SEGMENTATION AND DECONVOLUTION OF FLUORESCENCE MICROSCOPY VOLUMES

A Dissertation

Submitted to the Faculty

of

Purdue University

by

Soonam Lee

In Partial Fulfillment of the

Requirements for the Degree

of

Doctor of Philosophy

August 2019

Purdue University

West Lafayette, Indiana

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF DISSERTATION APPROVAL

Dr. Edward J. Delp, Co-Chair

School of Electrical and Computer Engineering

Dr. Paul Salama, Co-Chair

School of Electrical and Computer Engineering, IUPUI

Dr. Mary L. Comer

School of Electrical and Computer Engineering

Dr. Fengqing Zhu

School of Electrical and Computer Engineering

Approved by:

Dr. Dimitri Peroulis

Head of the Graduate Program

To my parents, Jang Won Lee and Bokyeon Hwang,
for their unconditional support, inspiration, and encouragement;
To my wife, So Yeon Kim,
for her endless trust and love.

ACKNOWLEDGMENTS

Upon completion of this thesis, there are many people who have earned my gratitude for their inspiration, contribution, and support to my time in Purdue. Without their support and advice, this thesis would not have been possible.

First and foremost, I would like to express my sincere appreciation to my advisor Professor Edward J. Delp. I am grateful to him for the opportunity and privilege of becoming a member of the Video and Image Processing Laboratory (VIPER). He has offered many opportunities regarding the research and development and I have learned a lot while I accomplished these tasks. He has also guided me to learn how to deal with the real problems and how to deliver the realistic solution efficiently. His practical, problem-oriented, and straightforward approaches helped me to become a better researcher and engineer. I feel proud and privileged to have worked with him.

I would like to thank to my co-advisor Professor Paul Salama. I appreciate his engagement with all details of my research project. He has contributed for our weekly research meeting, proposed thinking from different perspectives that I did not come up with, and performed paper revisions using countless hours. I could still recall my first year in VIPER lab that he called me and gave me valuable advice how to conduct research and deliver the results efficiently. I would also express my gratitude towards my other committee members, Professor Mary Comer and Professor Fengqing Zhu, for their valuable feedback.

I wish to thank Professor Kenneth W. Dunn for his dedication of projects and providing numerous 3D microscopy datasets. His contributions from a medical and biological perspective are highly appreciated and made this work truly valuable.

I would like to thank Dr. Neeraj Gadgil, Dr. David Joon Ho, Dr. Chichen Fu, and Shuo Han as a team of “microscopy boys and girl(s)”. We have often been tough time and stay until late night to address problems. Without teamwork and collaborations,

microscopy project may not go well like nowadays. In particular, I would like to thank Dr. Chichen Fu for bringing up deep learning techniques to our projects and his contribution of DINAVID system. I would also like to thank Dr. David Joon Ho for a good quality of deep learning work and his hard-working for given many tasks. I would also like to appreciate Shuo Han for her dedication of DINAVID system, working for paper together as a co-author and developing visualization together.

I would like to extend my appreciation to all other great former and current colleagues in the VIPER laboratory: Dr. Ye He, Dr. Chang Xu, Dr. Albert Parra Pozo, Dr. Bin Zhao, Dr. Khalid Tahboub, Dr. Joonsoo Kim, Dr. Yu Wang, Dr. Shaobo Fang, Dr. Javier Ribera, Jeehyun Choe, Dahjung Chung, Yuhao Chen, Qingshuang Chen, Di Chen, David Guera, Daniel Mas, Ruiting Shao, Sri Kalyan Yarlagadda, Blanca Delgado, He Li, Jiaju Yue, Chang Liu, Sriram Baireddy, Janos Horvath, Enyu Cai, Jiaqi Guo, Hanxiang Hao, Jiangpeng He, Han Hu, Changye Yang, Yifan Zhao, Emily Bartusiak, and Alain Chen. I would also like to thank the visiting students in the VIPER laboratory: Thitiporn (Bee) Pramoun and Kharittha (Poy) Thongkor.

I want to show my gratitude to Purdue ECE department and Professor Michael Melloch who supported me as a teaching assistant for a long time. Without their support, I could not finish my Ph.D study. As a lab instructor and teaching assistant in the ECE classes, I learned how to teach students and deal with their difficulties during the semester. In addition, I also want to thank the Engineering Projects In Community Service (EPICS) for offering me the position of a graduate teaching assistant. It was unique experience working with a group of people who has diverse technical expertise to fulfill the local community partners needs.

I wish to show my gratitude to all the members in "Purdue G6 group": Dr. Chung Hwan Kim, Dr. Kyubyung Kang, Dr. Jihwan Oh, Dr. Joonsoo Kim, and Dr. Haejun Chung for sharing great memory in Purdue since 2012 Fall. I am sure we will be a "big shot" in each field as a researcher, scientist, or engineer some day and proud of being Purdue together.

I would also like to show my gratitude to Purdue Korean Tennis Club (PKTC). I am really fortunate to meet good friends in Purdue and I would never forget the time we enjoyed. My another special thanks go to Purdue Electrical Engineering Korean Association (PEEKA), Hanyang University alumni association (Hanyang@Purdue), and whomever I met in Purdue but not belong to the groups mentioned earlier. I thank all my friends who gave me the strength and motivation whenever I was discouraged.

Finally, I would like to express my deep gratitude to my family. This thesis is made by their unconditioned support, love, and trust. I wish to show my appreciation to my father, Jang Won Lee, my mother, Bokyeon Hwang, and my sister, Sookyeong Lee for their unlimited support and engagement. I also want to thank my father-in-law, Dong Hun Kim, my mother-in-law, Soonjae Lee, and my sister-in law, Haeyeon Kim, for their belief in me. Without their unconditioned support and belief, I would never been able to complete my Ph.D. Last but not least, I would like to thank my wife, So Yeon Kim for her trust and love. I truly appreciate her love and cherish every single moment with her from the beginning we met in Purdue.

This work was partially supported by a George M. OBrien 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.

TABLE OF CONTENTS

Page

[1]The work demonstrated in this section was jointly done with Shuo Han of Purdue University.

Page

[2]The work demonstrated in this section was jointly done with Shuo Han and Chichen Fu of Purdue University.
[3]The work demonstrated in this section was jointly done with Shuo Han of Purdue University.
[4]The work described in this section was jointly done with David Ho of Purdue University.

LIST OF TABLES

LIST OF FIGURES

Figure Page

Figure Page

ABSTRACT

Lee, Soonam Ph.D., Purdue University, August 2019. Segmentation and Deconvolution of Fluorescence Microscopy Volumes. Major Professor: Edward J. Delp and Paul Salama.

Recent advances in optical microscopy have enabled biologists collect fluorescence microscopy volumes cellular and subcellular structures of living tissue. This results in collecting large datasets of microscopy volume and needs image processing aided automated quantification method. To quantify biological structures a first and fundamental step is segmentation. Yet, the quantitative analysis of the microscopy volume is hampered by light diffraction, distortion created by lens aberrations in different directions, complex variation of biological structures. This thesis describes several proposed segmentation methods to identify various biological structures such as nuclei or tubules observed in fluorescence microscopy volumes. To achieve nuclei segmentation, multiscale edge detection method and 3D active contours with inhomogeneity correction method are used for segmenting nuclei. Our proposed 3D active contours with inhomogeneity correction method utilizes 3D microscopy volume information while addressing intensity inhomogeneity across vertical and horizontal directions. To achieve tubules segmentation, ellipse model fitting to tubule boundary method and convolutional neural networks with inhomogeneity correction method are performed. More specifically, ellipse fitting method utilizes a combination of adaptive and global thresholding, potentials, z direction refinement, branch pruning, end point matching, and boundary fitting steps to delineate tubular objects. Also, the deep learning based method combines intensity inhomogeneity correction, data augmentation, followed by convolutional neural networks architecture. Moreover, this thesis demonstrates a new deconvolution method to improve microscopy image quality without knowing the 3D

point spread function using a spatially constrained cycle-consistent adversarial networks. The results of proposed methods are visually and numerically compared with other methods. Experimental results demonstrate that our proposed methods achieve better performance than other methods for nuclei/tubules segmentation as well as deconvolution.

1. INTRODUCTION

1.1 Background in Optical Microscopy

Traditional microscopy is considered as a tool for observing objects and areas of objects that cannot be seen with human naked eyes. However, due to recent advances in microscopy, optical microscopy also known as light microscopy is one of the most suitable modality to observe biological entities and study cell biology in cellular and subcellular structures from live specimens also known as *in vivo*. Microscopy has many types but optical microscopy and electron microscopy are the most popularly used for research purposes. Optical microscope uses visible light to produce a magnified image of an object that is projected onto the retina of human eye or onto the photosensitive surface of an imaging device whereas electron microscopes utilizes electrons and electromagnetic waves. Since the wavelength (λ) of an electron much shorter than the visible light photons, the electron contains higher energy than visible light photon by the Planck−Einstein relation shown in Eq (1.1):

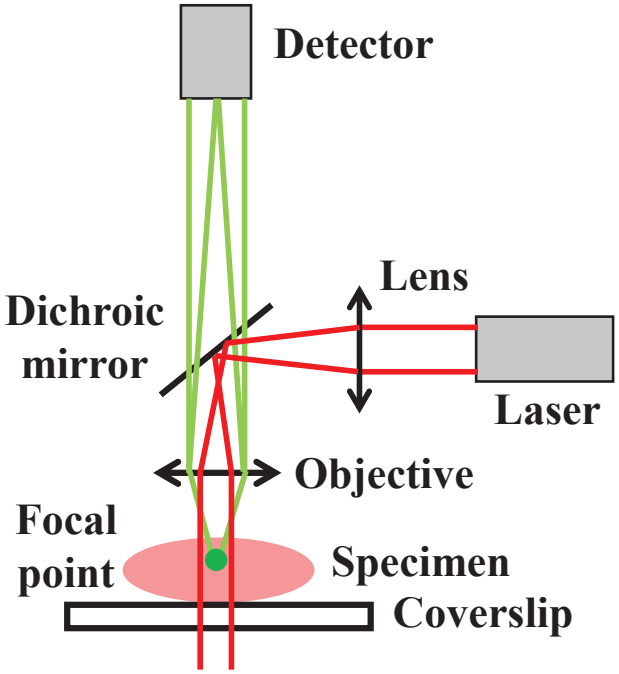
$$E = h\nu = h\frac{c}{\lambda} \tag{1.1}$$

where E is the energy of a photon, ν is a frequency, h is the Planck's constant ($6.626 \times 10^{-34} J \cdot s$), and c is a speed of light ($3 \times 10^{8} m/sec$). Since h and c are constant, E is inversely proportional to the wavelength of λ. Electron which has a shorter wavelength than visible light photon holds more energy so that electron microscope has much higher resolving power than an optical microscope. That is why electron microscope could be used to observe much smaller objects between $1nm$ and $10\mu m$ scale. However, owing to very short wavelength, electron microscope cannot be used for observing living cells from alive animals since this strong wavelength may damage the specimen's biological structures. Unless biologists need to observe a

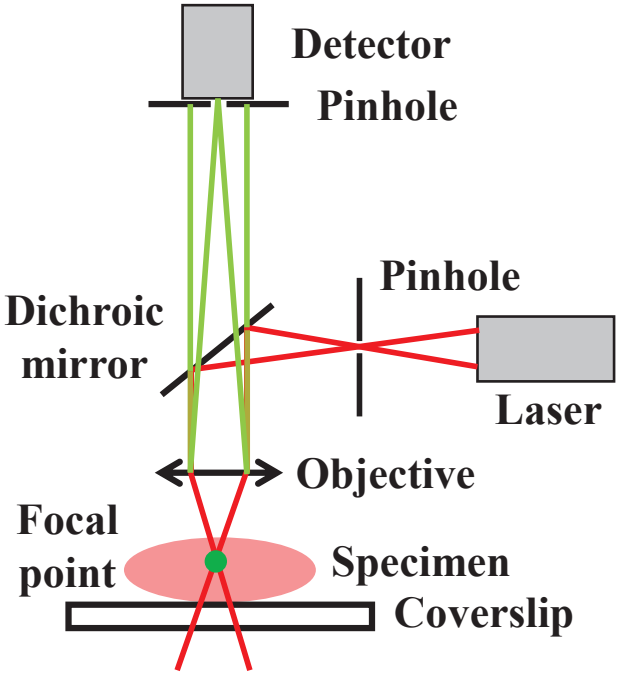
bacteria ($1\mu m$) or a virus ($100nm$) size of entities, the light microscope is the mainly used tool for biological structures observation and study.

To understand how light microscopy works we must know how this optical microscope was constructed. The objective and the condenser are regarded as two major components in forming the image from samples. The objective which located on top of the stage where a living specimen laid collects light diffracted by the specimen and forms a magnified real image. The condenser which located below of the stage focuses light from the illuminator onto a small area of the specimen [1]. In addition at least two lens are included in microscope at the objective and at the ocular in order to magnify the samples. Since many lenses are used for compound microscopy, illumination is a critical determinant of performance in light microscopy. Fundamental illumination technique was established in the 19^{th} century called the Koehler illumination in [2]. The main idea of the Koehler illumination is locating lens in front of detectors so that light rays that are passed focus in lens go parallel to each other. Due to these parallel light rays the Koehler illumination provides bright and even illumination in the specimen plane. Therefore, utilizing this method of illumination highly improves image quality.

Later, in the mid-20^{th} century, confocal microscope was invented [3] which added pinhole placed in front of electron detectors. After this invention conventional microscope without pinhole in front of detectors is called widefield microscope to distinguish between two different microscopy methods. Since the widefield microscopy follows the idea from the Koehler illumination, illumination of the specimen plane is brighter and even. One drawback of this widefield microscopy technique is that the images acquired from widefield microscopy are blurred by emission originating away from the focal plane. More specifically, thick fluorescent specimens may have challenges for collecting images using conventional widefield microscopy since bright signals from object lying outside the focal plane increases the background and yield low contrast images. Therefore, the widefield microscopy images are generally suffered from blur and out-of-focus which results in needs of deconvolution technique [4].



(a) Widefield microscope



(b) Confocal microscope

Fig. 1.1. The principle of widefield microscope and confocal microscope

As opposed to widefield microscope confocal microscope located a pinhole in front of detectors so that it successfully rejects signals from nearby sources above and below the focal plane. The principle of the widefield microscope and confocal microscope are shown in Figure 1.1.

First of all, the light from the laser that highlighted in red passes through the lens and reaches the dichroic (or dichromatic) mirror as shown in Figure 1.1a. Secondly, the laser is reflective from dichroic mirror and passes objective and arrives at live specimen. Then, the reflective light that highlighted in green from the section at the focal plane of specimen passes through the dichroic mirror and go to the detector. The detector catches the photons from the green light and form images. Note that the dichroic mirror is one of the example of nonlinear mirror which uses second harmonic generation [5] so that the dichroic mirror serves as a beamsplitter. More precisely, the dichroic mirror transmits rays that have certain angle range as well as reflects rays with different angle range [1]. This, in turn, can be used to split two different frequency light beams: One from the laser and another from the reflection of specimen at the focal plane. As demonstrated above confocal microscope adds pinhole right front of the detector to reject the out-of-focus rays. Also, another pinhole is located in front of the laser to gather the lights and make them to focus on focal point (or plane) of specimen. The background signal is reduced by these pinholes so that the signal-to-noise ratio (SNR) is improved. Thus, contrast and definition of the images obtained from the microscope are also improved [1].

Meanwhile, conventional light microscope based on specimen dependent properties of light absorption, optical path differences, phase gradients, and birefringence was combined with fluorescence microscope which uses fluorescence. With light microscope optics adjusted for fluorescence microscopy, it becomes possible to examine the distribution of a single molecular species in a living animal including individual fluorescent molecules. Typically, a specimen is first injected with fluorescent molecules also known as fluorophores and then it is tagged with a fluorescent dye or fluorochrome in order to be visible. The specimen is illuminated with light which is absorbed by

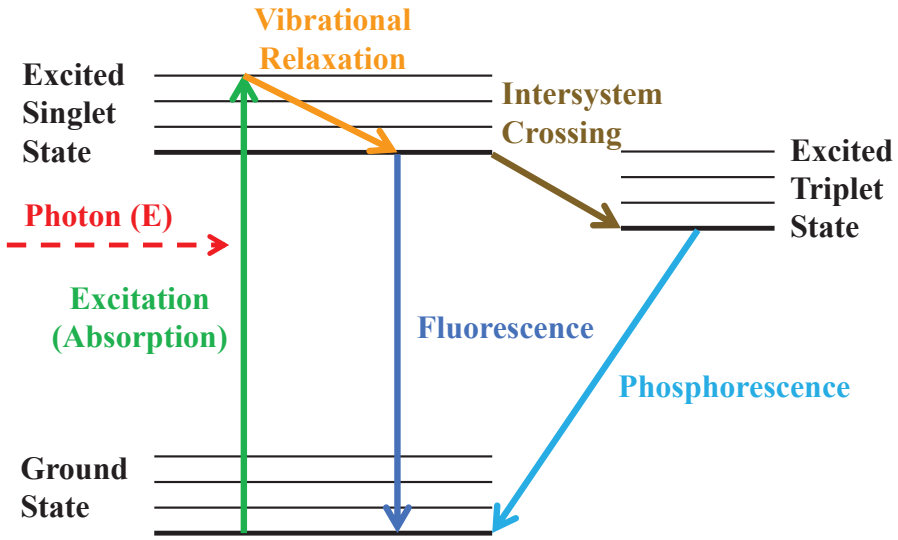


Fig. 1.2. The Jablonski diagram of the conventional microscope

and excites the fluorophores. This phenomenon is illustrated by Jablonski diagram, a diagram that explained electronic states of a molecule and the transitions among energy states [6]. From this illumination process the fluorophores emit photons so that fluorophores hold a longer wavelength than the absorbed light by the law of conservation of energy.

The Jablonski diagram of the conventional microscope including widefield and confocal microscope is displayed in Figure 1.2. Note that each horizontal line represent the energy levels occupied by an photon and upper line means higher energy level hold by an photon. There are three different states, for example, ground state, excitation singlet state, and excited triplet state. First of all, molecules absorb photons from energy such as rays from the laser and the photons inside of the molecules are excited. It results in state changes of the photons from ground state which is the lowest energy state to excited singlet state. This phenomenon is called absorption and exhibited in green up arrow shown in Figure 1.2. This excitation happens instantaneously such as in femtoseconds (10^{-15} seconds) [7]. Then, the high energy level photons return to the relatively low level in excited singlet state satisfying the Maxwell-Boltzmann

distribution. This process is called the vibrational relaxation exhibited in yellow diagonal down arrow in Figure 1.2 [8]. The vibrational relaxation from excited singlet state to the lower level is much slower than excitation and occurs in picoseconds (10^{-12} seconds) [7]. After the vibrational relaxation there are two ways to go back to the ground state. One way is molecules collapse to the ground state directly from the excited singlet state. Then, the energy is released as a fluorescence emission which occurs in the relatively longer time period in nanoscale (10^{-9} seconds) [7] highlighted in blue down arrow in Figure 1.2. The other way is molecules are not collapsing in a few nanoseconds but staying another excited state (excited triplet state) by intersystem crossing and then collapse to the ground state from the excited triplet state. In this emission pathway, the energy is released as a phosphorescence emission instead of fluorescence in milliseconds (10^{-3} seconds) exhibited in light blue arrow in Figure 1.2. Different from fluorescence, the phosphorescent material does not instantaneously emit after photon got excited [7].

The fluorescence microscopy shows the distribution of molecules in cells and tissues at great specificity and sensitivity. Yet, to collect 3D volume images of subcellular structures from a living specimen, fluorescence microscopy has to overcome several problems. First of all, utilizing short wavelength (bluish wavelength) for exciting enough photons to make subcellular structures visible often damages living tissues [9]. Secondly, due to strongly scattered visible wavelength, resulting in weak excitation of fluorophores, image collected from fluorescence microscopy is faint and blurred lacking in object details [10].

These are overcome through the two-photon excitation microscopy firstly introduced in [11]. The fundamental concept of two-photon microscopy is utilizing two photons to excite molecules by the simultaneous absorption of two-photons. More precisely, the fluorophore absorbs this combined energy by two photons, resulting in the same as traditional fluorescence microscopy. The Jablonski diagram of the two-photon microscope is shown in Figure 1.3. As observed in Figure 1.3, the Jablonski diagram is almost same as previous explained conventional microscope except

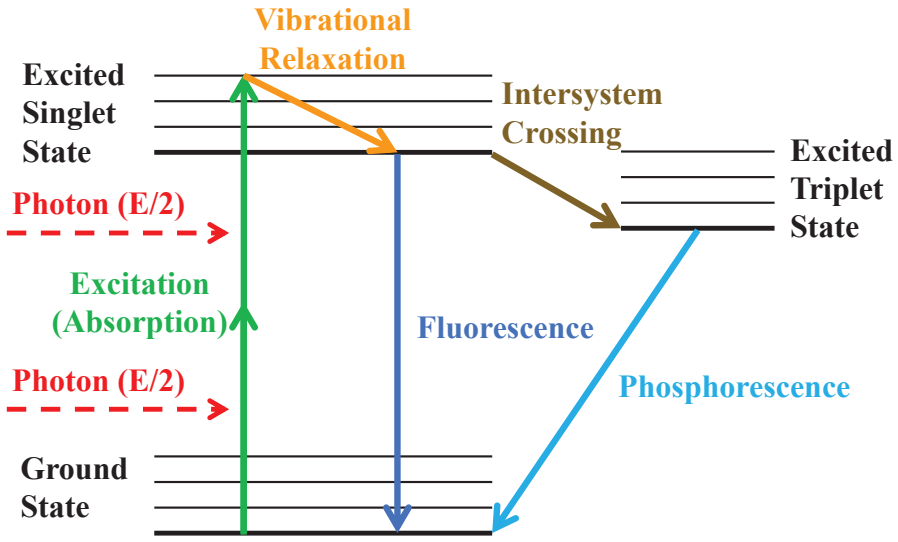


Fig. 1.3. The Jablonski diagram of the two-photon microscope

from utilizing two photons in the process of excitation. Note that each photon in two-photon microscope needs only half of the energy to excite molecules than conventional microscope. By Eq (1.1), the illumination and excitation wavelength is longer than the emission wavelength. Notify that longer wavelength contains lower energy which enables the imaging with less damage to living tissues [12].

Compared to confocal microscopy two-photon microscopy has several advantages specifically for deep tissue imaging. As indicated confocal microscopy places pinhole aperture in front of detector to reject unfocused scattered light rays so that signal strength is decreased [13]. In contrast, two-photon microscopy uses near infrared rays which are able to image the subcellular structures without damaging *in vivo* [14] since each molecule needs photon with less energy to be excited. Owing to preserving signal strength deep images collection within intact living tissues become plausible. Moreover, utilizing near infrared rays reduce the scattering effect so that fluorescence emission from the focal plane is efficiently collected on a wide area of detector . More specifically, the amount of scattering is inversely proportional to the fourth power of the wavelength in case that the scattering particle is much smaller compared to λ, for

example, smaller than 1/10 λ [13]. Then, the relationship between wavelength and the amount of scattering can be explained with the Rayleigh scattering as [15]:

$$\sigma_s = \frac{2\pi^5}{3}\frac{d^6}{\lambda^4}\left(\frac{n^2-1}{n^2+2}\right)^2 \tag{1.2}$$

where σ_s is the amount of scattering, d is one of the nearest spheres of diameter for the light scattering, and n is a refractive index. The Rayleigh scattering supports the fact that the longer wavelength has less scattered effect compared to the shorter wavelength as observed in Eq (1.2). Figure 1.4 portrays the relationship between λ and σ_s for the case of the visible light ($400nm - 700nm$) and infrared light ($700 - 1050nm$).

