IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received May 2, 2019, accepted June 14, 2019, date of publication June 21, 2019, date of current version July 11, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2924255

Deep-Learning-Based Label-Free Segmentation of Cell Nuclei in Time-Lapse Refractive Index Tomograms

JIMIN LEE[1], HYEJIN KIM[1], HYUNGJOO CHO[1], YOUNGJU JO[2], YUJIN SONG[3], DAEWOONG AHN[4], KANGWON LEE[3], YONGKEUN PARK[4,5,6], AND SUNG-JOON YE[1,7,8]

[1]Program in Biomedical Radiation Sciences, Department of Transdisciplinary Studies, Graduate School of Convergence Science and Technology, Seoul National University, Seoul 08826, South Korea
[2]Department of Applied Physics, Stanford University, Stanford, CA 94305, USA
[3]Program in Nano Science and Technology, Department of Transdisciplinary Studies, Graduate School of Convergence Science and Technology, Seoul National University, Seoul 08826, South Korea
[4]Tomocube Inc., Daejeon 34051, South Korea
[5]Department of Physics, Korea Advanced Institute of Science and Technology (KAIST), Daejeon 34141, South Korea
[6]KAIST Institute for Health Science and Technology, Korea Advanced Institute of Science and Technology (KAIST), Daejeon 34141, South Korea
[7]Department of Radiation Oncology, Seoul National University College of Medicine, Seoul 03080, South Korea
[8]Robotics Research Laboratory for Extreme Environments, Advanced Institutes of Convergence Technology, Suwon 16229, South Korea

Corresponding author: Sung-Joon Ye (sye@snu.ac.kr)

This work was supported in part by the National Research Foundation of Korea (NRF) Grant funded by the Korean Government (Ministry of Science and ICT) under Grant NRF-2017M2A2A6A01071214. The work of Y. Jo was supported in part by the KAIST Presidential Fellowship and in part by the Asan Foundation Biomedical Science Scholarship.

ABSTRACT We proposed a method of label-free segmentation of cell nuclei by exploiting a deep learning (DL) framework. Over the years, fluorescent proteins and staining agents have been widely used to identify cell nuclei. However, the use of exogenous agents inevitably prevents from long-term imaging of live cells and rapid analysis and even interferes with intrinsic physiological conditions. Without any agents, the proposed method was applied to label-free optical diffraction tomography (ODT) of human breast cancer cells. A novel architecture with optimized training strategies was validated through cross-modality and cross-laboratory experiments. The nucleus volumes from the DL-based label-free ODT segmentation accurately agreed with those from fluorescent-based. Furthermore, the 4D cell nucleus segmentation was successfully performed for the time-lapse ODT images. The proposed method would bring out broad and immediate biomedical applications with our framework publicly available.

INDEX TERMS Cell nucleus segmentation, deep learning, label-free segmentation, optical diffraction tomography, refractive index tomogram.

I. INTRODUCTION

The precise localization and segmentation of cell nucleus are crucial to understand the cell physiology in cell biology and to diagnose a malignant tumor in histopathology. In addition to its primary biological function as the carrier of genetic information, the characteristics of cell nuclei play a variety of roles in medicine. For instance, the volume ratio of the nucleus to the cytoplasm is a well-established indicator of cell malignancy [1]. Light scattering spectroscopy techniques for non-invasive cancer diagnosis are known to be closely related to this nucleus-based diagnostic marker [2], [3]. Furthermore, targeted dose enhancement of cell nuclei by gold nanoparticles has been shown to improve the therapeutic efficiency in radiotherapy of tumors [4]. However, despite these far-reaching implications, nucleus segmentation of live unlabeled cells has not been adequately addressed yet. Conventional approaches for cell identification and segmentation utilized exogenous agents such as fluorescence proteins or dyes to specifically label nucleus structures. However, these methods inevitably prevent long-term live cell imaging or rapid analysis.

Recently, various quantitative phase imaging (QPI) techniques have been developed and utilized for label-free imaging of live cells [5]. Optical diffraction tomography (ODT) is one of the 3D QPI techniques, which reconstructs the 3D

The associate editor coordinating the review of this manuscript and approving it for publication was Andrei Muller.

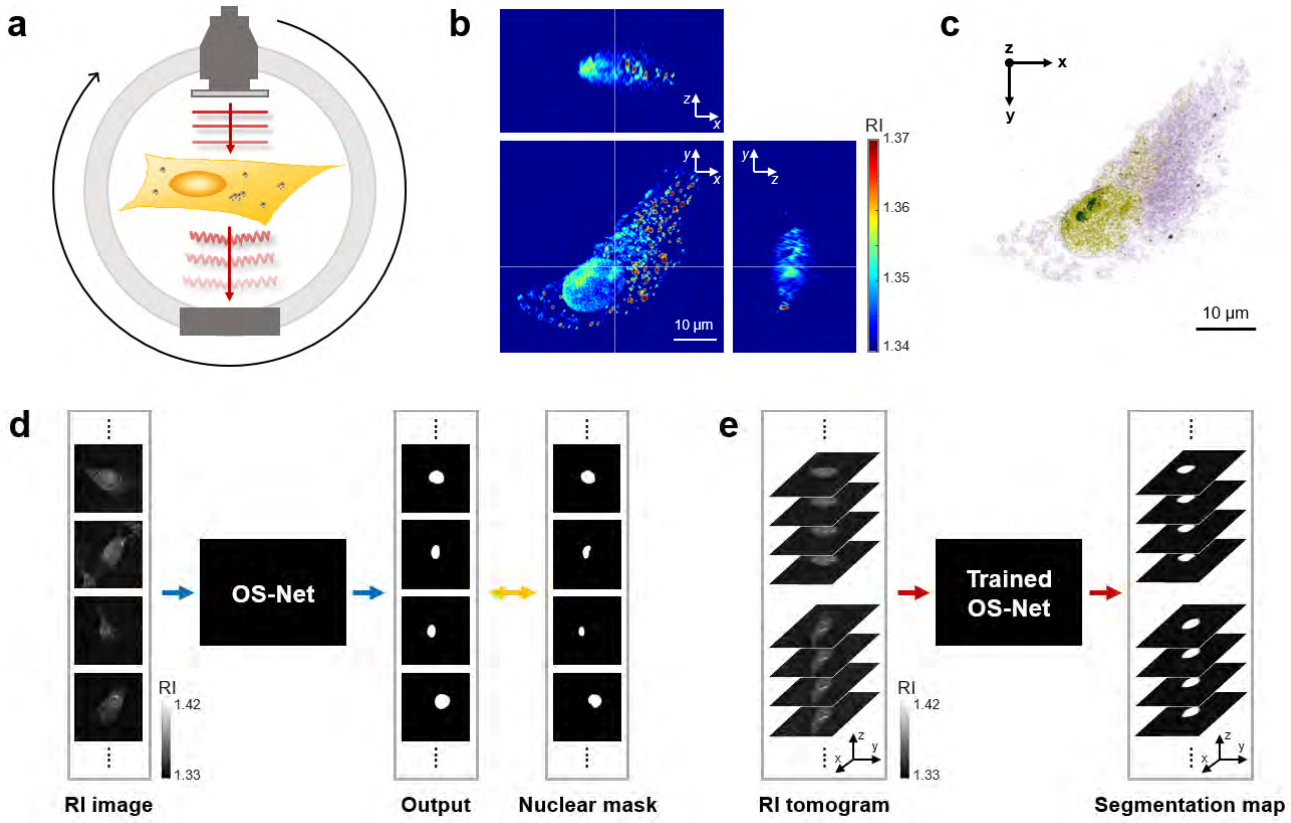

FIGURE 1. The overall scheme of OS-Net. (a) Schematic diagram representing the principle of 3D ODT. (b) 2D sections of a measured 3D tomogram. (c) 3D rendering of the tomogram. (d) Training of OS-Net using expert-annotated 2D RI images. (e) 3D cell nucleus segmentation using the trained OSNet.

refractive index (RI) distribution of a sample from multiple 2D holographic images measured at various illumination angles [6]–[8]. Due to its label-free and quantitative imaging capability, ODT has been utilized in various topics of studies including microalgae [9], hematology [10], infectious diseases [11], and yeast study [12]. In addition, RI is an intrinsic property of materials governing light-matter interaction (i.e., scattering potential) and its reconstructed tomogram provides abundant morphological information about cells [13]. However, the determination of the boundary of certain subcellular organelles (especially nucleus) based on RI is often an ill-posed inverse problem.

