

# Nuclei Counting in Microscopy Images With Three Dimensional Generative Adversarial Networks

Shuo Han<sup>a</sup>, Soonam Lee<sup>a</sup>, Chichen Fu<sup>a</sup>,  
Paul Salama<sup>b</sup>, Kenneth W. Dunn<sup>c</sup>, and Edward J. Delp<sup>a</sup>

<sup>a</sup>Video and Image Processing Laboratory, School of Electrical and Computer Engineering,  
Purdue University, West Lafayette, Indiana 47907 U.S.A.

<sup>b</sup>Department of Electrical and Computer Engineering, Indiana University-Purdue University,  
Indianapolis, Indiana 46202 U.S.A.

<sup>c</sup>Division of Nephrology, School of Medicine, Indiana University, Indianapolis, Indiana 46202  
U.S.A.

## ABSTRACT

Microscopy image analysis can provide substantial information for clinical study and understanding of biological structures. Two-photon microscopy is a type of fluorescence microscopy that can image deep into tissue with near-infrared excitation light. We are interested in methods that can detect and characterize nuclei in 3D fluorescence microscopy image volumes. In general, several challenges exist for counting nuclei in 3D image volumes. These include crowding and touching of nuclei, overlapping of nuclei, and shape and size variances of the nuclei. In this paper, a 3D nuclei counter using two different generative adversarial networks (GAN) is proposed and evaluated. Synthetic data that resembles real microscopy image is generated with a GAN and used to train another 3D GAN that counts the number of nuclei. Our approach is evaluated with respect to the number of groundtruth nuclei and compared with common ways of counting used in the biological research. Fluorescence microscopy 3D image volumes of rat kidneys are used to test our 3D nuclei counter. The accuracy results of proposed nuclei counter are compared with the ImageJ's 3D object counter (JACoP) and the 3D watershed. Both the counting accuracy and the object-based evaluation show that the proposed technique is successful for counting nuclei in 3D.

**Keywords:** nuclei counting , fluorescence microscopy, synthetic data generation, generative adversarial networks

## 1. INTRODUCTION

Quantitative information obtained from microscopy images is vital for biomedical research and clinical diagnosis. For example, analyzing the distribution of cells or extracellular structures can provide a better understanding of the physiological and pathological status of the tissue. Two-photon microscopy is a type of fluorescence microscopy that is favored for living tissue imaging. With the use of near-infrared excitation to increase penetration depth in tissue, large image volumes in 3D are produced. The large 3D image sizes and the high cost of manual processing require automated means to analyze quantitative biological information [1].

Many challenges exist for counting objects in 3D image volumes, which include crowding and touching of objects, overlapping of two or more objects, and variances in object shape and size. There are many real-life examples where counting is needed, for example, traffic surveillance, pedestrian counting, and cell density estimation.

Several semi-automatic techniques have been proposed for counting. In [2], a supervised machine learning counting method that can estimate the object count with manual annotation input was introduced. Another method that uses machine learning to count objects and estimates the density of the objects in images was described in [3]. The ImageJ toolset [4] has a 3D object counter, known as JACoP [5], that is a subcellular

---

This work was partially supported by a George M. O'Brien Award from the National Institutes of Health under grant NIH/NIDDK P30 DK079312 and the endowment of the Charles William Harrison Distinguished Professorship at Purdue University.

colocalization analysis tool that uses a statistical approach with a manually selected threshold to analyze intensity information to obtain the object count and location. In some counting approaches, segmentation is done prior to counting in order to distinguish the counting targets (foreground) from the background. The 3D nuclei segmentation presented in [6] classifies each voxel in the microscopy image volume as nuclei or non-nuclei using 3D active contours. Typically the nuclei count can be obtained from the segmentation results using 3D connected component analysis. However, multiple nuclei could be counted as one nucleus if they are close enough to be connected. Thus, using segmentation for counting is less accurate when nuclei are crowded. One way to address this problem is to use morphological operations [7]. Morphological watershed (2D or 3D) is one of the techniques commonly used for separating overlapping objects [8], [9]. In the watershed, the topographical distance transform is used to find the number of local minima in the image, which indicates the object count [10]. However, a drawback of the watershed is over-segmentation when more than one local minimum is found for each object, producing inaccurate counting results [11].

Deep learning has recently become a prominent approach to address the counting problem [12]. A way of counting that uses convolutional neural networks (CNN) to estimate the number of pedestrians in a video was presented in [13]. Similarly, a tumor cell counting CNN is trained to provide both cell count and the cell locations in [3]. In another example, cell counting using fully convolutional regression networks (FCRNs) was introduced [14]. The networks were trained to find the locations of the cells from which cell count was obtained. A combination of a CNN nuclei segmentation and 3D watershed was used for 3D segmentation and nuclei counting in [15].

Acquiring image groundtruth for the training of deep learning networks is labor intensive and in many applications difficult to obtain. Synthetic data has been used to train and test the networks [16]. An approach for generating synthetic microscopy images in 3D was presented in [17]. Recently, generative adversarial networks (GAN) [18] have been useful for generating realistic synthetic data. There are two different networks in a GAN, a generative network and a discriminative network. More specifically, the generative network is trained to generate sample images whereas the discriminative network is used to estimate the probability of a generated sample being a real image. Both networks are trained to minimize their loss functions simultaneously. A deep convolutional generative adversarial network (DCGAN) was described in [19] for unsupervised learning which uses a GAN together with a CNN for synthesizing images. Another example of using GANs to synthesize image is the cycle-consistent adversarial network (CycleGAN) [20]. However, the synthetic images generated by CycleGAN can be geometrically distorted which includes spatially shifting of objects [21]. More recently, 3D realistic synthetic microscopy volumes were generated by a spatially constrained CycleGAN (SpCycleGAN) and utilized with corresponding synthetic binary volumes to achieve 3D nuclei segmentation [22].