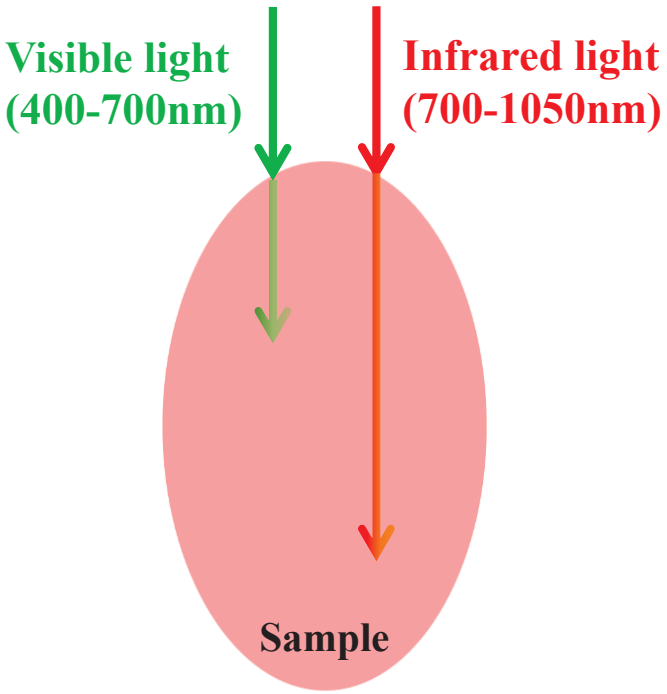


Fig. 1.4. The relationship between wavelength (λ) and the amount of scattering (σ_s)

Due to these advantages the two-photon excitation microscopy could offer high-resolution imaging of thick living samples as deep as $1mm$ [16]. Compared to the

two-photon microscopy the conventional microscopy technique only can image much shallower depth. In particular, the useful imaging depth in uncleared tissue sections for widefield microscopy is usually less than $10\mu m$ whereas imaging depth for the confocal microscopy is no more than $60\mu m$ [1, 17].

More recently, the two-photon microscopy technique is further improved using more than two photons called multiphoton microscopy [14, 18, 19]. Utilizing more photons provides lower probability of damaging living tissues because longer wavelength can be used compared to two-photon microscopy. For example, an infrared beam at $1050nm$ can produce three-photon excitation of a equivalent fluorophore absorbing ultraviolet light at $350nm$ while the same illumination could be done by two-photon utilizing green fluorophore at $525nm$. Moreover, three-photon excitation can be employed of useful imaging into deep ultraviolet. As an example, near infrared light at $720nm$ can be superseded to excite a fluorophore that normally required deep ultraviolet light of $240nm$. One difficulty of multiphoton microscopy is that multiphotons must interact simultaneously with the fluorophore to produce emission.

1.2 Notation

In order to describe our various methods efficiently, general notation for the microscopy volume analysis is introduced and used for this thesis. Without loss of generality, florescence microscopy datasets can be expressed in 5D as width (x), height (y), depth (z), time (t), and color channel (c). We use subscripts for the indices of the 5D volume and superscripts for the detail information of the 5D volume. We denote I_{z_p,t_m,c_n} as a 2D grayscale image size of $X \times Y$ in p^{th} focal plane image along z-direction in a volume, the m^{th} time sample, and the n^{th} color channel, where $p \in \{1, \ldots, Z\}$, $m \in \{1, \ldots, T\}$, and $n \in \{1, \ldots, C\}$, respectively. Here, X and Y are image width and height, Z is the number of focal planes that is captured and formed microscopy volume. Similarly, T is the number of time samples and C is the number of color channels. Note that the original fluorescence microscopy volume is always a grayscale

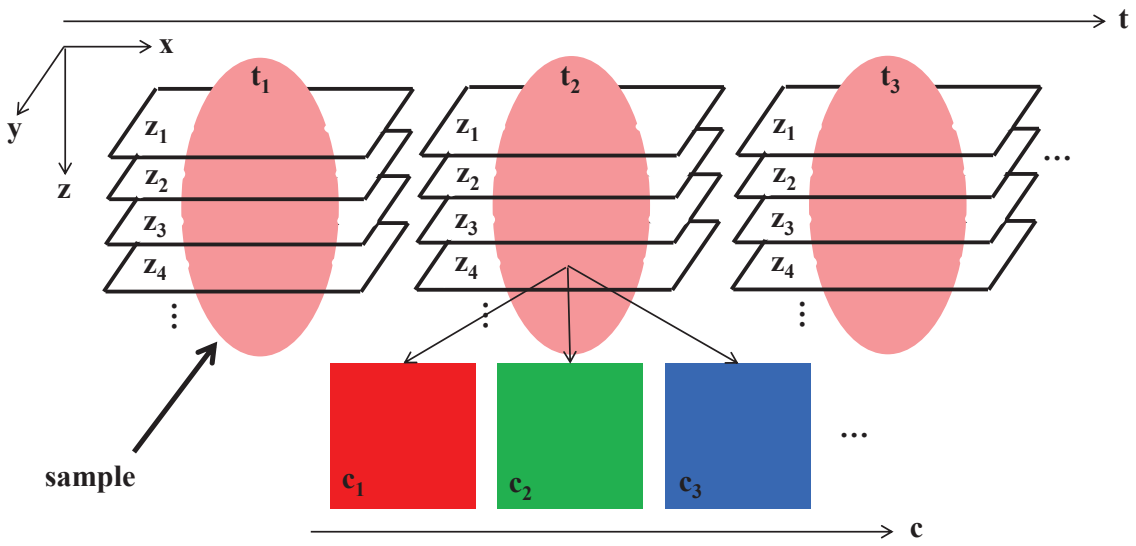


Fig. 1.5. The general notation for the microscopy volume in this thesis

volume in one time stamp. However, biologists often acquires volumes with multiple fluorescence with highlighting different structures simultaneously and save the volume with different channel volumes for the purpose of observation. In this thesis, we always separate channel first as preprocessing to get a 3D grayscale volume with structure of interest. Also, we use a 3D volume with single time stamp. Therefore, for brevity, we denote I as a 3D image volume of size $X \times Y \times Z$. The volume could be formed as a stack of multiple images in z-, y-, and x-direction. Therefore, we denoted I_{z_p} as a xy section with p^{th} focal plane along the z-direction in a volume, where $p \in \{1, \ldots, Z\}$. Similarly, I_{y_q} is a xz section with q^{th} focal plane along y-direction, where $q \in \{1, \ldots, Y\}$, and I_{x_r} is a yz section with r^{th} focal plane along x-direction, where $r \in \{1, \ldots, X\}$. For example, I^{orig}_{z23} is the 23^{rd} focal plane image of an original volume, I^{orig}. In addition, let $I_{(r_i:r_f,q_i:q_f,p_i:p_f)}$ be a subvolume of I, whose x-coordinate is $r_i \leq x \leq r_f$, y-coordinate is $q_i \leq y \leq q_f$, and z-coordinate is $p_i \leq z \leq p_f$, where $r_i, r_f \in \{1, \ldots X\}$, $q_i, q_f \in \{1, \ldots Y\}$, and $p_i, p_f \in \{1, \ldots Z\}$. For example, $I^{seg}_{(241:272,241:272,131:162)}$ is a subvolume of segmented volume, I^{seg}, where the subvolume is cropped between 241^{st} slice and 272^{nd} slice in x-direction, between

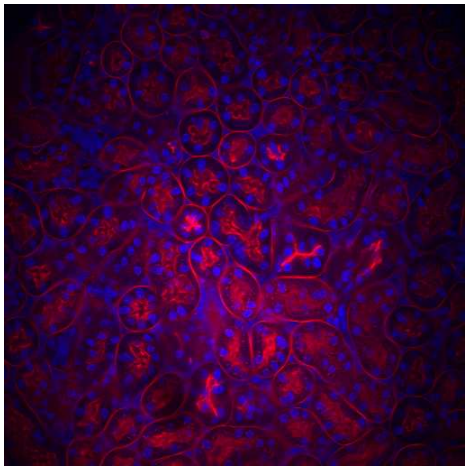
241^{st} slice and 272^{nd} slice in y-direction, and between 131^{st} slice and 162^{nd} slice in z-direction.

1.3 The Problem Formulation

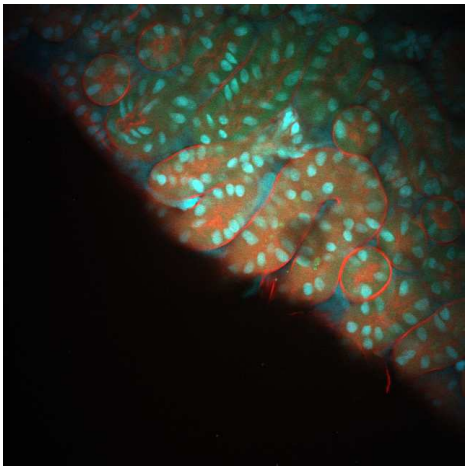
This thesis demonstrates various methods to analyze multiphoton microscopy volumes focusing on segmentation and deconvolution. Several image acquisition techniques were introduced from Chapter 1.1 and our datasets are mainly acquired via multiphoton microscopy. Intravital microscopy is a powerful tool for observing and studying physiological process in the living animal [20]. With recent development of microscopy techniques such as confocal and multiphoton microscopy, intravital microscopy has enabled to capture physiological process at cellular and subcellular levels. However, due to this enormous and complex data acquisition, digital image analysis becomes a necessary component in the quantitative analysis.

Fluorescence microscopy volumes have unique characteristics that complicate image segmentation. In particular, microscopy volumes are anisotropic, with aberrations and distortions that vary in different directions [21] giving rise to inhomogeneous intensities along the horizontal directions as well as poorer depth resolution, where the vertical resolution is approximately 3 to 4 times less than the horizontal resolution [22, 23]. Since biological structures often consist of non rigid shapes with varying orientations, fluorescent probes frequently fail to delineate correct boundaries. Due to drawing by probes fluorescence microscopy typically generates poorer resolution images compared to normal images taken from camera. This uniqueness of fluorescence microscopy makes segmentation results rely heavily on parameters and methods [24].

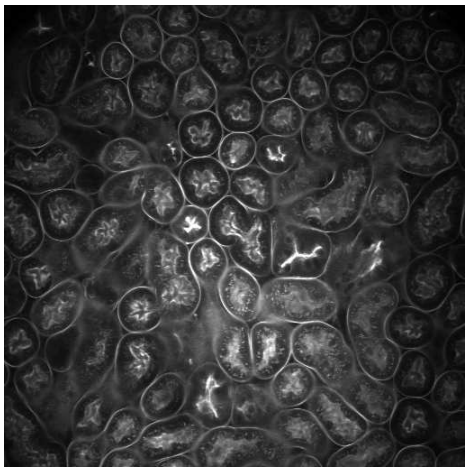
As previously indicated our datasets are collected from living animals. With live specimens imaging, motion artifacts will be introduced as a result of internal motions as well as external motions. For instance, respiration and heartbeat are examples of internal motions whereas small movements by living animals introduce external motions. Both motions introduce image distortions including translation, rotation,



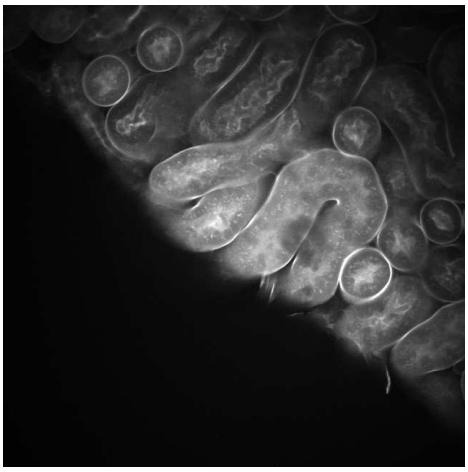
(a) $I_{z_{100}}$ of *WSM*



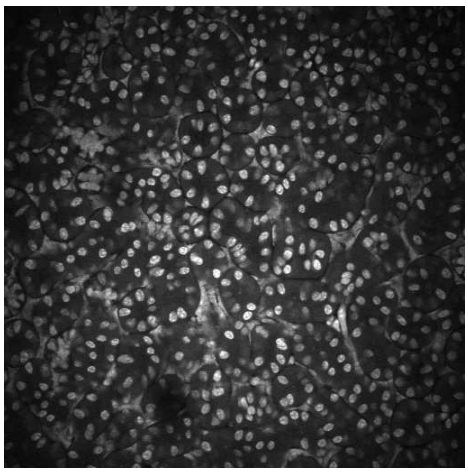
(b) $I_{z_{100}}$ of *Lectin*



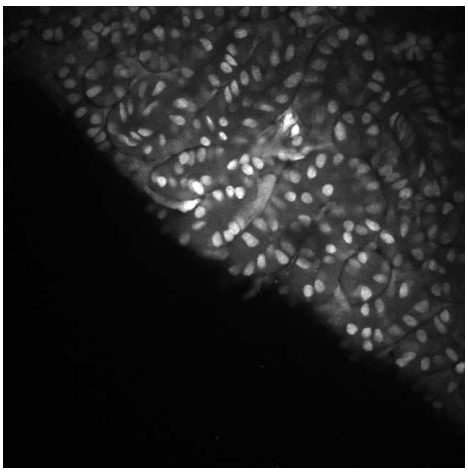
(c) I_{z_{100},c_1} of *WSM* (R channel)



(d) I_{z_{100},c_1} of *Lectin* (R channel)

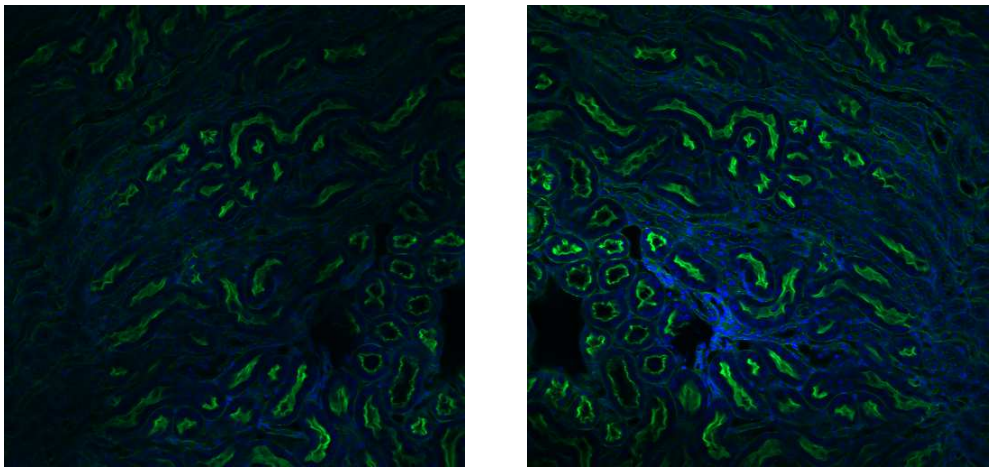


(e) I_{z_{100},c_3} of *WSM* (B channel)



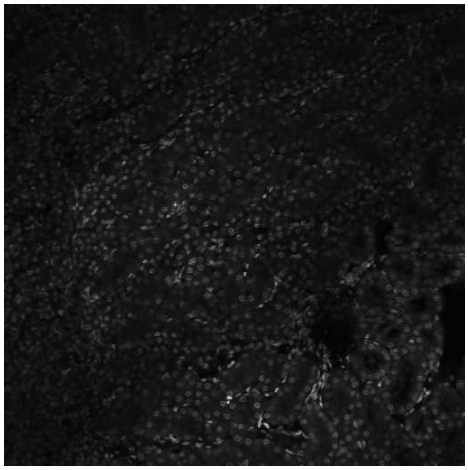
(f) I_{z_{100},c_3} of *Lectin* (B channel)

Fig. 1.6. Sample images of *WSM* and *Lectin* datasets

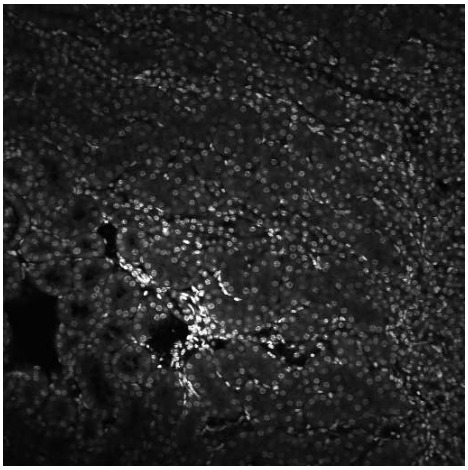


(a) $I_{z_{24}}$ of *FNa*

(b) $I_{z_{24}}$ of *FNa1*

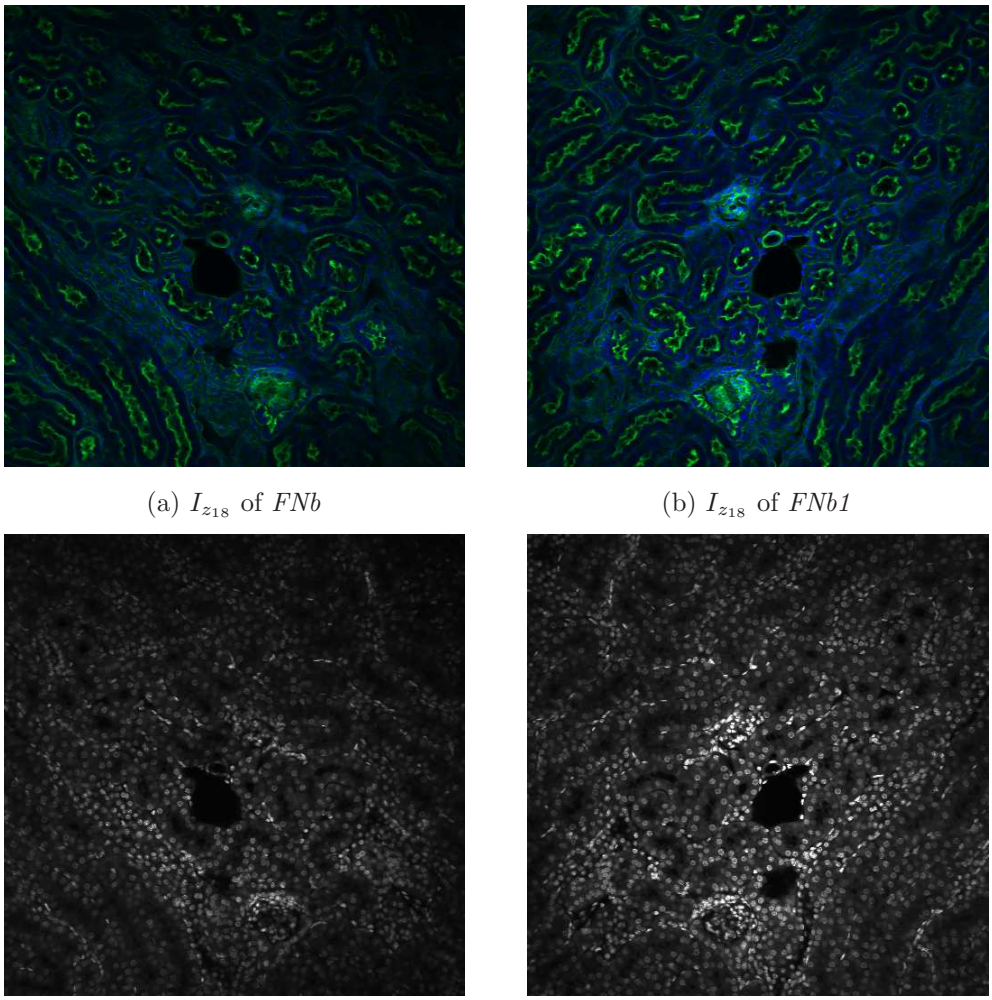


(c) I_{z_{24},c_3} of *FNa* (B channel)



(d) I_{z_{24},c_3} of *FNa1* (B channel)

Fig. 1.7. Sample images of *FNa* datasets



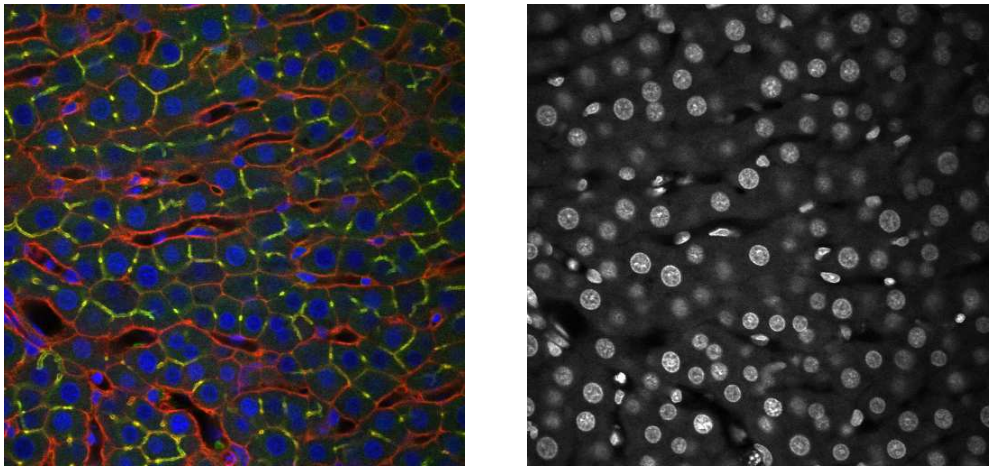
(a) $I_{z_{18}}$ of *FNb*

(b) $I_{z_{18}}$ of *FNb1*

(c) I_{z_{18},c_3} of *FNb* (B channel)

(d) I_{z_{18},c_3} of *FNb1* (B channel)

Fig. 1.8. Sample images of *FNb* datasets



(a) $I_{z_{16}}$ of *immuno* (b) I_{z_{16},c_3} of *immuno* (B channel)

Fig. 1.9. Sample images of *immuno* dataset

scaling, warping motions. These distortions can be corrected by rigid or non-rigid image registration methods. Note that image registration is aligning two or more images acquired at different time points, different information from different sensors, different viewpoints or perspectives [25]. In general, image registration should be performed before utilizing any segmentation methods to be well-aligned with adjacent images. In this thesis, however, we will not investigate any image registration techniques and assume all datasets that we use are well-aligned.

We will use five different 3D microscopy volume datasets to test proposed segmentation and deconvolution methods. Sample images of *WSM* and *Lectin* datasets[1] are shown in Figure 1.6 which each of datasets are acquired from a 3D volume of rat kidney. Our main goal here is delineating small nuclei dyed with blue as well as tubular structures dyed with red. As can be seen from Figure 1.6 image brightness are inhomogeneous and edge details are not clearly drawn which makes segmentation challenging. Other sample images are illustrated in Figure 1.7 and 1.8. Again, these images are sample images from 3D volume of rat kidney acquired from multiphoton

[1] *WSM* and *Lectin* datasets were provided by Malgorzata Kamocka of the Indiana Center for Biological Microscopy.

microscopy. Notify that *FNa* and *FNb* datasets[2] are collected two times with forward and reverse acquisition. More precisely, *FNa* is a 3D volume collected from top to bottom whereas *FNa1* is collected from bottom to top utilizing exact same volume. *FNa* and *FNa1* are mirror images at plane 24. Similarly, *FNb* and *FNb1* represents same object volume but different direction with mirror images at plane 18. Therefore, each of flipped pairs (Figure 1.7a, 1.7b and Figure 1.8a ,1.8b) portray theoretically exact same image plane but just shown as flipped. The reason of two ways collection is due to failure of probe drawing at deeper tissues. Relatively deep planes from forward acquisition can be turned into relatively shallow planes from reverse acquisition. Hence, segmentation performance of low contrast images can be compared with relatively high contrast images. As observed in Figure 1.7 and 1.8 right column images are flipped version of left column images but image contrasts are different each other. We use only the B channel from the RGB image of the *FNa* and *FNb* dataset and their flipped datasets for the nuclei segmentation. Lastly, sample images of *immuno* dataset[3] is displayed in Figure 1.9. Compared to *WSM*, *FNa* or *FNb* datasets, this dataset has well-defined nuclei image in B channel as observed in Figure 1.9b. Along with *WSM* dataset this *immuno* dataset is used for testing our segmentation and visualization including color labeling. Detail will be discussed more in Chapter 8.

