LOCAL IMAGE FITTING BASED ACTIVE CONTOUR LOSS WITH DEEP LEARNING FOR NUCLEI SEGMENTATION

Thi-Thao Tran*, Van-Truong Pham

School of Electrical and Electronic Engineering, Hanoi University of Science and Technology

*Corresponding author: thao.tranthi@hust.edu.vn

(Received: February 25, 2022; Accepted: April 01, 2022)

Abstract - This paper proposes an approach for segmentation of nuclei images based on deep learning. In particular, the recent TransUnet inspired from transformers' strong ability in modeling long-range context, is employed and adapted for the nuclei segmentation. For training the neural network, we propose a new loss inspired from active contour models with the guide of local image fitting. The loss when applied for the TransUnet has shown promising results over common Dice and Binary Cross Entropy loss functions. Our approach has been validated on the Data Science Bowl 2018 dataset, which includes 670 data folders for training model and 65 data folders for testing. State of the art models, such as FCN, SegNet, Unet, and DoubleU-Net are also conducted and evaluated. Quantitative assessments with high Dice similarity coefficient and Intersection over Union metrics demonstrate the performances of the proposed approach for nuclei segmentation.

Key words - Deep Learning; Nuclei Segmentation; Image Segmentation; The Fully Convolutional Neural Network (FCN); UNet; TransUnet.

1. Introduction

Segmentation of nuclei is an important step in biomedical microscopy image analysis [1]. From nuclei segmentation results, several medical analyses such as cell counting [2], cell type classification [3], could be derived. However, segmentation of nuclei is a challenging task because of the staining differences, tissue type, and different visual characteristics of cell type convey nuclei [1]. Manual segmentation of nuclei is tedious and might be unfeasible in the case of large data sets. In addition, the accurate of segmentation results depend on experience of the experts and lack of reproducibility. Therefore, automatic nuclei instance segmentation from microscopy images is high demand.

There are several methods to be developed to automatically segment nuclei in literature ranging from simple background subtraction to more sophisticated approaches, i.e., Otsu-based method [4], the watershed method [5], Grab Cut [6], active contour [7]. However, the traditional methods surfer from drawbacks such as sensitive to parameters and only effectiveness for one or a few precise categories of structural nuclei structures.

The advances of deep learning (DL) models have shown excellent performance for many aspects of medical image analysis such as super-resolution [8], classification [9], and segmentation [10, 11]. For cell and nuclei segmentation, DL-based approaches have been applied such as the works of Sommer et al. in [12], Pan et al. in [13], Zhou et al. [14]. Vuola et al. [15] employed Mask-RCNN for nuclei segmentation. Van et al. presented a method namely DeepCell to analysis cells in live-cell imaging [16]. Ronneberger et al. [17] presented the U-Net architecture proposed to segment the nuclei segmentation. Based on U-Net, Zeng et al. [18] presented a modified version developed for nuclei segmentation, in which residual blocks and channel attention mechanism are adapted in original U-Net model.

In the current study, inspired from transformers' strong ability in modeling long-range context, we propose to apply the TransUnet that incorporate the Unet and Transformers, for segmentation of nuclei images. In addition, for network training, we introduce a new loss based on both binary maps and image local fitting. Additionally, we conduct a feasibility study on DL-based technique for nuclei images convolution Network (FCN) [19], SegNet [20], UNet [17], and DoubleUNet [21].

2. Related Works

2.1. Fully Convolutional Neural Network

The Fully Convolutional Neural Network (FCN) by Long et al. [22] was one of the first convolutional neural networks proposed for image segmentation problems. The architecture has been extent to other segmentation tasks, such as ventricle segmentation [19]. The FCN network contains two main parts, encoder (contracting path), and decoder (expanding path). The encoder includes convolutional layers and max pooling layers that respectively account for feature extraction and downsampling. The decoder contains upsampling layers and a classifier layer accounts for recovering the input image resolution as well as outputting the segmentation map.