In this paper, we introduce a nuclei counting technique using two GANs, a SpCycleGAN and a 3D GAN for nuclei counting. The 3D GAN is trained with synthetic microscopy data generated by the SpCycleGAN and tested with both synthetic and real microscopy data. Our approach is evaluated with respect to the number of groundtruth nuclei and compared with common ways of counting used in the biological research. Fluorescence microscopy 3D image volumes of rat kidneys are used to test our 3D nuclei counter. The data contains fluorescent labeled (Hoechst 33342 stain) nuclei of kidney cells collected using two-photon microscopy.

## 2. PROPOSED METHOD

In this paper,  $I$  denotes a 3D image volume of size  $X \times Y \times Z$ . Figure 1(a) shows the block diagram of our proposed way of nuclei counting. There are two main steps in our nuclei counter: 3D synthetic data generation and 3D nuclei counting.  $I^{orig}$  denotes a subvolume of the original microscopy volume used for training, and  $I^{test}$  is a subvolume of original or synthetic microscopy volume that is used for testing. During the evaluation,  $I^{test}$  is used as an input to the 3D nuclei counting network to estimate the number of nuclei,  $N$ .  $I^{syn}$  denotes a synthetic microscopy volume generated with 3D synthetic data generation based on the features of  $I^{orig}$ .  $I^{dist}$  is a distance map volume that contains information regarding the locations of the nuclei of the microscopy image volumes. Here,  $I^{dist}$  is paired with  $I^{syn}$  to train the 3D nuclei counting network. Figure 1(b) and Figure 1(c) are the detailed block diagrams for 3D synthetic image generation and 3D nuclei counting, respectively.  $I^{bitr}$  and  $I^{bi}$  denote the synthetic binary volumes used for SpCycleGAN training and inference, respectively.  $I^{distg}$  is a

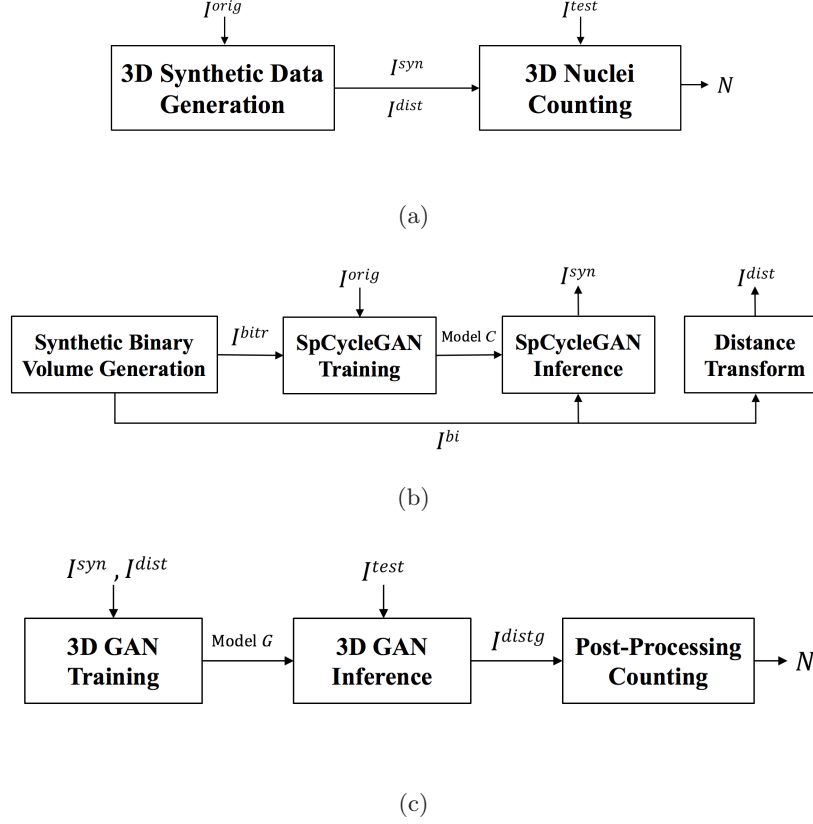


Figure 1. Block diagrams of (a) the proposed nuclei counter, (b) 3D synthetic data generation, (c) 3D nuclei counting

synthetic distance map volume generated by 3D nuclei counting network that contains nuclei location information of  $I^{test}$ .

As shown in Figure 1(b), synthetic binary volumes,  $I^{bitr}$ , are generated first.  $I^{bitr}$  is used together with  $I^{orig}$  to train a SpCycleGAN from which a generative model  $C$  is obtained. A different synthetic binary volume,  $I^{bi}$ , is then used to generate corresponding  $I^{syn}$  with the trained model  $C$ . For each  $I^{syn}$ , the corresponding  $I^{dist}$  is generated. In Figure 1(c), paired  $I^{syn}$  and  $I^{dist}$  are used to train a 3D GAN for nuclei counting and obtain a 3D GAN generative model  $G$ . Model  $G$  is used to transform  $I^{test}$  to its synthetic distance map volume,  $I^{distg}$ , for counting purposes. Then, post-processing steps which include thresholding and connected component analysis are used on  $I^{distg}$  to obtain the number of nuclei  $N$ .

## 2.1 3D Synthetic Data Generation

Three dimensional synthetic data generation includes synthetic binary volume generation, distance transformation of the synthetic binary volume, SpCycleGAN training, and SpCycleGAN inference. A synthetic binary volume is generated by drawing ellipsoidal shapes in 3D according to randomly selected locations of nuclei centroids and their orientations.