1.4 Contributions of This Thesis

In this thesis we developed new segmentation and methods to identify various biological structures including nuclei and tubules under unique properties and characteristics of fluorescence microscopy volumes. Also, we developed new deconvolution methods so that we can successfully restore the deeper tissues from noisy and out-of-focus volume to clean and focused volume. Additionally, we developed the Distributed and Networked Analysis of Volumetric Image Data (DINAVID) system for

[2] *FNa*, *FNa1*, *FNb*, and *FNb1* datasets were provided by Tarek Ashkar of the Indiana University School of Medicine.

[3] *immuno* dataset was provided by Kenneth W. Dunn of the Indiana University School of Medicine.

biologists to use deep learning inference with five pretrained model selection. Also, this DINAVID system provides a 3D interactive visualization tool so that biologists can understand and analyze 3D segmentation results more clearly. The main contributions regarding proposed methods are listed as follows:

- Wavelet Based Multistage Edge Detection

  We review current edge detection techniques using wavelet and contourlet transform. We introduce our multiscale edge detection method using dyadic wavelet transform. To be more specific, the proposed method utilizes a combination of highpass components in x and y directions passed from 1D dyadic filter, pointwise multiplications across two adjacent dyadic wavelet transform, the modulus and angle computation followed by thresholding, and nonmaximum suppression in 8 directions to get edge maps. This method is tested on two-photon microscopy volume acquired from a rat kidney labeled with Hoechst 33342. The edge detection results from different adjacent scales are compared with Canny edge detector side by side. Experimental results show ability to capture nuclei edges despite inhomogeneity background.

- Boundary Fitting Based Segmentation

  We conduct a literature review focusing on tubular shapes segmentation schemes. We demonstrate boundary fitting based segmentation technique for segmenting boundaries of tubular structures. In particular, this technique segments 2D cross-sections of tubules in microscopy images using a combination of adaptive and global thresholding, potentials, z direction refinement, branch pruning, end point matching, and curve fitting. The aim of this technique is to be able to segment the 2D cross-sections at various depth as an initial step to 3D segmentation of these objects. This technique is tested on our datasets that are 3D fluorescence microscopy volume of rat kidney labeled with fluorescent phalloidin dyes using two photon fluorescence excitation microscopy. The experimental re-

sults are compared to that of an active contours based method and show efficacy of our proposed boundary fitting based segmentation technique.

- Three Dimensional Active Contours with Inhomogeneity Correction
  We conduct a literature review regarding edge-based active contours and region-based active contours including 2D and 3D. We extend previous 3D region-based active contours method by utilizing 3D inhomogeneity correction. Prior work that incorporated 3D snakes did not account for intensity inhomogeneities. More particularly, our proposed method modified energy function of 3D snakes and add inhomogeneity field terms so that active contours find optimum 3D segmentation as well as 3D inhomogeneity field information. The goal of this method is to account for intensity inhomogeneity while incorporating 3D information so as to segment 3D nuclei successfully. This method is tested on five different datasets that contains 3D nuclei structures obtained from rat kidney labeled with Hoechst 33342 collected using two-photon microscopy. The segmentation results of proposed method are compared with five different methods visually and numerically. Also, we present 3D segmentation results using 3D visualization tool. Experimental results demonstrate that the proposed method achieves better performance than other reported methods.

- Tubule Synthesis and Segmentation Based on Deep Learning
  We review recent deep learning schemes for image segmentation. We also review multiple schemes for the image synthesis. We present tubular structure segmentation method using convolutional neural networks with data augmentation and inhomogeneity correction. Using 3D region-based active contours method as a preprocessing step to correct inhomogeneity background. After that data augmentation is performed to increase size of paired training set. This paired training set becomes convolutional neural networks input to train model. This trained model is then used for inference for the test set. Lastly, postprocessing is done for filling holes and small components removal. The goal

of this method is identifying and segmenting individual tubules enclosed by their membranes. In order to identify individual tubule regions, we use 3D connected component labeling to assign a color to each tubule region. This method is tested on two different datasets that contain 3D tubular structures imaged from rat kidney labeled with a phalloidin which labels filamentous actin collected using two-photon microscopy. The segmentation results of the proposed methods are evaluated qualitatively and quantitatively compared with other microscopy segmentation methods. Also, we introduce object-based metrics as well as pixel-based metric to evaluate segmentation correctness. Experimental results shows that our proposed method outperforms with identifying multiple tubules than other compared methods. Moreover, we present a 3D tubular structure synthesis method using a gradient noise model by generating random but smooth centerlines. The results demonstrate that our proposed method can successfully generate 3D synthetic binary tubular structures.

- Three Dimensional Blind Image Deconvolution Using Generative Adversarial Networks

  We review various methods from conventional methods to recent deep learning methods for image/volume deconvolution. We present a blind image deconvolution method for fluorescence microscopy volumes using the 3-Way SpCycleGAN. In particular, we present a new approach to restore various biological structures in 3D microscopy images in deeper tissue without knowing the 3D point spread function using a spatially constrained CycleGAN (SpCycleGAN). We train and inference the SpCycleGAN in three directions along with xy, yz, and xz sections (3-Way SpCycleGAN). These restored 3-way microscopy volumes are averaged to incorporate 3D information. The restored volumes of proposed deconvolution method and other well-known deconvolution methods, denoising methods, and an inhomogeneity correction method are visually and numerically evaluated. To produce numerical evaluations, three image quality metrics are used. We test on our proposed and compared methods to two datasets which consist of Hoechst

33342 labeled nuclei and phalloidin labeled filament actin collected from a rat kidney, respectively. Moreover, we use this restored volume as a preprocessing step to perform SpCycleGAN segmentation and visualize the 3D segmentation results. Experimental results indicate that our proposed blind deconvolution method can restore and improve the quality of blurred and noisy deep depth microscopy image visually and quantitatively.

- Microscopy Volume Visualization
  We develop the Distributed and Networked Analysis of Volumetric Image Data (DINAVID) system that provides image analysis and visualization tools to biologists. In particular, this DINAVID system provides pretrained deep learning model for the nuclei segmentation and biologist can choose area to perform deep learning inference. We develop this color labeling method such that biologist can distinguish adjacent nuclei as much as possible. We also develop 3D interactive visualization tools to help biologist observe the original volume as well as segmentation results. Along with color labeling, the 3D interactive visualization can provide better insight for biologists to evaluate segmentation results.

1.5 Publications Resulting from This Work

Journal Papers

1. C. Fu, S. Han, **S. Lee**, D. J. Ho, P. Salama, K. W. Dunn, and E. J. Delp, "Three dimensional nuclei synthesis and instance segmentation," *To be submitted to the IEEE Transactions on Medical Imaging.*

2. **S. Lee**, S. Han, C. Fu, P. Salama, K. W. Dunn, and E. J. Delp, "Three dimensional tubule synthesis and segmentation for fluorescence microscopy using generative adversarial networks," *To be submitted to the IEEE Transactions on Medical Imaging.*

3. S. Han, **S. Lee***, C. Fu*, D. J. Ho, P. Salama, K. W. Dunn, and E. J. Delp, "A high-performance distributed and networked image analysis system for volumetric image data (DINAVID)," *To be submitted to the IEEE Transactions on Medical Imaging.*

4. K. W. Dunn, C. Fu, D. J. Ho, **S. Lee**, S. Han, P. Salama, and E. J. Delp, "DeepSynth: Three-dimensional nuclear segmentation of biological images using neural networks trained with synthetic data," *Submitted to Scientific Reports.*

Conference Papers

1. **S. Lee**, S. Han, P. Salama, K. W. Dunn, and E. J. Delp, "Three dimensional blind image deconvolution for fluorescence microscopy using generative adversarial networks," *Proceedings of the IEEE International Symposium on Biomedical Imaging*, pp. 538-542, April 2019, Venice, Italy.

2. S. Han, **S. Lee**, C. Fu, P. Salama, K. W. Dunn, and E. J. Delp, "Nuclei counting in microscopy images with three dimensional generative adversarial networks," *Proceedings of the SPIE Conference on Medical Imaging*, pp. 109492Y-1-11, February 2019, San Diego, CA.

3. C. Fu, **S. Lee**, D. J. Ho, S. Han, P. Salama, K. W. Dunn, and E. J. Delp, "Fluorescence microscopy image segmentation using convolutional neural network with generative adversarial networks," *Proceedings of the IEEE International Conferences on Computer Vision and Pattern Recognition Workshop*, pp. 2302-2310, June 2018, Salt Lake City, UT.

4. **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*, vol. 2018, no. 15, pp. 199-1-199-8, January 2018, Burlingame, CA.

5. **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.

6. **S. Lee**, P. Salama, K. W. Dunn, and E. J. Delp, "Boundary fitting based segmentation of fluorescence microscopy images," *Proceedings of the IS&T / SPIE International Symposium on Electronic Imaging*, pp. 940805-1-10, February 2015, San Francisco, CA.

2. LITERATURE REVIEW

2.1 Review of the Segmentation Methods

Although fluorescence microscopy was traditionally used for observing biological entities, it has recently become a very powerful tool for studying cell biology when combined with digital image analysis, as the development of fluorescence microscopy has enabled the acquisition of image volumes deeper into tissue [9, 10, 26]. This has resulted in the generation of large datasets of 3D microscopy image volumes, which in turn need automatic image segmentation techniques for quantification [17, 27]. However, the quantitative analyses of these datasets still pose a challenge due to light scattering, distortion created by lens aberrations in different directions, and the complexity of biological structures [1]. The end result is blurry image volumes with poor edge details that become worse in deeper tissue depths.

To quantify biological structures a first and a vital step is segmentation. There have been efforts to develop automatic segmentation methods for biomedical image data sets that attempt to address these issues, that in general have relied on thresholding, the watershed method [28], and active contours [29]. Since fluorescence microscopy suffers from various noises and intensity inhomogeneity, thresholds need to be chosen manually to achieve reasonable segmentation results. To overcome this problem, Otsu's threshold [30] was proposed to automatically determine threshold that minimizes the intra-class variance and maximize inter-class variance. Instead using a global threshold proposed in Otsu's method, Niblack [31] and Sauvola [32] proposed local thresholding methods which are more useful for intensity inhomogeneity background. However, threshold-based method does not have an ability to identify overlapped nuclei. Thus, thresholding is normally used in conjunction with other methods such as gradient flow tracking [33]. The watershed method [28] is

widely used for nuclei segmentation to solve this problem. The watershed method combines regions growing and edge detection techniques: it finds local minima (basin) and groups adjacent voxels to form a cluster and build watersheds to separate neighbor clusters. However, since the selection of local minima is highly dependent on the shape of interest and noise, watershed method tends to segment larger regions than expected. To address this oversegmentation problem, marker controlled watershed was introduced in [34] that replaces local minima with predefined markers. Also, this marker controlled watershed was improved using mean shift and Kalman filter to automatically determine marker locations so that it can achieve better segmentation of time-lapse microscopy [35]. Additionally, the watershed method is usually used in conjunction with other methods, for example, gradient flow tracking [36] or region splitting and merging [37]. Extending watershed methods from 2D to 3D, [38] described a hybrid method to segment nuclei in 3D. Meanwhile, [39] demonstrated combining registration and segmentation methods. Performing registration first, utilizing more accurate 3D information in microscopy volumes becomes possible.

Another widely used class of methods is based on the active contours technique which minimizes an energy functional to fit contours to objects of interest [29, 40]. Early version of active contours [29] generally produced poor segmentation results since the segmentation results are noise sensitive and initial contour dependent. There are several variants of active contours. One of these is edge-based active contours [29, 41] that utilizes image gradient maps to aid in object identification. The segmented results of edge-based active contours tend to be sensitive to image noise and rely heavily on the placement of the initial contour. To circumvent the dependence of the final outcome on the initial contour, [42] described integrating a background removal model with a region-based active contour method that uses multiple contour initializations to perform segmentation. Meanwhile, [43] presented a new method based on introducing a new external force called gradient vector flow (GVF). An external energy term which convolves a controllable vector field kernel with an image edge map was presented in [44] to address the noise sensitive problem. Similarly,

the Poisson inverse gradient was introduced to determine initial contours locations to segment microscopy images in [45].

Moreover, active contours have also been integrated with region-based approaches in an attempt to seek an energy equilibrium between foreground and background regions [46, 47]. Region-based methods tend to generate better segmentation results than edge-based active contours because region-based methods are relatively independent of the initial contour generation and are more robust against noise. In [21, 48] a 3D version of region-based active contours [46] which results in improved segmentation by incorporating 3D information was presented. More specifically, in [21] a coupled active surface which is a 3D version of multiphase level sets [47] that utilize an indicator function was proposed to avoid multiple level set overlapping. Similarly, in [49] the coupled active surfaces based method was improved by incorporating a watershed method and the Radon transform. These region-based methods, however, fail to produce satisfactory results in images with intensity inhomogeneities [50, 51]. To address this [50] proposed a localized region-based active contour method that uses a new energy functional, means separation, and histogram separation to distinguish between foreground regions and inhomogeneous background areas. Similarly, a modified energy function was used in [52] to address image inhomogeneities. More recently, in [51] we utilized adaptive thresholding and vertical direction refinement followed by boundary fitting to segment microscopy volumes while taking into account intensity inhomogeneities. Alternatively, [53] implemented a 3D segmentation method with an energy function based on exponential B-splines, while [54] proposed combining edge-based and region-based energy functions to segment fluorescence microscopy images. More recently, the 3D region-based active contours introduced in [48] was combined with 3D inhomogeneity correction to provide better segmentation since this technique takes into consideration inhomogeneities in volume intensity [55].

Active contours have also been integrated with region-based approaches in an attempt to seek an energy equilibrium between foreground and background regions [46, 47]. Region-based methods tend to generate better results than edge-based ac-

tive contours because region-based methods are relatively independent of the initial contour generation and are more robust against noise. In [21, 48] a 3D version of region-based active contours [46] which results in improved segmentation by incorporating 3D information was presented. More specifically, in [21] a coupled active surface which is a 3D version of multiphase level sets [47] that utilize an indicator function was proposed to avoid multiple level set overlapping. Similarly, in [49] the coupled active surfaces based method was improved by incorporating a watershed method and the Radon transform. These region-based methods, however, fail to produce satisfactory results in images with intensity inhomogeneities [50, 51]. To address this [50] proposed a localized region-based active contour method that uses a new energy functional, means separation, and histogram separation to distinguish between foreground regions and inhomogeneous background areas. Similarly, a modified energy function was used in [52] to address image inhomogeneities. Additionally, we utilized adaptive thresholding and vertical direction refinement followed by boundary fitting to segment microscopy volumes while taking into account intensity inhomogeneities [51]. Alternatively, [53] implemented a 3D segmentation method with an energy function based on exponential B-splines while [54] proposed combining edge-based and region-based energy functions to segment fluorescence microscopy images.

Different from the methods that we discussed above, wavelets [56] or contourlets [57] also widely used for image segmentation especially focusing on multiscale information. The primitive idea of detecting edges utilizing wavelets coefficients was introduced in [56, 58]. Also, edge detection method utilizing the dyadic wavelet transforms was suggested in [59]. By identifying edges utilizing a multi-scale wavelet edge detection technique, [35] described method for segmenting confocal microscopy of neurons and dendrites. More recently, in [60] edge detection technique using directional wavelet transform [61] was described. Similarly, [62] presented directional multiscale edge detection schemes with the contourlet transform.

In contrast to previously discussed methods, a new segmentation method known as Squassh that couples image restoration and segmentation using a generalized linear

model and Bergman divergence was introduced [63], whereas a method that combined with detecting primitives based on nuclei boundaries and identifying nuclei region using region growing was demonstrated in [64]. Alternatively, combination of multiresolution, multiscale, and region growing methods using random seeds to perform multidimensional segmentation was described in [65]. Also, in [66] a combination of midpoint analysis, shape-based function optimization, and a Marked Point Process (MPP) simulation to quantify nuclei was presented. More specifically, two segmentation results based on a shape fitting function and an MPP were obtained and combined to identify individual nuclei from non-homogeneous intensity images.

The challenge of properly delineating coupled cells or nuclei is an ongoing research area. Identifying individual nuclei is a prerequisite step for a variety of cell cytometry approaches used in biological research. The methods we previously discussed did not take consideration of splitting these densely clustered problem. Only watershed method can automatically separate touching cells into two or multiples with straight lines. Several techniques have been developed to address cell splitting, particularly for overlapping nuclei. For example, in [67], a combination of curvature analysis and splitting method is presented. More specifically, maximum curvature points by curvature analysis was firstly detected and the Delaunay triangulation was used from each maximum curvature points. After utilizing geometry constraints, final edges to split coupled cells are acquired. Recently, [67] was integrated with a multiphase level set followed by voting function to produce better segmentation in [68]. Similarly, automated cell counting and cluster segmentation technique using ellipse fitting is presented in [69, 70]. These ellipse fitting methods split clustered nuclei into individual ellipses fitting given contours. Meanwhile, [64] described new models which imitate how a human locates a nucleus by identifying the nucleus boundaries and piecing them. In particular, this method primitively identified partial contours of each nucleus and closed boundaries using region-growing. Alternatively, an automated 3D detection and segmentation method for touching cells using a combination of concave points clustering and random walk is introduced in [71]. Additionally, [72] devel-

oped an integrative cut (iCut) method which uses information of spatial location, intervening, and concave contours based on normalized cut.

As indicated above, florescence image segmentation still remains a challenging problem. Tubule, a biological structure with a tubular shape, segmentation is even more challenging since tubular shape and orientation is varied without known patterns. Also, since typical tubular structures have hollow shapes with unclear boundaries, traditional energy minimization based methods such as active contours have failed to segment tubular structures [51]. A typical dataset we use in our studies consists of two tissue structures, the base membrane of the tubular structures and the brush border which is generally located interior to proximal tubules. Our goal here is to segment individual tubular structure membrane. There has been some work particularly focusing on tubular structure segmentation. A minimal path based approach was described in [73,74] where tubule shape is modeled as the envelope of a family of spheres (3D) or disk (2D). Similarly, a new approach for 3D human vessels segmentation and quantification using 3D cylindrical parametric intensity model was demonstrated in [75]. Also, multiple tubule segmentation technique that combined with level set methods and the geodesic distance transform was introduced in [76]. Also, [77] described a scheme using geodesic active contours to detect tubular structures. More recently, one method used to segment tubular structures was delineating tubule boundaries followed by ellipse fitting to close the boundaries while considering intensity inhomogeneity [51]. Another method known as Jelly filling [78] utilized adaptive thresholding, component analysis, and 3D consistency to achieve segmentation, whereas a method for tubule boundary segmentation used steerable filters to generate potential seeds from which to grow tubule boundaries followed by tubule/lumen separation and 3D propagation to generate segmented tubules in 3D [79]. Previous methods, however, focused on segmenting boundaries of tubule membrane. Since some tubule membranes are not clearly delineated in fluorescence microscopy image volume, finding tubule boundaries may not always result in identifying individual tubule regions.

Convolutional neural network (CNN) has been used to address segmentation problems [80] in biomedical imaging [81]. The fully convolutional network [82] introduced an encoder-decoder architecture for semantic segmentation. This encoder-decoder architecture was extended in SegNet [83] that utilizing VGG network architecture [84] as an encoder network and adding up corresponding decoder network with sharing pooling indices to perform better image segmentation. Using this SegNet [83], nuclei segmentation method [85] and tubular structure segmentation method [86] were described. Instead of sharing pooling indices locations, U-Net [87] transferred entire feature maps from encoder layer to corresponding decoder layer so that when expanding the 2D image size in encoder layer rich feature is available for the reconstruction. Similarly, in [88] a U-Net trained on cell objects and contours was used to identify tubular structures. Additionally, a multiple input and multiple output structure based on a CNN for cell segmentation in fluorescence microscopy images was demonstrated in [89]. Since these approaches [82, 83, 87, 89] are all 2D segmentation methods, they may fail to produce reasonable segmentation in 3D. More specifically, stacking these 2D segmentation images into 3D volume may result in misalignment in the depth direction [48]. Consequently, there are some paper to extend this 2D segmentation method to 3D. A nuclei segmentation method that combined with a 2D CNN and a 3D refinement process was introduced in [85]. A 3D U-Net [90] was introduced to identify 3D structures by extending the architecture of [87] to 3D. Similarly, V-Net used the Dice loss for the CNN training for volumetric medical image segmentation was introduced in [91]. Meanwhile, VoxResNet which is a voxelwise segmentation using ResNet [92] was introduced in [93]. However, this approach requires manually annotated groundtruth to train the network which is tedious and time-consuming process and implausible to obtain correct 3D groundtruth by 2D slices. One way to address this is to use synthetic ground truth data [94, 95]. A method that segments nuclei by training a 3D CNN with synthetic microscopy volumes was described in [96]. Note that the synthetic microscopy volume generated in [96] are generated by adding blur and noise.

Apart from the various methods to generate synthetic groundtruth data to overcome the needs of gigantic number of groundtruth, the style transfer methods which transfers between two different styles such as from summer scene to winter scene, one artist image to other artist image and vice versa. The style transfer methods needs to do two things. First of all, the methods should understand/capture the style of given set of training images. Also, the methods could generate high quality textures/images based on test images but learned styles from the set of the training images so that the final output should have the characteristics of test images with trained styles. To achieve this style transfer, [97, 98] uses CNN networks stacked similar to encoder-decoder structures using the VGG network [84] with minimizing averaging pixel differences between two generated set of Gram matrix representations which taken from one layer in encoder stage and corresponding layer in decoder stage. Similarly, the perceptual losses were introduced in [99] for the real-time the style transfer and super-resolution. This paper is the first paper discussed using this style transfer not only for image generation with artistic sense, but also solving image quality problem by increasing high quality. Alternatively, the texture network which utilizes a feed forward generation networks was introduced in [100]. The improved version of this texture networks by alternating architecture were presented in [101].

Generating realistic synthetic microscopy image volumes remains a challenging problem since various types of noise and biological structures with different shapes are present and need to be modeled. Recently, a generative adversarial network (GAN) was proposed to generate realistic images from random noise using two adversarial networks, a generative network and a discriminative network in [102]. This GAN can be used as an image-to-image transfer which is exact same roles of the style transfer with utilizing the generative network. More particularly, the discriminative network learns a loss function to distinguish whether the output image is real or fake whereas the generative network tries to minimize this loss function. This two networks optimization problems can be considered as saddle point optimization problem since the discriminative networks should do maximize its distinguishable

power and generative networks should minimize the loss. The problem of the GAN is both discriminative and generative networks may not converge. To circumvent this problem, a deep convolutional GAN (DCGAN) which uses logarithms to the discriminative network loss to improve its convergence was demonstrated in [103]. Also, the Wasserstein GAN (W-GAN) [104] utilizes the Earth-Mover distance instead of the Jensen-Shannon divergence distance to improve its training convergence. Another example of variation of discriminative network for stabilizing training is a boundary equilibrium GAN (BEGAN) [105] which replaces the discriminative network with an autoencoder architecture [106]. Meanwhile, two separate encoder networks instead of one generative network were used in [107, 108]. Alternatively, a variational autoencoder (VAE) is used instead of decoder to the architecture of generative networks (VAEGAN) was proposed in [109].

