the performance [29].

Most deep learning researchers believe that the deeper model is the most effective model [27,30]. Nevertheless, the deeper model faces popping/vanishing gradients problems, which obstruct convergence through training. Huang et al. [31] proposed densely connected networks (DenseNets) to address these issues. The architecture of DenseNets contains direct connections between the consecutive layers, and this

facilitates training and produces accurate results. DenseNets provides a direct connection among the entire layers, thereby improving flow of gradients and information throughout the layers. Furthermore, dense connections decrease the overfitting problem when working on a smaller training dataset. DenseNets exhibited good performances in image classification tasks. Therefore, many researchers extend the definition of DenseNets to segmentation problems [2,30,32–34].

Li et al. [2] presented a hybrid densely UNet (H-DenseUNet) architecture to segment liver and tumor. The H-DenseUNet extracts local and global features from slices via a cascaded 2D and 3D DenseUNet networks. The architecture addresses heavy computation and memory consumption problems where local and global features are extracted and fused together for precise liver and tumor segmentation.

3. Methodology

Based on the great success of DenseNets in medical images segmentation [2,30,35], we propose an efficient, 3D-DenseUNet-569, 3D deep learning model for liver and tumor semantic segmentation. However, the use of DenseNets for 3D image segmentation exhibits the following challenges.

1. The original DenseNet (DenseNet-161) [31] was developed for classification tasks instead of segmentation tasks.
2. Deep densely connected networks for object segmentation actually consist of several pooling and upsampling layers, that decreases the resolution (low level features).
3. The 3D deep neural networks suffer from high GPU memory consumption and computational cost. High memory utilization limits input size, network depth, and filter size, which are important factors in achieving high performance [15]. Additionally, the high computational cost of 3D convolutions limits training on a large-scale dataset.

Given the three aforementioned challenges, we designed an efficient 3D-DenseUNet-569 model. As shown in Fig. 1, the 3D-DenseUNet-569 structure exhibits the following advantages.

1. The proposed 3D-DenseUNet-569 model inherits both advantages of 3D UNet [28] and DenseNets [2,31]. The direct dense connections between the layers produce a growth in the information flow, and this leads to an easier and more accurate training. Additionally, there exists UNet links from the encoding to decoding layers to improve the resolution.
2. The 3D-DenseUNet-569 adopts a standard convolution with strides as opposed to the pooling layer. The strategy offers two significant benefits. Firstly, it eliminates the need for using the pooling operation. Subsequently, it increases the number of training features, which maintains the resolution. Secondly, it decreases the memory consumption.
3. The 3D-DenseUNet-569 replaces the standard convolution in each dense block by a Depthwise Separable Convolution (DS-Conv). The DS-Conv significantly decreases the memory consumption and computation cost while maintaining a comparable (or preferable) performance [16].

The DS-Conv consists of depthwise and pointwise (i.e., 1×1) convolutions, which significantly decreases the computation cost and memory consumption. A 3D depthwise convolution executes a 3D convolution on each input volume while a 3D pointwise convolution is utilized to merge the 3D depthwise convolution outputs [15].

As depicted in Fig. 1, the 3D-DenseUNet-569 model consists of iterative 3D DenseNet blocks. Each 3D densely block includes different output dimensions, and there are links among the consecutive layers. The connected path between the densely blocks guarantees the maximum flow of information, which enhances convergence to an optimal solution in deep networks [2].

We assume that $I \in R^{n\times(224\times224\times12)\times cn}$ denotes the 3D training samples (for $224 \times 224 \times 12$ 3D input volume) with ground-truth labels: $Y \in R^{n\times(224\times224\times12)\times1}$, such that n denotes the input batch size and the last dimension cn represents the channel. Each pixel (i, j, k)can be classified into a class c (liver, tumor, and background), i.e., $Y_{i,j,k} = c$, and thus the 3D-DenseUNet-569 conducts segmentation for liver and tumor as follows:

$$X = f(I, \theta); X \in R^{n\times(224\times224\times12)\times64},$$

$$\hat{y} = f(X, \theta), \hat{y} \in R^{n\times(224\times224\times12)\times3} \quad (1)$$

where X denotes the feature vector from the final up-sampling layer (layer 5, see Fig. 1) and $\hat{y}$ denotes the corresponding pixelwise probabilities for the input I. Let θ denote the model parameters (for e.g., convolution weights and bias terms of the rectified linear unit).