Following synthetic binary volume generation, the 3D Euclidean distance transform [23] is used to obtain the distance transformation of nuclei. We denote the distance transform as a function  $F_{DT} : I^{bi} \rightarrow I^{dist}$ . The  $I^{dist}$  volume is obtained by the distance transform of the binary synthetic nuclei volume  $I^{bi}$ .  $F_{DT}$  assigns a number to each voxel of the binary volume that is the distance from the voxel to its nearest background voxel. Considering an individual nucleus in  $I^{bi}$  and its corresponding  $I^{dist}$ , the voxels close to the center of the nucleus have high distance values and the voxels at the boundary have low distance values. Thus,  $I^{dist}$  can be considered as a distance map of the centroid locations of nuclei.

The SpCycleGAN described in our work [22] was an extension of the CycleGAN [20] by considering spatial loss during the training. The goal is to obtain a generative model  $C$  that can transfer a binary synthetic nuclei volume to a synthetic microscopy volume with accurate nuclei locations. During SpCycleGAN training, the objective function is to solve the minimax problem where the training loss ( $\mathcal{L}$ ) consists of the sum of an adversarial loss, a cycle consistency loss ( $\mathcal{L}_{cyc}$ ), and a spatial loss ( $\mathcal{L}_{spatial}$ ). The cycle consistent term and spatial constrained term are used in the training loss for regularization of the SpCycleGAN model. The training loss can be expressed as:

$$\mathcal{L} = \mathcal{L}_{GAN} + \lambda_1 \mathcal{L}_{cyc} + \lambda_2 \mathcal{L}_{spatial} \quad (1)$$

where  $\lambda_1$  and  $\lambda_2$  are the adjustable coefficients for  $\mathcal{L}_{cyc}$  and  $\mathcal{L}_{spatial}$ . To find the spatial loss, a separate generative model  $H$  is trained by an additional generative network. Model  $H$  can generate a binary mask from a synthetic volume generated by model  $C$ . More specifically, model  $H$  uses  $C(I^{bitr})$  as an input and generates  $H(C(I^{bitr}))$ . This generative model  $H$  minimizes  $\mathcal{L}_2$  loss between  $I^{bitr}$  and  $H(C(I^{bitr}))$ . The spatial loss is then formed as:

$$\mathcal{L}_{spatial}(C, H, I^{orig}, I^{bitr}) = \mathbb{E}_{I^{bitr}} [\|H(C(I^{bitr})) - I^{bitr}\|_2] \quad (2)$$

where  $\|\cdot\|_2$  denotes the  $L_2$  norm. Note that an unpaired set of training data,  $I^{orig}$  and  $I^{bitr}$ , is used for the training of model  $C$ .

During SpCycleGAN inference, the synthetic microscopy volume,  $I^{syn}$ , is generated from model  $C$  with the input as the synthetic binary volume,  $I^{bi}$ . Then,  $I^{bi}$  is used to obtain the distance map volume,  $I^{dist}$ , using distance transformation. Since the synthetic microscopy volume,  $I^{syn}$ , and the distance map volume,  $I^{dist}$ , are obtained from the same synthetic binary volume,  $I^{bi}$ ,  $I^{syn}$  and  $I^{dist}$  are now a paired set of volumes and can be used for 3D nuclei counting network training.

## 2.2 3D Nuclei Counting

Three dimensional counting consists of 3D GAN training, 3D GAN inference, and post-processing counting. As shown in Figure 1(c), we utilize a paired set  $I^{syn}$  and  $I^{dist}$  to train a 3D GAN and obtain a generative model  $G$ . Here, the groundtruth distance map volume  $I^{dist} = F_{DT}(I^{bi})$  provides information on the locations of the nuclei.

The 3D GAN generative model  $G$  is used to map nuclei in  $I^{test}$  to its synthetic distance map volume  $I^{distg}$ . This approach has the advantage over learning a direct mapping between original microscopy volume and nuclei count since the distance map volume preserves nuclei location information. The number of nuclei is then obtained from  $I^{distg}$  using thresholding and 3D connected component analysis. Thresholding of  $I^{distg}$  is done to keep voxels at or near the nuclei centroids and remove noises. By removing low distance transform voxels around nuclei boundaries, two or more connected nuclei can be identified as individual nuclei. Then, a 3D connected component analysis is done to assign different colors to individual nuclei at their centroid locations followed by obtaining  $N$ .

The architecture of the 3D GAN for nuclei counting is shown in Figure 2. For the generative network in Figure 2(a), a filter size of  $3 \times 3 \times 3$  is used for the 3D convolution. Each 3D convolution is followed by a batch normalization [24] and a rectified-linear units (ReLU) activation function. Similarly, for the discriminative network shown in Figure 2(b), a batch normalization and a Leaky ReLU activation function are used to follow each 3D convolution layer. Convolutions (stride 2) are used in replacement of pooling layers, while the number of features is doubled. Finally, a flatten layer, two dense layers, and a sigmoid activation function are used to produce the probability of classification  $P$ . This probability is used to discriminate groundtruth distance map volume  $I^{dist}$  from the generated synthetic distance map volume  $G(I^{syn})$ . The architecture of the 3D GAN for counting is designed to have a lightweight generative network for efficient inference. The discriminator is used to supervise the prediction of the 3D GAN for nuclei counting to have the features of a real distance map.