One of the extensions of GANs is Pix2Pix [110] which uses conditional GANs to learn the relationship between the input image and output image that can generate realistic images. One issue with Pix2Pix [110] is that it still requires paired training data to train the networks. To achieve the learning without having paired dataset, a coupled GANs (CoGAN) for learning the joint distribution of multi-domain images without having the corresponding groundtruth images was introduced in [111]. Additionally, a method that combines simulated and unsupervised learning to improve generating realistic synthetic images was demonstrated in [112]. Later, a cycle-consistent adversarial networks (CycleGAN) [113] employed a cycle consistent term in the adversarial loss function for image generation without using paired training data. Similarly, a method called DiscoGAN that uses a reconstruction loss with the two GAN losses for discover the cross-domain relations was presented in [114]. Apart from the CycleGAN [113] and DiscoGAN [114], combination of two VAEGAN [109] together to achieve unpaired image-to-image transform. More recently, a segmentation method using concatenating segmentation network to CycleGAN to learn the style of CT segmentation and MRI segmentation was described in [115]. Additionally, [116] presented a spatially constrained CycleGAN (SpCycleGAN) which adds

spatial loss term to regularize the location of synthetic nuclei to improve nuclei segmentation results. Using this SpCycleGAN along with the 3D GAN to perform a 3D nuclei counting method was demonstrated in [117]. However, aforementioned methods are all 2D based image-to-image transfer. Again, 2D based methods are limited for the volume analysis since the volume is naturally correlated with previous and next sections so that 3D information is always helpful to understand each section. Consequently, there are some literatures that try to utilize this 3D information in various ways. To address 2D based image-to-image transfer, we extended the SpCycleGAN to a 3-Way SpCycleGAN which uses the SpCycleGAN three times in axial (xy), sagittal (xz), and coronal (yz) section directions, stacks them to build three 3D volumes and performs weighted average on these 3D volumes was presented in [118]. More recently, The CycleGAN was extended to full 3D in [119] so that the volume-to-volume transfer such as CT and MRI volume synthesis was proposed. The typical 3D volume obtained from medical modality can be considered as set of 2D images which is correlated by depth (z) direction. Therefore, this 3D volume data is similar to video data. Recent approach to combine both spatial and temporal information at the same time to the loss function to minimize loss (ReCycleGAN) to solve the video retargetting problem was proposed in [120]. Similarly, video-to-video synthesis approach was introduced in [121].

2.2 Review of the Deconvolution Methods

Fluorescence microscopy is a modality that allows imaging of subcellular structures from live specimens [10, 122]. During this image acquisition process, large datasets of 3D microscopy image volumes are generated which in turn need automatic quantification process [123]. The quantitative analysis of the fluorescence microscopy volume is hampered by light diffraction, distortion created by lens aberrations in different directions, complex variation of biological structures [1]. The image acquisition process can be typically modeled as the convolution of the observed objects with a

3D point spread function (PSF) followed by degradation from noise such as Poisson noise and Gaussian noise [124]. Poisson noise can represent the stochastic nature of the photon counting process while Gaussian noise is occurred due to the intrinsic thermal and electronic fluctuations of the acquisition process [124]. Moreover, the resolution of fluorescence microscopy is often insufficient for biological research to visualize subcellular structures [125]. These limitations result in anisotropic, inhomogeneous background, blurry (out-of-focus), and noisy image volume with poor edge details which aggravate the image quality poorer in depth [86].

There has been various approaches to improve 3D fluorescence microscopy images quality. One popular approach is known as image deconvolution which "inverts" the convolution process to restore the original microscopy image volume [126]. Numerically inverting the blurring process to solve out-of-focus problem in the widefield microscopy was introduced in [127]. Richard-Lucy (RL) deconvolution [128, 129] which maximizes the likelihood distribution based on a Poisson noise assumption for confocal microscopy was proposed. This RL deconvolution was further extended in [130] that incorporated the total variation as a regularization term in the cost function. Since the PSF is usually not known, blind deconvolution which estimates the PSF and the original image simultaneously is favorable [131]. Blind deconvolution using RL deconvolution was described in [132]. A pupil model for the PSF was presented in [133] and the PSF was estimated using machine learning approaches in [134]. Sparse coding to learn 2D features for coarse resolution along the depth axis to mitigate anisotropic issues was presented in [135].

Another approach to achieve better image quality stems from image denoising research. One example is Poisson noise removal using a combination of the Haar wavelet and the linear expansion of thresholds (PURE-LET) proposed in [136]. This PURE-LET approach was extended further to 3D widefield microscopy in [137]. Additionally, a denoising and deblurring method for Poisson noise corrupted data using variance stabilizing transforms (VST) was described in [138]. Meanwhile, a 3D in-

homogeneity correction method that combines 3D active contours segmentation was presented in [55].

Convolutional neural network (CNN) has been popular to address various problems in medical image analysis and computer vision such as image denoising, image segmentation, and image registration [81]. There are a few papers to focus on image restoration and image deconvolution in natural images were presented using sparse coding and pretrained denoising autoencoder [139], outlier-rejection deconvolution CNN [140], and combination of deep neural network and general regression neural networks [141]. There are few papers that focus on image deconvolution in fluorescence microscopy using CNNs. One example is an anisotropic fluorescence microscopy restoration method using a CNN [142]. Later, semi-blind spatially-variant deconvolution in optical microscopy with a local PSF using a CNN was described in [131]. More recently, generative adversarial networks has gradually gained interest in medical imaging especially for medical image analysis [143]. One of the useful architectures for medical image is a cycle-consistent adversarial networks (CycleGAN) [113] which learns image-to-image translation without having paired images (actual groundtruth images). This CycleGAN was utilized for CT denoising by an analogy of mapping low dose phase images to high dose phase images to improve image quality [144]. Moreover, a synthesis and segmentation network using this CycleGAN was introduced in [115]. Additionally, this CycleGAN was further extended by [116] incorporating a spatial constrained term to minimize misalignment between synthetically generated binary volume and corresponding synthetic microscopy volume to achieve better segmentation results.

One of the problem on these synthetically generated images without having reference image is that there is no standard quality metric for measuring the quality of 3D microscopy images. Traditional image quality metrics such as PSNR, SSIM, and FSIM cannot be obtained without having reference images. Additionally, these traditional image quality metrics do not agree with human judgment [145] either. To address this problem, several no reference image quality assessments (NR-IQA) were

proposed using natural images. The blind/referenceless image spatial quality evaluator (BRISQUE) model [146] is a well-known NR-IQA that learned local statistics models from natural scene while oriented gradient image quality assessment (OG-IQA) [147] is another NR-IQA which obtained the quality assessment by relative gradient statistics and adaboosting neural network were proposed. This BRISQUE was modified to evaluate structural MRI images in [148]. More recently, a new assessment for microscopy image focus quality using deep learning was introduced [149].

3. WAVELET BASED MULTISCALE EDGE DETECTION

In this chapter, multiscale edge detection method using dyadic wavelet transform (DWT) is introduced. As can be observed from Chapter 3.2, this method provides reasonable edge detection results compared to the Canny edge detector [150]. However since the edges extracted using proposed multiscale wavelet method is vulnerable from noise and the results heavily depends on selection of threshold, we did not investigate more with multiscale wavelet based method. Instead, 3D active contours with inhomogeneity correction method is introduced in favor of capturing small nuclei shown in Chapter 5.4.

Wavelet transforms are widely used for image processing area including denoising, compression, and feature extraction. Wavelet transforms are broadly divided into three classes: continuous, discrete and multiresolution-based. Since wavelet transform can perform multiresolution, localization, and critical sampling properties from wavelet transform, wavelet based method can be suitable method for obtaining edge information from images [57].

The intuition behind this proposed edge detection is the property of wavelets which peaks (1D) or edges (2D) are preserved across the scales. This property is mathematically analyzed using local Lipschitz regularity of the signal from [58]. More specifically, signal across scales in wavelet domain depends on the local regularity obtained by Lipschitz exponents. Therefore highpass filtering from each scale contains peaks or edges information. However, different from peaks or edges, the Gaussian white noise is suppressed to half compared to previous scale [56]. Due to this property, the wavelet could be used for edge detection especially for the noisy images.

The overview of proposed wavelet based multiscale edge detection method is shown in Figure 3.1. First of all, we utilize dyadic wavelet transform in x and y directions to obtain highpass filtered outputs for each scale j. After passing highpass filters

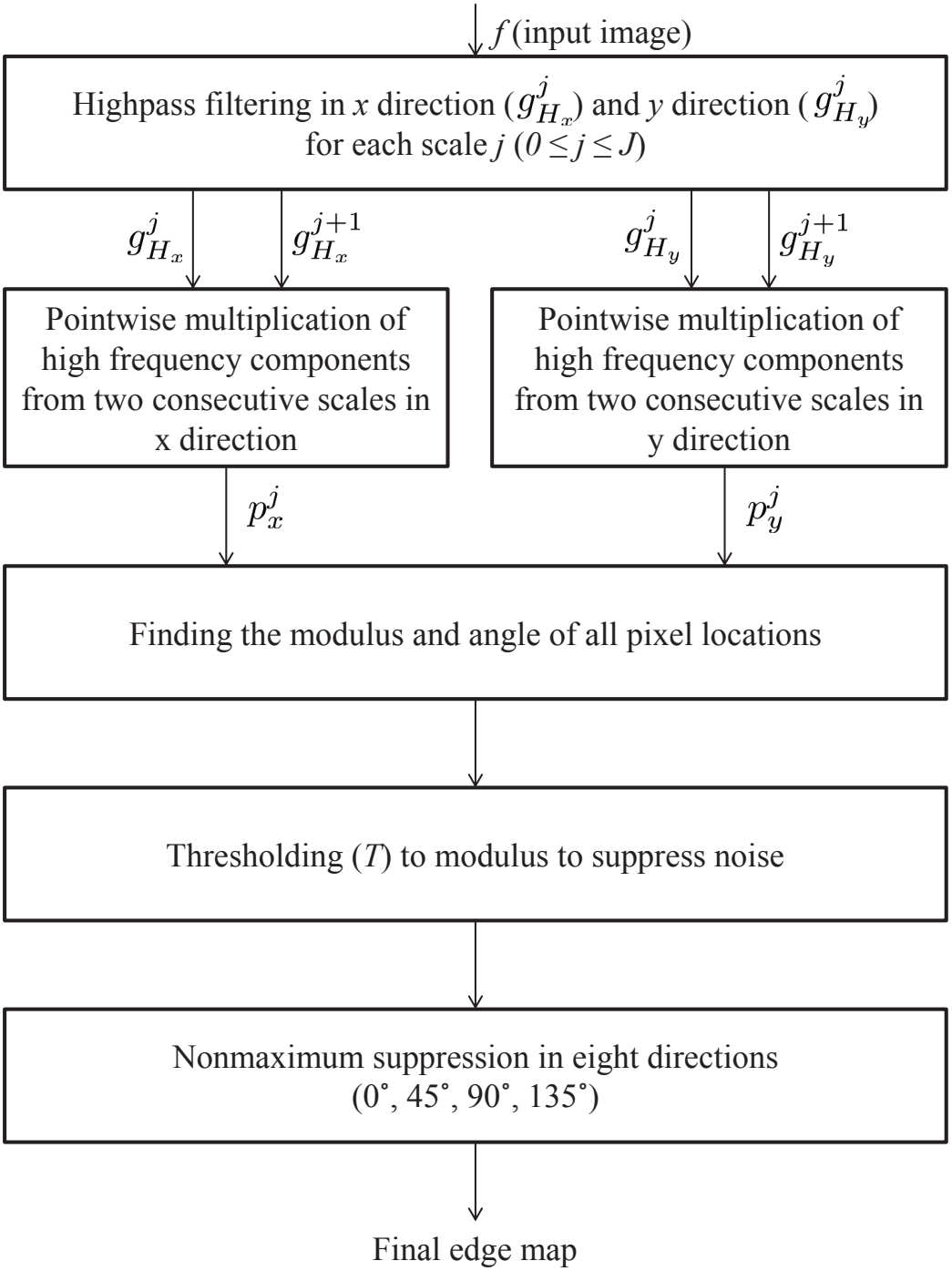


Fig. 3.1. Overview of the proposed wavelet based multiscale edge detection method

with different scales, pointwise multiplication is performed across two adjacent DWT scales. As indicated above, this scale multiplication should contain strong peaks or edges information. Using this product function across two adjacent scales, the modulus and angle of all pixel locations are obtained. After that thresholding followed by performing non-maxima suppression in four directions to find final edge map. Each step will be described in detail from the following section.

3.1 Proposed Technique

3.1.1 Dyadic Wavelet Transform

Suppose any smoothing function $\theta(x)$ which is differentiable and whose integral is equal to 1 and that converges to 0 at infinity. Let wavelet $\psi(x)$ be the first derivative of $\theta(x)$ such that

$$\psi(x) = \frac{d\theta(x)}{dx}. \tag{3.1}$$

In particular, $\theta(x)$ is widely chosen as a cubic spline function [56, 59] such that

$$\theta(x) = \begin{cases} 0 & |x| \geq 1 \\ \theta(-x) & 0 \leq x \leq 1 \\ -8x^3 - 8x^2 + 4/3 & -0.5 \leq x \leq 0 \\ 8(x+1)^3/3 & -1 \leq x \leq -0.5 \end{cases} \tag{3.2}$$

and the first order derivative of corresponding $\theta(x)$, $\psi(x)$, can be shown as a quadratic spline function

$$\psi(x) = \begin{cases} 0 & |x| \geq 1 \\ -\psi(-x) & 0 \leq x \leq 1 \\ -24x^2 - 16x & -0.5 \leq x \leq 0 \\ 8(x+1)^2 & -1 \leq x \leq -0.5 \end{cases}. \tag{3.3}$$

Note that the smooth function is even and the wavelet function is odd. To obtain wavelet coefficients from digital signal $\theta(x)$ and $\psi(x)$ are appropriately sampled so as

to construct lowpass and highpass filters, respectively. We utilize particular wavelet decomposition filters as [56] such that

$$\begin{aligned} L &= [0, \cdots, 0, 0.125, 0.375, 0.375, 0.125, 0, \cdots, 0] \\ H &= [0, \cdots, 0, -2, 2, 0, \cdots, 0] \end{aligned} \tag{3.4}$$

where L and H are lowpass and highpass filters, respectively. These filters are convolved with given signal f to produce wavelet coefficients.

Since main goal of our proposed method is to find edges from fluorescence microscopy images, all signals discussed after this section are 2D signals. However all the filters used in this chapter are 1D filters so that each 1D filters are applied two times in x and y directions to provide 2D wavelet coefficients. Denote L_x^j as the 1D lowpass filter in x direction at scale j where $0 \leq j \leq J$ and L_y^j as the 1D lowpass filter in y direction at scale j. Similarly, H_x^j defined as the 1D highpass filter in x direction at scale j and H_y^j defined as the 1D highpass filter in y direction at scale j.

Due to fundamental property of wavelet transform, the lowpass filtered output at the previous scale is passed to next scale as an input. Therefore lowpass and highpass filters should be constructed differently as scale changes. To address this problem, we use dyadic wavelet transform [56, 151].

To be more specific, each lowpass and highpass filters are constructed with 2^j dilation, which means $2^j - 1$ zeros are padded between each of the non-zero coefficients of filters [56]. For instance, at the scale 0, 1D lowpass filter in x, y directions (L_x^0, L_y^0) are the same as the lowpass filter (L) and 1D highpass filter x, y directions (H_x^0, H_y^0) are the same as the highpass filter (H) in Eq (3.4). However, at the scale 1, 1D lowpass filter L_x^1 and L_y^1 are constructed as

$$L_x^1 = L_y^1 = [0, \cdots, 0, 0.125, 0, 0.375, 0, 0.375, 0, 0.125, 0, \cdots, 0] \tag{3.5}$$

owing to $2^1 - 1$ zero-padded between non-zero coefficients. Similarly, at the scale 1, 1D highpass filter H_x^1 and H_y^1 are used such that

$$H_x^1 = H_y^1 = [0, \cdots, 0, -2, 0, 2, 0, \cdots, 0] \tag{3.6}$$

since $2^1 - 1$ zero was padded between -2 and 2.

Denote f^j as an input image at the scale j. Also, denote $g^j_{H_x}$ as high frequency components in x direction after passing highpass filter in x direction (${H_x}^j$). Likewise f_j is convolved with highpass filter in y direction (${H_y}^j$) produces $g^j_{H_y}$. Once obtaining two highpass components at the scale j, perform 2D lowpass filtering (${L_x}^j {L_y}^j$) to get $g^j_{L_{xy}}$. This lowpass filtered output is passed over to the next scale. Since lowpass components passed through L^0_x and L^0_y always handed over to the next scale, new input at the scale $j+1$ ($f^{(j+1)}$) is always the same as the 2D lowpass filtered output ($g^j_{L_{xy}}$) from previous input (f^j). Detail procedures are shown in Figure 3.2. Note that f^0 is given fluorescence microscopy image and considered as an input at the scale 0.

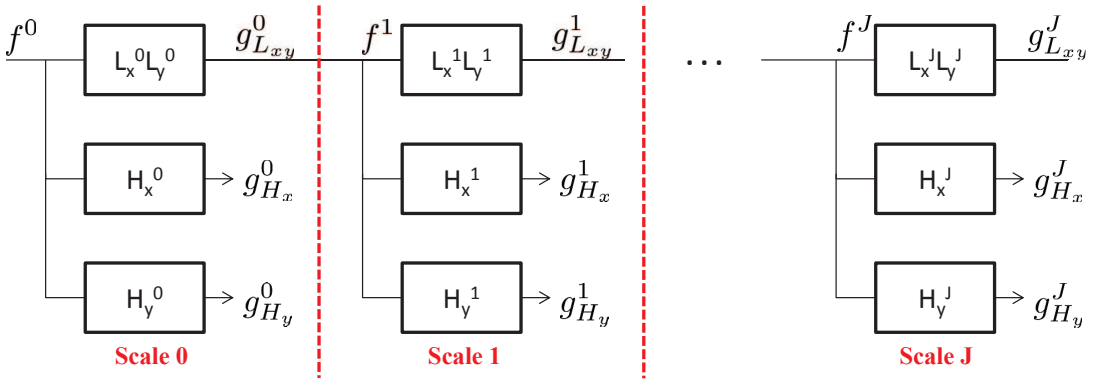


Fig. 3.2. Dyadic wavelet transform across the scale j $(0 \leq j \leq J)$

Starting from the first stage, this process is performed iteratively and pass lowpass filtered information until last scale J $(0 \leq j \leq J)$. This process is similar to general wavelet transform, but our proposed method does not have 2D downsampling step. In general, wavelet transform used lowpass and highpass filtering across x and y direction followed by 2D downsampling with coefficient 2 across the scale. Due to this fact, the size of coefficients keeps decreasing by a factor 2. However, since our proposed method did not perform 2D downsampling, the size of highpass filtering results across the scale is remained the same as original input image size. Instead

of 2D downsampling, we utilize 2^j dilated lowpass and highpass filters so that the filtering results look similar to wavelet coefficients. Moreover, our proposed method is similar to stationary wavelet transform for the fact that preserving the size of input passed to next scale. Yet, our method obtains 1D filtered results in x and y directions at each scale whereas stationary wavelet transform used 1D filter two times to acquire four different wavelet coefficients.

3.1.2 Pointwise Multiplication

Once all highpass filtered components are obtained from each scale, pointwise multiplication of high frequency components from two consecutive scales in x and y direction is performed. Let the pointwise multiplication between at scale j and $j+1$ in x and y direction as P_x^j and P_y^j, respectively such that

$$\begin{aligned} P_x^j &= g_{H_x}^j \circ g_{H_x}^{j+1} \\ P_y^j &= g_{H_y}^j \circ g_{H_y}^{j+1} \end{aligned} \tag{3.7}$$

Since our method does not have downsampling step in decomposition process, the size of highpass filtered components is the same across the scale. As indicated from previous section, edges are preserved across the scale by Lipschitz regularity. Hence, pointwise multiplication of high frequency components between two consecutive scales should preserve edge information.

3.1.3 Modulus and Angle Computation and Thresholding

After performing the pointwise multiplication between two consecutive scales, modulus and angle matrix are found. Modulus and angle of all pixel locations are defined such that [59]

$$M_{xy}^{j} = \sqrt{P_x^j + P_y^j} \tag{3.8}$$

$$A_{xy}^{j} = \arctan\left(\frac{\operatorname{sgn}(g_{H_y}^{j}) \circ \sqrt{P_y^j}}{\operatorname{sgn}(g_{H_x}^{j}) \circ \sqrt{P_x^j}}\right). \tag{3.9}$$

Since the modulus matrix should contain noise, appropriate threshold (T) is selected to suppress noise. Applying T to modulus matrix produced a corresponding matrix $\bar{M}_{xy}^{j}$:

$$\bar{M}_{xy}^{j}(x,y) = \begin{cases} M_{xy}^{j}(x,y) & \text{if } M_{xy}^{j}(x,y) > T \\ 0 & \text{if } M_{xy}^{j}(x,y) \leq T \end{cases}. \tag{3.10}$$

Similar to the Canny edge detector [150], an edge pixel is determined wherever $\bar{M}_{xy}^{j}$ are local maximum in the direction of gradient by A_{xy}^{j}.

3.1.4 Nonmaximum Suppression

The last step of wavelet based multiscale edge detection method is nonmaximum suppression. The main idea of nonmaximum suppression is locating edges at the points of maxima and suppress non-maxima points. Canny edge detector [150] also utilizes this to detect the primary edge candidates followed by hysteresis thresholding.

As portrayed in Figure 3.3, our proposed method performs nonmaximum suppression in eight directions including horizontal, vertical and diagonal directions. For instance, if $A^j(x,y) = 0°$, the edge is in the north-south direction. Then, the point is considered as edge if its $\bar{M}^j(x,y)$ is greater than the pixel of north ($\bar{M}^j(x,y-1)$) and south direction ($\bar{M}^j(x,y+1)$).

1. If $A^j(x,y) = 0°$, $E^j(x,y) = 1$ when $\bar{M}^j(x,y) > \bar{M}^j(x,y-1)$ and $\bar{M}^j(x,y) > \bar{M}^j(x,y+1)$.

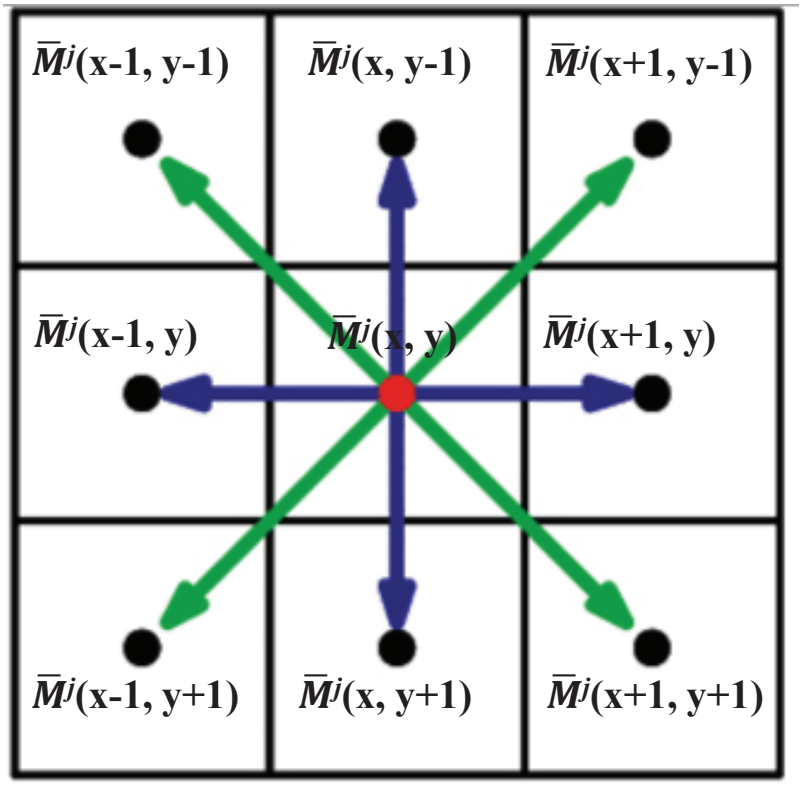


Fig. 3.3. Nonmaximum suppression in eight directions

2. If $A^j(x,y) = 45°$, $E^j(x,y) = 1$ when $\bar{M}^j(x,y) > \bar{M}^j(x-1,y-1)$ and $\bar{M}^j(x,y) > \bar{M}^j(x+1,y+1)$.