Why is it difficult to segment the cell nucleus in ODT? As shown in Fig. 1, the nuclei of eukaryotic cells in the RI tomograms measured using ODT can be readily recognized by trained biologists, at least in two dimensions (2D). However, automating this process toward 3D (three-dimensional) high-throughput or time-lapse 3D (i.e., 4D) is not straightforward and generally challenging due to the following reasons: (i) There are significant cell type dependence, cell cycle dependence, and even cell-to-cell variations in the nuclear RI threshold; (ii) overlapped or similar RI ranges of various intracellular structures further complicate the problem [13]. Previously proposed algorithms for ODT-based segmentation of the nucleus, or subcellular organelles typically have been the case-by-case design of image processing steps, such as thresholding, filtering, and various transforms [14]. This *rule-based* approaches are laborious and require significant domain knowledge and assumptions. In short, this task is easy-to-human but difficult-to-machine, and thus, would benefit from *learning-based* approaches instead of explicit design [15].

Here, we proposed a deep learning framework for label-free segmentation of cell nuclei in ODT. While determining a region of certain organelles (here, nucleus) or their chemical identity using RI in a *pointwise* manner is challenging, we hypothesized that certain patterns in the *spatial distribution* of RI might facilitate the chemical identification [15]–[21]. We implemented this strategy through end-to-end training of convolutional neural networks (CNN) that detect local and global spatial correlations. We performed extensively comparative experiments exploring a variety of network architectures and training strategies in terms of various evaluation metrics. Then we rigorously tested the trained networks via cross-modality and cross-laboratory validation. To our knowledge, the present work, named OS-Net (ODT-based Segmentation Network), was the first deep learning approach to biomedical applications of ODT.

The rest part of the paper is organized as follows. Section II provides the details of the overall scheme of the OS-NET, data

preparation process, elaborate description of the OS-NET architecture, the network training strategy and the evaluation of the trained network. Section III discusses the experimental results and Section IV gives conclusion and the contribution of this work.

II. MATERIALS AND METHODS

A. THE OVERALL SCHEME OF OS-NET

Fig. 1 illustrates the overall scheme of the proposed deep-learning-based label-free cell nucleus segmentation in ODT. ODT provides 3D RI tomograms of eukaryotic cells, in which the nucleus can be visualized (Figs. 1a-c). In order to emulate trained biologists who can readily recognize nuclear regions, first, we built an expert-annotated training dataset with the x-y cross-sectional images of the 3D tomograms. The annotated dataset was utilized for training OS-Net in a supervised manner (Fig. 1d). Once trained, OS-Net can automatically infer 3D nuclear regions of previously unseen cells through section-wise segmentation (Fig. 1e).

Note that we harnessed four-fold cross-validation of the dataset in order to compare the performance of different architectures and training strategies. Then, the trained 2D segmentation capability of OS-Net was rigorously evaluated by cross-modality and cross-laboratory validations based on simultaneous ODT and 3D fluorescence imaging. Finally, we demonstrated the 4D cell nucleus segmentation by frame-wise 3D segmentation of time-lapse ODT data. A detailed description of each step of OS-Net is presented below.

B. OPTICAL DIFFRACTION TOMOGRAPHY

ODT is essentially an inverse imaging problem of the Helmholtz equation that governs light propagation in matter. In the weak scattering regime, first-order scattering can be assumed, and the 3D RI tomogram of a sample is reconstructed from multiple 2D optical field images acquired with various illumination angles (Fig. 1a). In this study, a commercial ODT system (HT-2H; Tomocube Inc., Republic of Korea) was used, which also enables 3D fluorescence imaging. This system employs the digital mirror device (DMD) to control the illumination angle of a laser beam impinging onto a sample [22]. The voxel size of the tomograms obtained by this system was $0.098 \times 0.098 \times 0.195\ \mu m^3$ which was finer than its default optical resolution ($0.110 \times 0.110 \times 0.160\ \mu m^3$). For cross-laboratory validation, we used a separate ODT with the same specification, which was installed at a different institution.

C. SAMPLE PREPARATION AND IMAGING PROTOCOLS

For the acquisition of training and validation data, human breast cancer cells (MDA-MB-231, Korean Cell Line Bank) were cultured in Roswell Park Memorial Institute 1640 medium (RPMI-1640; Welgene, Republic of Korea), supplemented with 10% fetal bovine serum (FBS; CellSera, Australia) and 1% Penicillin-Streptomycin (Welgene, Republic of Korea) at 37°C in a humidified 5% CO2 atmosphere for 24 hours. The cells were fixed with 4% paraformaldehyde (PFP; Biosesang Inc., Republic of Korea) treatment for less than 10 minutes and then, their 3D RI tomograms were obtained using ODT.

For the cross-modality validation, we implemented the same cell culture and fixation protocols, and stained the cells with 4′ ,6-diamidino-2-phenylindole (DAPI; 1 μg/mL; Sigma Aldrich, MO) for 3 minutes. DAPI is a DNA-specific fluorescent probe that strongly binds to adenine-thymine rich regions of the double-stranded DNA [23]. For these cells, simultaneous ODT and 3D fluorescence imaging (z-stacked epi-fluorescence microscopy combined with 3D deconvolution) were performed.

For cross-laboratory validation, the same cell line was prepared with slightly different protocols. The cells were maintained in Dulbecco's Modified Eagle's Medium (DMEM; High Glucose, Pyruvate; Gibco, Thermo Fisher Scientific, MA), supplemented with 10% FBS and 1% Penicillin-Streptomycin at 37 °C in a humidified 10% CO2 atmosphere. Then, the cells were stained with the DNA-staining fluorescent dye Hoechst 33432 (0.1 μg/mL; Thermo Fisher Scientific, MA) and washed with fresh growth medium prior to ODT and fluorescence imaging. Note that no fixation was performed in this case.

For the time-lapse imaging, we prepared unlabeled live cells following the former preparation protocol only without fixation process. The ODT measurement of the live cells, maintained in a stable imaging chamber (37° and 5%; TomoChamber; Tomocube Inc., Republic of Korea), were performed every 10 minutes for a total of 1 hour.

D. DATASET PREPARATION

Tomographic reconstruction was done using a commercial software (TomoStudio, Tomocube Inc., Republic of Korea). Then, image processing was performed with the custom codes written in MATLAB (R2018a; MathWorks, MA). First, the 3D RI tomograms were decomposed into multiple 2D z-sections. Then, the sections were resized into 448 × 448 pixels. For the training and validation sets, we measured 3D RI tomograms of 50 cells including 934 2D RI cross-sections of the nucleus. For these RI images, we generated ground truth masks of the nucleus through manual binary annotation cross-confirmed by 3 trained biologists (see Fig. 1d). For four-fold cross-validation, the labeled dataset was divided into four equally-sized subsamples. Among the four subsamples, three of them were used as training set while the single remaining subsample was utilized as a validation set to evaluate the model. This validation process was repeated four times using each subsample as the validation set one after the other. Note that the 2D RI images from each cell were put in a single subsample to avoid overfitting.