The generative model  $G$  learns to transform  $I^{syn}$  into  $I^{dist}$ , whereas the discriminative model  $D$  distinguishes between  $I^{dist}$  and  $G(I^{syn})$ . As shown in Equation 3, the training loss function of the 3D GAN for nuclei counting is defined as the sum of an adversarial loss  $\mathcal{L}_{GAN}$  (Equation 4) and a content loss  $\mathcal{L}_{MSE}$  (Equation 5). Here, we used the mean square error between the synthetically generated distance map volume  $G(I^{syn})$  and the

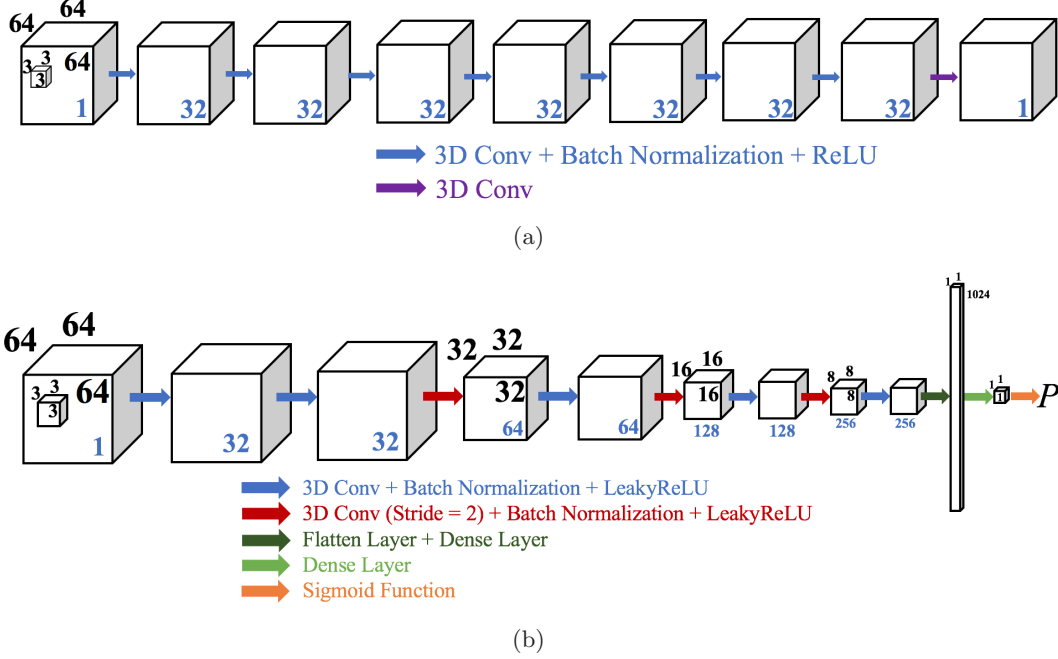


Figure 2. Architecture of 3D GAN for nuclei counting (a) generative network structure, (b) discriminative network structure

groundtruth distance map volume  $I^{dist}$  as the content loss.  $\lambda$  is the weight coefficient used to control the relative importance of the adversarial loss and the content loss.

$$\mathcal{L}(G, D, I^{syn}, I^{dist}) = \mathcal{L}_{GAN}(G, D, I^{syn}, I^{dist}) + \lambda \mathcal{L}_{MSE}(G, I^{syn}, I^{dist}) \quad (3)$$

$$\mathcal{L}_{GAN}(G, D, I^{syn}, I^{dist}) = \mathbb{E}_{I^{dist}}[\log(D(I^{dist}))] + \mathbb{E}_{I^{syn}}[\log(1 - D(G(I^{syn})))] \quad (4)$$

$$\mathcal{L}_{MSE}(G, I^{syn}, I^{dist}) = \mathbb{E}_{I^{dist}, I^{syn}}[\|G(I^{syn}) - I^{dist}\|_2] \quad (5)$$

Table 1. Experimental Settings

	Data-I		Data-II	
Network	SpCycleGAN	3D GAN counting	SpCycleGAN	3D GAN counting
Training volume size	$128 \times 128 \times 128$	$64 \times 64 \times 64$	$128 \times 128 \times 32$	$64 \times 64 \times 64$
Number of pairs of training volumes	1	80	4	80
Inference volume size	$128 \times 128 \times 128$	$128 \times 128 \times 64$	$128 \times 128 \times 128$	$128 \times 128 \times (64 \text{ or } 32)$
Number of inference volumes (total)	15	20	15	20
Model	$C^{Data-I}$	$G^{Data-I}$	$C^{Data-II}$	$G^{Data-II}$

### 3. EXPERIMENTAL RESULTS

The testing of our counting approach involves two different rat kidney data sets which are denoted as Data-I \* and Data-II. Data-I has size of  $X \times Y = 512 \times 512$  pixels with  $Z = 512$  (grayscale images), whereas Data-II consists of  $Z = 32$  with the same size in  $X$  and  $Y$  as Data-I. The experimental settings are listed in Table 1.

\*Data-I was provided by Malgorzata Kamocka of Indiana University and was collected at the Indiana Center for Biological Microscopy.

The size of the training and inference volumes were selected according to the original size of the volumes. The total training volume size of SpCycleGAN is  $128 \times 128 \times 128$  for both data sets. Then, 10 synthetic microscopy volumes of size  $128 \times 128 \times 128$  were generated and divided into 80 pairs of synthetic microscopy volumes of size  $64 \times 64 \times 64$ . These volumes were used together with their corresponded distance map volumes to train a 3D GAN for nuclei counting for each data set. We selected  $\lambda = 10$  in Equation 4 for both data sets. Our network architecture is implemented in TensorFlow [25] using Adam optimizer [26] with a learning rate of 0.005.