3. If $A^j(x,y) = 90°$, $E^j(x,y) = 1$ when $\bar{M}^j(x,y) > \bar{M}^j(x-1,y)$ and $\bar{M}^j(x,y) > \bar{M}^j(x+1,y)$.

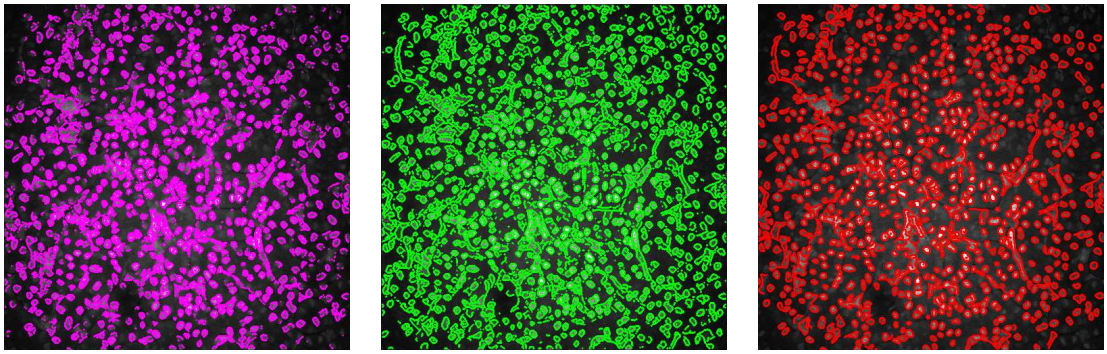
4. If $A^j(x,y) = 135°$, $E^j(x,y) = 1$ when $\bar{M}^j(x,y) > \bar{M}^j(x+1,y-1)$ and $\bar{M}^j(x,y) > \bar{M}^j(x-1,y+1)$.

Final edge map (E^j) using DWT coefficients at the scale j and $j+1$ is attained after performing nonmaximum suppression in eight directions.

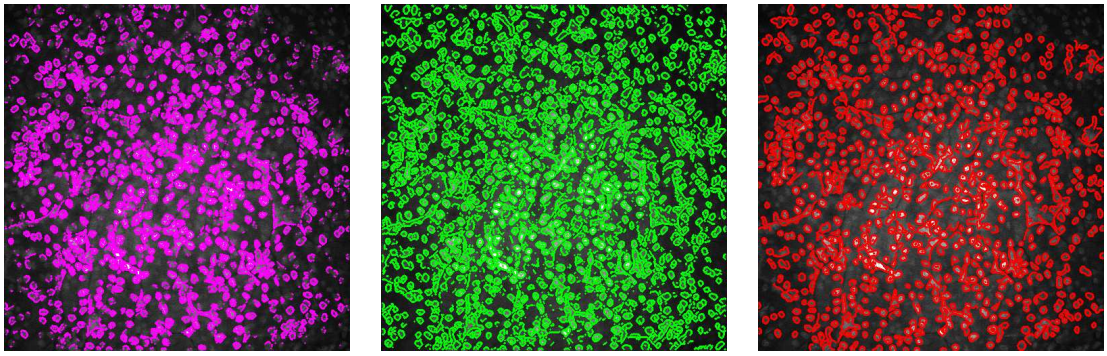
3.2 Experimental Results

Our proposed wavelet based multiscale edge detection method was tested on two different datasets, namely *WSM* and *Lectin*. The first set of images (*WSM*) was comprised of 512 images each 512×512 pixels in size, while the second set (*Lectin*) consists of 821 images each 640×640 pixels. Our datasets consist of 3D volumes of a rat kidney labeled with Hoechst 33342 collected using two-photon microscopy. The main goal of our proposed multiscale edge detection is to detect edges of nuclei mainly appeared in the B channel of each dataset. Therefore we separated the B channel from each dataset and tested using three images from different depths. In particular, 100^{th}, 200^{th}, and 300^{th} from the stack of images are used from *WSM* and *Lectin*, respectively.

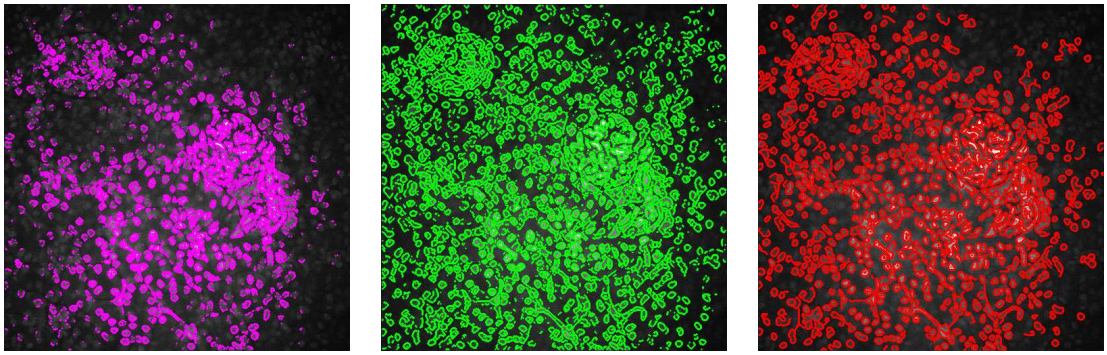
Figure 3.4 and Figure 3.5 portrays the proposed edge detection results and comparison with Canny edge detector. Each figure contains 9 image results with various depth. The first column indicates proposed edge detection results using highpass components at scale 1 and 2 highlighted in magenta. Similarly, the second column shows the overlaid results onto original highlighted in green using highpass components at scale 2 and 3. Last column depicts Canny edge detection results overlaid onto three different depth images. The threshold values for each images are determined by a function of scale j. Using highpass filter results at scale 1 and 2, the threshold value we take is $T = 38$ whereas with scale 2 and 3 highpass components, the threshold is just $T = 14$. Be advised that these thresholds are not directly compared with image intensities but compared with highpass filtered results. Scale increases in wavelet domain, highpass components contain finer edge information compared to lower scale filtered components. Therefore, different thresholds should be utilized with respect to different scale level for our proposed edge detection method. Due to this fact, higher thresholds are used at the first column than second column which used scale 1 and 2 highpass components and used scale 2 and 3 highpass components, respectively.



(a) $T = 38$ (b) $T = 14$ (c) $T_L = 18$, $T_H = 44$

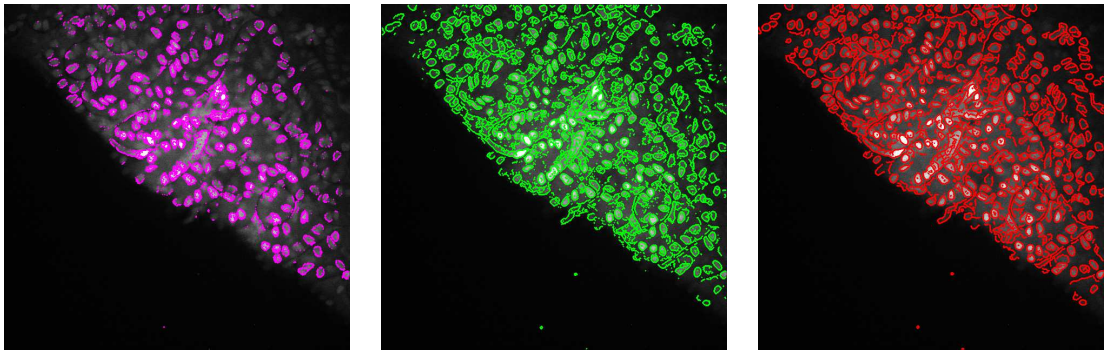


(d) $T = 38$ (e) $T = 14$ (f) $T_L = 14$, $T_H = 36$

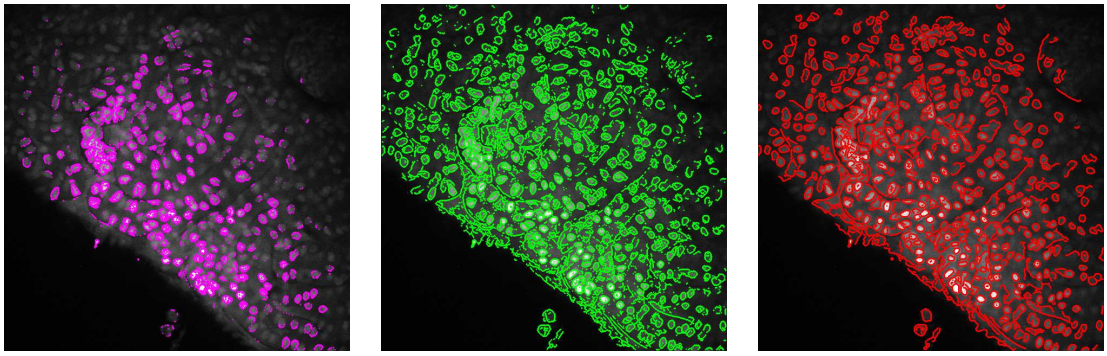


(g) $T = 38$ (h) $T = 14$ (i) $T_L = 10$, $T_H = 24$

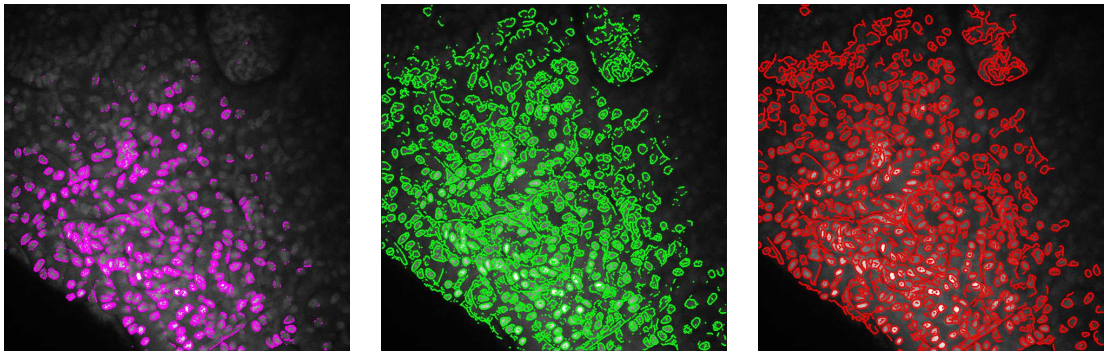
Fig. 3.4. Edge detection results from proposed wavelet based method at various depth for *WSM* dataset. Left column: Wavelet edge detection results using highpass components at scale 1 and 2 overlaid to 100^{th}, 200^{th}, 300^{th} images. Middle column: Wavelet edge detection results using highpass components at scale 2 and 3 overlaid to 100^{th}, 200^{th}, 300^{th} images. Right column: Corresponding overlaid Canny edge detection images



(a) $T = 38$ (b) $T = 14$ (c) $T_L = 6$, $T_H = 16$



(d) $T = 38$ (e) $T = 14$ (f) $T_L = 6$, $T_H = 16$



(g) $T = 38$ (h) $T = 14$ (i) $T_L = 8$, $T_H = 20$

Fig. 3.5. Edge detection results from proposed wavelet based method at various depth for *Lectin* dataset. Left column: Wavelet edge detection results using highpass components at scale 1 and 2 overlaid to 100^{th}, 200^{th}, 300^{th} images. Middle column: Wavelet edge detection results using highpass components at scale 2 and 3 overlaid to 100^{th}, 200^{th}, 300^{th} images. Right column: Corresponding overlaid Canny edge detection images

Note that our proposed edge detection method needs only one threshold used for modulus matrix values to attain thresholded modulus matrix, $\bar{M}^j_{xy}(x, y)$. Different from our method, Canny edge detector uses a nonmaximum suppression followed by hysteresis thresholding. The hysteresis thresholding utilizes two thresholds, low threshold (T_L) and high threshold (T_H). If compared modulus matrix values using Canny edge detection method are higher than T_H, that corresponding pixel locations are considered as strong edges and preserve the locations as edges. Conversely, if compared values are smaller than T_H but larger than T_L, that corresponding pixel locations are called weak edges. A rule of thumb to select T_L is 0.4 times less than T_H. Once appropriate high threshold value T_H, T_L is selected, T_L is determined by automatically. This T_H values are determined empirically as our proposed method does. Each threshold value is shown in Figure 3.4 and Figure 3.5.

As mentioned from Chapter 1 the microscopy volumes are inherently anisotropic. To be more specific, image contrast decreases with respect to depth of image volume. Due to this characteristic, successful segmentation or edge detection in deeper tissue still poses a challenge. In addition, microscopy images contrast varies even in same depth. For example, we can clearly observe that the center of each sliced image contains high contrast compared to the boundary of image. This also makes correct segmentation difficult.

Despite these microscopy characteristics, our proposed method successfully detects most nuclei edges as can be seen from Figure 3.4 and Figure 3.5. In general, using scale 2 and 3 highpass component results shown in the second column produces better performance than using scale 1 and 2 highpass component results in the first column. In particular, middle column results in both datasets have more robust to noise and capture more nuclei at the corner of images than left column's one. Compared with Canny edge detector presented in the right column, results from our method detects more nuclei but it may not be robust enough to noise at the same time. Also since both methods are edge detection methods, they often fail to detect entire nuclei boundaries but parts of them.

Our future work after segmentation will be quantifying characteristics, such as volume, shape, and orientation of each biology entities presented in the volume. In this perspective detecting closed boundaries of each different nuclei is vital since connected components will be mainly used for quantification. Additionally, both methods used threshold and the edge detection results heavily rely on thresholding values. Selecting thresholds empirically in each image is time consuming and inefficient. Due to these drawbacks we abandoned this wavelet multiscale edge detection approach later and came back to active contours method. Since active contours method uses the level set function, the segmentation results produced by this method will be closed path. We will demonstrate nuclei segmentation method with 3D active contours with 3D inhomogeneity correction method in Chapter 5.4.

4. BOUNDARY FITTING BASED SEGMENTATION

In this chapter boundary fitting based segmentation technique is described[1]. More specifically, we describe a technique that segments 2D cross-sections of tubules in microscopy images using a combination of adaptive and global thresholding, potentials, z direction refinement, branch pruning, end point matching, and curve fitting. The aim here is to be able to segment the 2D cross-sections at various depths as an initial step to 3D segmentation of these objects. The performance of the proposed technique is compared to that of an active contours based method [46]. The tubular structures are successfully delineated by proposed method but an active contour based method does not overcome the inhomogeneity of given microscopy images. We discuss more about this problem in the Chapter 5.

4.1 Proposed Scheme

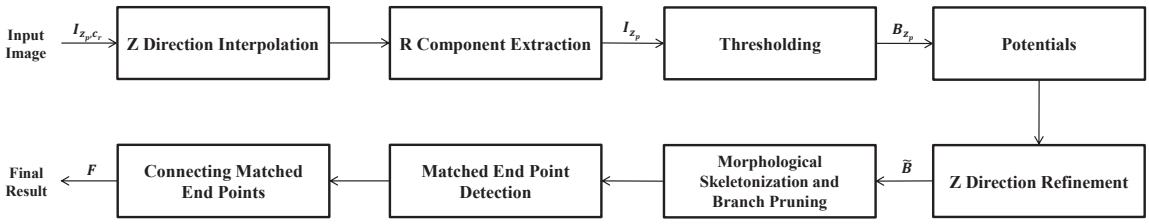


Fig. 4.1. Overview of the proposed tubule boundary segmentation technique

[1]The boundary fitting based segmentation technique described in this chapter is appeared in [51].

4.1.1 Z Direction Interpolation and R Component Extraction

As mentioned above the objective of this work is to segment tubular structures present in microscopy images. Figure 4.1 highlights the various stages of the proposed technique. Representative images from two of the datasets being processed are provided in Figure 4.4a and Figure 4.5a. Figure 4.4a shows an image from a 3D volume of rat kidney labeled with Hoechst 33342 (blue) and fluorescent phalloidin (red) dyes whereas Figure 4.5a presents an image from a 3D volume of rat kidney labeled with Hoechst 33342 (blue) and a fluorescent lectin (red) dye. Both datasets were imaged using multiphoton fluorescence excitation microscopy [9] and we denote the microscopy volume as I_{z_p,c_r} as introduced from the Chapter 1.2. Note that there is no time-lapse volume so time index is omitted. In both cases the structures of interest reside mostly in the red (R) component of the data. Thus, as shown in Figure 4.1, the R component of each image is initially extracted denoted as I_{z_p,c_1}. This is preceded however by an interpolation step where we use cubic interpolation to compensate for the fact that the resolution along the z direction is smaller than along the x and y direction. Here, we drop the subscript c_1 for the red channel images in this chapter for the simplicity.

4.1.2 Thresholding

Figure 4.4b and Figure 4.5b depict the R components of the images shown in Figure 4.4a and 4.5a, respectively. As can be expected and observed both images have non-homogeneous intensities. Due to this non-homogeneity, we first utilize an adaptive thresholding scheme [32] that obtains local thresholds T_{z_p} at each location (x, y) based on a local neighborhood W centered at (x, y):

$$T_{z_p} = m_{z_p} \cdot \left[1 + k \cdot \left(\frac{s_{z_p}}{R} - 1\right)\right], \tag{4.1}$$

where m_{z_p} and s_{z_p} are the empirical mean and variance of the pixel values within region W respectively, k a tuning parameter, and R the dynamic range of s_{z_p}. The

values of W, R, and k used in this chapter are different from those proposed in [32] and will be provided in Chapter 4.2.

Due to the fluctuations in pixel intensities, adaptive thresholding produces false contours especially in dark regions. To address this problem, we add another global threshold T_G on top of the T_{z_p} to produce a final threshold T'_{z_p} defined as:

$$T'_{z_p} = \max(T_{z_p}, T_G). \tag{4.2}$$

The value of T_G is empirically chosen that most effectively reduces the number of false contours. Utilizing the threshold T'_{z_p} to each red channel image (I_{z_p}), the binary image (B_{z_p}) is obtained as:

$$B_{z_p} = \begin{cases} 1 & \text{if } I_{z_p} > T'_{z_p} \\ 0 & \text{if } I_{z_p} \leq T'_{z_p} \end{cases}. \tag{4.3}$$

We also investigated using conventional edge detection methods such as the Canny operator [150] to detect object boundaries. However, due to the nature of the microscopy images, the Canny operator did not produce good enough results compared to the modified adaptive thresholding method used here. To be more specific, the results of the Canny operator wrongly detected noise region as a part of the object boundary.

4.1.3 Potentials

B_{z_p} generally highlights object boundaries as well as lumen that appear as edges interior to an object, as shown in Figure 4.4c and 4.5c, respectively. The objective is to retain the object boundaries only. This is achieved by using potential functions and z direction refinement. We employ two different potential functions: potential for position (P^P) and potential for occupancy (P^O), weighted by two coefficients λ_1 and λ_2, respectively, to create total potential $P^T = \lambda_1 P^P + \lambda_2 P^O$. This potential can be used to determine foreground and background regions. The description of each potential function follows:

1. Potential for position, P^P.

   Since B_{z_p} contains unwanted edges, suppressing them should be a necessary step. To do this, we first stack all the binary images B_{z_p} along the z direction to form a binary volume denoted as B. P^P is then obtained by convolving the binary volume with a 3D Gaussian filter $h(x, y, z)$:

   $$P^P = B * h(x, y, z). \tag{4.4}$$

   The Gaussian filter used here is $h(x, y, z) = \frac{1}{\sqrt{(2\pi)^3}} \exp\left(-\frac{x^2+y^2+z^2}{2}\right)$ and $*$ represents 3D convolution. Since B is a binary volume, the value of P^P at each voxel is always between 0 and 1.

2. Potential for occupancy, P^O.

   Since B_{z_p} highlights both objects and lumens, P^O is designed such that pixels on the outer boundaries are assigned larger values and small values given to interior lumens. This is achieved by first labeling B_{z_p} using 2D connected components with an 8-neighborhood. In this case a connected component with an 8-neighborhood is a set of pixels connected horizontally, vertically, or diagonally, that have the same characteristics or features such as belonging to the same edge. To distinguish the different connected components, each component is assigned an unique label. Starting from an unlabeled non-zero pixel in B_{z_p}, we assign the same label to all pixels that are connected to it within its 8-neighborhood. In addition, the neighbors of all the labeled current pixel's neighbors are given the same label. This process repeated until all the connected neighbors have been identified. Once a single connected component has been identified, we locate the next non-zero unlabeled pixel in raster scan order and assign a different label to all the pixels in its associated neighborhood. The process of identifying and labeling another connected component is then repeated. This is continued until all the non-zero pixels in B_{z_p} have been labeled. At this stage all the pixels in B_{z_p} have been grouped into separate connected components that are identified by their unique labels. In addition, the size and

the location of each connected component are recorded. Here, the size of each connected component denoted as s_{CC} is defined to be the number of pixels that constitute the connected component. Moreover, the smallest bounding rectangle containing all the pixels of each connected component is constructed for each connected component. The fraction denoted as S of how much each bounding rectangle is occupied by its corresponding connected component's pixels is defined as

$$S \triangleq \frac{s_{CC}}{s_{BB}}, \tag{4.5}$$

where s_{BB} denotes the area of the bounding rectangle. The ratio S for each labeled connected component is obtained and normalized such that the smallest S from B_{z_p} maps to 0 and the largest S maps to 1. We assign $P^O_{z_p}$ to have a large value when S is small and vice-versa so that the value of the $P^O_{z_p}$ has relatively larger value at pixels on tubule boundaries than pixels interior to the object. We define $P^O_{z_p}$ to be $P^O_{z_p} = 1 - S$ at each x, y and z corresponding to the connected component's location. Once each image's corresponding $P^O_{z_p}$ is obtained, we stack them up to form a 3D potential for occupancy denoted as P^O.

3. Total potential, P^T.

   Once we obtain P^P and P^O, total potential (P^T) is obtained for each voxel location using

$$P^T = \lambda_1 P^P + \lambda_2 P^O, \tag{4.6}$$

   where λ_1 and λ_2 are non-negative weights such that $\lambda_1 + \lambda_2 = 1$. The value of weights λ_1, λ_2 are empirically determined and appeared in Chapter 4.2. The value of P^T is thresholded using a threshold T_P to produce a thresholded total potential $\tilde{P}^T$ where

$$\tilde{P}^T = \begin{cases} 1 & \text{if } P^T > T_P \\ 0 & \text{if } P^T \leq T_P \end{cases}. \tag{4.7}$$

If $\tilde{P}^T$ is 1, the corresponding voxel is considered to be part of an object/tubule boundary. If however, $\tilde{P}^T$ is 0, then the voxel is deemed to be belonging to lumen. This combination of two potentials successfully removes most of the lumen and retains boundary.