For cross-modality and cross-laboratory validation, the nuclear masks were directly obtained by thresholding the fluorescence images that were simultaneously obtained

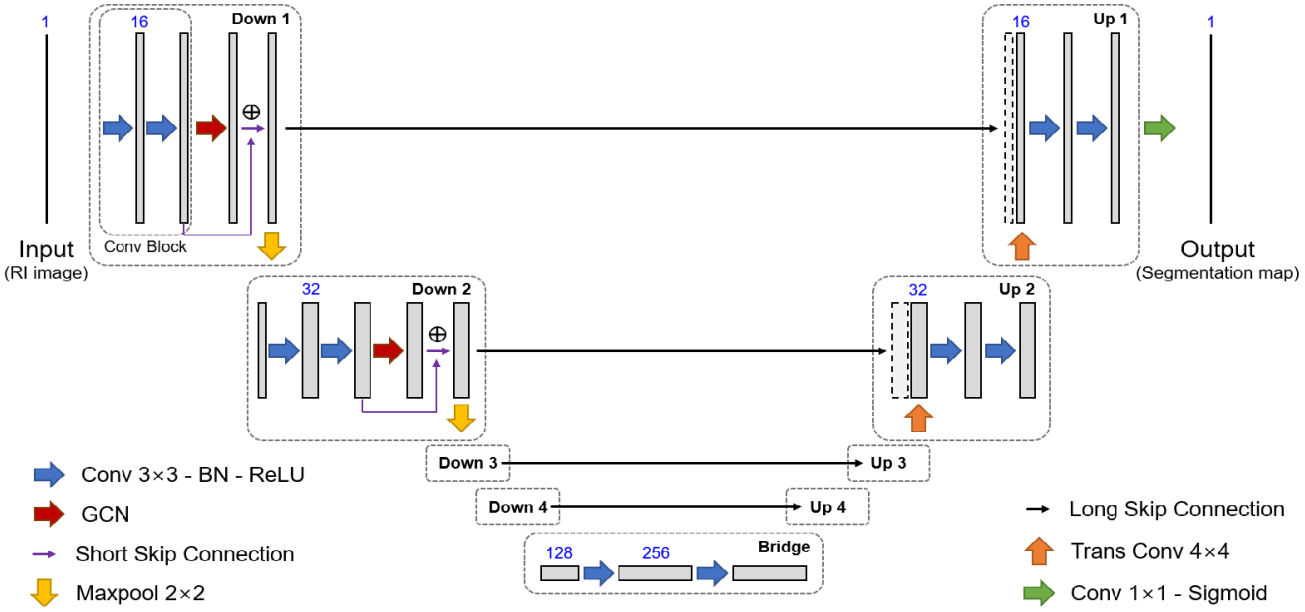

FIGURE 2. The architecture of OS-Net.

with ODT. The cross-modality validation data consisted of 122 2D RI sections and corresponding masks from 20 cells. The cross-laboratory validation data was composed of 181 sections and masks from 16 cells.

E. ARCHITECTURE OF OS-NET

Our proposed model, OS-Net has distinctive components such as GCN layers and SSC based on the encoder-decoder structure of U-Net [24]. Fig. 2 illustrates the overall architecture of OS-Net. First, OS-Net is a network that generates 2D cell nucleus segmentation map (448 × 448) by receiving 2D RI images (448 × 448) as input. It is divided into the feature extraction stage (encoder part) and the spatial resolution recovery stage (decoder part). There are a total of four Down modules in the former stage and four Up modules in the latter stage. OS-Net has four-times-reduced number of feature maps from 16 to 256 (Fig. 2, blue numbers above the feature maps), compared to original U-Net containing 64 to 1024 feature maps in each stage.

One Down module has a Conv (convolution) Block containing two sets of the convolutional layer with 3 × 3 filters followed by batch normalization [25] and rectified linear unit (ReLU) activation function [26]. After the Conv Block, a GCN (global convolutional network) layer extracts different features with other characteristics along the axes. (Details of GCN layer are well described in 2.6.) The feature maps after the GCN layer are combined with those before the GCN layer, by being added together. This process is called SSC (short skip connection), and is also important to convey meaningful information to the next step. At the last part of the Down module, there is a 2 × 2 max-pooling layer which reduces the dimension of the feature maps to half (N×N to N/2 × N/2).

After the four Down modules, there is a bridge part with only one Conv Block, followed by the four Up modules. In one of the Up modules, there is a Trans Conv (transposed convolutional) layer with 4 × 4 filters to increase the dimension of feature maps (N ×N to 2N × 2N). The upsampled feature maps are concatenated to those in the same level as the Up module, which is the LSC (long skip connection) to transfer spatial information across the level. Then, a Conv Block follows LSC again. Finally, the final segmentation map is released through a 1 × 1 convolution and sigmoid function, after 4 Up modules. The detailed OS-Net architecture and dimension of the feature maps after each component are also summarized in Table 1 (Conv Block: Convolutional Block containing 2 sets of convolutional layers with 3 × 3 filters followed by batch normalization and rectified linear unit (ReLU) activation function, GCN: Global Convolutional Network, SSC: Short Skip Connection, Trans Conv: Transposed Convolutional layer, LSC: Long Skip Connection).

In addition, the entire code for OS-Net was implemented using a deep learning open framework (PyTorch 0.4.1, Facebook, CA) and can be found in the given link https://github.com/ljm861/OSNet/blob/master/osnet.py.

F. GLOBAL CONVOLUTION NETWORK

The main purpose of the GCN structure is to enlarge the receptive field, because the segmentation task generally requires a larger receptive field compared to the classification task [27]. For more details on the structure of the GCN used above for this purpose, two strategies were used, rather than simply adding filters of size 7×7× 7 instead of 3 × 3 (Fig. 3).

The first strategy is to decompose a 7×7 filter into 1×7 and 7 × 1, which has two advantages. First, splitting large filters into small filters is advantageous from a learning point of

TABLE 1. Architecture and dimension of feature maps of OS-Net.

	Components	Dimension of feature maps
Input		448×448×1
Down 1	Conv Block – GCN – SSC	448×448×16
	Maxpool	224×224×16
Down 2	Conv Block – GCN – SSC	224×224×32
	Maxpool	112×112×32
Down 3	Conv Block – GCN – SSC	112×112×64
	Maxpool	56×56×64
Down 4	Conv Block – GCN – SSC	56×56×128
	Maxpool	28×28×128
Bridge	Conv Block	28×28×256
Up 1	Trans Conv – LSC	56×56×256
	Conv Block	56×56×128
Up 2	Trans Conv – LSC	112×112×128
	Conv Block	112×112×64
Up 3	Trans Conv – LSC	224×224×64
	Conv Block	224×224×32
Up 4	Trans Conv – LSC	448×448×32
	Conv Block	448×448×16
Output	1×1 Convolution – Sigmoid	448×448×1

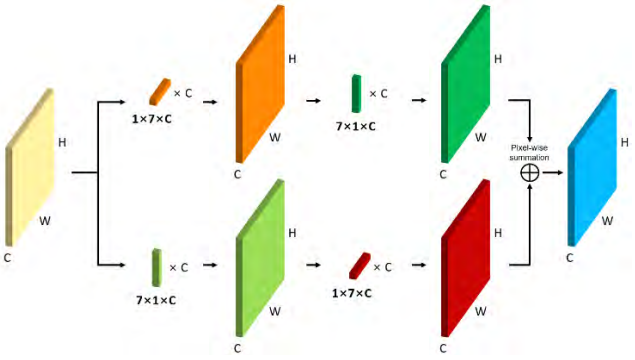

FIGURE 3. Schematic diagram of the GCN layer.

view, because the smaller filter size results in more distinctive features rather than some dead or useless features [28]. Secondly, from a memory point of view, the number of parameters is reduced from 49 (7 × 7 filter) to 14 (1 × 7 and 7 × 1 filters), so there is a large benefit of reducing the model complexity. Furthermore, it is also possible to extend the receptive field to the width or height direction and extract better features, rather than just simply see and only calculate a 3 × 3 region.

The second strategy is that GCN operates on two parallel paths. In the first path, feature maps are computed through 1 × 7 filters and then, 7 × 1 filters. In another path, feature maps are computed conversely. By using two paths, GCN can have similar advantage to Group convolutions [29]. After that, the pixel-wise summation results of two feature maps from each path become the final output of GCN. In doing so, we are able to utilize more suitable features by combining the features learned from the two paths.

G. TRAINING OF OS-NET

To train the deep-learning based models including OS-Net, we implemented binary cross entropy (BCE) loss between outputs from the model and corresponding nuclear masks (Fig. 1d, yellow double arrow). The influence of each weight in the model with respect to the loss function was computed by the backpropagation method [30]. Then, weights were updated by ADAM optimizer which is a first-order gradient-based optimization method based on adaptive estimates of lower-order moments [31]. For ADAM optimizer, we set the learning rate to 0.0005, β_1 to 0.5 and β_2 to 0.999. Furthermore, to reduce overfitting on the RI images in the training set, we artificially augmented the training set by using elastic deformation, flip and random crop methods [32]. To apply the elastic deformation method, we set the alpha and sigma as the following four pairs (1, 1), (5, 2), (1, 0.5), (1, 3). For the flip method, we flipped the training set images both horizontally and vertically. In case of the random crop method, we enlarged the image to 1.2, 1.3 or 1.4 times and then, randomly cropped the enlarged image into 448 × 448 pixels. Then, we used all the images in the training set as well as the randomly selected 30% of the images which were augmented randomly with these augmentation parameters in every training epoch. The training batch size was 32.