The depth of the proposed 3D-DenseUNet-569 is extended to 569 main layers. The layers consist of convolution layers, depthwise convolution layers, pointwise convolution layers, Batch Normalization (BN) layers, activation layers, transition blocks, and upsampling blocks.

The transition block is employed to decrease the size of the feature vectors. Each transition block contains a BN layer, a 1x1 convolution layer and a convolution layer with stride 2 as opposed to a pooling layer. Furthermore, a compression factor is set to 0.5 during the experiments to prevent expansion of feature maps.

The upsampling block consists of a bilinear interpolation layer, and this is followed by a combination of low-level features from the opposite dense block (i.e., UNet connections) and a 3x3 convolutional layer. Specifically, BN layer and activation layer were applied after each convolution layer. The BN aids in creating deeper and wider networks, and it is an important factor in the considerable success of ResNet. The activation layer implements a rectified linear unit (ReLU) layer. The ReLU is employed to optimize the proposed architecture and consequently enhance the performance [36].

With the proposed architecture, the number of parameters corresponds to 3.6 Million training parameters, and the network depth is extended to 569 layers.

3.1. Loss function

The proposed model learns parameters/features from CT images by minimizing the loss function. The loss function is adapted as weighted cross entropy as follows:

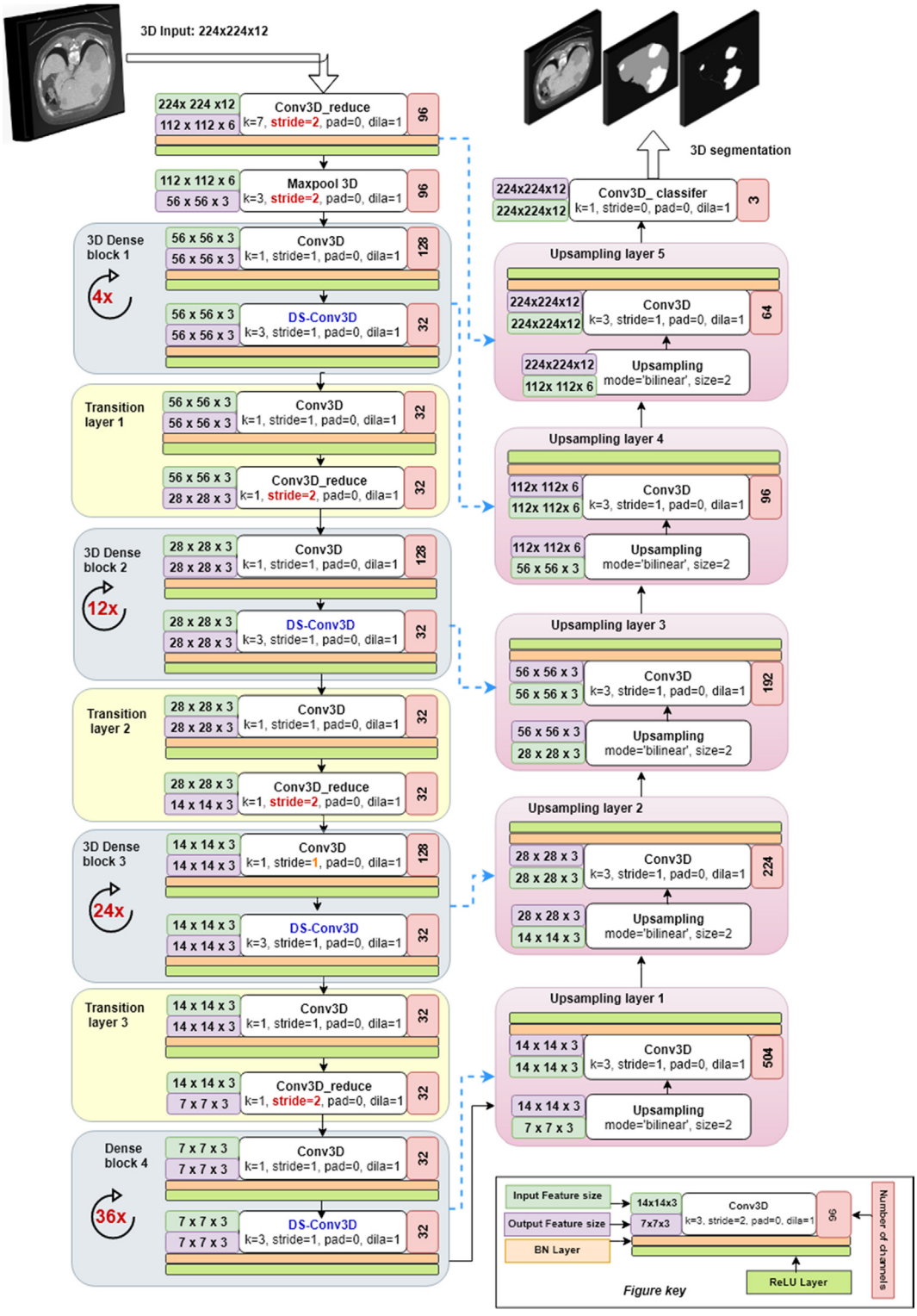

Fig. 1 Proposed 3D-DenseUNet-569 Architecture for Semantic Liver and Tumor Segmentation.

$$L(y,\hat{y}) = -\frac{1}{N}\sum_{i=1}^{N}\sum_{c=1}^{3} w_i^c y_i^c \log \hat{y}_i^c \quad (2)$$

where y_i^c is probability of the ground truth, $\hat{y}_i^c$ is the predicted probability of voxel i belongs to specific class c (liver, tumor, or background), and w_i^c denotes a weighting factor for each class. Empirically, the weights are set as 1.2 for liver, 2.2 for liver lesion, and 0.2 for background.

3.2. Training scheme

Many extant studies proved the effectiveness of knowledge transfer to boost performance [37,38]. Unfortunately, a

reliable pre-trained 3D model is absent, and this limits knowledge transfer and thereby the performance.

To address the problem, we adopted cascaded training methodology, which yields good performance and fast convergence. In the methodology, the model was initially trained for 100 epochs, and each epoch contains 10 steps (sub-epochs). The optimal weights obtained in the initial training phase are used as a base for the next/final phase. In next training phase, the 3D-DenseUNet-569 was trained for 1000 epochs in which each epoch contains 10 steps.