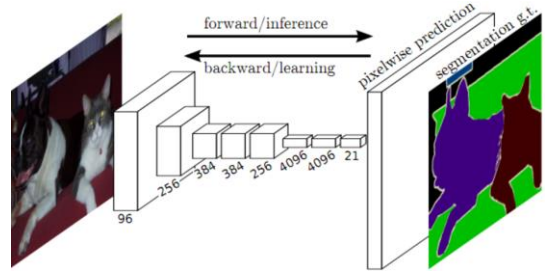


Figure 1. *Basic structure of the FCN*

2.2. U-Net

The U-Net network architecture [17] has been considered as the standard architecture for medical image segmentation. Inspired from FCN [19], the U-Net, as shown in Figure 2, also consists two symmetric paths, the encoder and decoder. The convolutional (conv) and max pooling (max pool) layers in the encoder account for

reducing the feature resolution. The decoder contains upsampling (up-conv) and convolutional layers. To maintain spatial information that may be lost during downsampling, skip connections (copy and crop paths in Figure 2) are used to concatenate the encoder feature maps to the equivalent layers of the decoder. Inspired from U-Net architecture, more and more network architecture has been proposed.

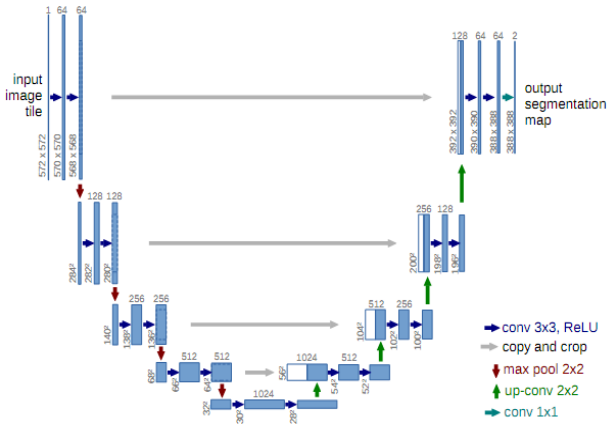


Figure 2. *Basic structure of the U-Net*

2.3. Vision transformer

Transformers, presented by Vaswani et al. [23] first for the natural language processing (NLP) tasks. Since their inception, the transformer-based models have demonstrated excellent performance on NLP problems and become state-of-the-art methods in some computer vision tasks [23-25]. Built on the transformer model, the Vision Transformers (ViTs) by Dosovitskiy et al. [24] showed favorable performance in image classification tasks by cascading multiple transformer layers in order to capture input image global contexts. The ViT is one of the first transformer-based methods to surpass the conventional CNN-based works. Inspired by ViTs, a number of modern methods have been developed not only for image classification but also image segmentation tasks [25].

2.4. Active Contour Loss

Active contour (AC) loss is built based on the active contour models [26] which have been extensively used for energy based image segmentation for the last two decades. brief introduction on the active contour loss is given as follows.

Denote $u \in [0,1]$ be the prediction- the output map of the neural network, and $v \in [0,1]$ be the ground truth- the map manually labelled by the experts. The active contour loss adapted from the work in [27] as

$$L_{\text{AC}} = \left(\mu Length + \lambda Region\right) \tag{1}$$

where μ and λ are hyper parameters. Length is associated the boundary length of the prediction contour, defined as follows.

$$Length = \sum_{i=1}^{M}\sum_{j=1}^{N}\sqrt{\left|(\nabla u_{x_{i,j}})^2 + (\nabla u_{y_{i,j}})^2\right|} \tag{2}$$

with M is the image width and N is the image height; x and y are respectively horizontal and vertical directions of the image. $u_{x_{i,j}}$ and $u_{y_{i,j}}$ are pixels with respect to those directions of the prediction mask.

The Region term in Eq. 1 is computed as the following equation:

$$Region = \sum_{i=1}^{M}\sum_{j=1}^{N}\left(u_{i,j}(v_{i,j}-c_1)^2\right) + \left((1-u_{i,j})(v_{i,j}-c_2)^2\right) \tag{3}$$

where c_1, and c_2 are defined as:

$$c_1 = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N} v_{i,j}u_{i,j}}{\sum_{i=1}^{M}\sum_{j=1}^{N} u_{i,j}}, c_2 = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N} v_{i,j}(1-u_{i,j})}{\sum_{i=1}^{M}\sum_{j=1}^{N}(1-u_{i,j})} \tag{4}$$

3. Method

3.1. TransUnet

Inspired by ViTs, the TransUNet by Chen et al. [25] is one of the first transformer-based models proposed for image segmentation. The TransUNet adopts the transformers as the encoder and incorporated with U-Net for the image segmentation task. Similar to other U-Net variants, the TransUnet includes encoder, decoder, and the encoder-decoder skip connections paths, as shown in Figure 3. In the TransUnet, the encoder is a hybrid CNN-Transformer architecture where CNN is an account for feature extractionr to create feature maps.