We utilized GPUs (Nvidia Tesla V100 32 GB) in Kakao Brain Cloud for efficient training. When performing four-fold cross-validation in the training stage, we allocated each fold into a single GPU, and it took almost 4 hours to complete 300 epochs in one fold.

H. EVALUATION OF THE TRAINED OS-NET

In order to evaluate the trained models, we utilized the datasets which were not used in the training stage. To measure how the model segmented the nucleus region well, we needed to compare two binary maps, a segmentation map and its corresponding nuclear mask. Every pixel in the generated segmentation map can be classified correctly or incorrectly to nuclear region or background. Thus, the generated map may contain four cases: true positive (TP), false positive (FP), true negative (TN) and false negative (FN) [33]. From these values, Dice similarity coefficient (DICE), Jaccard similarity coefficient (Jaccard), F0.5 and precision and recall are defined as follows:

$$\text{DICE} = \frac{2|X \cap Y|}{|X| + |Y|} = \frac{2 \times TP}{2 \times TP + FP + FN} \quad (1)$$

$$\text{Jaccard} = \frac{|X \cap Y|}{|X \cup Y|} = \frac{TP}{TP + FP + FN} \quad (2)$$

$$\text{F0.5 score} = \frac{1.25 \times TP}{1.25 \times TP + FP + 0.25 \times FN} \quad (3)$$

$$\text{Precision} = \frac{TP}{TP + FP} \tag{4}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{5}$$

Here, X and Y are the segmentation map from the segmentation model and the corresponding nuclear mask, respectively. In addition, area under curve (AUC) of precision and recall (PR) curve was utilized to evaluate the overall performance of a particular model. In the PR curve, the x-axis is the precision and y-axis is the recall which is the same as the true positive rate.

The overall evaluation process was divided into four parts: design and analysis of OS-Net (four-fold cross-validation), cross-modality validation between ODT and fluorescence imaging in 2D and 3D, cross-laboratory validation and application to 4D segmentation. First, we performed four-fold cross-validation conducted with a different validation set for each fold to compare the performance of numerous network architectures and optimize the OS-Net structure. In order to quantitatively evaluate the performance of various trained models, we calculated DICE, Jaccard, F0.5 score and AUC of PR curves between the segmentation maps obtained from the trained model and the corresponding nuclear masks annotated by experts (Table 2). To calculate these four metrics, the segmentation map should be revised into a binary image, since the nuclear mask is also a binary image only containing 0 and 1. Thus, after the 1×1 convolution and sigmoid function (the last part of OS-Net), we applied a threshold value of 0.5, changed the map into the binary image and compared it to the label. Then, we extensively compared the performance of various architectures to demonstrate the superior performance of our proposed model, OS-Net. We chose the baseline model as Unet64, the original Unet [24]. In Unet64, the number of feature maps increases from 64 to 1,024 as the level increases. We also experimented with Unet16 which is four times lighter than Unet64, to see if it would be possible to reduce the number of parameters in the model with the same performance. Then, we added the GCN layer and SSC to Unet16 to improve the segmentation performance. In the next step, various data augmentation (Aug) techniques described above were applied in order to increase the performance and reduce overfitting. Furthermore, we also compared the segmentation results from FusionNet [34] which is an end-to-end image segmentation model for electron microscopy images. Note that we implemented FusionNet containing 16 to 256 feature maps like Unet16 and OS-Net for a fair comparison.

Next, we performed the cross-modality validation to confirm that the cell nucleus segmentation results from OS-Net, which was trained only with expert-annotated 2D RI images, would be comparable to the nuclear masks obtained from fluorescence imaging. The cross-modality validation data obtained with DAPI was utilized to evaluate the segmentation performance of all the trained models that we already compared in the four-fold cross-validation. Furthermore, to compare the segmentation performance of learning-based method with the conventional and rule-based method, we also developed an edge-based method, which has been traditionally and widely used in cell image processing [35], [36]. (In the edge-based method, the gradient of the image pixels was first calculated using the 3×3 Sobel operator. Then, a dilated gradient mask was created from the calculated gradient mask, and the interior gap was filled. After that, the surrounding diamond structuring elements were removed using a smoothing kernel, and the segmentation map was generated.) Then, the difference between the segmentation maps produced by various trained models or edge-based method, and the corresponding nuclear masks were quantified by calculating four metrics, DICE, Jaccard, F0.5 and AUC of PR (Table 3). We also compared the 2D segmentation results from OS-Net and edge-based method with the nuclear masks in the image domain (Fig. 5b). In addition, 3D cell nucleus segmentation was performed via section-wise segmentation and the results were rendered to 3D nuclear volume (Fig. 5c).

Furthermore, the cross-laboratory validation was also performed by using the dataset taken from an external institute, in order to confirm the robustness of OS-Net (Fig. 6). Finally, we also applied OS-Net to time-lapse ODT data to demonstrate the feasibility of 4D cell nucleus segmentation using our proposed method (Fig. 7).

III. RESULTS AND DISCUSSION

A. DESIGN AND ANALYSIS OF OS-NET

The nucleus segmentation performance of various network architectures (deep-learning based models) are summarized in Table 2. In order to quantitatively compare and analyze the results, DICE, Jaccard, F0.5 score and AUC of PR curves between the generated segmentation maps and the corresponding nuclear masks annotated by experts were calculated. Because the results were obtained through four-fold cross-validation, the mean values and standard deviations of the four folds' results were calculated together. During the evaluation process, 2D RI images in the validation set were inserted into the trained model to infer the nucleus segmentation maps, which took only 0.91 seconds to get one segmentation map from its corresponding RI image.

In Table 2, the results on DICE, Jaccard, and F0.5 of Unet64 were slightly improved than those from Unet16. However, the threshold was fixed to a certain value (here, 0.5) when making the final segmentation map (after sigmoid function) binary. Since DICE, Jaccard, and F0.5 were all threshold dependent values, the threshold of 0.5 had worked a little better for the output from Unet64 than Unet16. In the case of AUC of PR, which was independent of the threshold value and indicative of the overall performance of a particular model, the result of Unet16 was slightly higher than that of Unet64. When all four metrics were considered, the performances of Unet64 and Unet16 were nearly comparable. However, the most important point here is that the compression of the model for the cell nucleus segmentation task was well performed, because the performance was nearly similar,

TABLE 2. Comparison of the segmentation performance among various architectures and training strategies.

Model	DICE	Jaccard	F0.5	PR (AUC)	The number of parameters
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	
Unet64	0.8506 (0.0687)	0.7360 (0.0781)	0.9666 (0.0921)	0.8768 (0.0687)	8,638,592
Unet16	0.8496 (0.0322)	0.7330 (0.0662)	0.9531 (0.0217)	0.8822 (0.0607)	2,159,648
Unet16 + GCN	0.8694 (0.0256)	0.7529 (0.0592)	0.9918 (0.0888)	0.9025 (0.0380)	2,770,848
Unet16 + GCN + SSC	0.8692 (0.0299)	0.7610 (0.0555)	0.9907 (0.0889)	0.9039 (0.0244)	2,770,848
Unet16 + Aug	0.8870 (0.0275)	0.8038 (0.0561)	1.0393 (0.0746)	0.9229 (0.0115)	2,159,648
Unet16 + GCN + Aug	0.8951 (0.0231)	0.8113 (0.0520)	1.0478 (0.0743)	0.9235 (0.0183)	2,770,848
FusionNet + Aug	0.9001 (0.0261)	0.8194 (0.0532)	1.0560 (0.0798)	0.9325 (0.0171)	4,913,344
FusionNet + GCN + SSC + Aug (FusionNet v2)	0.9033 (0.0258)	0.8267 (0.0466)	1.0601 (0.0798)	0.9366 (0.0132)	5,524,544
Unet16 + GCN + SSC + Aug (OS-Net)	**0.9131 (0.0131)**	**0.8452 (0.0211)**	**1.1016 (0.0345)**	**0.9344 (0.0124)**	**2,770,848**

even though the number of parameters in Unet16 was reduced by a factor of four. Moreover, when the GCN layers were added to Unet16 model, all the metric values were increased. The similar situation was repeated when SSC were added together. This was because the GCN layer and SSC helped to extract improved feature maps that were advantageous for this task.