4. Experimental results and analysis

4.1. Dataset

The 3D-DenseUNet-569 was trained and evaluated on the well-known Liver Tumor Segmentation Challenge (LiTS-2017) dataset [17]. The LiTS-2017 dataset contains 201 3D CT scans of the abdomen, 131 for training, and 70 for testing. The LiTS-2017 dataset was collected from many different clinical sites via different CT scanners and methodologies. The dataset exhibits a different plane resolution in the range [0.55,1.0] mm and slice spacing in range [0.45,6.0] mm. To remove any irrelevant details, each image intensity out of the range [-200, 250] HU was truncated. To augment the dataset and alleviate the overheating problem, we adopted a random scaling and mirror from 0.8 to 1.2 for all the training data.

4.2. Evaluation metrics

Based on the LiTS benchmark [23], we utilized the Dice per case score (DICE) and Dice global score to evaluate segmentation effectiveness. The DICE is an F-measures score, which measures the average of recall and precision. For a segmentation task, it measures the overlap degree between the reference and predicted segmentation masks. We assume that A and B are two segmentation masks that the DICE defines as follows:

$$\mathrm{DICE}(A,B) = \frac{2|A \cap B|}{|A| + |B|} \tag{3}$$

The DICE final scores are located in the range [0,1] the value 1 represents an accurate segmentation. The DICE score is represent an average score per each case. The Dice global score is an average Dice score for the entire dataset. The Dice global score is more sensitive to larger tumors than small tumors [23].

4.3. Implementation details

The 3D-DenseUNet-569 model was implemented by TensorFlow [39] with backend Keras packages [40]. Momentum was selected as 0.5, and the initial learning rate is 0.01, and then decreased by a value of 2 after every 10 epochs. The training of the 3D-DenseUNet-569 model required approximately 42 h on a single NVIDIA GTX 1080 with 8 GB memory. It is noted that the 3D-DenseUNet-569 is an end-to-end model. The total time required for the segmentation ranged from 34 s to 130 s for the LiTS dataset. The required time depends on the resolution and the slice number for each scan.

4.4. Results and discussion

Fig. 2 displays the training losses of the 3D-DenseUNet-569. The advantage of our model is that we can train it in a cascaded mode, which yields fast convergence to an optimal solution. As shown in Fig. 2, the convergence of second training phase is faster and realizes lower loss values. Moreover, the results indicate that the 3D-DenseUNet-569 utilizes the transfer learning strategy.