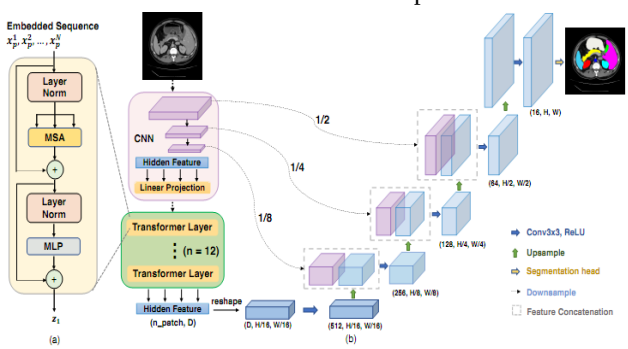


Figure 3. *Basic structure of the TransUNet. (a) presents the Transformer layer; and (b) is the architecture of the TransUnet*

The decoder includes upsampling layers, call Cascaded Upsampler (CUP) with multiple upsampling steps to decode hidden features to output final segmentation mask. By using the CUP with the hybrid CNN-transformer encoder, the model allows feature aggregation at various resolutions levels via skip connections.

3.2. The proposed loss

Inspired from the active contour models and advances of training losses for neural networks, we proposed, in this study, a new active contour loss function for nuclei image segmentation as:

$$L_{FIT\_ac} = \alpha L_{\text{Region}} + \beta L_{\text{FIT}} + \mu L_{\text{Rec}} \tag{5}$$

where α, β, and μ are hyper-parameters; $L_{\text{Region}}, L_{\text{FIT}}, L_{\text{Rec}}$ are respectively the region active contour loss, the Local Image Fitting (LIF), and the Regularization loss terms, defined as the following.

The region based active contour loss, adapted from the works in [28] and [29]. Let **Ω** be the spatial domain of the input image *I*. Denote v_x as the pixel value located at **x**th pixel of the ground truth *v*, and u_x as the pixel value located at **x**th pixel of the prediction *u*. Let M be the image width and *N* be the image height; the region loss is defined as:

$$L_{\text{Region}} = \frac{-1}{M \times N} \sum_{\mathbf{x} \in \Omega} \left(\log\left(u_{\mathbf{x}}\right)\left(v_{\mathbf{x}} - d_1\right)^2 + \log\left(1 - u_{\mathbf{x}}\right)\left(v_{\mathbf{x}} - d_2\right)^2 \right) \tag{6}$$

where

$$d_1 = \frac{\sum_{\mathbf{x} \in \Omega} u_{\mathbf{x}}\left(1 - v_{\mathbf{x}}\right)}{\sum_{\mathbf{x} \in \Omega} u_{\mathbf{x}}}; d_2 = \frac{\sum_{\mathbf{x} \in \Omega} \left(1 - u_{\mathbf{x}}\right) v_{\mathbf{x}}}{\sum_{\mathbf{x} \in \Omega} \left(1 - u_{\mathbf{x}}\right)} \tag{7}$$

For the Local Image Fitting loss, we exploit the energy function in [30] that originally developed for classical level set based active contour models. First, denote $I_{\mathbf{x}}$ the pixel value located at ***x***th of image *I*, and consider a circular neighborhood with a radius of σ centered at each pixel $\mathbf{y} \in \Omega$, defined by $\Omega_{\mathbf{y}} \triangleq \left\{\mathbf{x} : |\mathbf{x} - \mathbf{y}| \le \sigma\right\}$. Furthermore, we also introduce a nonnegative window function $\mathrm{K}\left(\mathbf{y} - \mathbf{x}\right)$, with $\mathrm{K}\left(\mathbf{y} - \mathbf{x}\right) = 0$ for $\mathbf{x} \notin \Omega_{\mathbf{y}}$. Following [31], the kernel function has the following properties:

i) $K\left(-\mathbf{z}\right) = K\left(\mathbf{z}\right)$;

ii) $K\left(\mathbf{z}_1\right) \ge K\left(\mathbf{z}_2\right)$ if $|\mathbf{z}_1| < |\mathbf{z}_2|$, and $\lim_{|\mathbf{z}| \to \infty} K\left(\mathbf{z}\right) = 0$;

iii) $\int K\left(\mathbf{x}\right) d\mathbf{x} = 1$.

