

# TDC-Unet: Triple Unet with Dilated Convolution for Medical Image Segmentation

Song-Toan Tran<sup>1,2</sup>, Thanh-Tuan Nguyen<sup>1</sup>, Minh-Hai Le<sup>1,2</sup>, Ching-Hwa Cheng<sup>3</sup>, and Don-Gey Liu<sup>1,3</sup>

<sup>1</sup>Feng Chia University, Taichung, Taiwan

<sup>2</sup>Tra Vinh University, Tra Vinh, Viet Nam

<sup>3</sup>Department of Electronic Engineering, Feng Chia University, Taichung, Taiwan

Email: tstoan1512@tvu.edu.vn, nttuan@kgc.edu.vn, lmhai@tvu.edu.vn, chengch@fcu.edu.tw, dgliu@fcu.edu.tw

**Abstract**—Medical image segmentation is one of the research directions that are interested in recent years. The Unet model is one of the most architecture commonly used for medical image segmentation. However, Unet and Unet-based models still have a drawback that is concentrating only on the last feature output of the convolution unit and forgetting the feature of the previous convolution in the node. In this paper, we propose a new model based on Unet model, called by TDC-Unet that would exploit the intra-feature of the nodes in the Unet architecture. We also apply the Dilated Convolution (DC) and dense connection in the nodes structure. We used four datasets, that cover different modalities of medical image: colonoscopy, dermoscopy, and Magnetic Resonance Imaging (MRI) to evaluate the proposed model. The applications in our experiment are: nuclei segmentation, polyp segmentation, left atrium segmentation, and skin lesion segmentation. The experimental results show that our model achieves better results than the current models.

**Index Terms**—medical image segmentation, nuclei segmentation, polyp segmentation, left atrium segmentation, skin lesion segmentation, Unet structure, dilated convolution

## I. INTRODUCTION

Medical image segmentation is an important step in diagnosis and pre-surgery. It greatly aids the doctor in diagnosis and decision-making before surgery. Medical image processing is one of the challenges that are of concern for the researchers. There are many approaches to the medical image segmentation challenge. However, deep learning is showing outstanding advantages compared to traditional image processing methods. Two deep learning architectures that are commonly used for segmentation tasks are Fully Convolutional Network (FCN) [1] and Unet [2]. Deep learning is used for many medical imaging segmentation applications such as liver and liver tumor [3], [4], brain tumor [3], [5], [6], polyp [3], [7]-[9], nuclei [3], [5], [10].

The Unet is most commonly used as a platform architecture for medical image segmentation models [11]. The advantages of Unet are to combine low-level, coarse-grained features of the encoder with high-level, fine-

grained features of the decoder node. However, Unet still has some disadvantages such as: optimizing for different applications with different numbers and sizes of datasets is not easy, the skip connection structure still does not take full advantage of the features from the encoder node. The traditional Unet models and Unet-based models are used in many medical image segmentation applications. These studies focus on solving the shortcomings of Unet by modifying the node structure of the encoder and decoder blocks [12], [13], or changing the structure of the skip connection [3]-[5], or using the cascade structure [10], [14], [15].

The conventional Unet and the new approaches based on Unet only focus on the last output feature of the convolution node. The features from the previous nodes were forgotten. To address this problem, in this paper, we propose a new architecture based on Unet, called TDC-Unet. TDC-Unet consists of 3 sub-Unets arranged in parallel. With the new architecture, the intra-features of the nodes are explored. All of the features in the nodes would be utilized. The skip connections also use all the features from the encoder node to combine with the decoder node. The convolution was used in the nodes is Dilated Convolution (DC) [16]. The advantage of DC is that it covers a large feature area without using the pooling function. With the same number of DC parameters, it is possible to extend the region of interest larger than traditional convolution. The dense structure [17] is also used in the nodes, tackle the vanishing-gradient problem, and use the feature more efficient.

In summary, the contributions of this study are: (i) introduce the new Unet structure network called TDC-Unet, for medical image segmentation. The new model exploits the intra-feature of the convolution node more efficient; (ii) verify the efficiency of DC and dense structure for the medical image segmentation challenges; (iii) evaluate the effectiveness of the model on 4 datasets that cover different modalities of medical image, thereby proving the high generality of the proposed model.

## II. TDC-UNET NETWORK

Fig. 1(a) illustrates the network structure of the TDC-Unet model. Similar to conventional Unet, TDC-Unet consists of two main parts: encoder and decoder. Each part includes four convolution nodes. The transition node

is at the bottom of the network. There are three convolution units in the nodes of the TDC-Unet. The convolution unit is composed of two convolutional functions followed by the ReLU function and batch normalization (BN) (Fig. 1(b)). The connections of the nodes are detailed in Fig. 1(c).