In addition, various Aug techniques were applied. Our proposed model, OS-Net showed the highest segmentation performance again, even though the other models with Aug also showed increased performance. FusionNet with GCN, SSC, and Aug (FusionNet v2) showed comparable performance to our proposed model, OS-Net. However, even if the number of parameters in FusionNet v2 was two times more than that in OS-Net, FusionNet v2 failed to show improved performance. Thus, we concluded that increasing the depth of the model did not simply improve the performance depending on the task. Therefore, our proposed model, OS-Net showed superior nucleus segmentation performance when considering the number of parameters.

So how did our proposed framework, OS-Net, produce such superior cell nucleus segmentation results? The addition of GCN layer and SSC, which were the main components of OS-Net, enhanced the performance because the network structure could be trained to extract better feature maps for the cell nucleus segmentation task. This was verified through the visualization results of the feature maps created after each module. Fig. 4 shows the visualization results of OS-Net and the baseline model, Unet16, which was trained with data augmentation, at each level. In OS-Net, as the input image passed through the Down modules, the features that could represent the nuclear region were gradually extracted out. Then, as the extracted features were combined with the Up module at the same level, the nuclear region became clearer. Since the confidence of the nuclear region in the final output after the sigmoid function was very high, the nuclear region could be segmented with high accuracy when converted to a binary image at a threshold of 0.5. However, the results of Unet16 showed that the Down module could not extract the effective features compared to OS-Net, and the feature maps after Up modules exhibited checkerboard patterns. Even though the final output of Unet16 also had high pixel values in the nuclear region, it did not give strong confidence to the whole region. In this case, the nuclear region was likely to be under-segmented depending on the threshold value chosen for the binary image in the final stage.

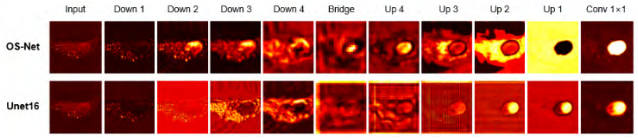

FIGURE 4. Visualization of the feature maps in the different levels from OS-Net and Unet16.

B. CROSS-MODALITY VALIDATION IN 2D AND 3D USING FLUORESCENCE IMAGING

The cross-modality validation results of the conventional edge-based method and various deep-learning based models are summarized in the upper part of Table 3. First, the edge-based method showed poor performance compared to all deep-learning based models. Since the Sobel operator had only 3×3 dimension to calculate the gradient, it would be extremely difficult for the rule-based method to distinguish the nuclear region, simply based on the difference among neighboring pixels. For the deep-learning based models, the metric calculation results showed very similar tendency with those in Table 2. Again, OS-Net showed the best segmentation performance on the cross-modality validation data, when considering the lightness of the model.

Figure 5a is 2D RI images of the cross-modality validation data. Figure 5b is the same images with nuclear masks by fluorescence imaging and the nucleus segmentation results from OS-Net and the conventional edge-based method. The segmentation results from OS-Net trained only with RI images (yellow contour) accurately agreed with ground truth mask obtained from fluorescence imaging with DAPI (red contour). However, the edge-based method (cyan contour), falsely identified the nucleolus or organelles with higher RI values than surrounding pixels as part of the nuclear region. It clearly showed the limitation of the conventional rule-based methods.

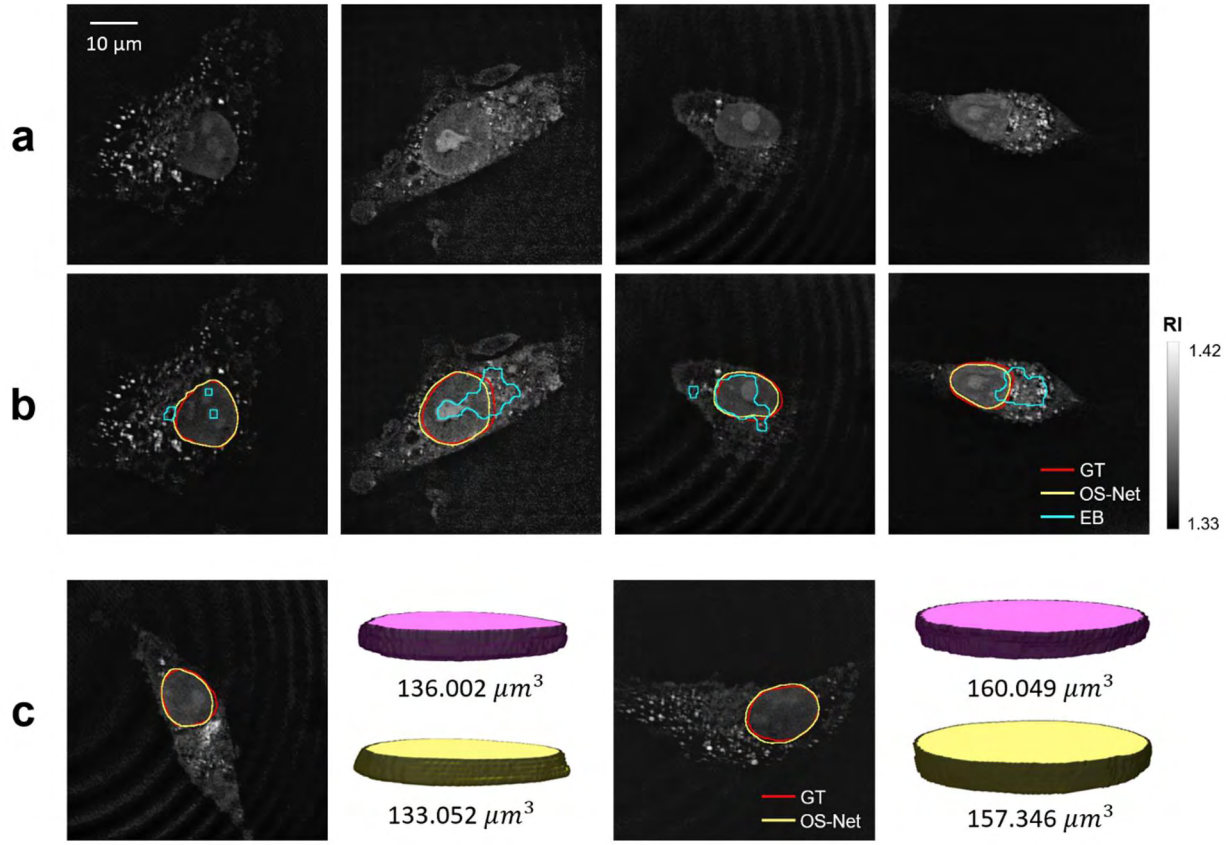

FIGURE 5. The cell nucleus segmentation performance of OS-Net in cross-modality validation using fluorescence imaging. (a) 2D RI images in the cross-modality validation set. (b) The corresponding ground truth (GT) nuclear regions (red contour); and segmentation by OS-Net (yellow contour) and by conventional edge-based (EB) method (cyan contour). (c) 3D cell nucleus segmentation. RI section images with ground truth nuclear regions (red contour) and the corresponding segmentation by OS-Net (yellow contour) with the corresponding ground truth 3D nuclear volumes (light purple) and the corresponding 3D segmentation by OS-Net via section-wise segmentation (light yellow).