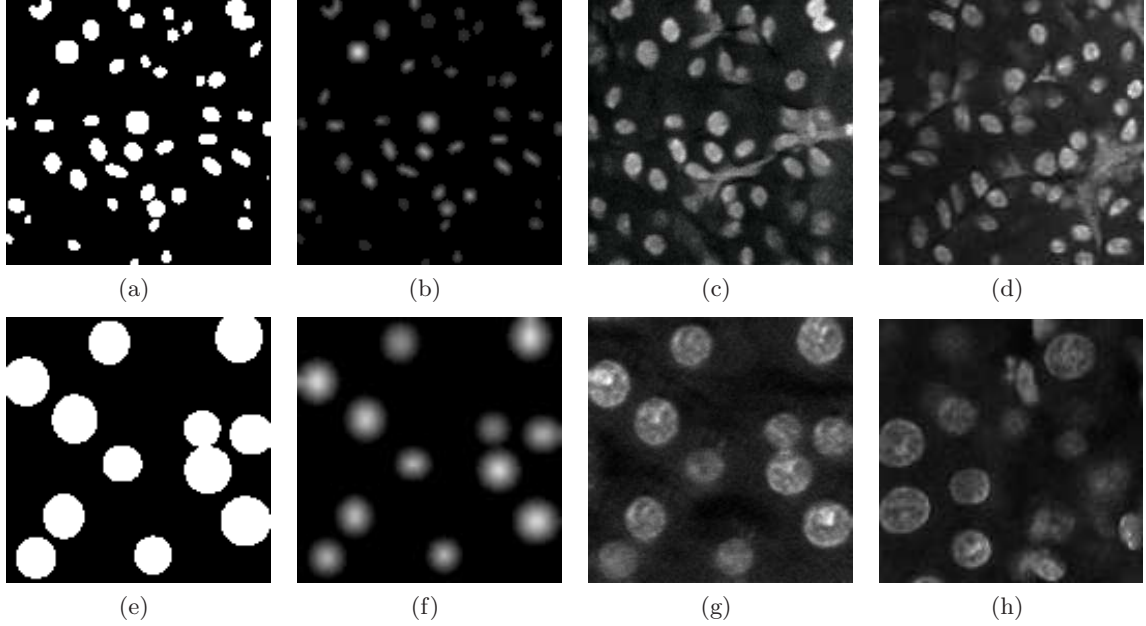


Figure 3. Comparison of slices of the synthetic binary volume, the distance map volume, the synthetic microscopy volume, and the original volume (a) Data-I synthetic binary image, (b) Data-I distance map image, (c) Data-I synthetic microscopy image, (d) Data-I original image, (e) Data-II synthetic binary image, (f) Data-II distance map image, (g) Data-II synthetic microscopy image, (h) Data-II original image

Figure 3 shows the synthetic binary images, their corresponding distance map images, the synthetic microscopy images, and the original microscopy images. The synthetic images are similar to the original microscopy images in terms of nuclei size and shape, nuclei distribution, intensity, and noise level. Using the trained model  $C$ , we can generate a paired set of training data for the 3D nuclei counting network. The trained model  $G$  is used to transform previously generated synthetic microscopy volumes into its corresponding synthetic distance map volumes, which contains nuclei centroid location information. Post-processing is done by first thresholding each voxel  $v$  with value  $v_i$  in  $I^{distg}$  as in Equation 6:

$$v_i = \begin{cases} 0 & \text{if } v_i \leq T \text{ or } 245 \leq v_i \\ v_i & \text{if } T \leq v_i \leq 245 \end{cases} \quad (6)$$

The threshold  $T$  is selected experimentally where  $T = 45$  for Data-I and  $T = 55$  for Data-II. Connected component analysis is then used to count and label each nucleus into a different color. For better visualization and evaluation on the object level, the labeled nuclei are dilated in 3D with a spherical mask to match the original nuclei sizes. In Figure 4, we show the post-processing steps of 3D nuclei counting with example slices from both synthetic and real microscopy image volumes.

The evaluation of microscopy image counting is difficult since groundtruth information is hard to obtain. We evaluate our results by comparing the results of nuclei counting  $N_i$  for the  $i$ th subvolume where  $i \in \{1, \dots, n\}$  with two other common ways of counting.  $N_i^{gt}$  is the groundtruth nuclei count for the  $i$ th subvolume where  $i \in \{1, \dots, n\}$ . Here,  $n$  stands for the number of subvolumes. The mean absolute percentage error (MAPE) is



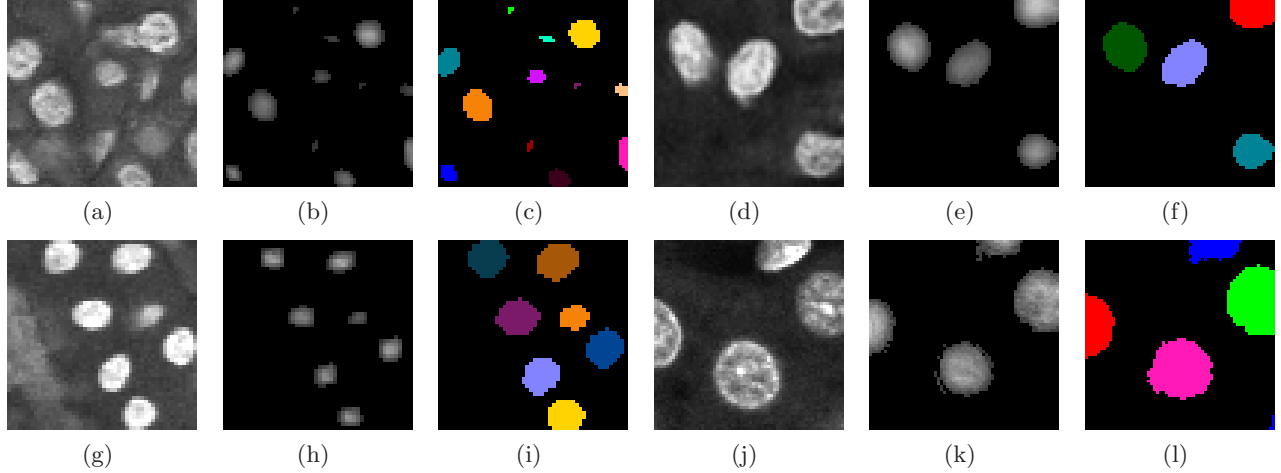


Figure 4. Example of slices of the testing volume, the synthetic distance map volume with thresholding, and the color-coded volume from 3D nuclei counting (a) Data-I synthetic image, (b) corresponding  $I^{distg}$  with thresholding for (a), (c) color-coded image for (a), (d) Data-II synthetic image, (e) corresponding  $I^{distg}$  with thresholding for (d), (f) color-coded image for (d), (g) Data-I real image, (h) corresponding  $I^{distg}$  with thresholding for (g), (i) color-coded image for (g), (j) Data-II real image, (k) corresponding  $I^{distg}$  with thresholding for (j), (l) color-coded image for (j)

then measured with Equation 7 and reported for each method.