To validate the effectiveness, the proposed 3D-DenseUNet-569 is compared with Kaluva et al. (2D densely connected model) [32], H-DenseUNet (hybrid densely UNet model) [2], and RA-UNet (3D UNet model) [41]. These models were chosen as it had high accuracy results, also it represents the different alternatives to the proposed model. From Table 1, the proposed method achieves a DICE score of 96.2% and DICE global score of 96.7%. For tumor burden evaluation, our method reached a DICE score of 69.6%, DICE global score of 80.7% corresponds to a desirable performance in the LiTS challenge for liver and tumor segmentation.

The number of trainable parameters depends on the model structure (i.e. number and type of the layers). The 3D DS-Conv reduces the number of training parameters than the standard 3D convolution, which reduces the required memory and computational power while maintaining a good performance [42].

As in Fig. 3, the proposed 3D-DenseUNet-569 outperformed the related methods for liver segmentation. Unfortunately, the H-DenseUNet [2] outperformed our proposed model for tumor segmentation (see Fig. 4). A suitable explanation can be that the H-DenseUNet [2] utilizes local and global slices features through a cascaded 2D/3D DenseUNet. It is worth noting that, the proposed 3D-DenseUNet-569 is more efficient than the H-DenseUNet [2] with a significantly deeper network (569 layers) and lower trainable parameters (3.6 Million).

Fig. 5 shows examples of LiTS tumor segmentation results of the 3D-Denseunet-569. It demonstrates that the proposed 3D-Denseunet-569 model can successfully segment liver and tumors in large regions. The proposed model can segment tumors that are small and hard to detect. Given the very small

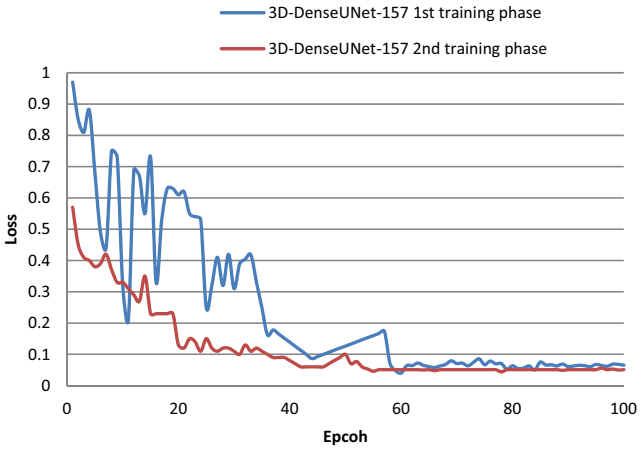

Fig. 2 Training Losses of the 3D-Denseunet-569.

Table 1 The 3D-DenseUNet-569 Effectiveness (Dice: %) Compared with Other Related Methods on the LiTS Dataset.

	Dimension	No. of trainable parameters(Million)	LiTS liver		LiTS tumor	
			DICE	DICE global	DICE	DICE global
Kaluva et al. [32]	2D	–	91.2	92.3	49.2	62.5
H-DenseUNet [2]	2.5 D	4 M	96.1	96.5	72.2	82.4
3D DenseUNet-65 [2]	3D	–	93.6	92.9	59.4	78.8
RA-UNet [41]	3D	4 M	96.1	96.3	59.5	79.5
The proposed 3D-DenseUNet-569	**3D**	**3.6 M**	**96.2**	**96.7**	**69.6**	**80.7**