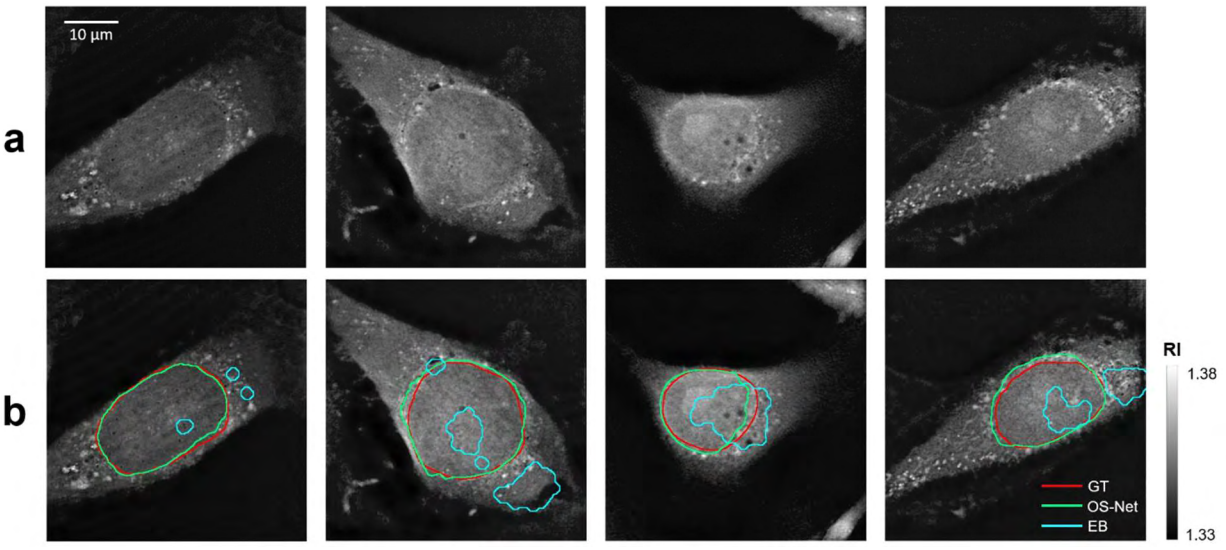

FIGURE 6. The cell nucleus segmentation performance of OS-Net in cross-laboratory validation. (a) 2D RI images measured in a different lab oratory (cross-laboratory-validation set). (b) The corresponding ground truth nuclear regions (red contour); segmentation by OS-Net (green contour) and edge-based method (cyan contour).

In addition, when testing the segmentation performance of the trained OS-Net described in Fig. 1e, the RI sections of the tomogram were sequentially inserted along the z-direction, and the nucleus segmentation results were subsequently obtained, allowing 3D segmentation of RI tomogram (section-wise segmentation). Fig. 5c shows the

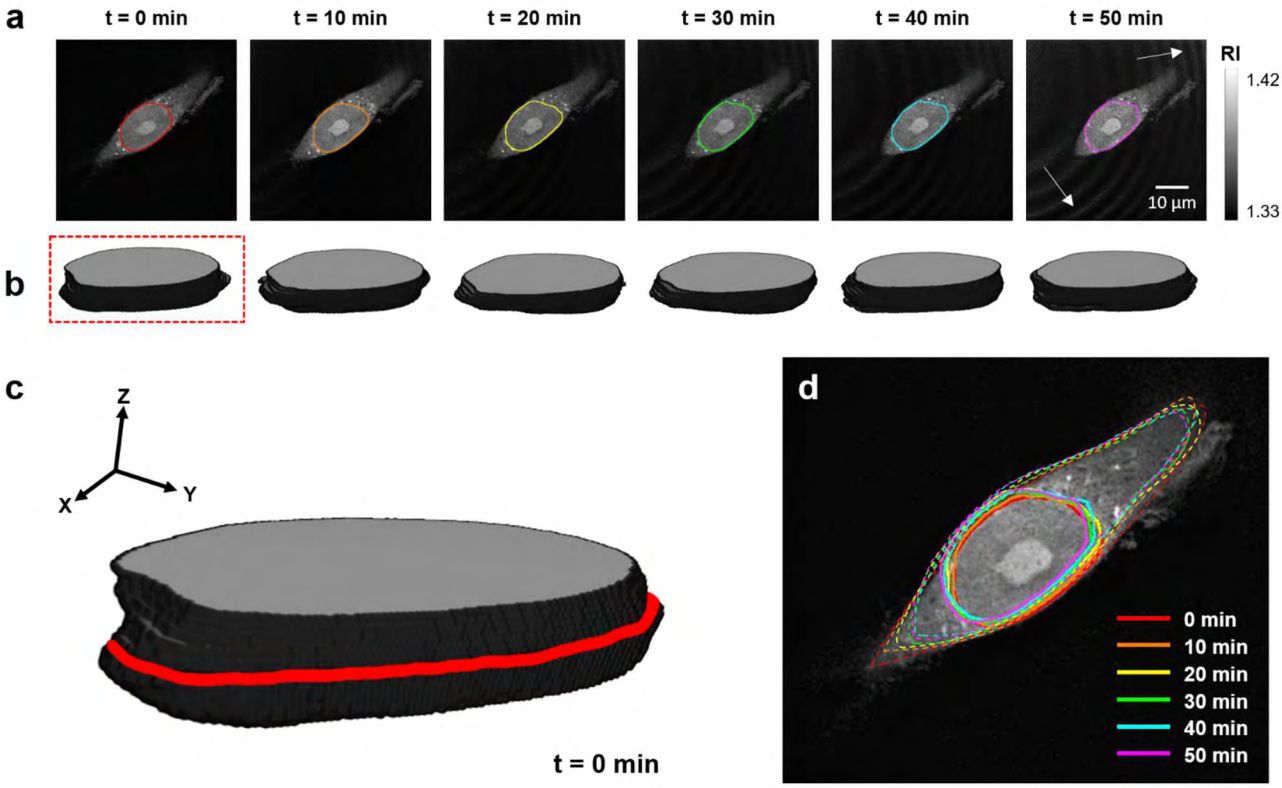

FIGURE 7. 4D cell nucleus segmentation by frame-wise 3D segmentation. (a) 2D RI section images of time-lapse 3D RI tomograms with segmentat ion by OS-Net. (b) 3D rendering result of 15 sections. (c) 3D segmentation at the first frame. (d) 3D nuclear shape dynamics visualized in a middle section (the red contour in c).

3D volume rendering results of 6 sections containing the nuclear region. The RI sections with nuclear masks (red contour) and segmentation results from OS-Net (yellow contour), and their 3D volume rendering results of the nuclear masks and segmentation from OS-Net are shown respectively. Since we performed 3D rendering only with 6 sections, which could be confirmed with the corresponding nuclear masks obtained with DAPI, the nuclear volume might look like a cylinder. Furthermore, it was also feasible to calculate the volume of the nucleus based on the voxel size. There was a slight difference between the sizes of pink volume (ground truth) and yellow volume (segmentation results from OS-Net). OS-Net tended to undersegment the nucleus region compared to ground truth slightly. (Fig. 5c) However, since the average volume intersection ratio was 90.372% and the absolute difference was in a scale of sub-μm, we could conclude that 3D segmentation was successfully performed using OS-Net.

C. CROSS-LABORATORY VALIDATION

The bottom parts of Table 3 and Fig. 6 show the results of cross-laboratory validation using the data obtained from a different institution to evaluate the robustness of the OS-Net. In Fig. 6a, RI images of the same cell line had quite different distribution from the images taken in our laboratory (shown in Fig. 5a). Figure 6b are the 2D RI images with the nuclear mask obtained from fluorescence imaging with Hoechst dye (red contour), and the segmentation results (green contour) from OS-Net and edge-based method (cyan contour). Even though OS-Net had never seen the data from the other institution before, the segmentation maps produced from OS-Net were almost similar with the nuclear masks. In addition, the average volume intersection ratio for cross-laboratory validation set was 81.002%. This meant that the features extracted from the spatial distribution of RI seemed to work well. Therefore, we could confirm that OS-Net had the highly robust performance of nucleus segmentation for any RI image. With training data from the external institute added to the original training set, it is expected that the OS-Net performance of nucleus segmentation will be further improved.

D. 4D CELL NUCLEUS SEGMENTATION

Finally, we also performed frame-wise 3D segmentation for the time-lapse ODT data taken at an interval of 10 minutes. Figs. 7a and 7b show the RI sections from t = 0 to t = 50 min and their 3D rendering results of 15 sections, respectively. Fig. 7b includes the segmented nuclear region to visualize the changes in the nuclear shape along the z-axis. Fig. 7c depicts the 3D volume at the first frame (t = 0 min, red

TABLE 3. Comparison of the segmentation performance on the cross-modality and cross-laboratory validation data.