The Local Image Fitting (LIF) loss is expressed as:

$$L_{\text{LIF}} = \frac{1}{M \times N} \sum_{\mathbf{x} \in \Omega} \left(u_{\mathbf{x}}\left(I_{\mathbf{x}} - f_1\right)^2 + \left(1 - u_{\mathbf{x}}\right)\left(I_{\mathbf{x}} - f_2\right)^2 \right) \tag{8}$$

where

$$f_1 = \frac{\sum_{\mathbf{y} \in \Omega_{\mathbf{x}}} K\left(\mathbf{x} - \mathbf{y}\right) I_y u_y}{\sum_{\mathbf{y} \in \Omega_{\mathbf{x}}} K\left(\mathbf{x} - \mathbf{y}\right) u_y}; f_2 = \frac{\sum_{\mathbf{y} \in \Omega_{\mathbf{x}}} K\left(\mathbf{x} - \mathbf{y}\right) I_y \left(1 - u_y\right)}{\sum_{\mathbf{y} \in \Omega_{\mathbf{x}}} K\left(\mathbf{x} - \mathbf{y}\right)\left(1 - u_y\right)} \tag{9}$$

The regularization term, also called the length term, is defined as

$$L_{\text{Rec}} = \frac{1}{M \times N} \sum_{\mathbf{x} \in \Omega} |\nabla u_{\mathbf{x}}| \approx \frac{1}{M \times N} \sum_{\mathbf{x} \in \Omega} |u_{\mathbf{x}+1} - u_{\mathbf{x}}| \tag{10}$$

It is noted that in this study, the proposed loss includes two main terms, the Region and the Local Image Fitting (LIF) terms. The Region one is a supervised term that measures the dissimilarity between the binary masks of ground truth (*v*) and prediction (*u*) of the image *I*. Whereas the LIF is an unsupervised term that measures the dissimilarity between the image intensity values inside and outside the prediction *u* of the image *I*. The difference between the two terms is that in the LIF term, the ground truth *v* is not taken into consideration. Besides, the Regularization term is just used to control the boundary smoothness of the prediction mask.

3.3. Evaluation Metrics

In image segmentation, the Dice Similarity Coefficient (DSC) is commonly used to evaluate the segmentation performance by the neural network. The DSC metric measures the overlap between *A* and *B*, described as

$$DSC\left(A, B\right) = \frac{2 \times |A \cap B|}{|A| + |B|} \tag{11}$$

Besides the Dice Similarity Coefficient, the Intersection over Union (IoU) index is also often used. The IoU is defined as

$$IoU\left(A, B\right) = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B| - |A \cap B|} \tag{12}$$

4. Experiments

4.1. Dataset and Training

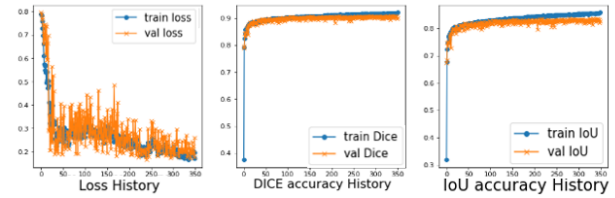
The Data Science Bowl 2018 (DSB2018) dataset [32] presented a challenge to sci-entists globally to automatically identify cells in a series of micro images. The main goal is to find image segmentation strategies that can be applied to multiple experiments automatically without additional human intervention. This approach could reduce image quantification time, allow future scientists to adopt and test more methods and applications. The dataset includes 670 image pairs with each pair contains image and the corresponding nuclei mask. The dataset is split in the ratio of 80:20 for respectively training and testing sets. In the training phase, 10 percent of training set is used for validation. The training data is augmented with some operations such as random rotations, shift, scaling. Hypermeters are set as α = 0.5, $\beta = 10^{-2}$, and $\lambda = 10^{-5}$ for all experiments. All network architectures are trained using PyTorch framework on a NVIDIA RTX 2080 GPU.