In this study, we apply the dense connection to the nodes of the TDC-Unet. For the convolution function, we employ the dilated convolution with the dilation rate is 3, 2, and 1 for the first, the second, and the third convolution units, respectively. Formally, let  $x_{E/D}^{i,j}$  is the output feature of the  $j^{th}$  convolution unit in the  $i^{th}$  node and  $C^r(\cdot)$  is a dilated convolution function, that has the dilation rate equal to  $r$ , followed by the ReLU function and BN. The subscript letter, "E/D", denotes the encoder or the decoder. The output features of the node are described as:

$$X_{E/D}^i = \{x_{E/D}^{i,1}, x_{E/D}^{i,2}, x_{E/D}^{i,3}\} \quad (1)$$

$$X_T = \{x_T^1, x_T^2, \dots, x_T^3\} \quad (2)$$

where the  $X_{E/D}^i$  is the output feature of the  $i^{th}$  node,  $X_T$

and  $x_T^i$  are the output features of the transition node and the  $i^{th}$  convolution unit, respectively. The output features of the first encoder node are computed as follows,

$$x_E^{1,j} = \begin{cases} C^{4-j}(input) & j = 1 \\ C^{4-j}([\![x_E^{1,k}]\!]_{k=1}^{j-1}) & j > 1 \end{cases} \quad (3)$$

where the  $[\![\cdot]\!]$  is the concatenation function. From the second to the fourth encoder node, the output features are computed as follows,

$$x_E^{i,j} = \begin{cases} C^{4-j}(P(x_E^{i-1,j})) & j = 1 \\ C^{4-j}([\![x_E^{i,k}]\!]_{k=1}^{j-1}, P(x_E^{i-1,j})]) & j > 1 \end{cases} \quad (4)$$

where the  $P(\cdot)$  depicts a max pooling function. The output features of the transition node are obtained by getting the feature map of the fourth encoder node, are described as:

$$x_T^j = \begin{cases} C^{4-j}(P(x_E^{4,j})) & j = 1 \\ C^{4-j}([\![x_T^k]\!]_{k=1}^{j-1}, P(x_E^{4,j})]) & j > 1 \end{cases} \quad (5)$$

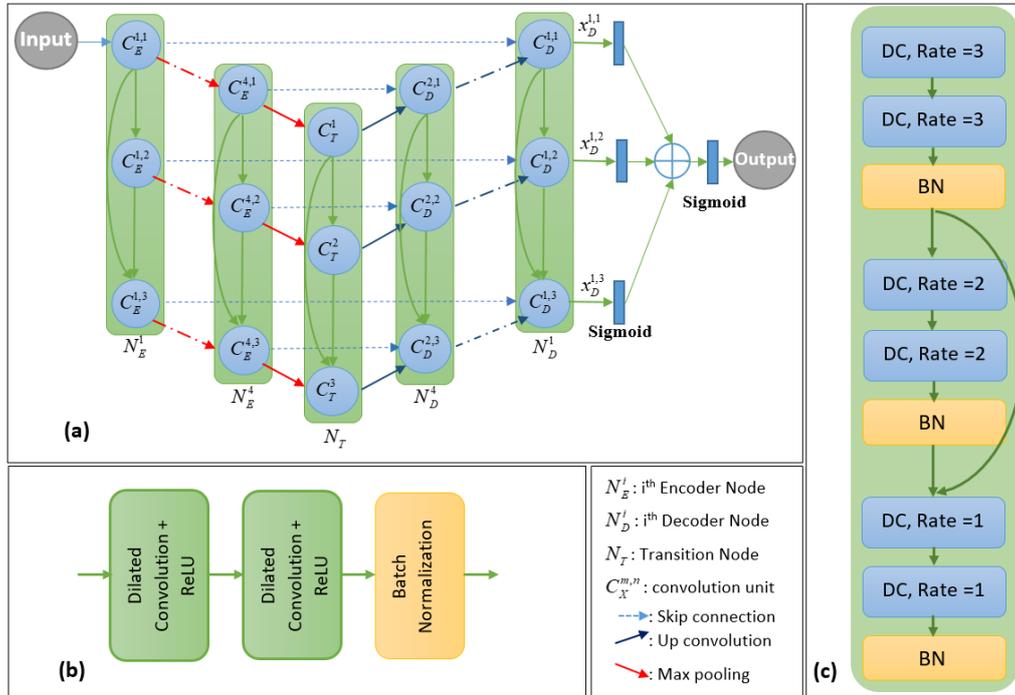


Figure 1. (a) The architecture of TDC-Unet model, (b) the convolution unit, and (c) the node structure.