4.1.4 Z Direction Refinement

To further refine $\tilde{P}^T$, we employ z direction information. Object/tubules boundaries are generally continuous and consistent along the z direction whereas lumen are not. We thus use a 1D Gaussian filter ($\psi(z)$) to convolve $\tilde{P}^T$ along the z direction to obtain the final potential P^F as:

$$P^F = \tilde{P}^T * \psi(z), \tag{4.8}$$

where $\psi(z) = \frac{1}{\sqrt{2\pi\sigma_z^2}} \exp\left(-\frac{z^2}{2\sigma_z^2}\right)$. P^F is then thresholded again using threshold T_P to produce a binary volume. The image obtained after applying the potentials and z direction refinement is denoted by $\tilde{B}_{z_p}$. The parameter values that we use for the potentials and z direction refinement are provided in Chapter 4.2.

4.1.5 Morphological Skeletonization and Branch Pruning

While the previous procedure removes interior edges, as seen from Figure 4.4d and Figure 4.5d, not all resulting boundaries are fully connected and many have random protruding branches that are pruned through the following procedure:

1. Create a morphological skeleton. In this case we define a morphological skeleton to be the outcome of performing morphological thinning on $\tilde{B}_{z_p}$. In particular, morphological thinning is achieved by iteratively performing morphological erosion using a 2×2 square structuring element to achieve a single-pixel-width skeleton [152].

2. Identify end points and branch points along each boundary. An end point is a point that has one or fewer neighbors in the morphological skeleton. A branch

point is defined as a point that has 3 or more neighbors in the morphological skeleton. Record the number of end points before branch pruning and denote it as E_p^{i-1} where i is the number of branch pruning iterations.

3. Trace back a maximum of 7 edge pixels from each end point. While tracing back, obtain the locations of the pixels on the traversed path and check whether they are branch points.

4. For each end point, retain all pixels on the back path if there are no branch points. However, if there is a branch point among the pixels on the back path, then remove all pixels connecting the end point to the branch point but keep the branch point.

5. Repeat for all end points and count the number of end points after branch pruning. Denote it as E_p^i.

6. Perform branch pruning steps $(2-5)$ iteratively until a stopping criterion is met. This criterion is based on the relative change in number of end points at each iteration. In particular, define Q as

$$Q \triangleq \frac{|E_p^{i-1} - E_p^i|}{E_p^{i-1}}. \tag{4.9}$$

The iterative branch pruning is terminated when Q falls below 0.1.

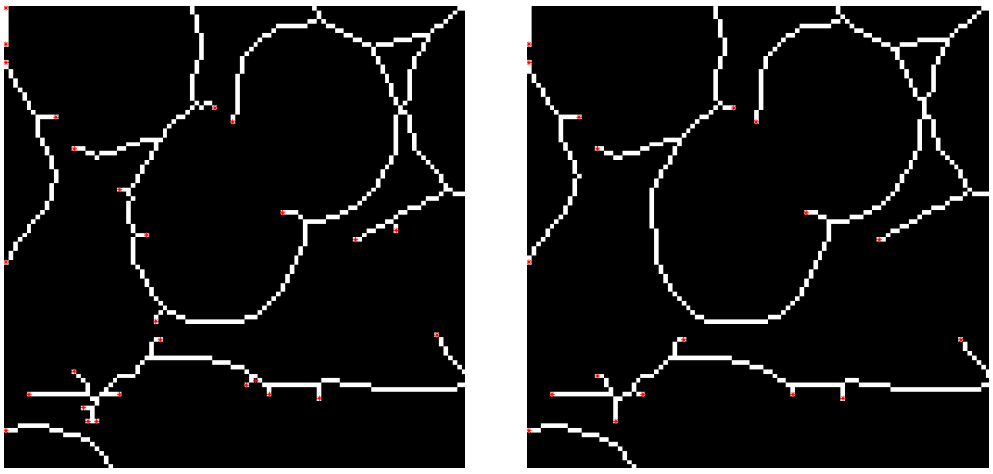
Figure 4.2 shows the outcome of applying the above procedure to the image in Figure 4.4d. As observed, small branches are successfully removed.

4.1.6 Matched End Point Detection and Reconnection

The next step is to reconnect the entire boundary. This is accomplished as follows:

1. Identify all end points. This is done as described above.

2. Obtain the Euclidean distance between all pairs of end points. If two end points are closer than a certain threshold τ_D, they are considered to be points on a



(a) Before pruning small branches (b) After pruning small branches

Fig. 4.2. Example of pruning small branches from a morphological skeleton

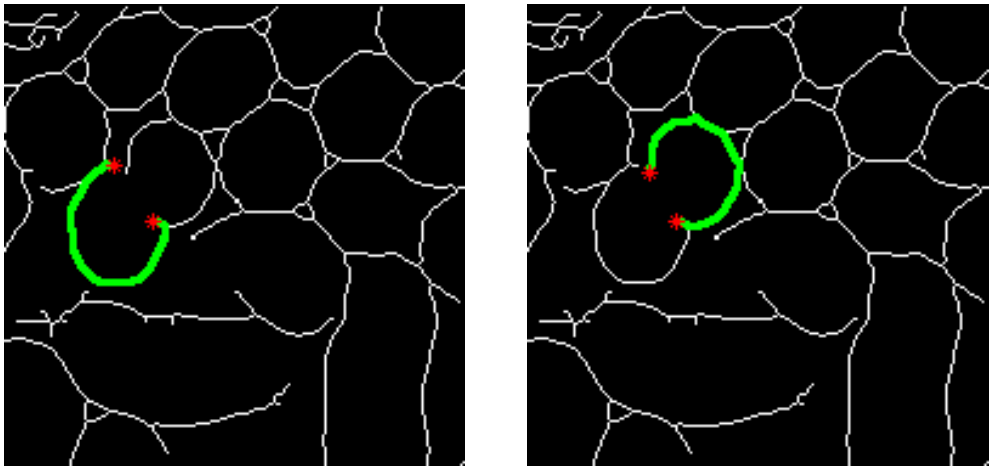
common boundary and need to be connected to close the boundary. Such points are denoted as matching end points.

3. Having found all pairs of matching end points, obtain the "shortest path", based on geodesic distance [153], between any two matching end points while traversing that part of the boundary that already joins them. Figure 4.3 depicts two examples of shortest paths between two end points. In Figure 4.3, the matching end points are marked in red and the corresponding "shortest paths," based on geodesic distance, are highlighted in green.

4. Find an ellipse that best (in the least square sense) fits the shortest path between the two matching end points, while passing through the end points. This is formulated as a constrained least squares curve fitting problem. In particular, using the general form for conic sections

$$a\mathbf{x} \circ \mathbf{x} + 2b\mathbf{x} \circ \mathbf{y} + c\mathbf{y} \circ \mathbf{y} + 2d\mathbf{x} + 2f\mathbf{y} + g = 0, \tag{4.10}$$

where $\mathbf{x}$ and $\mathbf{y}$ are column vectors representing the x and y values of each pixel location on the shortest path and $\circ$ represents Hadamard product. Di-



(a) (b)

Fig. 4.3. Two "shortest paths" between two matching end points. The end points are indicated in red and the "shortest paths" in green.

viding both sides by a and re-arranging, we obtain $M\mathbf{p} = \mathbf{q}$, where $M = \begin{bmatrix} 2\mathbf{x} \circ \mathbf{y} & \mathbf{y} \circ \mathbf{y} & 2\mathbf{x} & 2\mathbf{y} & \mathbf{1} \end{bmatrix}$, $\mathbf{p} = \begin{bmatrix} b/a & c/a & d/a & f/a & g/a \end{bmatrix}^T$, and $\mathbf{q} = -\mathbf{x} \circ \mathbf{x}$. Note that $\mathbf{1}$ is a vector whose entries are all 1. In addition, to ensure this formulation will form an ellipse we impose $b^2 - 4ac = -1$ as in [154] rather than imposing the general constraint $b^2 - 4ac < 0$. This leads to an overdetermined system if all the points lying on the shortest path between the matching end points are taken into account. Thus, to solve for the unknown coefficients in $\mathbf{p}$, we find $\mathbf{p}$ that minimizes the squared error $\frac{1}{2}\|M\mathbf{p} - \mathbf{q}\|_2^2$, while ensuring that the ellipse passes through the end points. Thus, we pose this as follows:

$$\min_{\mathbf{p}} \quad \frac{1}{2}\|M\mathbf{p} - \mathbf{q}\|_2^2 \tag{4.11}$$

$$\text{subject to} \quad G\mathbf{p} = \mathbf{h}$$

$$b^2 - 4ac = -1, \tag{4.12}$$

$$\text{where} \quad G = \begin{bmatrix} 2x_{e_0}y_{e_0} & y_{e_0}^2 & 2x_{e_0} & 2y_{e_0} & 1 \\ 2x_{e_1}y_{e_1} & y_{e_1}^2 & 2x_{e_1} & 2y_{e_1} & 1 \end{bmatrix}$$

$$\mathbf{h} = \begin{bmatrix} -x_{e_0}^2 & -x_{e_1}^2 \end{bmatrix}^T.$$

Here (x_{e_0}, y_{e_0}), (x_{e_1}, y_{e_1}) are the locations of the two matching end points, e_0 and e_1, respectively. Once the coefficients have been found, they are used to join the matching end points using an elliptical curve.

After having completed the above procedures, there remain some exceptions. They are:

1. The solution to the constrained optimization problem does not result in an ellipse, or

2. The elliptical arc between two matching end points extends beyond the image boundary, or

3. The elliptical arc between two matching end points interferes with boundaries already identified.

In these three cases, the matched end points are left unconnected.

In addition, there are some end points that do not belong to the same connected component, yet are within a distance τ_D of each other. Such points are joined by a straight line unless the line intersects another already existing boundary. In the case of the latter, the points are left unconnected. Finally, a search is carried out for all remaining unconnected end points. These are connected to the closest object boundary via a straight line if the distance between the unconnected end points and closest object boundary is within $\tau_D/2$. In particular, the straight line is extended in one of three directions: $-45°$, $0°$, $45°$ relative to the direction of the end point. However, if the distance is larger than $\tau_D/2$, the end points are left unconnected. The resulting image is denoted by F_{z_p}.

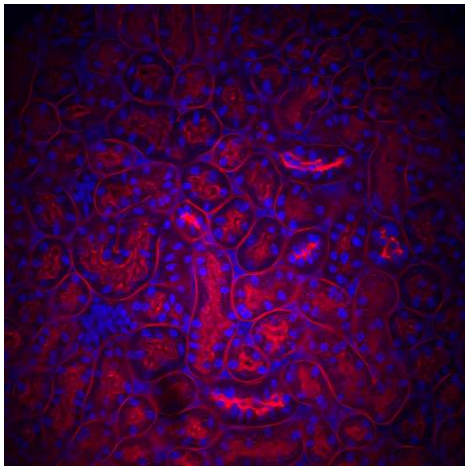
4.2 Experimental Results

Our proposed boundary fitting based segmentation technique was evaluated using two different 3D microscopy datasets. As indicated from the beginning of this chapter, capturing fluorescent phalloidin, tubular shape resided in R channel, is our main goal of this technique. The results of applying the proposed technique to two datasets with representative images shown in Figure 4.4a and Figure 4.5a, respectively, are given here. Note that these datasets are the same datasets introduced from Chapter 3 but different color channel was used for this proposed technique.

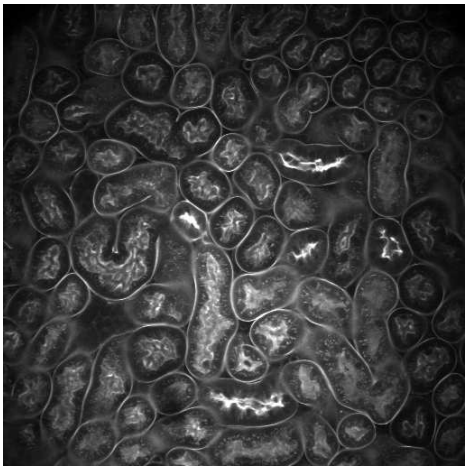
Table 4.1.
Parameters used in our proposed boundary fitting based tubule segmentation technique

Parameters	Description	*WSM*	*Lectin*
k	Sauvola's parameter	-0.1	-0.1
T_G	Global threshold used to modify the local threshold	0	15
λ_1	Coefficient corresponding to P^P	0.25	0.25
λ_2	Coefficient corresponding to P^O	0.75	0.75
T_P	Threshold used for the total potential P^T and z direction refinement	0.8	0.8
σ_z	Standard deviation used for z direction refinement	0.5	0.5
τ_D	Matched end points distance threshold	35	40

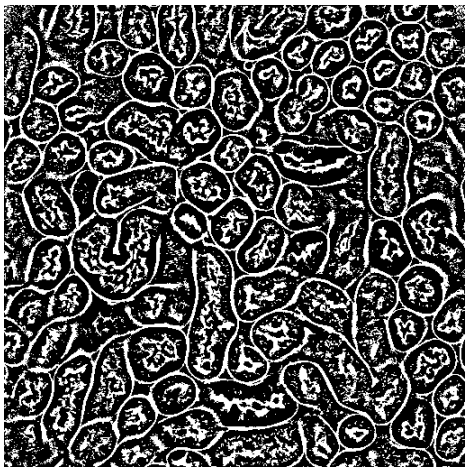
The structures of interest reside mostly in the R components shown in Figures 4.4b and 4.5b, respectively. These were then thresholded as depicted in Figures 4.4c and 4.5d, respectively. In this case the window W used for an adaptive thresholding for the first dataset was 16×16 pixels in size, whereas it was 20×20 for the second dataset. In addition, the value of R used was the maximum empirical variance and ensured that $\frac{s_{z_p}}{R} \leq 1$. Finally, the value of the parameter k was empirically determined to



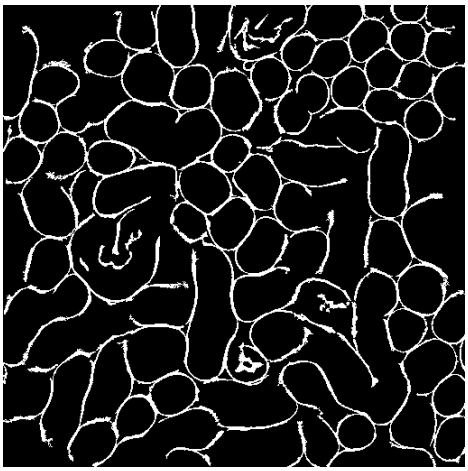
(a) Original image, I_{z_{81},c_r}



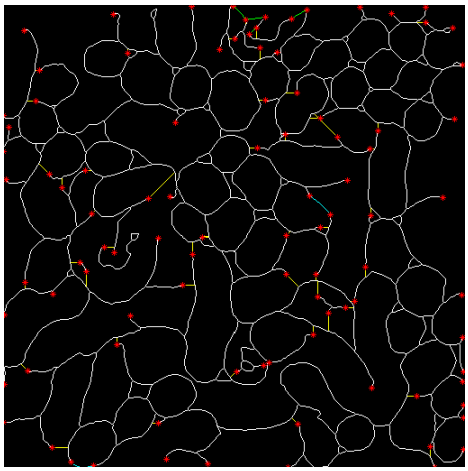
(b) Red component, $I_{z_{81}}$



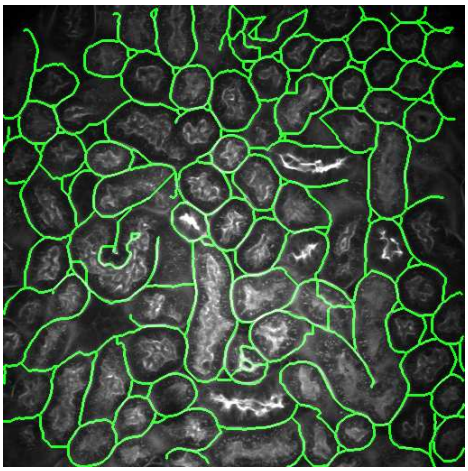
(c) $B_{z_{81}}$



(d) $\tilde{B}_{z_{81}}$

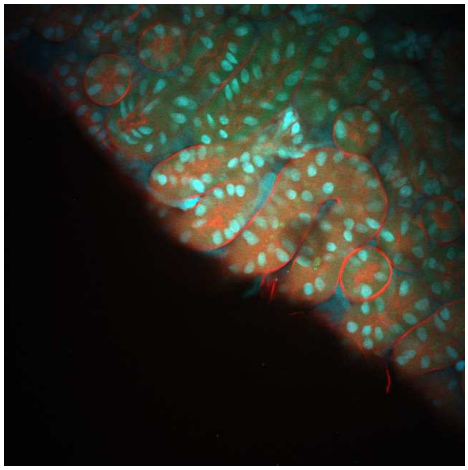


(e) $F_{z_{81}}$

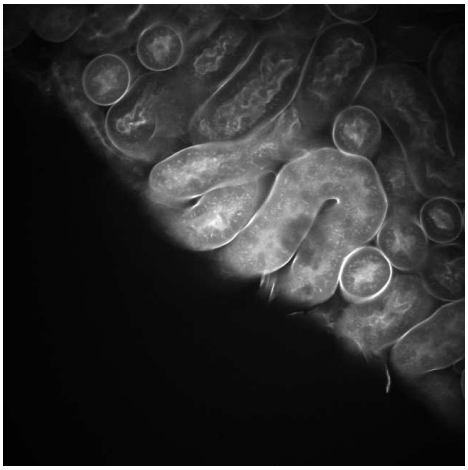


(f) $F_{z_{81}}$ overlaid onto $I_{z_{81}}$

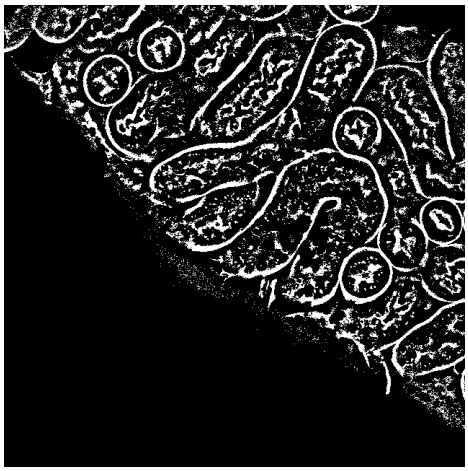
Fig. 4.4. Results of applying the proposed technique to the $I_{z_{81}}$ of the *WSM* dataset



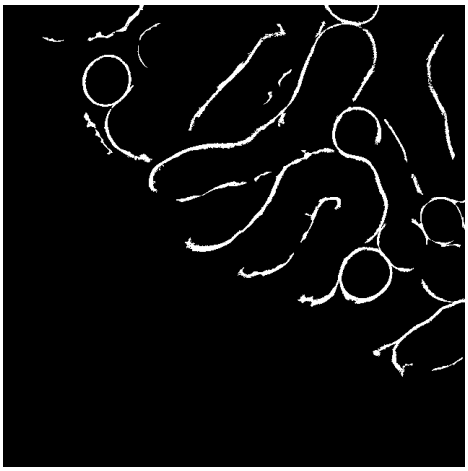
(a) Original image, I_{z_{101},c_r}



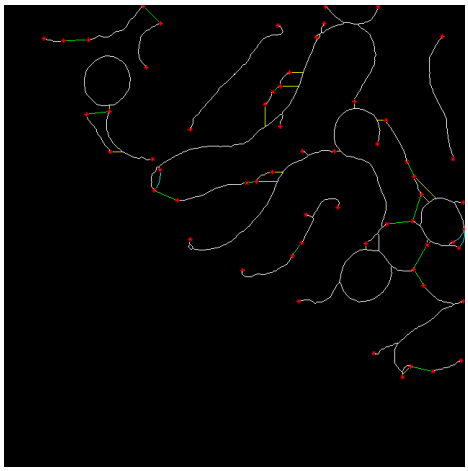
(b) Red component $I_{z_{101}}$



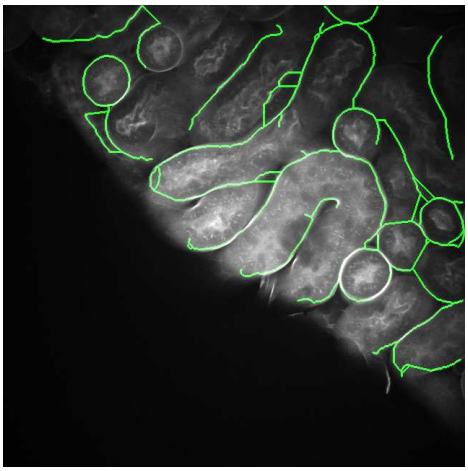
(c) $B_{z_{101}}$



(d) $\tilde{B}_{z_{101}}$



(e) $F_{z_{101}}$



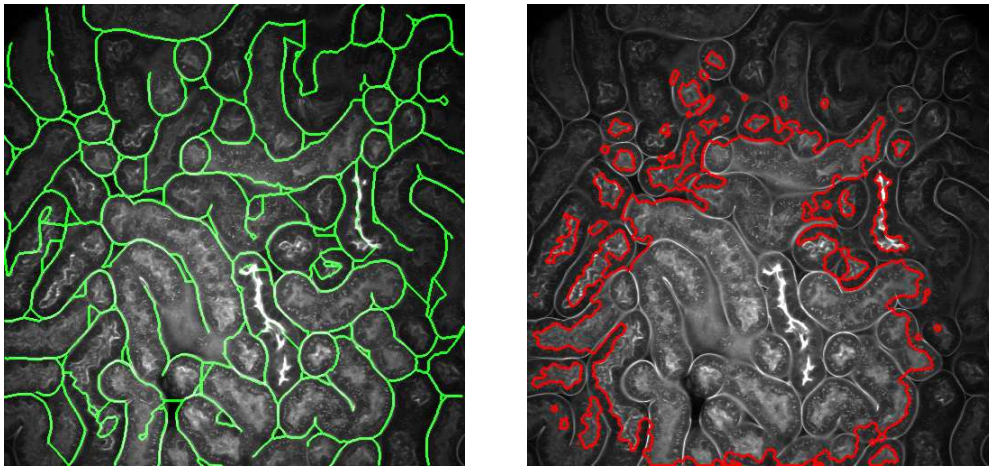
(f) $F_{z_{101}}$ overlaid onto $I_{z_{101}}$

Fig. 4.5. Results of applying the proposed technique to the $I_{z_{101}}$ of the *Lectin* dataset

be -0.1. Table 4.1 lists the various parameters used and their corresponding values. Note that different values are used for different datasets. Again, the values used were empirically determined.

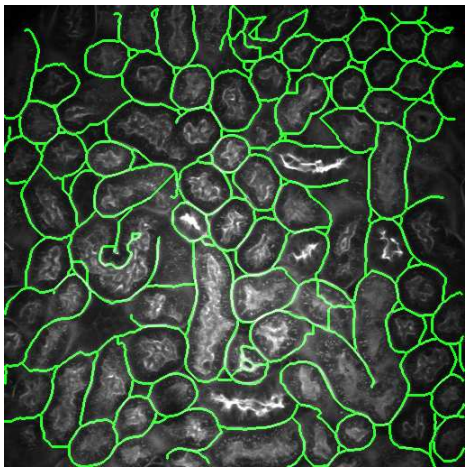
As described above the interiors were cleared out using connected component labeling and the boundaries pruned of any protruding branches. The outcome of these two steps is depicted in Figures 4.4d and 4.5d, respectively. Subsequently, the boundaries were closed. In Figure 4.4e and Figure 4.5e, red dots delineate the locations of all the end points, cyan curves indicate boundaries that were closed using ellipse fitting based on constrained least squares, green curves highlight the connection between two end points that do not belong to the same connected component, and the yellow straight lines are the extensions from non-matched end points to the closest boundary. The final boundaries are overlaid onto the original images as shown in Figures 4.4f and 4.5f, respectively.

In Figure 4.6 and Figure 4.7, we compare the results of the proposed scheme to the 2D region-based active contour technique described in [46], when applied to images at various depths. The first and third rows exhibit the results of proposed scheme whereas the second and the fourth rows show the corresponding results using [46]. The 2D region-based active contour technique [46] needed 1000 iterations to converge to a solution as well as multiple circles to be used as initial contours. Furthermore, compared to the proposed technique, it did not capture as many boundaries of the various tubules. In particular, we observe that it successfully groups brighter regions in the images into one large object, but fails to delineate smaller objects or any objects for that matter in the darker regions.