Validation	Model	DICE	Jaccard	F0.5	PR (AUC)
		Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
Cross-modality validation	Edge-based method (Conventional)	0.3716	0.2325	0.3773	
	Unet64	0.8136 (0.0148)	0.6819 (0.0203)	1.0211 (0.0464)	0.8651 (0.0242)
	Unet16	0.8461 (0.0301)	0.7328 (0.0537)	1.0479 (0.0217)	0.8663 (0.0087)
	Unet16 + GCN	0.8622 (0.0244)	0.7496 (0.0440)	1.0806 (0.0228)	0.8730 (0.0413)
	Unet16 + GCN + SSC	0.8691 (0.0143)	0.7584 (0.0269)	1.0932 (0.0137)	0.8813 (0.0116)
	Unet16 + Aug	0.8841 (0.0177)	0.8012 (0.0217)	1.1157 (0.0041)	0.9098 (0.0056)
	Unet16 + GCN + Aug	0.8936 (0.0199)	0.8100 (0.0304)	1.1371 (0.0099)	0.9147 (0.0099)
	FusionNet + Aug	0.8973 (0.0051)	0.8127 (0.0083)	1.1237 (0.0061)	0.9063 (0.0201)
	FusionNet v2	0.8980 (0.0084)	0.8178 (0.0065)	1.1241 (0.0205)	0.9191 (0.0052)
	OS-Net	**0.9094 (0.0053)**	**0.8352 (0.0091)**	**1.1391 (0.0090)**	**0.9192 (0.0019)**
Cross-laboratory validation	Edge-based method (Conventional)	0.1759	0.1095	0.2812	
	OS-Net	**0.7981 (0.0416)**	**0.6699 (0.0711)**	**1.0281 (0.0349)**	**0.8022 (0.0251)**

box in Fig. 5b) and the red line corresponds to the sections in Fig. 7a. In Fig. 7d, the boundaries of the cell and nucleus segmentation at each time step are shown on top of the cell image to demonstrate the changes in the cellular and nuclear shapes over time.

As shown in Fig. 7, the cell nucleus segmentation was successfully performed at each time step. However, noise was observed in the RI image for a long period (i.e., Fig. 7a, $t = 50$ min, the fringe noise denoted by white arrows). This noise stemmed from the interference that could occur over time due to the movement of the cell. When using the time-lapse function of the ODT, the RI images were continuously acquired using the settings chosen for the first time step, which could have resulted in the fringe noise. Despite the existence of noise in the images, OS-Net accurately performed the nucleus segmentation as a whole. Therefore, with its feasibility of 4D segmentation of the nucleus, OS-Net can be applied in various studies such as real-time observation of changes in the nuclear region.

IV. CONCLUSION

In order to segment a cell nucleus from label-free ODT images, a deep learning framework was developed. A novel architecture with a lightweight encoder-decoder structure and specialized substructures, and optimized training strategies were carefully designed to enrich the spatial information from RI distributions. Once trained with expert-annotated data, the proposed network presented accurate cell nucleus segmentation in 2D, 3D, and even 4D label-free ODT images. We rigorously validated this network via cross-modality and cross-laboratory experiments. The results indicate that certain patterns in the 3D spatial distribution of RI tomograms might facilitate the identification of biological substances. The proposed framework is ready for broad biomedical applications. We also made OS-Net publicly available to facilitate its applications in other research areas.

CONFLICT OF INTEREST

YoungJu Jo and YongKeun Park has financial interests in Tomocube Inc., a company that commercializes optical diffraction tomography and quantitative phase imaging instruments.

ACKNOWLEDGMENT

The authors would like to thank Dr. Hyun-seok Min and Dr. Wonmo Sung for useful comments. They would also like to thank Kakao Brain corporations to provide Kakao Brain Cloud. Y. Jo was with the Department of Physics, Korea Advanced Institute of Science and Technology (KAIST), Daejeon 34141, South Korea, also with Tomocube Inc., Daejeon 34051, South Korea, and also with the KAIST Institute for Health Science and Technology, KAIST, Daejeon 34141, South Korea.

REFERENCES

[1] M. L. Turgeon, *Clinical Hematology: Theory and Procedures*. Philadelphia, PA, USA: Lippincott Williams & Wilkins, 2005.
[2] V. Backman *et al.*, "Detection of preinvasive cancer cells," *Nature*, vol. 406, no. 6791, pp. 35–36, 2000.
[3] L. T. Perelman, V. Backman, M. Wallace, G. Zonios, R. Manoharan, A. Nusrat, S. Shields, M. Seiler, C. Lima, T. Hamano, I. Itzkan, J. Van Dam, J. M. Crawford, and M. S. Feld, "Observation of periodic fine structure in reflectance from biological tissue: A new technique for measuring nuclear size distribution," *Phys. Rev. Lett.*, vol. 80, no. 3, p. 627, 1998.
[4] W. Sung, Y. Jeong, H. Kim, H. Jeong, C. Grassberger, S. Jung, G-O. Ahn, I. H. Kim, J. Schuemann, K. Lee, and S.-J. Ye, "Computational modeling and clonogenic assay for radioenhancement of gold nanoparticles using 3D live cell images," *Radiat. Res.*, vol. 190, no. 5, pp. 558–564, 2018.
[5] E. Wolf, "Three-dimensional structure determination of semi-transparent objects from holographic data," *Opt. Commun.*, vol. 1, no. 4, pp. 153–156, 1969.
[6] Y. Park, C. Depeursinge, and G. Popescu, "Quantitative phase imaging in biomedicine," *Nature Photon.*, vol. 12, no. 10, pp. 578–589, 2018.
[7] K. Lee, K. Kim, J. Jung, J. Heo, S. Cho, S. Lee, G. Chang, Y. Jo, H. Park, and Y. Park, "Quantitative phase imaging techniques for the study of cell pathophysiology: From principles to applications," *Sensors*, vol. 13, no. 4, pp. 4170–4191, 2013.
[8] K. Kim, J. Yoon, S. Shin, S. Lee, S.-A. Yang, and Y. Park, "Optical diffraction tomography techniques for the study of cell pathophysiology," *J. Biomed. Photon. Eng.*, vol. 2, no. 2, 2016, Art. no. 020201.

[9] J. Jung, S.-J. Hong, H.-B. Kim, G. Kim, M. Lee, S. Shin, S. Lee, D.-J. Kim, C.-G. Lee, and Y. Park, "Label-free non-invasive quantitative measurement of lipid contents in individual microalgal cells using refractive index tomography," *Sci. Rep.*, vol. 8, 2018, Art. no. 6524.

[10] Y. Kim, H. Shim, K. Kim, H. Park, S. Jang, and Y. Park, "Profiling individual human red blood cells using common-path diffraction optical tomography," *Sci. Rep.*, vol. 4, Oct. 2014, Art. no. 6659.

[11] H. Park, S.-H. Hong, K. Kim, S.-H. Cho, W.-J. Lee, Y. Kim, S.-E. Lee, and Y. Park, "Characterizations of individual mouse red blood cells parasitized by *Babesia microti* using 3-D holographic microscopy," *Sci. Rep.*, vol. 5, Jun. 2015, Art. no. 10827.

[12] S. Abuhattum, K. Kim, T. M. Franzmann, A. Esslinger, D. Midtvedt, R. Schluessler, S. Moellmert, H. S. Kuan, S. Alberti, V. Zaburdaev, and J. Guck, "Intracellular mass density increase is accompanying but not sufficient for stiffening and growth arrest of yeast cells," *BioRxiv*, Jan. 2018, Art. no. 388876.

[13] D. Kim, S. Lee, M. Lee, J. Oh, S.-A. Yang, and Y. Park, "Refractive index as an intrinsic imaging contrast for 3-D label-free live cell imaging," *BioRxiv*, Jan. 2017, Art. no. 106328.

[14] K. Kim, S. Lee, J. Yoon, J. Heo, C. Choi, and Y. Park, "Three-dimensional label-free imaging and quantification of lipid droplets in live hepatocytes," *Sci. Rep.*, vol. 6, Nov. 2016, Art. no. 36815.

[15] Y. Jo, H. Cho, S. Y. Lee, G. Choi, G. Kim, H.-S. Min, and Y. Park, "Quantitative phase imaging and artificial intelligence: A review," *IEEE J. Sel. Topics Quantum Electron.*, vol. 25, no. 1, Jan./Feb. 2019, Art. no. 6800914.

[16] C. L. Chen, A. Mahjoubfar, L.-C. Tai, I. K. Blaby, A. Huang, K. R. Niazi, and B. Jalali, "Deep learning in label-free cell classification," *Sci. Rep.*, vol. 6, Mar. 2016, Art. no. 21471.

[17] Y. Jo, S. Park, J. Jung, J. Yoon, H. Joo, M.-H. Kim, S.-J. Kang, M. C. Choi, S. Y. Lee, and Y. Park, "Holographic deep learning for rapid optical screening of anthrax spores," *Sci. Adv.*, vol. 3, no. 8, 2017, Art. no. e1700606.