In the decoder part, the nodes get the features from the encoder node and lower decoder nodes. The output features of the decoder nodes are computed as

$$x_D^{4,j} = \begin{cases} C^{4-j}([\![U(x_T^j), x_E^{4,j}]\!]]) & j = 1 \\ C^{4-j}([\![x_D^{4,k}]\!]_{k=1}^{j-1}, U(x_T^j), x_E^{4,j}]) & j > 1 \end{cases} \quad (6)$$

$$x_D^{i,j} = \begin{cases} C^{4-j}([\![U(x_D^{i+1,j}), x_E^{i,j}]\!]]) & j = 1 \\ C^{4-j}([\![x_D^{i,k}]\!]_{k=1}^{j-1}, U(x_D^{i+1,j}), x_E^{i,j}]) & j > 1 \end{cases} \quad (7)$$

$i \in [1, 3]$

where  $U(\cdot)$  indicates an up-convolution operation.

The output of the convolution units of the top decoder node  $\{x_d^{1,1}, x_d^{1,2}, x_d^{1,3}\}$  is followed by the sigmoid function to achieve the results  $\{O_{v_1}, O_{v_2}, O_{v_3}\}$ . The final output of the TDC-Unet model  $O_{TDC}$ , which is obtained by using the below equation.

$$O_{TDC} = \text{sigmoid}\left(\left[O_{v_1}, O_{v_2}, O_{v_3}\right]\right) \quad (8)$$

Table I shows the architecture of TDC-Unet in detail. The convolution kernel, which is used for the convolution units, is the size of 3x3.

TABLE I. THE ARCHITECTURE OF TDC-UNET

Nodes	Encoder	Decoder
1	16 x (DiConv 3 + ReLU) <sup>2</sup> + BN + [(2x2) maxpooling] 16 x (DiConv 2 + ReLU) <sup>2</sup> + BN + [(2x2) maxpooling] 16 x (DiConv 1 + ReLU) <sup>2</sup> + BN + [(2x2) maxpooling]	[(2x2) deconv] + 16 x (DiConv 3 + ReLU) <sup>2</sup> + BN [(2x2) deconv] + 16 x (DiConv 2 + ReLU) <sup>2</sup> + BN [(2x2) deconv] + 16 x (DiConv 1 + ReLU) <sup>2</sup> + BN
2	32 x (DiConv 3 + ReLU) <sup>2</sup> + BN + [(2x2) maxpooling] 32 x (DiConv 2 + ReLU) <sup>2</sup> + BN + [(2x2) maxpooling] 32 x (DiConv 1 + ReLU) <sup>2</sup> + BN + [(2x2) maxpooling]	[(2x2) deconv] + 32 x (DiConv 3 + ReLU) <sup>2</sup> + BN [(2x2) deconv] + 32 x (DiConv 2 + ReLU) <sup>2</sup> + BN [(2x2) deconv] + 32 x (DiConv 1 + ReLU) <sup>2</sup> + BN
3	64 x (DiConv 3 + ReLU) <sup>2</sup> + BN + [(2x2) maxpooling] 64 x (DiConv 2 + ReLU) <sup>2</sup> + BN + [(2x2) maxpooling] 64 x (DiConv 1 + ReLU) <sup>2</sup> + BN + [(2x2) maxpooling]	[(2x2) deconv] + 64 x (DiConv 3 + ReLU) <sup>2</sup> + BN [(2x2) deconv] + 64 x (DiConv 2 + ReLU) <sup>2</sup> + BN [(2x2) deconv] + 64 x (DiConv 1 + ReLU) <sup>2</sup> + BN
4	128 x (DiConv 3 + ReLU) <sup>2</sup> + BN + [(2x2) maxpooling] 128 x (DiConv 2 + ReLU) <sup>2</sup> + BN + [(2x2) maxpooling] 128 x (DiConv 1 + ReLU) <sup>2</sup> + BN + [(2x2) maxpooling]	[(2x2) deconv] + 128 x (DiConv 3 + ReLU) <sup>2</sup> + BN [(2x2) deconv] + 128 x (DiConv 2 + ReLU) <sup>2</sup> + BN [(2x2) deconv] + 128 x (DiConv 1 + ReLU) <sup>2</sup> + BN
Transition		256 x (DiConv 3 + ReLU) <sup>2</sup> + BN 256 x (DiConv 2 + ReLU) <sup>2</sup> + BN 256 x (DiConv 1 + ReLU) <sup>2</sup> + BN

“DiConv/d” denotes the 3x3 dilated convolution, ‘d’ is the dilation rate. The superscript number depicts the number of the functions. All convolutional layers include the dropout.

### III. EXPERIMENTS

#### A. Datasets and Data Pre-Processing

We used four datasets that cover different modalities of medical images, such as MRI, colonoscopy, fluorescence, and dermoscopy. Fig. 2 shows some examples of the datasets used in our experiments.