4.2. Segmentation Results

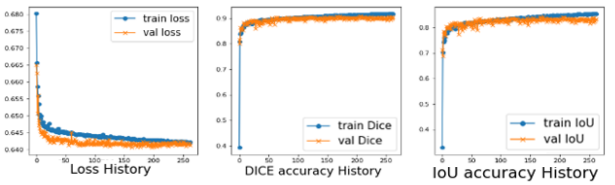
To assess the performance of the proposed loss function, we trained the TransUNet and compare with results on the Data Science Bowl 2018 by training other common losses including the binary cross entropy (BCE), and the Dice loss. The learning curves including the loss values, the DSC, and IoU scores versus epochs on the training and validation sets are given in Figure 4. As can be shown in Figure 4c, compared to the results by BCE and Dice loss functions, the proposed loss function converges faster with highest evaluation scores.

To further demonstrate advantages of the proposed approach, we show some representative segmentation samples from the testing data in Figure 5. Representative results when trained by BCE loss, Dice loss and the proposed loss are shown in consecutive columns of Figure 5. The last column of this Figureure shows the correspond ground truths. By visualization, we can observe that the results by the proposed loss are close to the ground truths than those by other losses, especially in the regions inside the boxes.

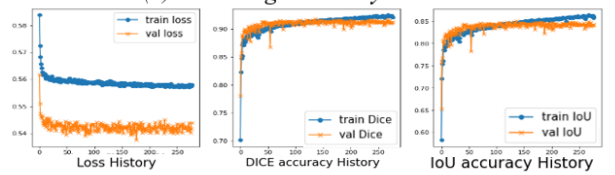
For further evaluation, the metrics including the DSC and IoU scores by the proposed and other losses are reported in Table 1. As one can see from this table, the results obtained when trained by the proposed loss are with highest scores for both DSC and IoU metrics.



(a). Learning curves by Dice loss



(b). Learning curves by BCE loss



(c). Learning curves by proposed loss

Figure 4. *Learning curves by the TransUnet when trained by (a) Dice loss, (b) BCE loss, and (c) proposed loss functions. Left: Loss values; Middle: Dice similarity score (DSC); Right: IoU.*

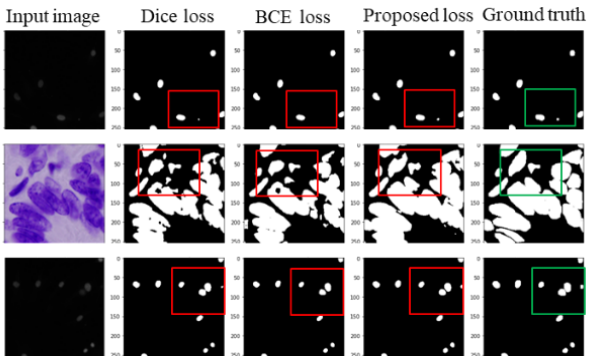


Figure 5. *Results by the TransUnet when trained by different losses. The boxes show the comparative agreements between predictions and ground truths*

Table 1. *Evaluation metrics on the Data Science Bowl 2018 by TransUnet obtained by using the Dice loss, Binary cross entropy (BCE) and the proposed loss*

Loss	DSC	IoU
Dice loss	0.909	0.839
BCE loss	0.902	0.825
Proposed	0.921	0.856

For comparison, we also reimplemented other networks including FCN, SegNet, U-Net, and Double-UNet, and evaluated on the Data Science Bowl 2018 with the same training protocol as TransUnet. The evaluation scores including the DSC and IoU by the TransUnet using the proposed loss together with those be comparative models are given in Table 2. As shown in Table 2, the TransUNet with proposed loss achieves the best DSC and IoU scores for both cases, using the LIF and without using the LIF term in the proposed loss. It is also noted that in our loss, the LIF is used as an augmented term that considers the image intensity information and does not use the ground truth. The use of LIF helps an increase of DSC from 0.918 to 0.821, and an increase of IoU from 0.852 to 0.856 as in the last two rows of Table 2.