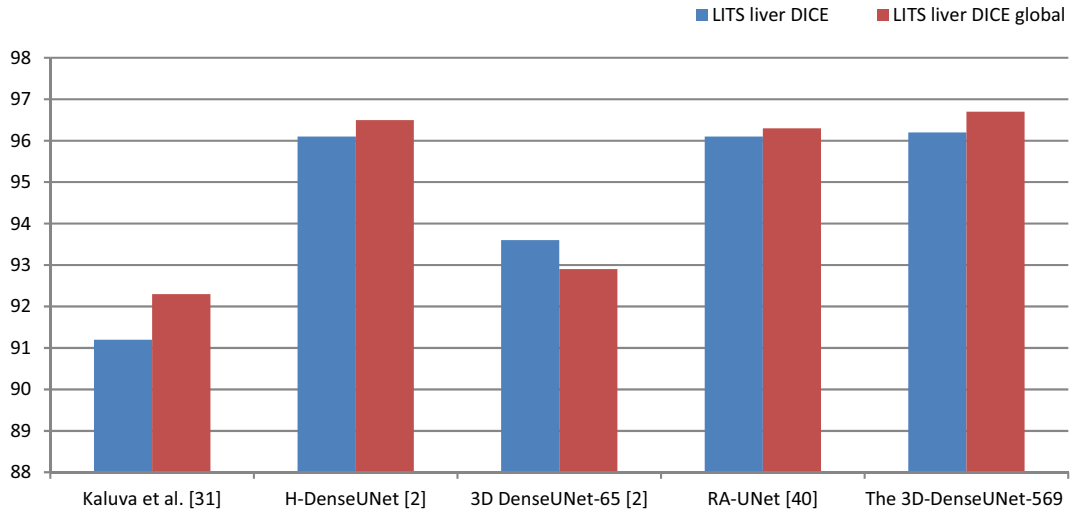

Fig. 3 Liver Segmentation Evaluations for the 3D-DenseUNet-569 Model Compared with the Related Models.

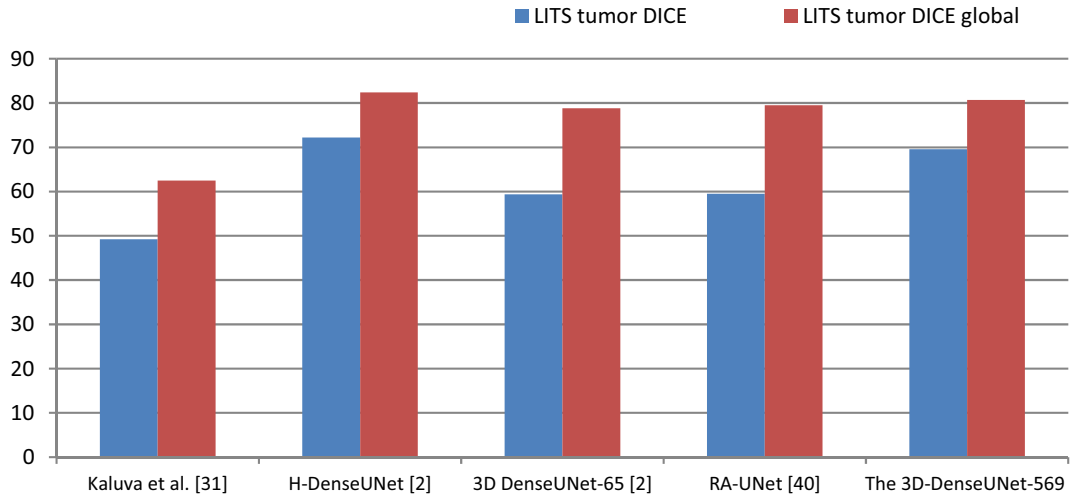

Fig. 4 Liver Tumor Segmentation Evaluations for the 3D-DenseUNet-569 Model Compared with the Related Models.

size and low contrast of a few tumors, the 3D-DenseUNet-569 model still exhibits some false results for tumor segmentation.

In order to evaluate the performance of using the 3D DS-Conv, experiments were conducted on the 3D-DenseUNet-569 model. As in the table 2, the model was trained using 3D DS-Conv and 3D standard convolution. These experiments were conducted on various numbers of input cases (e.g. 5, 10, 15 and 20). As a result, the 3D DS-Conv requires an average 10.4 s for each epoch/step opposed to 38.85 s for the 3D standard convolution. Therefore, these results show that the proposed model with 3D DS-Conv achieves better performance. It's worth to mention that, the processing time for the first training epoch/step is relatively large because it includes the model compiling time.

Table 3 shows that, the number of total parameters as well as trainable parameters of the 3D-DenseUNet-569 using 3D