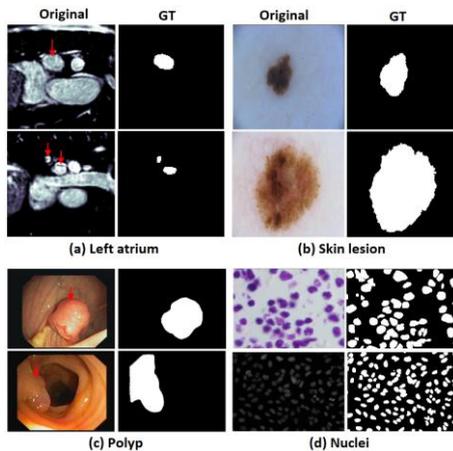


Figure 2. Some examples of the datasets used in this study. The first and the third columns present the original images while the second and the fourth columns indicate the ground truths (GT).

The first dataset is the nuclei dataset, is provided by Data Science Bowl 2018 segmentation challenge (DSB Challenge 2018). It consists of 670 nuclei images with the size of 256x256 from two modalities: brightfield and fluorescence. The image data is accompanied by segmentation results (ground truth). In our experiments, the Nuclei dataset is divided into three parts: 423 images for training, 108 images for validation, and 130 images

for testing. In training and testing progress, we resize the image to the size of 128x128.

The second is the CVC-clinicDB dataset, which comes from the 2015 MICCAI sub-challenge on automatic polyp detection. There is a total of 612 images with a size of 384x288 in the dataset. The images are extracted from 25 different colonoscopy videos and contain several examples of polyps. The ground truths are also included. For training, the dataset is separated as follows: 489 images for training and validation (20% used for validation), 123 images for testing. The images are re-scaled to the size of 224x224 for training and testing the models.

The third dataset is the skin lesion dataset, which is provided by the ISIC-2018 Challenge. The dataset includes 2954 high-resolution dermoscopy images with different sizes. For training, we divide the dataset into three parts: training part (1660 images), validation part (415 images), and testing part (519 images). Both training and testing are done based on 224x224 images.

The last dataset is a CT volumes dataset for the left atrium segmentation challenge. It is provided by Medical Segmentation Decathlon Challenge 2018 (MSD 2018). There are 20 MRI volumes for training and 10 volumes for testing. We only use the training part of the dataset in our experiments. There are 2271 2D slices with a size of 320x320 pixels. The dataset is divided into two parts: 15 volumes for training and validation (1702 slices, 20% used for validation), 5 volumes (569 slices) for testing. In order to reduce the calculation time, we crop the slices to a size of 128x128 after applied the Hounsfield unit in a range of [500,1500]. Another advantage of cropping is to reduce the case of false positives. Table II summarizes the dataset used in our experiments.

All of the datasets are trained with data augmentation techniques. We applied the same data augmentation techniques for each dataset, including shearing, rotation, zoom, flip, and shift.

TABLE II. SUMMARY OF THE DATASET USED IN THIS PAPER

Segmentation Application	No images /slices	Size	Modality	Provider
Nuclei	670	256x256	brightfield and fluorescence	DSB Challenge 2018
Polyp (CVC-clinicDB)	612	384x288	Colonoscopy	MICCAI Challenge 2015
Left atrium	2271	320x320	MRI	MSD Challenge 2018
Skin lesion	2594	Variable	Dermoscopy	ISIC Challenge 2018

### B. Experiments Setting

In our study, we used the hybrid loss to address the imbalance problem. The hybrid loss is a combination between Dice loss and cross-entropy loss. We use two types of cross-entropy loss that are Binary Cross-Entropy (BCE) and Weighted Cross-Entropy (WCE). The formulas of the hybrid loss are expressed by

$$L = L_{DC} + L_{CE} \quad (9)$$

where  $L_{CE}$  and  $L_{DC}$  represent the cross-entropy loss and the dice loss, respectively. The  $L_{CE}$  consists of two types  $L_{WCE}$  and  $L_{BCE}$ . The  $L_{WCE}$ ,  $L_{BCE}$  and  $L_{DC}$  are computed as:

$$L_{WCE} = -\frac{1}{N} \sum_{i=1}^N ((1-w)g_i \log p_i + w(1-g_i) \log(1-p_i)) \quad (10)$$

$$L_{BCE} = -\frac{1}{N} \sum_{i=1}^N (g_i \log p_i + (1-g_i) \log(1-p_i)) \quad (11)$$

$$L_{DC} = 1 - \frac{2 \sum_{i=1}^N (g_i p_i) + \varepsilon}{\sum_{i=1}^N (g_i + p_i) + \varepsilon} \quad (12)$$

where  $p_i$  is the probability that voxel  $i$  is foreground, and  $g_i$  indicates the probability of voxel  $i$  that is the background. The  $N$  is the number of the voxels that are predicted, the  $w$  denotes the weight, and the  $\varepsilon$  is a very small value to prevent the denominator is zero.