Table 2*. Evaluation metrics on the Data Science Bowl 2018 with some state of the arts DL- based approaches, compared with those by TransUnet trained by the proposed loss without using the LIF loss term, and when using the LIF term*

Methods	DSC	IoU
FCN	0.901	0.822
SegNet	0.899	0.819
U-Net	0.895	0.815
Double-UNet	0.913	0.840
TransUnet (w/o LIF)	0.918	0.852
TransUnet (with LIF)	0.921	0.856

5. Conclusions

This study has proposed a novel loss function and also presented a feasibility on the deep learning-based approaches for nuclei image segmentation. Several state-of- the- art models have been studied and conducted on the Data Science Bowl 2018 database. From the experiments and evaluation, it showed that the TransUnet when trained by the proposed loss obtains high segmentation performance as compared to state-of-the-art alternatives. In the future, we intend to focus on how to decrease the training time or increase the accuracy of the TransUnet for nuclei image segmentation as well as other segmentation tasks. Other approaches to improve the performance of the approach such as changing the pre-trained encoder, using more modern engine or find a better loss function can be taken into consideration.

Acknowledgement: This research is funded by the Hanoi University of Science and Technology (HUST) under project number T2021-PC-005.

REFERENCES

[1] J. C. Caicedo, J. Roth, A. Goodman, T. Becker, K. W. Karhohs, C. McQuin, *et al.*, "Evaluation of deep learning strategies for nucleus segmentation in fluorescence images", *Cytometry A*, vol. 95, pp. 952-965, 2019.

[2] T. Falk, D. Mai, R. Bensch, Ö. Çiçek, A. Abdulkadir, Y. Marrakchi, *et al.*, "U-Net: Deep learning for cell counting detection and morphometry", *Nature Methods*, vol. 16, pp. 67-70, 2019.

[3] Y. Liu and F. Long, "Acute lymphoblastic leukemia cells image analysis with deep bagging ensemble learning", in *in CNMC Challenge: Classification in Cancer Cell Imaging*, 2019, pp. 113-121.

[4] N. Otsu, "A threshold selection method from gray-level histograms", *IEEE Trans Syst Man Cybern*, vol. 9, pp. 62-66, 1979.

[5] C. Wählby, I.-M. Sintorn, F. Erlandsson, G. Borgefors, and E. Bengtsson, "Combining intensity, edge and shape information for 2D and 3D segmentation of cell nuclei in tissue sections", *Journal of Microscopy*, vol. 215, pp. 67-76, 2004.

[6] C. Rother, V. Kolmogorov, and A. Blake, "Grabcut: Interactive foreground extraction using iterated graph cuts", *ACM Transactions on Graphics (TOG)*, vol. 23, pp. 309-314, 2004.

[7] T. Hayakawa, V. B. Surya Prasath, H. Kawanaka, B. J. Aronow, and S. Tsuruoka, "Computational Nuclei Segmentation Methods in Digital Pathology: A Survey", *Archives of Computational Methods in Engineering*, vol. 28, pp. 1-13, 2021.

[8] Jebaseeli. T.J., C. A. D. Durai, and J. D. Peter, "Retinal blood vessel segmentation from diabetic retinopathy images using tandem PCNN model and deep learning based SVM", *Optik- International Journal for Light and Electron Optics*, vol. 199, p. 163328, 2019.

[9] A. Esteva, B. Kuprel, R. A. Novoato, J. Ko, S. M. Swetter, H. M. Blau, *et al.*, "Dermatologist-level classification of skin cancer with deep neural networks", *Nature*, vol. 542, pp. 115–118, 2017.

[10] V.-T. Pham, T.-T. Tran, P.-C. Wang, and M.-T. Lo, "Tympanic membrane segmentation in otoscopic images based on fully convolutional network with active contour loss", *Signal, Image and Video Processing*, vol. 15, pp. 519–527, 2021.

[11] Van-Truong Pham, Thi-Thao Tran, Pa-Chun Wang, Po-Yu Chen, and Men-Tzung Lo, "EAR-UNet: A deep learning-based approach for segmentation of tympanic membranes from otoscopic images", *Artificial Intelligence In Medicine*, vol. 115, pp. 1-12, 2021.