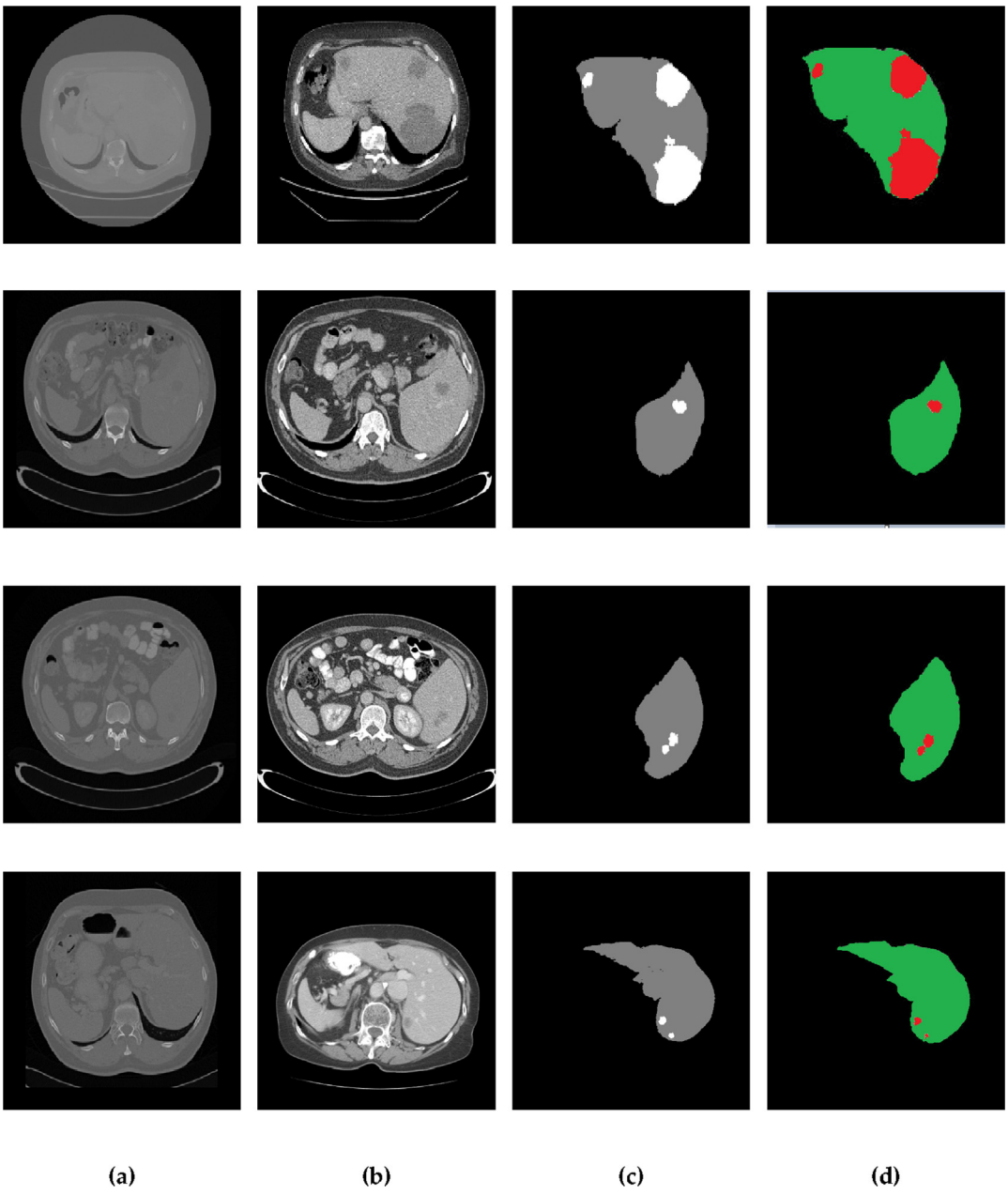

Fig. 5 Examples of Segmentation Results of the 3D-DenseUNet-569. **(a)** Original slice from the LiTS dataset, **(b)** corresponding enhanced intensity, **(c)** corresponding segmentation mask from the LiTS Dataset, **(d)** segmentation results of the 3D-Denseunet-569 where the liver is represented by green region and the tumors are represented by red regions.

DS-Conv is much lower than the 3D-DenseUNet-569 using 3D standard convolution. Therefore, the required memory size has been reduced by about 5 times. All the performance experiments were conducted on GPU equipped with 8 GB memory.

5. Conclusion

This paper proposed an efficient 3D segmentation deep learning model (named 3D-DenseUNet-569) for liver and tumor semantic segmentation from CT volumes. The proposed 3D-DenseUNet-569 model combines advantages of both DensNet connections and UNet links, which preserves low-level features and produces more accurate results. The 3D-DenseUNet-569 adapted the 3D depthwise separable convolution to overcome memory and computational cost limitations of 3D FCNs. Therefore, the 3D-DenseUNet-569 is an efficient fully 3D segmentation model with a significantly deeper network (569 main layers) and a lower number of trainable parameters (3.6 Million). The experimental results demonstrated that, the 3D-DenseUNet-569 model achieves a DICE of 96.2% and DICE global score of 96.7% for liver segmentation. With respect to tumor segmentation, the proposed model achieves a DICE

Table 2 The 3D-DenseUNet-569 performance evaluation on the LiTS Dataset.