Our experimental are conducted by the Keras package with Tensorflow as the backend. The he-normal distribution initializer, which is proposed by He *et al.* [18], is used to initialize the weights. The optimizer is used in the model is Adam optimizer, and the networks are trained with an initial learning rate of  $3e-4$  for all of the datasets except the left atrium dataset initialize with a learning rate of  $1e-3$ . The learning rate is updated by the learning rate scheduler, which is described as:

$$lr = lr \times (0.9^{\text{epochs}/10}) \quad (13)$$

And a dropout rate of 0.2 is applied for preventing the over-fitting. The early stopping mechanism is also used. All experiments are conducted by a workstation with Intel Xeon Silver 4114 CPU, GRID Virtual GPU V100D-8Q, and 32GB RAM memory. Table III shows details of the learning setting of the models for the segmentation application.

TABLE III. THE DETAILS OF THE LEARNING SETTINGS OF THE MODELS

App.	Loss function	Initial Learning rate	No. of Epochs	Batch size	Input size
Nuclei	Dice+BCE	3e-4	200	16	128x128
Polyp	Dice+WCE	3e-4	200	8	224x224
Skin lesion	Dice+BCE	3e-4	200	8	224x224
Left atrium	Dice+WCE	1e-3	100	16	128x128

## IV. RESULT AND DISCUSSION

To evaluate the effectiveness of the proposed model, we also employed Unet and Unet++ models with the same strategies and settings. The evaluation metrics are used in this study include three metrics: Dice Coefficient (DSC), F1-score, and mean Intersection over Union (mIoU). These metrics are described as follow:

$$DSC(G, P) = \frac{2 \times |G \cap P|}{|G| + |P|} \quad (14)$$

$$mIoU(G, P) = \frac{|G \cap P|}{|G \cup P|} \quad (15)$$

$$F1 - score = \frac{TP}{TP + \frac{1}{2}(FP + FN)} \quad (16)$$

where G denotes the ground truth, P denotes the predicted values. TP, FP, TN, FN depict the number case of true positives, false positives, true negatives, and false negatives, respectively.

Fig. 3 presents the segmentation results on the datasets. As seen in Fig. 3, our model performance is better than Unet and Unet++ on all evaluation metrics. The number of parameters for TDC-Unet is more than Unet (8.4 million against 7.8 million), but less than Unet++ (8.4 million against 9.5 million). In comparison with the Unet, our models significantly improve segmentation performance. For nuclei segmentation, the DSC, F1-score and mIoU values of TDC-Unet are greater 0.39%, 0.33% and 0.68%, respectively. The evaluation metrics on polyp segmentation increased 0.64%, 0.82% and 1.18%. The increments of skin lesion applications are 2.62%, 3.04%, and 3.94%. For left atrium segmentation, the values are improved by 1.06%, 1.42%, and 1.66%.

Observing the results, we can see that the segmentation results on skin lesion are more improved compared to the rest. The reason for this difference is the effect of DC, the segmented objects in skin lesion images usually have a large proportion in the image. The result is the opposite of nuclei segmentation, the least improvement due to the generally small size of the objects in the image. This

proves the influence of DC on the segmentation result of the large object in the image. This is the advantage of DC in the image segmentation task. However, DC has not shown efficacy for small objects segmentation. Furthermore, the computation time for DC is larger than that for traditional convolution.