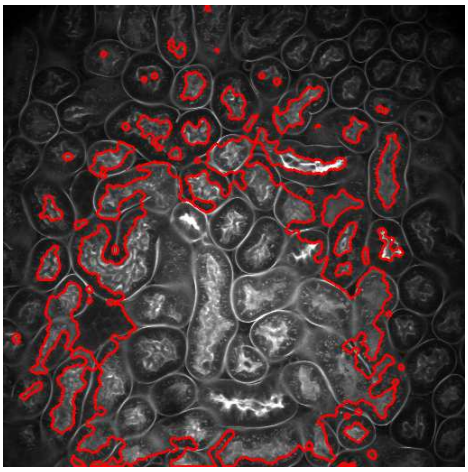


(a) $I_{z_{31}}$ of *WSM* (Proposed)

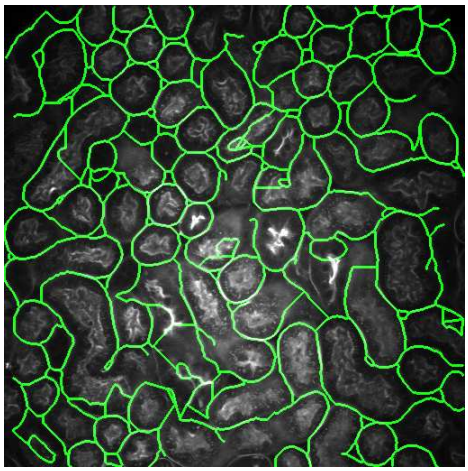
(b) $I_{z_{31}}$ of *WSM* [46]



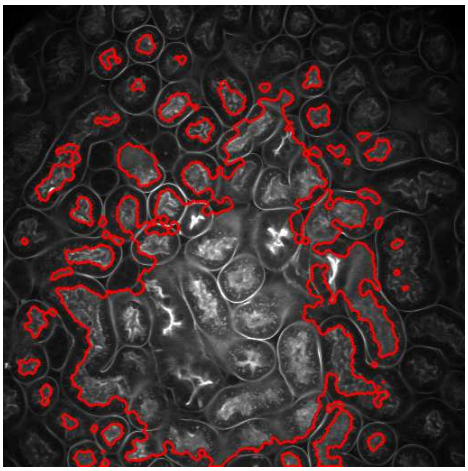
(c) $I_{z_{81}}$ of *WSM* (Proposed)



(d) $I_{z_{81}}$ of *WSM* [46]

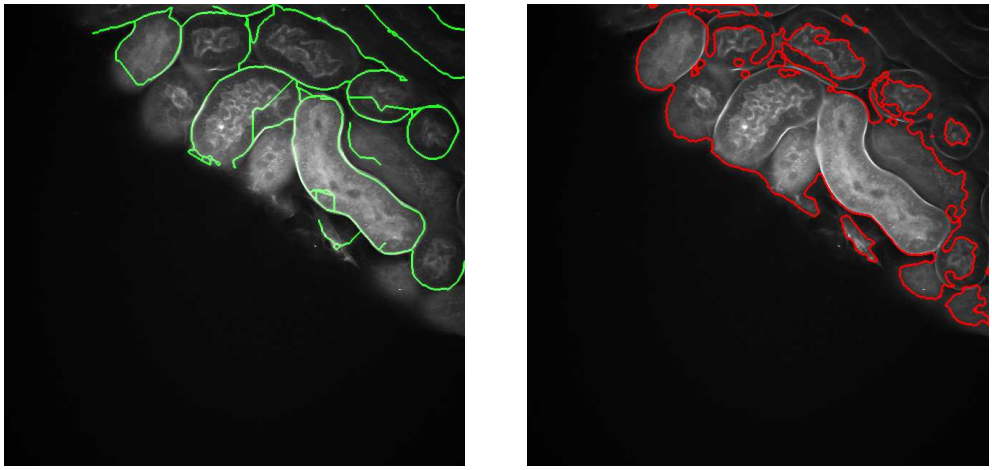


(e) $I_{z_{131}}$ of *WSM* (Proposed)



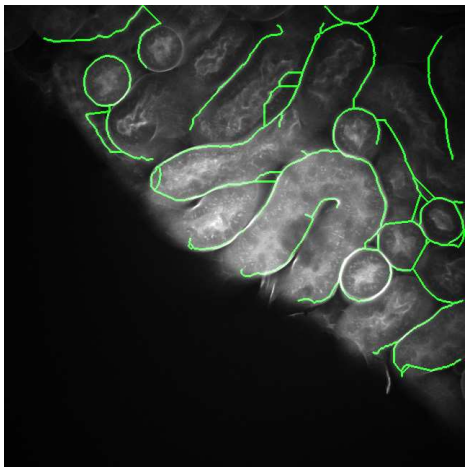
(f) $I_{z_{131}}$ of *WSM* [46]

Fig. 4.6. Segmentation results comparison with 2D active contours and proposed schemes of the *WSM* dataset at various depth

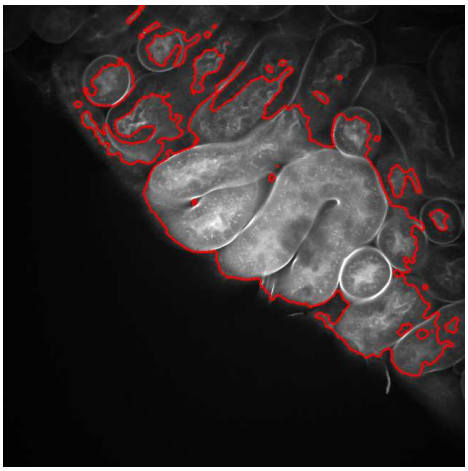


(a) $I_{z_{41}}$ of *Lectin* (Proposed)

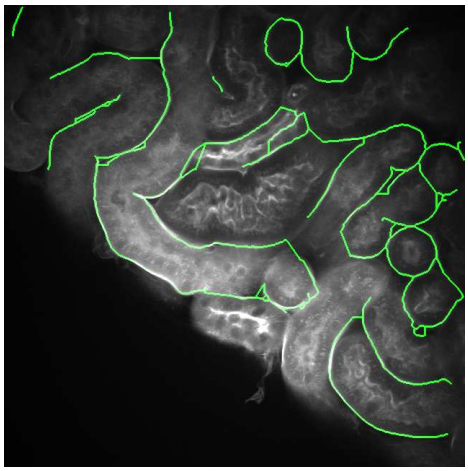
(b) $I_{z_{41}}$ of *Lectin* [46]



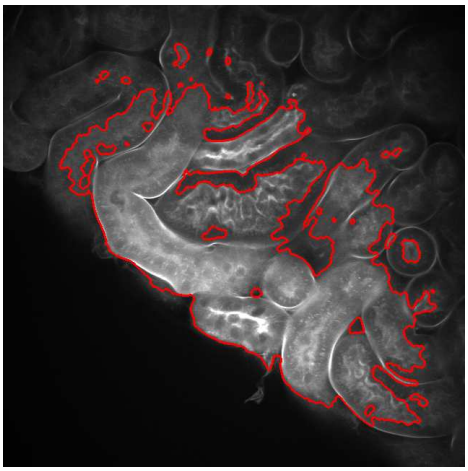
(c) $I_{z_{101}}$ of *Lectin* (Proposed)



(d) $I_{z_{101}}$ of *Lectin* [46]



(e) $I_{z_{161}}$ of *Lectin* (Proposed)



(f) $I_{z_{161}}$ of *Lectin* [46]

Fig. 4.7. Segmentation results comparison with 2D active contours and proposed schemes of the *Lectin* dataset at various depth

5. THREE DIMENSIONAL ACTIVE CONTOURS WITH INHOMOGENEITY CORRECTION

Various active contours methods are introduced in this chapter. More specifically, our proposed 3D active contours with inhomogeneity correction[1] as well as traditional edge-based and region-based active contours are demonstrated.

Active contours method is one of the well-known methods especially for biomedical images segmentation. Active contours, also called snakes, firstly introduced in [29], has been investigated for several decades and still ongoing method for research. The main idea of snakes is to deform an initial contour and fit the contour to desired segmentation regions. To be more specific, having recursively evolving an initial contour, fit a desired object based on the minimization of an energy functional. Depending on formulating an energy functional there exists two different categories so called edge-based and region-based active contour methods. We demonstrate the fundamental concepts of two different kinds of active contours and their limitations. After that, we introduce our proposed 3D active contours with inhomogeneity correction method to effectively identify nuclei in microscopy volumes.

5.1 Edge-Based Active Contours

Edge-based active contours method obtains gradients from an image to formulate energy function so as to detect object boundaries. Let $\mathcal{I}_{z_p}^O$ be the p^{th} 2D image slices in 3D microscopy volume and $\phi_{z_p}(s) = [x_{z_p}(s), y_{z_p}(s)]$ be a deformable curve in $\mathbb{R}^2$. In

[1]The 3D active contours with inhomogeneity correction method introduced in this chapter is based on our paper in [55]

the primitive edge-based active contour models [29] introduced the energy function (E_{edge1}) as

$$\begin{aligned} E_{edge1} &= E_{internal} + E_{external} \\ &= \alpha \int_0^1 |\phi_{z_p}(s)'|^2 ds + \beta \int_0^1 |\phi_{z_p}(s)''|^2 ds - \gamma \int_0^1 |\nabla \mathcal{I}^O_{z_p}(\phi(s))| ds \end{aligned} \tag{5.1}$$

where α, β, and γ are adjustable positive parameters.

The first two terms are known as the internal energy which represent the energy of the smoothness of the curve whereas the third term is known as external energy which represents the energy that is desired to minimize local minimum. In particular, the external energy forces curve to desired object edges or boundaries. The goal of edge-based active contours is finding the curve $\phi(s)_{z_p}$ that minimizes the energy function defined by Eq (5.1) so as to find the curve fitting to boundaries. More specifically, by minimizing the energy functional, the curve is drawn to the maxima of $|\nabla \mathcal{I}^O_{z_p}|$, essentially acting as an edge-detector while simultaneously maintaining a smoothness of curve.

The idea of edge detector part can be extended by a positive and decreasing function $g(\cdot)$ so that $-|\nabla \mathcal{I}^O_{z_p}|$ may be replaced by $g(|\nabla \mathcal{I}^O_{z_p}|)^2$. Utilizing this $g(\cdot)$ function, Eq (5.1) can be expressed as [155]

$$E_{edge1} = \alpha \int_0^1 |\phi_{z_p}(s)'|^2 ds + \beta \int_0^1 |\phi_{z_p}(s)''|^2 ds + \gamma \int_0^1 g(|\nabla \mathcal{I}^O_{z_p}(\phi(s))|)^2 ds \tag{5.2}$$

where

$$g(|\nabla \mathcal{I}^O_{z_p}|) = \frac{1}{1 + |\nabla G^\sigma_{z_p} * \mathcal{I}^O_{z_p}|^s}, \quad s \geq 1. \tag{5.3}$$

Note that $G^\sigma_{z_p} * \mathcal{I}^O_{z_p}$ is a smoothed version of given original image slice where convolved with Gaussian kernel such that

$$G^\sigma_{z_p}(x_{z_p}, y_{z_p}) = \frac{1}{\sigma^2} \exp\left(-\frac{|x^2_{z_p} + y^2_{z_p}|}{4\sigma} \right).$$

The function $g(|\nabla \mathcal{I}^O_{z_p}|)$ is positive in homogeneous regions and zero at the edges.

However, the primitive approach of snakes fails to detect edges with changes in topology. Alternatively, the geodesic active contour models are described in [41] where

formulating energy function as a problem of geodesic curve in a Riemannian space induced from the image $\mathcal{I}^O_{z_p}$. The energy function (E_{edge2}) of geodesic model is

$$E_{edge2} = 2\int_0^1 |\phi_{z_p}(s)'|^2 g(|\nabla \mathcal{I}^O_{z_p}(\phi_{z_p}(s))|)ds \tag{5.4}$$

where we used same $g(\cdot)$ function mentioned above.

Since these edge-based active contours rely heavily on the edge function $g(\cdot)$, a function of $|\nabla \mathcal{I}^O_{z_p}|$, it easily fails to identify poorly defined objects. As indicated in Chapter 1 microscopy volumes suffer from poor edge details so that the evolved curve of gradient dependent methods may not stop the boundary. Moreover, the evolving curves of edge-based active contours are highly dependent on initial contour placement and too sensitive to image noise [50]. Due to the limitation of microscopy volumes resolution, these methods does not produce adequate segmentation results. Instead of edge-based approach we use region-based active contours method.

5.2 Region-Based Active Contours

5.2.1 2D Active Contours

The region-based active contours are inspired by the region competition idea introduced in [156]. The region competition method combines the geometrical features of snakes and statistical techniques of region growing. Instead of defining boundaries using gradient, the region-based active contours determines foreground and background regions depending on the location of evolving curves. This idea is further improved by [46] where incorporating these foreground and background regions in energy functional. The energy function can be expressed as

$$\begin{aligned} E_{region} &= \lambda_1 E_{in} + \lambda_2 E_{out} + \mu \cdot Length(\phi_{z_p}(\mathbf{x})) \\ &= \lambda_1 \int_{in(\phi_{z_p})} (\mathcal{I}^O_{z_p}(\mathbf{x}) - c_1)^2 d\mathbf{x} + \lambda_2 \int_{out(\phi_{z_p})} (\mathcal{I}^O_{z_p}(\mathbf{x}) - c_2)^2 d\mathbf{x} \\ &+ \mu \cdot Length(\phi_{z_p}(\mathbf{x})) \end{aligned} \tag{5.5}$$

where $\mathcal{I}^O_{z_p}(\mathbf{x})$ is the original 2D image slices to be segmented, $\phi_{z_p}(\mathbf{x})$ a zero-level contour of level-set function obtained from $\mathcal{I}^O_{z_p}(\mathbf{x})$, c_1 the sample mean intensity of $\mathcal{I}^O_{z_p}$ inside of $\phi_{z_p}(\mathbf{x})$, c_2 the sample mean intensity of $\mathcal{I}^O_{z_p}$ outside of $\phi_{z_p}(\mathbf{x})$, and λ_1, λ_2, μ are weighting coefficients for each term, respectively. In particular, μ represents a stepsize, λ_1 and λ_2 is the coefficient of E_{in} and E_{out}, respectively. These coefficients can be adjusted appropriately. Note that we assume given input is an 2D image slice, the dimension of $\mathbf{x}$ should be $\mathbf{x} \in \mathbb{R}^2$. As observed from Eq (5.5) the total energy can be expressed by the summation of energy toward inside (E_{in}), energy toward outside(E_{out}), and regularization term related to length of the contour. To fit this contour to targeted segmented region we need to seek a zero-level contour $\phi_{z_p}(\mathbf{x})$ that satisfied the equilibrium energies induced by foreground and background. Therefore, this region-based segmentation problem turns into energy minimization problem as discussed in Chapter 5.1.

Utilizing Heaviside function and Dirac delta function, $\delta(\cdot)$, the energy functional shown in Eq (5.5) can be rewritten as:

$$\begin{aligned} E_{2Dac} = \lambda_1 \int_\Omega |\mathcal{I}^O_{z_p}(\mathbf{x}) - c_1|^2 H(\phi_{z_p}(\mathbf{x}))d\mathbf{x} + \lambda_2 \int_\Omega |\mathcal{I}^O_{z_p}(\mathbf{x}) - c_2|^2 (1 - H(\phi_{z_p}(\mathbf{x}))d\mathbf{x} \\ + \mu \int_\Omega \delta(\phi_{z_p}(\mathbf{x}))|\nabla \phi_{z_p}(\mathbf{x})|d\mathbf{x} \end{aligned} \tag{5.6}$$

where Ω is the domain of the image $\mathbb{R}^2$. From now on, we use E_{2Dac} instead of E_{region} for energy functional of 2D region-based active contours to avoid confusion from suggested by different models. The Heaviside function H used here can be defined as

$$H(\phi_{z_p}) = \begin{cases} 1 & \text{if } \phi_{z_p} \geq 0 \\ 0 & \text{if } \phi_{z_p} < 0. \end{cases} \tag{5.7}$$

and $\delta(\cdot)$ is defined as the derivative of the Heaviside function. Since the Heaviside function in Eq (5.7) is mathematically impossible to differentiate, we use approximate version of $H(\cdot)$ described in [46, 157] and corresponding $\delta(\cdot)$ as

$$H(\phi_{z_p}) = \frac{1}{2}\left(1 + \frac{2}{\pi}\tan^{-1}(\phi_{z_p})\right), \quad \delta(\phi_{z_p}) = \frac{1}{\pi}\left(\frac{1}{1+\phi^2_{z_p}}\right). \tag{5.8}$$

By introducing the Heaviside function, given energy functional has three terms that each of terms are integrated over the entire image Ω.

Further, centroids c_1 and c_2 can be obtained with Heaviside function as

$$c_1(\phi_{z_p}) = \frac{\int_\Omega \mathcal{I}^O_{z_p}(\mathbf{x}) H(\phi_{z_p}(\mathbf{x}))d\mathbf{x}}{\int_\Omega H(\phi_{z_p}(\mathbf{x}))d\mathbf{x}}, \quad c_2(\phi_{z_p}) = \frac{\int_\Omega \mathcal{I}^O_{z_p}(\mathbf{x})(1 - H(\phi_{z_p}(\mathbf{x})))d\mathbf{x}}{\int_\Omega (1 - H(\phi_{z_p}(\mathbf{x})))d\mathbf{x}}. \tag{5.9}$$

As shown in Eq (5.9) we can easily observe that $c_1(\phi_{z_p})$ is the average intensity inside the ϕ_{z_p} and $c_2(\phi_{z_p})$ is the average intensity outside the ϕ_{z_p}.

Having found $c_1(\phi_{z_p})$ and $c_2(\phi_{z_p})$ we are ready to solve the energy minimization problem. To solve this energy minimization problem, we take derivatives of the energy function and find the ϕ_{z_p} from setting $\frac{\partial E_{2Dac}}{\partial \phi_{z_p}} = 0$. Parameterizing the descent direction by an artificial time $t \geq 0$, the associated partial differential equation (PDE) involved with Euler-Lagrange equation for ϕ_{z_p} is [46]

$$\frac{\partial E_{2Dac}}{\partial \phi_{z_p}} = -\frac{\partial \phi_{z_p}}{\partial t} = \delta(\phi_{z_p}) \left[\lambda_1 (\mathcal{I}^O_{z_p} - c_1)^2 - \lambda_2 (\mathcal{I}^O_{z_p} - c_2)^2 - \mu \text{div} \left(\frac{\nabla \phi_{z_p}}{|\nabla \phi_{z_p}|} \right) \right] \tag{5.10}$$

where the third term of the equation contains the 2D mean curvature $\kappa = \text{div}\left(\frac{\nabla \phi_{z_p}}{|\nabla \phi_{z_p}|}\right)$. 2D mean curvature can be expressed as

$$\kappa = \frac{\phi_{xx}\phi_y^2 - 2\phi_{xy}\phi_x\phi_y + \phi_{yy}\phi_x^2}{(\phi_x^2 + \phi_y^2)^{3/2}}. \tag{5.11}$$

Note that have dropped subscript z_p from ϕ for simplicity. After obtaining mean curvature, we can update ϕ_{curr} from ϕ_{prev} using following update equation [158]:

$$\begin{aligned} \phi_{curr} &= \phi_{prev} + \Delta\phi = \phi_{prev} + \Delta t \left(\frac{\partial \phi}{\partial t} \right) \\ &= \phi_{prev} + \Delta t \left[\lambda_2 (\mathcal{I}^O_{z_p} - c_2)^2 - \lambda_1 (\mathcal{I}^O_{z_p} - c_1)^2 + \mu \text{div} \left(\frac{\nabla \phi_{prev}}{|\nabla \phi_{prev}|} \right) \right] \end{aligned} \tag{5.12}$$

where ϕ_{prev} and ϕ_{curr} denote the previous and current ϕ, respectively, and Δt an artificial time step. The active contours method is iteratively deformed an initial contour and stop when either ϕ is converged or it reaches given number of iterations.

One drawback of this 2D region-based active contours is that this method cannot utilize the information across z direction. For example, we can easily think similar

biological structures are observed from adjacent slices of 3D volume. However, this method used for each slice images independently so that the segmentation results did not take consideration of any z direction information. Therefore, we prefer using an entire 3D volume as an input to this snakes method described in [48].

5.2.2 3D Active Contours

The energy function used in 3D region-based active contours [48], which is an extension of the 2D case, is given by:

$$E_{3Dac} = \lambda_1 \int_{in(\phi)} |\mathcal{I}^O(\mathbf{x}) - c_1|^2 d\mathbf{x} + \lambda_2 \int_{out(\phi)} |\mathcal{I}^O(\mathbf{x}) - c_2|^2 d\mathbf{x} + \mu \cdot Surface(\phi(\mathbf{x})) \tag{5.13}$$

where $\mathbf{x} \in \mathbb{R}^3$, $\phi(\mathbf{x})$ a zero-level surface (Lipschitz function) of same size as $\mathcal{I}^O(\mathbf{x})$, $\mathcal{I}^O(\mathbf{x})$ the original acquired image volume, c_1 the average intensity inside (centroid) $\phi(\mathbf{x})$, c_2 the average intensity outside $\phi(\mathbf{x})$, and λ_1, λ_2, μ are weighting coefficients for each term, respectively. The first and second terms represent internal and external energy terms respectively, whereas the third term is a regularization term for the given energy function. Note that we did not have subscript z_p in $\mathcal{I}(\mathbf{x})$ and $\phi(\mathbf{x})$ since this method is performed with entire 3D microscopy image volume. Previously, z_p was used only if the methods are used for the p^{th} 2D image slice. We denote this energy functional as E_{3Dac} in order to distinguish between 2D and 3D region-based active contours. Compared to Eq (5.5) this energy functional changes only for input from image ($\mathbb{R}^2$) to volume ($\mathbb{R}^3$) and regularization term from contour length to surface. As shown in Chapter 5.2.1 given energy functional can be rewritten using Heaviside function, $H(\cdot)$, and Dirac delta function, $\delta(\cdot)$, as:

$$E_{3Dac} = \lambda_1 \int_{\Omega} |\mathcal{I}^O(\mathbf{x}) - c_1|^2 H(\phi(\mathbf{x})) d\mathbf{x} + \lambda_2 \int_{\Omega} |\mathcal{I}^O(\mathbf{x}) - c_2|^2 (1 - H(\phi(\mathbf{x})) d\mathbf{x} + \mu \int_{\Omega} \delta(\phi(\mathbf{x})) |\nabla \phi(\mathbf{x})| d\mathbf{x} \tag{5.14}$$

and centroids c_1 and c_2 are obtained from

$$c_1(\phi) = \frac{\int_\Omega \mathcal{I}^O(\mathbf{x}) H(\phi(\mathbf{x})) d\mathbf{x}}{\int_\Omega H(\phi(\mathbf{x})) d\mathbf{x}}, \quad c_2(\phi) = \frac{\int_\Omega \mathcal{I}^O(\mathbf{x})(1 - H(\phi(\mathbf{x}))) d\mathbf{x}}{\int_\Omega (1 - H(\phi(\mathbf{x}))) d\mathbf{x}}. \tag{5.15}$$

Note that since given input is 3D volume, $\mathbf{x} \in \mathbb{R}^3$.