$$MAPE = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{N_i - N_i^{gt}}{N_i^{gt}} \right| \quad (7)$$

The techniques compared with the proposed include the 3D watershed [10] and the ImageJ’s 3D object counter (JACoP) [5]. Both counting approaches are semi-automatic since a threshold needs to be selected before counting. We initially selected the thresholding parameters based on the Otsu’s method [27] and adjusted the parameters to best fit the volumes. Since the original microscopy volume suffers from a large variation of intensity and noise, we chose subvolumes from different locations of the original volume and processed each of them to obtain the nuclei count.

Table 2. Comparison of mean absolute percentage error (MAPE) of the proposed 3D nuclei counter with 3D watershed [10] and JACoP [5]

n=10	Synthetic Data-I	Synthetic Data-II	Microscopy Data-I	Microscopy Data-II
3D Watershed	14.69%	13.78%	19.76%	10.88%
JACoP	11.86%	16.10%	12.67%	18.72%
Proposed 3D Nuclei Counter	5.64%	2.13%	6.68%	6.53%

For evaluation, we generated  $n = 10$  synthetic microscopy image subvolumes for both data sets denoted as  $I^{synData-I}$  and  $I^{synData-II}$ , respectively.  $I^{synData-I}$  has a size of  $128 \times 128 \times 64$  with the average number of nuclei being 256.  $I^{synData-II}$  has a size of  $128 \times 128 \times 64$  with the average number of nuclei being 64. We processed these volumes with the ImageJ’s 3D object counter (JACoP), the 3D watershed, and the proposed 3D nuclei counter to compare the counting results. The quantitative evaluations for the subvolumes are shown in Table 2. In Table 2, we can see that the 3D watershed and the JACoP method produce less accurate counting result comparing to our proposed method. Based on our observation for the JACoP, nuclei are often missing around darker boundaries of the volume. Note that the JACoP has much lower detection accuracy for microscopy volume of large size due to inhomogeneity of the microscopy volume, so small subvolumes of microscopy image volume are used for counting. Our proposed nuclei counter achieves lower mean absolute percentage error in counting than the other two techniques.

We also selected  $n = 10$  subvolumes from original microscopy data for the evaluation purpose. More specifically, we selected 10 subvolumes from  $I^{Data-I}$  with size of  $128 \times 128 \times 64$  with the average number of nuclei

being 298 and 10 subvolumes of  $I^{Data-II}$  with size of  $128 \times 128 \times 32$  with the average number of nuclei being 25, respectively. Note that the evaluation size of  $I^{Data-II}$  is selected according to its total size  $512 \times 512 \times 32$ . From Table 2, the proposed 3D nuclei counter achieves the lowest error among all three different counting approaches. Figure 5 shows the comparison of the color-coded images obtained from the 3D watershed, the JACoP, and our proposed method. It is observed that the 3D watershed and the JACoP method cannot distinguish between nuclei and non-nuclei structures. Our proposed method achieves better performance in finding nuclei and separating touching nuclei. The color-coded images from the proposed method have less noise from the inhomogeneity of the nucleus compared with the results from the other two techniques.