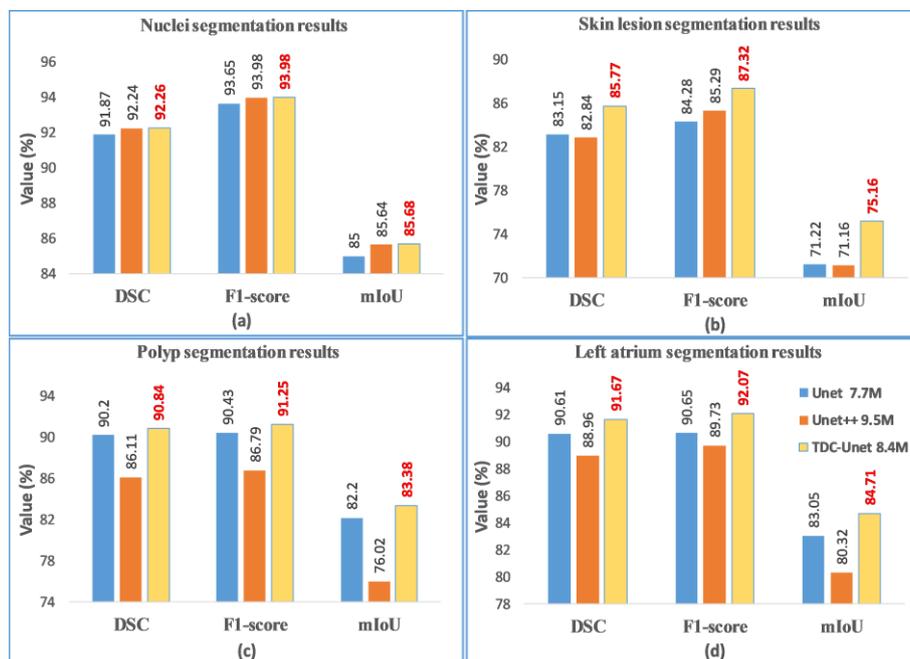


Figure 3. The quantitative segmentation results of the models that are implemented in our experiments.

Fig. 4 shows some examples of segmentation results of the models that were performed in our experiments. We can observe that TDC-Unet model got better results than Unet and Unet++. Table IV compares our model with some popular models on four datasets. The results indicate that our model achieves better results than other networks except for the Deeplab V3+ on skin lesion segmentation.

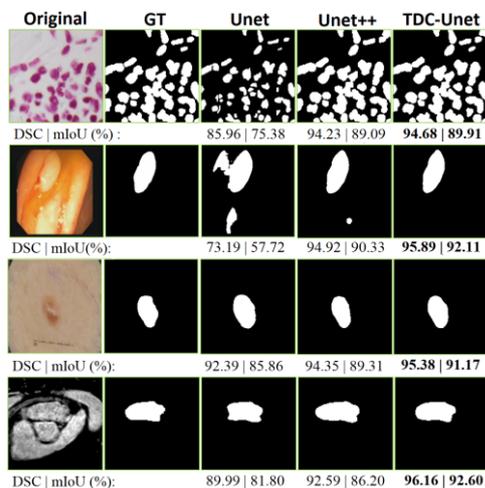


Figure 4. Some examples of the segmentation results. The first row shows the nuclei segmentation results. The second row illustrates the polyp segmentation results. The third one presents the skin lesion segmentation results. The last one shows the segmentation results of the left atrium.

TABLE IV. THE COMPARISON WITH POPULAR MODELS

	Models	Evaluation metrics (%)		
		DSC	F1-score	mIoU
Nuclei	Unet	91.87	93.65	85.00
	Unet++	92.24	93.98	85.64
	Double-Unet [10]	91.33	-	84.07
	TDC-Unet	<b>92.26</b>	<b>93.98</b>	<b>85.68</b>
Polyp	Unet	90.20	90.43	82.20
	Unet++	86.11	86.79	76.02
	GAN [9]	88.48	-	81.27
	TDC-Unet	<b>90.84</b>	<b>91.25</b>	<b>83.38</b>
Skin lesion	Unet	83.15	84.28	71.22
	Unet++	82.84	85.29	71.16
	Deeplab V3+ [19]	<b>85.90</b>	-	<b>77.40</b>
	TDC-Unet	85.77	87.32	75.16
Left atrium	Unet	90.61	90.65	83.05
	Unet++	88.96	89.73	80.32
	TDC-Unet	<b>91.67</b>	<b>92.07</b>	<b>84.71</b>

## V. CONCLUSION

In conclusion, this paper proposed the new deep learning model, TDC-Unet, for medical image segmentation. The TDC-Unet model exploited the multi-feature from three sub-Unet models to improve the segmentation performance. We also applied DC and dense structures to the nodes of the model, and the performance of the model improves with DC and dense connection. The efficiency of the model was evaluated on four datasets that cover different modalities of medical images. The TDC-Unet achieves better segmentation

results than the popular network models. This demonstrates the generality of the proposed model.

#### CONFLICT OF INTEREST

The authors declare no conflict of interest.

#### AUTHOR CONTRIBUTIONS

Song-Toan Tran, Minh-Hai Le, and Don-Gey Liu built up the models; Ching-Hwa Cheng and Thanh-Tuan Nguyen collected and pre-processed the data; Song-Toan Tran and Ching-Hwa Cheng and Minh-Hai Le analyzed the results; Song-Toan Tran, Thanh-Tuan Nguyen and Don-Gey Liu wrote the paper; all authors had approved the final version.

#### ACKNOWLEDGMENT

This work was supported in part by the Ministry of Science and Technology (MOST), Republic of China. The technical supports of the computer center of Feng Chia University on GPU resources is also acknowledged.