[12] C. Sommer, C. Straehle, U. Köthe, and F. A. Hamprecht, "Ilastik: Interactive learning and segmentation toolkit", *In: IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, 2011, pp. 230-233.

[13] X. Pan, L. Li, D. Yang, Y. He, Z. Liu, and H. Yang, "An Accurate Nuclei Segmentation Algorithm in Pathological Image Based on Deep Semantic Network", *IEEE Access*, vol. 7, pp. 110674 - 110686, 2019.

[14] Y. Zhou, O. F. Onder, Q. Dou, E. Tsougenis, H. Chen, and P. A. Heng, "CIA-net: Robust nuclei instance segmentation with contour-aware information aggregation", in *In: International Conference on Information Processing in Medical Imaging*, 2019, pp. 682-693.

[15] A. O. Vuola, S. U. Akram, and J. Kannala, in *IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, 2019, pp. 208-212.

[16] D. A. Van Valen, T. Kudo, K. M. Lane, D. N. Macklin, N. T. Quach, M. M. DeFelice, *et al.*, "Deep Learning Automates the Quantitative Analysis of Individual Cells in Live-Cell Imaging Experiments", *PLoS Comput. Biol.*, vol. 12, p. e1005177, 2016.

[17] 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*, 2015, pp. 234-241.

[18] Z. Zeng, W. Xie, Y. Zhang, and Y. Lu, "RIC-Unet: An improved neural network based on Unet for nuclei segmentation in histology images", *IEEE Access*, vol. 7, pp. 21420-21428, 2019.

[19] P. V. Tran, "A fully convolutional neural network for cardiac segmentation in short-axis MRI", *arXiv preprint arXiv:1604.00494*, 2016.

[20] 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, pp. 2481-2495, 2017.

[21] D. Jha, M. Riegler, D. Johansen, P. Halvorsen, and H. Johansen, "Doubleu-net: A deep convolutional neural network for medical image segmentation", in *IEEE 33rd International Symposium on Computer-Based Medical Systems (CBMS)*, 2020, pp. 558-564

[22] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation", *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 3431–3440, 2015.

[23] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, *et al.*, "Attention is all you need", in *Advances in neural information processing systems*, 2017, pp. 5998–6008.

[24] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, *et al.*, "An image is worth 16x16 words: Transformers for image recognition at scale", *arXiv:2010.11929*, 2020.

[25] J. Chen, Y. Lu, Q. Yu, X. Luo, E. Adeli, Y. Wang, *et al.*, "TransUNet: Transformers Make Strong Encoders for Medical Image Segmentation", *arXiv:2102.04306*, 2021.

[26] T. Chan and L. Vese, "Active contours without edges", *IEEE Trans. Image Process.*, vol. 10 (2), pp. 266-277, 2001.

[27] X. Chen, B. M. Williams, S. R. Vallabhaneni, G. Czanner, R. Williams, and Y. Zheng, "Learning Active Contour Models for Medical Image Segmentation", *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 11623-11640, 2019.

[28] X. Chen, B. M. Williams, S. R. Vallabhaneni, G. Czanner, R. Williams, and Y. Zheng, "Learning active contour models for medical image segmentation", in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2019, pp. 11632-11640.

[29] M.-N. Trinh, N.-T. Nguyen, T.-T. Tran, and V.-T. Pham, "A Semi-supervised Deep Learning-Based Approach with Multiphase Active Contour Loss for Left Ventricle Segmentation from CMR Images", in *The third International Conference on Sustainable Computing*, 2021, pp. 13-23.

[30] K. Zhang, H. Song, and L. Zhang, "Active contours driven by local image fitting energy " *Pattern Recognition*, vol. 43, pp. 1199-1206, 2010.

[31] C. Li, C. Kao, J. Gore, and Z. Ding, "Minimization of region-scalable fitting energy for image segmentation", *IEEE Trans. Image Process.*, vol. 17, pp. 1940-1949, 2008.

[32] J. C. Caicedo, A. Goodman, K. W. Karhohs, B. A. Cimini, J. Ackerman, M. Haghighi, *et al.*, "Nucleus segmentation across imaging experiments: the 2018 data science bowl", *Nature Methods*, vol. 16, pp. 1247–125, 2019.