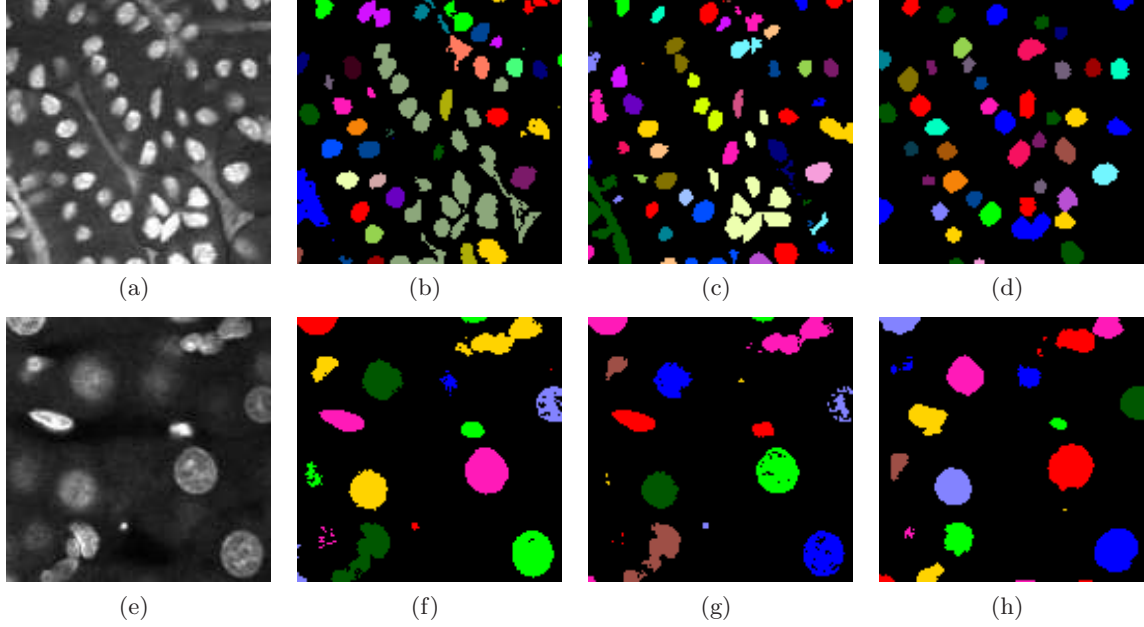


Figure 5. Comparison of slices of the original image volume and results obtained from 3D watershed, JACoP, and our proposed method (a) Data-I original image, (b) Data-I color-coded image from 3D watershed, (c) Data-I color-coded image from JACoP, (d) Data-I color-coded image from our method, (e) Data-II original image, (f) Data-II color-coded image from 3D watershed, (g) Data-II color-coded image from JACoP, (h) Data-II color-coded image from our method

Table 3. Object-based comparison of the proposed 3D nuclei counter with 3D watershed [10] and JACoP [5] for Data-I

Microscopy Data-I						
	$N_{TP}$	$N_{FP}$	$N_{FN}$	$P$	$R$	$F1$
3D Watershed	121	9	153	93.08%	44.16%	59.90%
JACoP	146	19	128	88.48%	53.28%	66.51%
Proposed 3D Nuclei Counter	212	18	56	92.17%	79.10%	85.14%

Table 4. Object-based comparison of the proposed 3D nuclei counter with 3D watershed [10] and JACoP [5] for Data-II

Microscopy Data-II						
	$N_{TP}$	$N_{FP}$	$N_{FN}$	$P$	$R$	$F1$
3D Watershed	171	13	50	92.93%	77.38%	84.44%
JACoP	151	16	62	90.42%	70.89%	79.47%
Proposed 3D Nuclei Counter	193	15	33	92.79%	85.40%	88.94%

To analyze the detection accuracy of the locations of the nuclei found with the counting methods, we use the object-based evaluation as described in [28]. The evaluation is based on manually annotated groundtruth of



two subvolumes from  $I^{Data-I}$  with size of  $128 \times 128 \times 64$  and from  $I^{Data-II}$  with size of  $512 \times 512 \times 16$ . The groundtruth is obtained using ITK-SNAP [29] where each nucleus is manually labeled individually. If a nucleus overlaps equal or more than 50% with its corresponding groundtruth nucleus, it is counted as a true-positive,  $N_{TP}$ . If a nucleus overlaps less than 50% with its corresponding groundtruth nucleus or there is no corresponding groundtruth, it is counted as a false-positive,  $N_{FP}$ . A false negative,  $N_{FN}$ , is defined as when a nucleus is present in the groundtruth but no corresponding nucleus is found by the counting method. Then, the F1 score (F1), Precision (P) and Recall (R) described in [30] can be calculated as:

$$P = \frac{N_{TP}}{N_{TP} + N_{FP}}, R = \frac{N_{TP}}{N_{TP} + N_{FN}}, F1 = \frac{2PR}{P + R} \quad (8)$$

For fair comparisons, small objects removal is done to remove objects with sizes less than 5% of the average nuclei sizes in voxels. In Table 3 and Table 4, our proposed 3D nuclei counter achieves higher F1 score than the compared methods, which shows that the proposed method can successfully find nuclei at the correct locations. Both the 3D watershed and the JACoP method produce large numbers of  $N_{FN}$ , causing low R and F1 score. Overall, our proposed method outperforms the other two methods in both the MAPE evaluation and the object-based evaluation.

## 4. CONCLUSION

In this paper, we described an approach for counting nuclei in 3D microscopy image volumes efficiently and accurately using synthetic training data. We first generated synthetic microscopy image volumes using a SpCycleGAN. This helps to solve data insufficiency and the lacking of groundtruth annotations for the analysis of microscopy images. We then developed a 3D GAN for nuclei counting that can predict the distance map volume for an input microscopy image volume with an arbitrary number of nuclei. The 3D GAN trained entirely on synthetic data was able to give predictions of nuclei count for microscopy images without fine-tuning. It is able to do fast inference and count the number of nuclei in microscopy image volumes. Our 3D nuclei counter shows the potential of counting and locating nuclei without segmentation and classification of each voxel within the image volume. This greatly reduced the burden of computation and directly provide quantitative information to biologists for further analysis of the images. We noticed that nuclei that are on the boundary of the volume are hard to detect, limiting the detection accuracy. In the future, we plan to improve our model for detecting nuclei more accurately on the boundary.

## REFERENCES

- [1] K. W. Dunn and P. A. Young, "Principles of multiphoton microscopy," *Nephron Experimental Nephrology*, vol. 103, no. 2, pp. e33–e40, March 2006.
- [2] V. Lempitsky and A. Zisserman, "Learning to count objects in images," *Proceedings of the Advances in Neural Information Processing Systems*, pp. 1324–1332, December 2010, Whistler, Canada.
- [3] Y. Xue, N. Ray, J. Hugh, and G. Bigras, "Cell counting by regression using convolutional neural network," *Proceedings of the European Conference on Computer Vision*, pp. 274–290, October 2016, Amsterdam, Netherlands.
- [4] C. T. Rueden, J. Schindelin, M. C. Hiner, B. E. DeZonia, A. E. Walter, E. T. Arena, and K. W. Eliceiri, "ImageJ2: ImageJ for the next generation of scientific image data," *BMC Bioinformatics*, vol. 18, no. 1, pp. 529–1–26, November 2017.
- [5] S. Bolte and F. P. Cordelieres, "A guided tour into subcellular colocalization analysis in light microscopy," *Journal of Microscopy*, vol. 224, no. 3, pp. 213–232, December 2006.
- [6] S. Lee, P. Salama, K. W. Dunn, and E. J. Delp, "Segmentation of fluorescence microscopy images using three dimensional active contours with inhomogeneity correction," *Proceedings of the IEEE International Symposium on Biomedical Imaging*, pp. 709–713, April 2017, Melbourne, Australia.
- [7] H. Tulsani, S. Saxena, and N. Yadav, "Segmentation using morphological watershed transformation for counting blood cells," *International Journal of Computer Applications and Information Technology*, vol. 2, no. 3, pp. 28–36, January 2013.