#### REFERENCES

- [1] E. Shelhamer, J. Long, and T. Darrell, "Fully convolutional networks for semantic segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 4, pp. 640-651, Apr. 2017.
- [2] O. Ronneberger, P. Fischer, and T. Brox. (2015). U-Net: Convolutional networks for biomedical image segmentation. [Online]. Available: [https://doi.org/10.1007/978-3-319-24574-4\\_28](https://doi.org/10.1007/978-3-319-24574-4_28)
- [3] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh, and J. Liang, "UNet++: Redesigning skip connections to exploit multiscale features in image segmentation," *IEEE Trans. Med. Imaging*, vol. 39, no. 6, pp. 1856-1867, Jun. 2020.
- [4] H. Huang, *et al.*, "UNet 3+: A full-scale connected UNet for medical image segmentation," in *Proc. 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Barcelona, Spain, 2020, pp. 1055-1059.
- [5] N. Ibtehaz and M. S. Rahman, "MultiResUNet: Rethinking the U-Net architecture for multimodal biomedical image segmentation," *Neural Netw.*, vol. 121, pp. 74-87, Jan. 2020.
- [6] L. Chen, P. Bentley, K. Mori, K. Misawa, M. Fujiwara, and D. Rueckert, "DRINet for medical image segmentation," *IEEE Trans. Med. Imaging*, vol. 37, no. 11, pp. 2453-2462, 2018.
- [7] M. Akbari, *et al.*, "Polyp segmentation in colonoscopy images using fully convolutional network," in *Proc. 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Honolulu, HI, 2018, pp. 69-72.
- [8] H. A. Qadir, Y. Shin, J. Solhusvik, J. Bergsland, L. Aabakken, and I. Balasingham, "Polyp detection and segmentation using mask R-CNN: Does a deeper feature extractor CNN always perform better?" in *Proc. 2019 13th International Symposium on Medical Information and Communication Technology*, Oslo, Norway, 2019, pp. 1-6.
- [9] J. M. Poorneshwaran, S. S. Kumar, K. Ram, J. Joseph, and M. Sivaprakasam, "Polyp Segmentation using Generative Adversarial Network," in *Proc. 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Berlin, Germany, 2019, pp. 7201-7204.
- [10] D. Jha, M. A. Riegler, D. Johansen, P. Halvorsen, and H. D. Johansen, "DoubleU-Net: A deep convolutional neural network for medical image segmentation," in *Proc. 2020 IEEE 33rd International Symposium on Computer-Based Medical Systems*, Rochester, USA, 2020, pp. 558-564.
- [11] A. S. Panayides, *et al.*, "AI in medical imaging informatics: Current challenges and future directions," *IEEE J. Biomed. Health Inform.*, vol. 24, no. 7, pp. 1837-1857, Jul. 2020.

- [12] H. Seo, C. Huang, M. Bassenne, R. Xiao, and L. Xing, "Modified U-Net (mU-Net) with incorporation of object-dependent high level features for improved liver and liver-tumor segmentation in CT images," *IEEE Trans. Med. Imaging*, vol. 39, no. 5, pp. 1316-1325, May 2020.
- [13] Y. Chen, *et al.*, "Channel-Unet: A spatial channel-wise convolutional neural network for liver and tumors segmentation," *Front. Genet.*, vol. 10, p. 1110, Nov. 2019.
- [14] 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 *Proc. 2019 IEEE International Conference on Bioinformatics and Biomedicine*, San Diego, CA, USA, Nov. 2019, pp. 1416-1423.
- [15] X. F. Xi, L. Wang, V. S. Sheng, Z. Cui, B. Fu, and F. Hu, "Cascade U-ResNets for simultaneous liver and lesion segmentation," *IEEE Access*, vol. 8, pp. 68944-68952, 2020.
- [16] F. Yu and V. Koltun. (2016). Multi-scale context aggregation by dilated convolutions. [Online]. Available: <https://arxiv.org/abs/1511.07122>
- [17] G. Huang, Z. Liu, L. V. D. Maaten, and K. Q. Weinberger. (Jan. 2018). Densely connected convolutional networks. [Online]. Available: <https://arxiv.org/abs/1608.06993v5>
- [18] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification," in *Proc. 2015 IEEE International Conference on Computer Vision*, Santiago, Chile, Dec. 2015, pp. 1026-1034.
- [19] R. Ali, R. C. Hardie, B. N. Narayanan, and S. D. Silva, "Deep learning ensemble methods for skin lesion analysis towards melanoma detection," in *Proc. 2019 IEEE National Aerospace and Electronics Conference*, Dayton, OH, USA, Jul. 2019, pp. 311-316.