[18] T. H. Nguyen, S. Sridharan, V. Macias, A. Kajdacsy-Balla, J. Melamed, M. N. Do, and G. Popescu, "Automatic Gleason grading of prostate cancer using quantitative phase imaging and machine learning," *Proc. SPIE*, vol. 22, no. 3, 2017, Art. no. 036015.

[19] J. Yoon, Y. Jo, M.-H. Kim, K. Kim, S. Lee, S.-J. Kang, and Y. Park, "Identification of non-activated lymphocytes using three-dimensional refractive index tomography and machine learning," *Sci. Rep.*, vol. 7, no. 1, 2017, Art. no. 6654.

[20] N. Pavillon, A. J. Hobro, S. Akira, and N. I. Smith, "Noninvasive detection of macrophage activation with single-cell resolution through machine learning," *Proc. Nat. Acad. Sci. USA*, vol. 115, no. 12, pp. E2676–E2685, 2018.

[21] G. Kim, Y. Jo, H. Cho, H.-S. Min, and Y. Park, "Learning-based screening of hematologic disorders using quantitative phase imaging of individual red blood cells," *Biosensors Bioelectron.*, vol. 123, pp. 69–76, Jan. 2019.

[22] S. Shin, K. Kim, T. Kim, J. Yoon, K. Hong, J. Park, and Y. Park, "Optical diffraction tomography using a digital micromirror device for stable measurements of 4D refractive index tomography of cells," *Proc. SPIE*, vol. 9718, Mar. 2016, Art. no. 971814.

[23] J. Kapuscinski, "DAPI: A DNA-specific fluorescent probe," *Biotechnic Histochem.*, vol. 70, no. 5, pp. 220–233, 1995.

[24] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* Cham, Switzerland: Springer, 2015, pp. 234–241.

[25] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," 2015, *arXiv:1502.03167*. [Online]. Available: https://arxiv.org/abs/1502.03167

[26] G. E. Dahl, T. N. Sainath, and G. E. Hinton, "Improving deep neural networks for LVCSR using rectified linear units and dropout," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2013, pp. 8609–8613.

[27] C. Peng, X. Zhang, G. Yu, G. Luo, and J. Sun, "Large kernel matters—Improve semantic segmentation by global convolutional network," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 1743–1751.

[28] M. D. Zeiler and R. Fergus, "Visualizing and understanding convolutional networks," in *Proc. Eur. Conf. Comput. Vis.* Springer, 2014, pp. 818–833.

[29] S. Xie, R. Girshick, P. Dollár, Z. Tu, and K. He, "Aggregated residual transformations for deep neural networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 5987–5995.

[30] M. Riedmiller and H. Braun, "A direct adaptive method for faster backpropagation learning: The RPROP algorithm," in *Proc. IEEE Int. Conf. Neural Netw.*, Mar./Apr. 1993, pp. 586–591.

[31] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*. [Online]. Available: https://arxiv.org/abs/1412.6980

[32] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2012, pp. 1097–1105.

[33] K. Hu, Z. Zhang, X. Niu, Y. Zhang, C. Cao, F. Xiao, and X. Gao, "Retinal vessel segmentation of color fundus images using multiscale convolutional neural network with an improved cross-entropy loss function," *Neurocomputing*, vol. 309, pp. 179–191, Oct. 2018.

[34] T. M. Quan, D. G. C. Hildebrand, and W.-K. Jeong, "FusionNet: A deep fully residual convolutional neural network for image segmentation in connectomics," 2016, *arXiv:1612.05360*. [Online]. Available: https://arxiv.org/abs/1612.05360

[35] S.-F. Yang-Mao, Y.-K. Chan, and Y.-P. Chu, "Edge enhancement nucleus and cytoplast contour detector of cervical smear images," *IEEE Trans. Syst. Man, Cybern. B, Cybern.*, vol. 38, no. 2, pp. 353–366, Apr. 2008.

[36] S. F. Kemeny and A. M. Clyne, "A simplified implementation of edge detection in MATLAB is faster and more sensitive than fast Fourier transform for actin fiber alignment quantification," *Microsc. Microanal.*, vol. 17, no. 2, pp. 156–166, 2011.

JIMIN LEE received the bachelor's degree in nuclear and quantum engineering from KAIST, South Korea, in 2015. She is currently pursuing the Ph.D. degree in the Biomedical Radiation Sciences Program of the Department of Transdisciplinary Studies, Seoul National University, South Korea. She is also with Prof. Ye's Research Group. Her research interest includes machine/deep learning approaches for solving biomedical imaging problems.

HYEJIN KIM received the bachelor's degree in nuclear engineering from Seoul National University, South Korea, where she is currently pursuing the Ph.D. degree in the Biomedical Radiation Sciences Program of the Department of Transdisciplinary Studies. She is also with Prof. Ye's Research Group. Her research interest includes biological modeling for NP-mediated radiation sensitization.

HYUNGJOO CHO received the bachelor's degree in electronic engineering from Dongguk University, South Korea. He is currently pursuing the M.S. degree in the Biomedical Radiation Sciences Program of the Department of Transdisciplinary Studies, Seoul National University, South Korea. He was a Software Engineer with LG Electronics, where he was in charge of developing gesture recognition algorithm using optical devices. He is also with the Prof. Ye's Research Group. His research interest includes machine learning approaches for solving biomedical imaging problems.

YOUNGJU JO received the bachelor's degree in physics and mathematics from KAIST, in 2018, under the supervision of Prof. Y. Park. He is currently pursuing the Ph.D. degree with the Department of Applied Physics, Stanford University. He is currently interested in systems and computational neurobiology based on novel techniques from optics and machine learning. In particular, he has led the group's machine learning approach to biomedical quantitative phase imaging, which is now commercialized by the start-up company Tomocube. He has been supported by the Asan Foundation Biomedical Science Scholarship, in 2018, the SPIE Optics and Photonics Education Scholarship, in 2014, and the KAIST Presidential Fellowship, in 2013. He was a recipient of the KAIST Creativity and Challenge Award, in 2018, the Samsung HumanTech Paper Award, in 2017, and the Talent Award of Korea, in 2015.

YUJIN SONG received the bachelor's degree in materials science and engineering from Kyunghee University, South Korea, in 2016. She is currently pursuing the M.S. degree in the Biomaterial Sciences Program of the Department of Transdisciplinary Studies, Seoul National University, South Korea. She is also with the Prof. Lee's Research Group. Her research interest includes colorimetric biosensor using nanoparticles for the prevention and detection of disease.

DAEWOONG AHN received the bachelor's degree in computer science from Seokyeong University, South Korea, in 2019. He was an Intern with the Prof. Ye's Research Group. He is currently a Researcher with Tomocube. His current research interest is to apply deep learning techniques to biomedical area.

KANGWON LEE received the Ph.D. degree in MSE from the University of Michigan, Ann Arbor. He has further studied cell and tissue engineering at Harvard University as a Postdoctoral Fellow under the supervision of Prof. D. J. Mooney. He was a Senior Research Scientist with the Korea Institute of Science and Technology (KIST) and an Adjunct Associate Professor of biomedical engineering with the University of Science and Technology (UST), South Korea. He is currently an Assistant Professor with the Graduate School of Convergence Science and Technology, Seoul National University (SNU), South Korea. His expertise spans interdisciplinary fields of tissue engineering and biosensors.

YONGKEUN (PAUL) PARK received the Ph.D. degree from Harvard-MIT Health Sciences and Technology. He is currently an Associate Professor of physics with the Korea Advanced Institute of Science and Technology, South Korea. Two start-up companies with over 30 employees have been created from his research (Tomocube Inc. and The.Wave.Talk). He has published over 120 peer-reviewed papers, including three *Nature Photonics*, two *Nature Communications*, four *Physical Review Letters*, and four PNAS papers. His area of research is wave optics and its applications for biology and medicine. His awards and honors include the Fellow Membership (Optical Society of America), the Jinki Hong Creative Award, and the Medal of Honor in Science and Technology (President of South Korea).

SUNG-JOON YE received the Ph.D. degree in nuclear engineering from Purdue University. Since 2018, he has been leading the Graduate School of Convergence Science and Technology at the forefront of multidisciplinary research as its Dean. He is currently a Professor with the Graduate School of Convergence Science and Technology, Seoul National University, South Korea. He has published more than 90 peer-reviewed papers and holds six patents. His research interests include innovative convergence research about radiation in medicine, space, and power.