	The 3D-DenseUNet-569 using 3D DS-Conv				The 3D-DenseUNet-569 using 3D standard Conv			
Epoch NO	**5-cases**	**10-cases**	**15-cases**	**20-cases**	**5-cases**	**10-cases**	**15-cases**	**20-cases**
1	40	55	59	45	192	241	225	228
2	8	6	7	7	18	17	12	17
3	6	5	6	6	14	10	19	18
4	6	6	6	6	25	14	19	17
5	6	6	7	6	13	15	17	18
6	6	5	6	6	13	14	19	20
7	6	6	6	6	16	13	18	27
8	6	5	6	6	25	17	21	22
9	6	6	6	6	15	32	19	17
10	6	5	6	6	27	30	20	20
Avarage	**9.6**	**10.5**	**11.5**	**10**	**35.8**	**40.3**	**38.9**	**40.4**
Total Avarage	**10.4**				**38.85**			

Table 3 Number of parameters of the 3D-DenseUNet-569 using the 3D DS-Conv compared with using the 3D standard convolution.

	The 3D-DenseUNet-569 using 3D DS-Conv	**The 3D-DenseUNet-569 using 3D standard Conv**
Total params	36,433,587	181,270,071
Trainable params	36,270,875	180,615,711
Non-trainable params	162,712	654,360

of 69.6% and DICE global score of 80.7%. These results represent a high performance in the LiTS challenge.

As a future work, we plan to improve the 3D-DenseUNet-569 architecture to be more general to other medical imaging segmentation tasks such as COVID-19 lesion segmentation of lung CT images. Also, we aim to apply it in real CT clinical cases.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This work was funded by the Researcher Supporting Project (RSP-2020/157), King Saud University, Riyadh, Saudi Arabia.

References

[1] A.S. Lundervold, A. Lundervold, An overview of deep learning in medical imaging focusing on MRI, Zeitschrift für Medizinische Physik (2018).
[2] X. Li et al, H-denseunet: Hybrid densely connected unet for liver and tumor segmentation from ct volumes, IEEE Trans. Med. Imag. 37 (12) (2018) 2663–2674.
[3] P.F. Christ, et al., Automatic liver and tumor segmentation of CT and MRI volumes using cascaded fully convolutional neural networks. arXiv preprint arXiv:1702.05970, 2017.
[4] T. Heimann et al, Comparison and evaluation of methods for liver segmentation from CT datasets, IEEE Trans. Med. Imag. 28 (8) (2009) 1251–1265.
[5] J. Ferlay et al, Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008, Int. J. Cancer 127 (12) (2010) 2893–2917.
[6] L. Sun, et al., Modelling liver cancer initiation with organoids derived from directly reprogrammed human hepatocytes, 21 (8) (2019) pp. 1015–1026.
[7] O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation, International Conference on Medical image computing and computer-assisted intervention, Springer, 2015.
[8] H. Chen et al, VoxResNet: Deep voxelwise residual networks for brain segmentation from 3D MR images, NeuroImage 170 (2018) 446–455.
[9] H.R. Roth et al, An application of cascaded 3D fully convolutional networks for medical image segmentation, Comput. Med. Imaging Graph. 66 (2018) 90–99.
[10] M. Kolařík et al, Optimized High Resolution 3D Dense-U-Net Network for Brain and Spine Segmentation, Appl. Sci. 9 (3) (2019) 404.
[11] G. Litjens et al, A survey on deep learning in medical image analysis, Med. Image Anal. 42 (2017) 60–88.
[12] J. Long, E. Shelhamer, T. Darrell, Fully convolutional networks for semantic segmentation, Proceedings of the IEEE conference on computer vision and pattern recognition, 2015.
[13] A. Ben-Cohen et al, Fully convolutional network for liver segmentation and lesions detection, in: Deep Learning and Data Labeling for Medical Applications, Springer, 2016, pp. 77–85.
[14] C. Sun et al, Automatic segmentation of liver tumors from multiphase contrast-enhanced CT images based on FCNs, Artif. Intell. Med. 83 (2017) 58–66.
[15] K. Simonyan, A.J.A.P.A. Zisserman, Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556, 2014.
[16] L.-C. Chen et al, Encoder-decoder with atrous separable convolution for semantic image segmentation, Proceedings of the European conference on computer vision (ECCV), 2018.
[17] LiTS - Liver Tumor Segmentation Challenge. [cited 10 Sep 2019 10 Sep 2019]; Available from: Available online: https://competitions.codalab.org/competitions/17094.
[18] C. Li et al, Supervised variational model with statistical inference and its application in medical image segmentation, IEEE Trans. Biomed. Eng. 62 (1) (2015) 196–207.