Copyright © 2022 by the authors. This is an open access article distributed under the Creative Commons Attribution License ([CC BY-NC-ND 4.0](https://creativecommons.org/licenses/by-nc-nd/4.0/)), which permits use, distribution and reproduction in any medium, provided that the article is properly cited, the use is non-commercial and no modifications or adaptations are made.

**Song-Toan Tran** was born in Tra Vinh City, Vietnam, in 1984. He received the B.S. degree from Can Tho University (CTU), Can Tho City, Vietnam, in 2007, and M.S. degree from Ho Chi Minh University of Technology (HCMUT), Ho Chi Minh City, Vietnam, in 2013. From 2017 until now, he is in PhD. Program in Electrical and Communication Engineering, Feng Chia University (FCU), Taichung, Taiwan. His research interested include medical image processing, deep learning, computer vision, and virtual reality-augmented reality and applications.

**Thanh-Tuan Nguyen** was born in Kien Giang Province, Vietnam, in 1980. He received the B.S. degree from Can Tho University (CTU), Can Tho City, Vietnam, in 2003, and M.S. degree from Ho Chi Minh University of Technology (HCMUT), Ho Chi Minh City, Vietnam, in 2014. From 2018 until now, he is in PhD. Program in Electrical and Communication Engineering, Feng Chia University (FCU), Taichung, Taiwan. His research interested include measuring technique, LiDAR design and application, image processing and machine learning.

**Minh-Hai Le** was born in Tra Vinh City, Vietnam, in 1994. He received the B.S. degree from Ho Chi Minh University of Science (HCMUS), Ho Chi Minh City, Vietnam, in 2016, and M.S. degree from Feng Chia University (FCU), Taichung, Taiwan, in 2020. Now, he is in PhD. Program in Electrical and Communication Engineering, Feng Chia University (FCU), Taichung, Taiwan. His research interested include image processing, deep learning, self-driving car, and IoT applications.

**Ching-Hwa Cheng** is currently a full professor of electronic engineering at Feng-Chia University. His research interests include endoscope system design, image process and low-power VLSI Design/EDA/testing related issues, such image guiding surgery and low-power chip designed by multi-voltage. His studies have included theoretical analysis model, design flow development, cell library generation, physical chip implementation and system validation. His work differs from that of other researchers in this field in that most of his research designs are silicon proven by a practical system.

**Don-Gey Liu** was born in Tainan, Taiwan, in 1963. He received the B.S. degree from Department of Electrical Engineering, National Taiwan University (NTU), in 1986; the M.S. degree from Institute of Electrical Engineering, National Tsing Hua University (NTHU), in 1988; and the Ph. D. degree from Institute of Electronics, National Chiao Tung University (NCTU), in 1992, respectively. Thereafter, he served the Chinese Air Forces as Lieutenant from 1992 to 1994. He was an officer in charge of maintaining wireless equipment. In 1994, he joined the Electronics Research and Service Organization, Industrial Technology Research Institute (ERSO/ITRI) in developing flat-panel displays by Field-Emission Devices (FEDs). Then he went to Department of Electronic Engineering at Feng Chia University (FCU), Taichung, Taiwan, as Associate Professor. He has been a professor since 2001 and head of the department from 2004 to 2008. From 2014, he serves as the director of the Ph. D. Program of Electrical and Communications Engineering. And from 2017, he is also the director of the Master Program of Biomedical Informatics and Biomedical Engineering. His current research interests are focused on high speed integrated circuit design for RF and analog applications.

Prof. Liu has filed 6 patents, and published over 40 journal papers, and 70 conference papers. He has been awarded more than 30 research

contracts and grants from government agencies such as National Science Council (NSC) and Ministry of Education, (MOE), and industries in Taiwan. In 1998, he was invited as Consultant for process of semiconductor wastes in research laboratories and hi-tech companies in Taiwan. From 2007 to 2010, he served as a member of Technical Program Committee for the International Meeting on Information Display (IMID) held in Daegu, South Korea. From 2010, he was invited as a keynote speaker for several international conferences held in Mainland China. In addition, Dr. Liu serves as a technical reviewer for several papers in Institute of Electrical and Electronic Engineering (IEEE) and Chinese Institute of Electrical Engineering (CIEE) journals. He also was a member of the Board of Reviewers for proposals in NSC from 2004 to 2007.

Prof. Liu received several awards in teaching and research from Feng Chia University and MOE. From 2004 to 2006, he was Co-chairperson in holding National Competition of Communication Researches. In his term of chairperson, he conducted the department to obtain the accreditation of Institute of Engineering Education, Taiwan (IEET) which has joined in Washington Covenants. In 2017, he also helped the advisory review of the IEET accreditation for Zhongshan Polytechnic University, Guan Dong, China.