- [8] G. Lin, U. Adiga, K. Olson, J. F. Guzowski, C. A. Barnes, and B. Roysam, "A hybrid 3D watershed algorithm incorporating gradient cues and object models for automatic segmentation of nuclei in confocal image stacks," *Cytometry Part A*, vol. 56, no. 1, pp. 23–36, November 2003.
- [9] M. S. Rahman and M. R. Islam, "Counting objects in an image by marker controlled watershed segmentation and thresholding," *Proceedings of International Advance Computing Conference*, pp. 1251–1256, May 2013, Ghaziabad, India.
- [10] J. B. Roerdink and A. Meijster, "The watershed transform: Definitions, algorithms and parallelization strategies," *Fundamenta Informaticae*, vol. 41, no. 1, 2, pp. 187–228, April 2000.
- [11] A. El Allaoui and M. Nasri, "Medical image segmentation by marker-controlled watershed and mathematical morphology," *The International Journal of Multimedia & Its Applications*, vol. 4, no. 3, pp. 1–9, June 2012.
- [12] E. Walach and L. Wolf, "Learning to count with CNN boosting," *Proceedings of the European Conference on Computer Vision*, pp. 660–676, October 2016, Amsterdam, Netherlands.
- [13] C. Zhang, H. Li, X. Wang, and X. Yang, "Cross-scene crowd counting via deep convolutional neural networks," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 833–841, June 2015, Boston, MA.
- [14] W. Xie, J. A. Noble, and A. Zisserman, "Microscopy cell counting with fully convolutional regression networks," *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, vol. 6, no. 3, pp. 283–292, 2018.
- [15] C. Fu, D. J. Ho, S. Han, P. Salama, K. W. Dunn, and E. J. Delp, "Nuclei segmentation of fluorescence microscopy images using convolutional neural networks," *Proceedings of the IEEE International Symposium on Biomedical Imaging*, pp. 704–708, April 2017, Melbourne, Australia.
- [16] X. Zhang, Y. Fu, S. Jiang, L. Sigal, and G. Agam, "Learning from synthetic data using a stacked multichannel autoencoder," *Proceedings of the IEEE International Conference on Machine Learning and Applications*, pp. 461–464, December 2015, Miami, FL.
- [17] D. J. Ho, C. Fu, P. Salama, K. W. Dunn, and E. J. Delp, "Nuclei segmentation of fluorescence microscopy images using three dimensional convolutional neural networks," *Proceedings of the IEEE Conferences on Computer Vision and Pattern Recognition Workshop*, pp. 834–842, July 2017, Honolulu, HI.
- [18] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," *Proceedings of the Advances in Neural Information Processing Systems*, pp. 2672–2680, December 2014, Montreal, Canada.
- [19] A. Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," *Proceedings of the International Conference on Learning Representations*, pp. 1–16, May 2016, San Juan, Puerto Rico.
- [20] J. Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," *Proceedings of the IEEE International Conference on Computer Vision*, pp. 2242–2251, October 2017, Venice, Italy.
- [21] Z. Zhang, L. Yang, and Y. Zheng, "Translating and segmenting multimodal medical volumes with cycle-and shapeconsistency generative adversarial network," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 9242–9251, June 2018, Salt Lake City, UT.
- [22] C. Fu, S. Lee, D. J. Ho, S. Han, P. Salama, K. W. Dunn, and E. J. Delp, "Three dimensional fluorescence microscopy image synthesis and segmentation," *Proceedings of the IEEE Conferences on Computer Vision and Pattern Recognition Workshop*, pp. 2302–2310, June 2018, Salt Lake City, UT.
- [23] C. R. Maurer, R. Qi, and V. Raghavan, "A linear time algorithm for computing exact Euclidean distance transforms of binary images in arbitrary dimensions," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 2, pp. 265–270, February 2003.
- [24] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," *Proceedings of the International Conference on Machine Learning*, vol. 37, pp. 448–456, July 2015, Lille, France.
- [25] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, and M. Isard, "TensorFlow: A system for large-scale machine learning," *Proceedings of the USENIX Symposium on Operating Systems Design and Implementation*, vol. 16, pp. 265–283, November 2016, Savannah, GA.

- [26] D. P. Kingma and J. L. Ba, “Adam: A method for stochastic optimization,” *arXiv preprint arXiv:1412.6980*, pp. 1–15, December 2014.
- [27] N. Otsu, “A threshold selection method from gray-level histograms,” *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62–66, January 1979.
- [28] S. Lee, C. Fu, P. Salama, K. W. Dunn, and E. J. Delp, “Tubule segmentation of fluorescence microscopy images based on convolutional neural networks with inhomogeneity correction,” *Proceedings of the IS&T International Symposium on Electronic Imaging*, January 2018, Burlingame, CA.
- [29] P. A. Yushkevich, J. Piven, H. C. Hazlett, R. G. Smith, S. Ho, J. C. Gee, and G. Gerig, “User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability,” *NeuroImage*, vol. 31, no. 3, pp. 1116–1128, July 2006.
- [30] D. M. W. Powers, “Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation,” *Journal of Machine Learning Technologies*, vol. 2, no. 1, pp. 37–63, December 2011.