[19] G. Li et al, Human lesion detection method based on image information and brain signal, IEEE Access, 2019.
[20] G. Li et al, Automatic liver segmentation based on shape constraints and deformable graph cut in CT images, IEEE Trans. Image Process. 24 (12) (2015) 5315–5329.
[21] A.H. Foruzan, Y.-W. Chen, Improved segmentation of low-contrast lesions using sigmoid edge model, Int. J. Comput. Assist. Radiol. Surg. 11 (7) (2016) 1267–1283.
[22] M. Moghbel et al, Review of liver segmentation and computer assisted detection/diagnosis methods in computed tomography, Artif. Intell. Rev. 50 (4) (2018) 497–537.
[23] P. Bilic, et al., The Liver Tumor Segmentation Benchmark (LiTS). arXiv preprint arXiv:1901.04056, 2019.
[24] P.F. Christ et al, Automatic liver and lesion segmentation in CT using cascaded fully convolutional neural networks and 3D conditional random fields, International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2016.
[25] Christ P., et al., LiTS - Liver Tumor Segmentation Challenge. 2017 [cited 10 Sep 2019 10 Sep 2019]; Available from: https://competitions.codalab.org/competitions/17094.
[26] X. Han, Automatic liver lesion segmentation using a deep convolutional neural network method. arXiv preprint arXiv:1704.07239, 2017.
[27] K. He et al, Deep residual learning for image recognition, Proceedings of the IEEE conference on computer vision and pattern recognition, 2016.
[28] Ö. Çiçek et al, 3D U-Net: learning dense volumetric segmentation from sparse annotation, International conference on medical image computing and computer-assisted intervention, Springer, 2016.
[29] K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556, 2014.
[30] J. Dolz et al, HyperDense-Net: A hyper-densely connected CNN for multi-modal image segmentation, IEEE Trans. Med. Imag. 38 (5) (2018) 1116–1126.
[31] G. Huang et al, Densely connected convolutional networks, Proceedings of the IEEE conference on computer vision and pattern recognition, 2017.
[32] K.C. Kaluva, et al., 2D-densely connected convolution neural networks for automatic liver and tumor segmentation. arXiv preprint arXiv:1802.02182, 2018.
[33] S. Li et al, Cascade Dense-Unet for Prostate Segmentation in MR Images, International Conference on Intelligent Computing, Springer, 2019.
[34] J.P. Vigueras-Guillén, et al. Automatic detection of the region of interest in corneal endothelium images using dense convolutional neural networks, in: Medical Imaging 2019: Image Processing. 2019. International Society for Optics and Photonics.
[35] Chen, L., et al. MRI tumor segmentation with densely connected 3D CNN. in Medical Imaging 2018: Image Processing. 2018. International Society for Optics and Photonics.
[36] G. Zhao, et al., Rethink ReLU to Training Better CNNs. arXiv preprint arXiv:.06247, 2017.
[37] N. Tajbakhsh et al, Convolutional neural networks for medical image analysis: Full training or fine tuning?, IEEE Trans Med. Imag. 35 (5) (2016) 1299–1312.
[38] V. Cheplygina, M. de Bruijne, J.P. Pluim, Not-so-supervised: a survey of semi-supervised, multi-instance, and transfer learning in medical image analysis, Med. Image Anal. 54 (2019) 280–296.
[39] TensorFlow: An end-to-end open source machine learning platform. [cited 10 Sep 2019 10 Sep 2019]; Available from: Available online: https://www.tensorflow.org/.
[40] Keras: The Python Deep Learning library. [cited 10 Sep 2019]; Available from: Available online: https://keras.io/.
[41] Q. Jin, et al., RA-UNet: A hybrid deep attention-aware network to extract liver and tumor in CT scans. arXiv preprint arXiv:1811.01328 2018.
[42] R. Ye, F. Liu, L. Zhang, 3D depthwise convolution: Reducing model parameters in 3D vision tasks, Canadian Conference on Artificial Intelligence, Springer, 2019.