Similarly from 2D case, we utilize the Euler-Lagrange equation to obtain PDE such that

$$\frac{\partial E_{3Dac}}{\partial \phi} = -\frac{\partial \phi}{\partial t} = \delta(\phi) \left[\lambda_1 (\mathcal{I}^O - c_1)^2 - \lambda_2 (\mathcal{I}^O - c_2)^2 - \mu \text{div} \left(\frac{\nabla \phi}{|\nabla \phi|} \right) \right]. \tag{5.16}$$

Since the input is not 2D image but 3D volume, 3D mean curvature should be needed from same analogy in previous section Chapter 5.2.1. The 3D mean curvature can be obtained using following equation [159]:

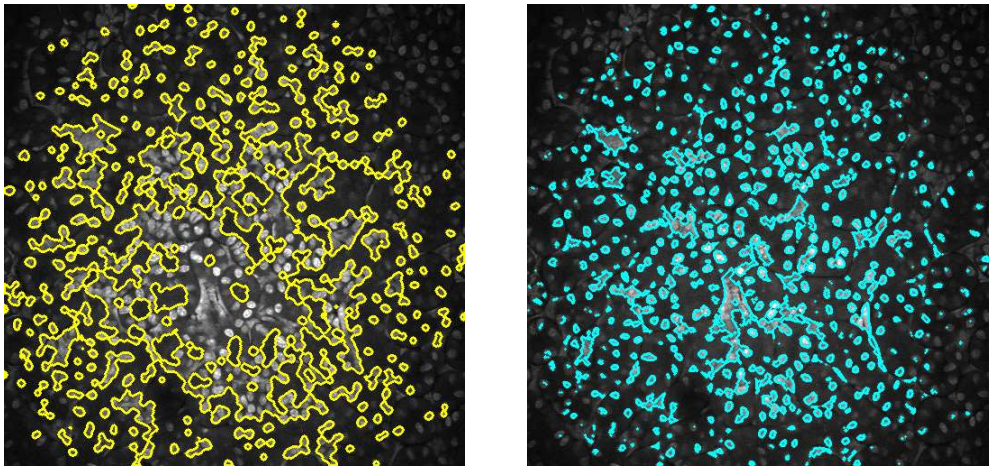
$$\kappa = \frac{\left\{ \begin{array}{r} \phi_x^2(\phi_{yy} + \phi_{zz}) + \phi_y^2(\phi_{xx} + \phi_{zz}) + \phi_z^2(\phi_{xx} + \phi_{yy}) \\ -2\left(\phi_x \phi_y \phi_{xy} + \phi_x \phi_z \phi_{xz} + \phi_y \phi_z \phi_{yz}\right) \end{array} \right\}}{\left(\phi_x^2 + \phi_y^2 + \phi_z^2\right)^{3/2}}. \tag{5.17}$$

Having found 3D mean curvature ϕ is updated as 2D method:

$$\phi_{curr} = \phi_{prev} + \Delta t \left[\lambda_2 (\mathcal{I}^O - c_2)^2 - \lambda_1 (\mathcal{I}^O - c_1)^2 + \mu \text{div} \left(\frac{\nabla \phi_{prev}}{|\nabla \phi_{prev}|} \right) \right] \tag{5.18}$$

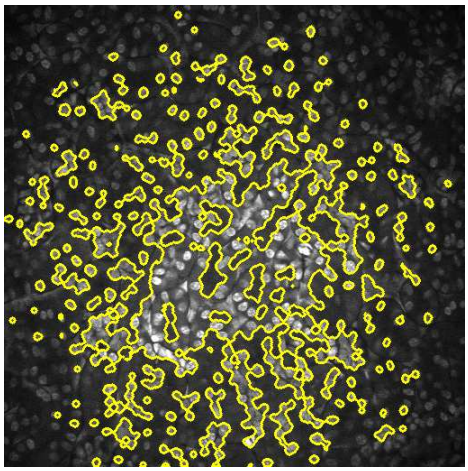
with same notations ϕ_{prev}, ϕ_{curr}, and Δt are used as Eq (5.12).

Different from the edge-based active contours, these region-based one provide robust to local noise and the segmentation results are relatively independent of an initial contour placement. However, these approaches may not be well-suited for our microscopy volumes due to inhomogeneity. In particular, region-based approaches may lead to erroneous segmentation when foreground and background regions are not clearly separated [50] as a function of intensity. The segmentation results with 2D and 3D region-based active contours are shown in Figure 5.1. We choose $\lambda_1 = \lambda_2 = 1$, $\mu = 1$ and set an initial contours as multiple circles (2D) or spheres (3D) with radius 10 distributed evenly to cover entire image or volume. Also, we set 1000 iterations for both 2D and 3D methods.

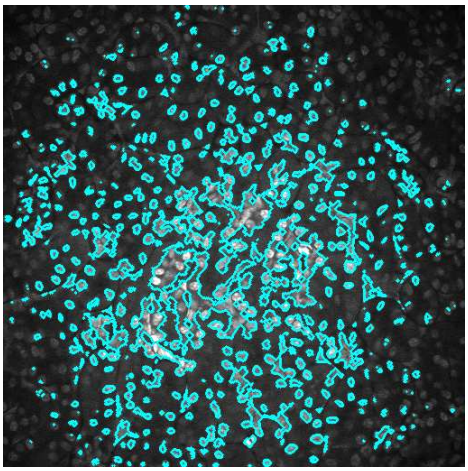


(a) 100th image of *WSM* [46]

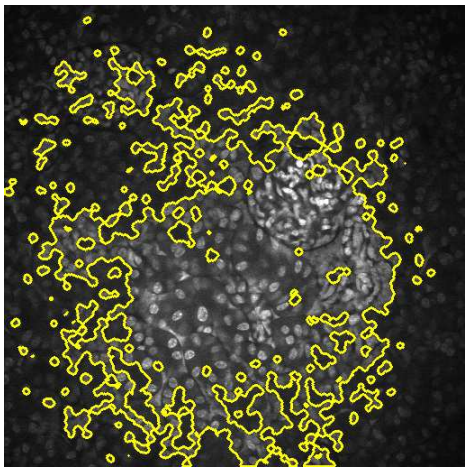
(b) 100th image of *WSM* [48]



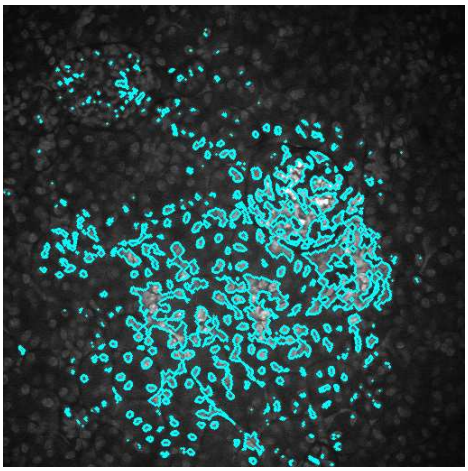
(c) 200th image of *WSM* [46]



(d) 200th image of *WSM* [48]



(e) 300th image of *WSM* [46]



(f) 300th image of *WSM* [48]

Fig. 5.1. Segmentation results at various depth using 2D and 3D region-based active contours methods to *WSM* dataset

As observed from Figure 5.1 3D region-based active contours (cyan) successfully captures more nuclei than 2D method (yellow). This is mainly due to utilizing volume information rather than using single sliced image. However, both methods miss a lot of nuclei resided in the boundary of images owing to inhomogeneous characteristic of microscopy volumes. At the same time, they fail to segment nuclei in brighter region mainly at the center part. Instead of capturing nuclei 2D method makes big blobs at the brighter regions whereas 3D method shrinks the contours to fit locally. To address this failure we propose new energy functional involved with inhomogeneity correction so that the method fits an initial contour to objects and corrects inhomogeneity field simultaneously.

5.3 3D Active Contours With Inhomogeneity Correction

In this section, we describe a method that segments nuclei in 3D microscopy volumes based on a combination of 3D region-based active contours and 3D inhomogeneity correction. As indicated from previous chapter, incorporated 3D snakes did not account for intensity inhomogeneities. The method described here extends the previous 3D region-based active contour method introduced in Chapter 5.2.2 by taking into account 3D intensity inhomogeneities. This is achieved by utilizing a multiplicative model where the original volume $\mathcal{I}^O(\mathbf{x})$ is modeled as

$$\mathcal{I}^O(\mathbf{x}) = \mathcal{W}(\mathbf{x}) \circ \mathcal{I}^C(\mathbf{x}) + \mathcal{N}(\mathbf{x}) \tag{5.19}$$

where $\mathbf{x} \in \mathbb{R}^3$, $\mathcal{I}^C(\mathbf{x})$ is a homogeneous volume, $\mathcal{W}(\mathbf{x})$ is a 3D weight matrix referred to as the inhomogeneity field that accounts for the degree of intensity inhomogeneity at each voxel location, and $\mathcal{N}(\mathbf{x})$ is zero-mean 3D Gaussian noise. The $\circ$ operator represents voxelwise multiplication (Hadamard product). Assuming $\mathcal{W}(\mathbf{x})$ is slowly varying, Eq (5.19) can be approximated as [52]

$$\mathcal{I}^O(\mathbf{x}) \approx \begin{cases} \mathcal{W}(\mathbf{y})c_1 + \mathcal{N}(\mathbf{x}) & \text{when } \phi(\mathbf{x}) > 0 \\ \mathcal{W}(\mathbf{y})c_2 + \mathcal{N}(\mathbf{x}) & \text{when } \phi(\mathbf{x}) < 0 \end{cases} \tag{5.20}$$

where $\mathbf{y}$ is any point in $B(\mathbf{x}, \boldsymbol{\rho})$, the 3D spherical neighborhood of $\mathbf{x}$ with radius $\boldsymbol{\rho}$, and c_1, c_2 are determined by whether $\mathbf{x}$ is inside or outside of the surface $\phi(\mathbf{x})$. Specifically, if the location of $\mathbf{x}$ is inside $\phi(\mathbf{x})$, c_1 is chosen as a centroid. However, if the location of $\mathbf{x}$ is outside of $\phi(\mathbf{x})$, c_2 is taken as the centroid. Using this relationship, a new energy function with a 3D kernel $\mathcal{K}$ is obtained:

$$\begin{aligned} E_{3DacIC} = \int_{\Omega} \Bigg(& \lambda_1 \int_{in(\phi)} \mathcal{K}(\mathbf{y}-\mathbf{x}) \cdot |\mathcal{I}^O(\mathbf{x}) - \mathcal{W}(\mathbf{y}) c_1|^2 d\mathbf{x} \\ & + \lambda_2 \int_{out(\phi)} \mathcal{K}(\mathbf{y}-\mathbf{x}) \cdot |\mathcal{I}^O(\mathbf{x}) - \mathcal{W}(\mathbf{y}) c_2|^2 d\mathbf{x} \Bigg) d\mathbf{y} \\ & + \mu \cdot Surface(\phi(\mathbf{x})). \end{aligned} \tag{5.21}$$

Note that $\mathbf{y}$ is integrated over the entire image volume Ω. Note also that the kernel function is chosen such that $\mathcal{K}(\mathbf{y}-\mathbf{x}) = 0$ when $\mathbf{y}$ is outside the neighborhood of $\mathbf{x}$. Details of the derivation are found in [160].

Since the kernel function $\mathcal{K}$ should account for a slow varying $\mathcal{W}$, we choose a modified normalized truncated 3D Gaussian function:

$$\mathcal{K}(\mathbf{u}) = \begin{cases} \frac{1}{C} e^{-\left| \frac{u_x^2}{2\rho_x^2} + \frac{u_y^2}{2\rho_y^2} + \frac{u_z^2}{2\rho_z^2} \right|} & \text{when } |\mathbf{u}| \leq \boldsymbol{\rho} \\ 0 & \text{otherwise} \end{cases} \tag{5.22}$$

where $\mathbf{u} = [u_x, u_y, u_z]^T$ and $\boldsymbol{\rho} = [\rho_x, \rho_y, \rho_z]^T$. Due to the difference in resolutions across the horizontal and vertical directions, and since the vertical resolution is approximately 4 times less than the horizontal resolution, we alter the neighborhood $B(\mathbf{x}, \boldsymbol{\rho})$ by setting $\rho_x = \rho_y = 4\rho_z$. In addition, C is chosen to be a normalizing constant so that $\int \mathcal{K}(\mathbf{u}) d\mathbf{u}$ is always 1. Without considering intensity inhomogeneities, for example, $\mathcal{W}(\mathbf{y}) = 1$ for the entire volume, the energy function from (5.21) is easily converted back to (5.13) using the fact that $\int \mathcal{K}(\mathbf{y}-\mathbf{x}) d\mathbf{y} = 1$.

Utilizing Heaviside function, $H(\cdot)$, the Dirac delta function, $\delta(\cdot)$, and swapping the order of the integrals, Eq (5.21) can be rewritten as [46]:

$$\begin{aligned} E_{3DacIC} = \lambda_1 \int_\Omega \left(\int_\Omega \mathcal{K}(\mathbf{y}-\mathbf{x}) \cdot |\mathcal{I}^O(\mathbf{x}) - \mathcal{W}(\mathbf{y})c_1|^2 d\mathbf{y} \right) H(\phi(\mathbf{x}))d\mathbf{x} \\ + \lambda_2 \int_\Omega \left(\int_\Omega \mathcal{K}(\mathbf{y}-\mathbf{x}) \cdot |\mathcal{I}^O(\mathbf{x}) - \mathcal{W}(\mathbf{y})c_2|^2 d\mathbf{y} \right) (1 - H(\phi(\mathbf{x}))d\mathbf{x} \\ + \mu \int_\Omega \delta(\phi(\mathbf{x}))|\nabla\phi(\mathbf{x})|d\mathbf{x} \end{aligned} \tag{5.23}$$

where Heaviside function and Dirac delta function are obtained by Eq (5.8). This Eq (5.23) can be equivalently written as:

$$\begin{aligned} E_{3DacIC} = \lambda_1 \int_\Omega \left((\mathcal{I}^O)^2 \circ 1_\mathcal{K} - 2\mathcal{I}^O \circ (\mathcal{W} * \mathcal{K})c_1 + (\mathcal{W}^2 * \mathcal{K})c_1^2 \right) H(\phi)d\mathbf{x} \\ + \lambda_2 \int_\Omega \left((\mathcal{I}^O)^2 \circ 1_\mathcal{K} - 2\mathcal{I}^O \circ (\mathcal{W} * \mathcal{K})c_2 + (\mathcal{W}^2 * \mathcal{K})c_2^2 \right) (1 - H(\phi))d\mathbf{x} \\ + \mu \int_\Omega \delta(\phi)|\nabla\phi|d\mathbf{x} \end{aligned} \tag{5.24}$$

where $*$ is the 3D convolution operator and $1_\mathcal{K}(\mathbf{x})$ is a 3D volume of same size as $\mathcal{I}^O(\mathbf{x})$ whose entries are all 1 except near the volume boundary. More specifically, $1_\mathcal{K}(\mathbf{x})$ is obtained by convolving a 3D matrix of ones with 3D kernel $\mathcal{K}$. Note that both c_1 and c_2 are vectors with three elements (3×1 vectors). For brevity we have omitted the explicit argument $\mathbf{x}$.

By minimizing the energy function shown in Eq (5.24) with respect to ϕ, $\mathcal{W}$, c_1 and c_2, we can obtain the 3D segmentation result (ϕ) and an estimated 3D inhomogeneity field ($\mathcal{W}$).

One way to achieve this is to first minimize the energy function with respect to c_1 and c_2 for given ϕ and $\mathcal{W}$. Denoting the optimal values for c_1 and c_2 by $\hat{c_1}$ and $\hat{c_2}$, respectively, it can be shown that

$$\hat{c_1} = \frac{\int_{in(\phi)} \mathcal{I}^O \circ (\mathcal{W} * \mathcal{K})d\mathbf{y}}{\int_{in(\phi)} (\mathcal{W}^2 * \mathcal{K})d\mathbf{y}}, \; \hat{c_2} = \frac{\int_{out(\phi)} \mathcal{I}^O \circ (\mathcal{W} * \mathcal{K})d\mathbf{y}}{\int_{out(\phi)} (\mathcal{W}^2 * \mathcal{K})d\mathbf{y}}. \tag{5.25}$$

Having found $\hat{c_1}$ and $\hat{c_2}$ the next step is to minimize the energy function with respect to ϕ for given c_1, c_2, and $\mathcal{W}$. In this case the values used for c_1 and c_2 are $\hat{c_1}$ and

$\hat{c_2}$, respectively. Since the partial differential equation (PDE) $\frac{\partial E_{3DacIC}}{\partial \phi} = 0$ does not have a closed form solution, an approximate solution $\hat{\phi}$ can be obtained by iteratively using the Euler-Lagrange equation as follows [48]:

$$\frac{\partial E_{3DacIC}}{\partial \phi} = -\frac{\partial \phi}{\partial t} = \delta(\phi)\left[\lambda_1 f_1 - \lambda_2 f_2 - \mu \text{div}\ \left(\frac{\nabla \phi}{|\nabla \phi|}\right)\right] \tag{5.26}$$

where f_1 and f_2 are expressed as

$$f_1 = (\mathcal{I}^O)^2 \circ 1_{\mathcal{K}} - 2\mathcal{I}^O \circ (\mathcal{W} * \mathcal{K})c_1 + (\mathcal{W}^2 * \mathcal{K})c_1^2$$
$$f_2 = (\mathcal{I}^O)^2 \circ 1_{\mathcal{K}} - 2\mathcal{I}^O \circ (\mathcal{W} * \mathcal{K})c_2 + (\mathcal{W}^2 * \mathcal{K})c_2^2$$

respectively. A solution to Eq (5.26) is obtained through numerical techniques, where the approximation solution $\hat{\phi}$ can be obtained iteratively as follows:

$$\hat{\phi}_{curr} = \hat{\phi}_{prev} + \Delta t\left[\lambda_2 f_2 - \lambda_1 f_1 + \mu \text{div}\ \left(\frac{\nabla \hat{\phi}_{prev}}{|\nabla \hat{\phi}_{prev}|}\right)\right]. \tag{5.27}$$

Here, $\hat{\phi}_{prev}$ and $\hat{\phi}_{curr}$ denote the previous and current estimates of ϕ, respectively, Δt a time step used to control the evolution speed of ϕ, and μ the surface weight coefficient, chosen based on the size of the desired object to be detected. Higher μ values are used for detecting all objects whereas smaller μ values are used for detecting only larger objects [46]. This is similar form as we already discussed from 2D active contours (Eq (5.12)) and 3D active contours (Eq (5.18)) but associated with f_1 and f_2. Note that the term $\kappa = \text{div}\ \left(\frac{\nabla \phi}{|\nabla \phi|}\right)$ is the curvature of the level set function ϕ and same 3D mean curvature shown in Eq (5.17) is used. Lastly, for given c_1, c_2, and ϕ, E is minimized with respect to the 3D inhomogeneity field $\mathcal{W}$. The optimal 3D inhomogeneity field, $\hat{\mathcal{W}}$, is the solution to the PDE arising from setting $\frac{\partial E_{3DacIC}}{\partial \mathcal{W}} = 0$ and is given by

$$\hat{\mathcal{W}} = \frac{(\mathcal{I}^O \circ \mathcal{J}^{(1)}) * \mathcal{K}}{\mathcal{J}^{(2)} * \mathcal{K}} \tag{5.28}$$

where $\mathcal{J}^{(1)} = \int_{in(\phi)} c_1 d\mathbf{y} + \int_{out(\phi)} c_2 d\mathbf{y}$ and $\mathcal{J}^{(2)} = \int_{in(\phi)} c_1^2 d\mathbf{y} + \int_{out(\phi)} c_2^2 d\mathbf{y}$ [52]. Using the above solution it is possible to iteratively arrive at a segmentation result as described in "Method 1". Note that N_I denotes the number of (inner loop) iterations needed to arrive at a solution for ϕ as indicated by Eq (5.27), whereas N_O is the number of (outer loop) iterations required to update c_1, c_2 and $\mathcal{W}$.

Method 1 3D active contours with inhomogeneity correction

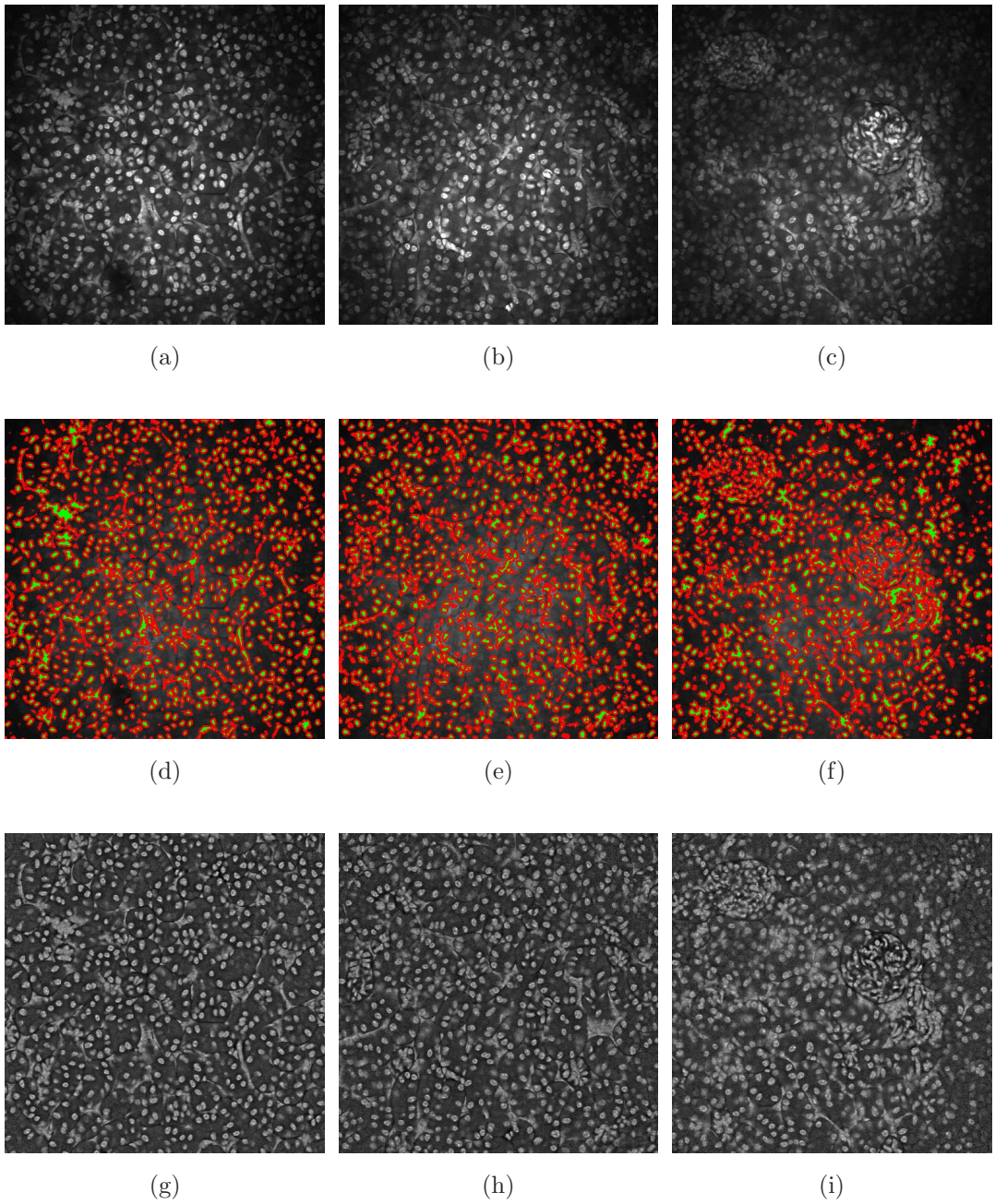
Set initial contour ϕ, initialize $\mathcal{W}$ to be a 3D matrix of ones, and c_1, c_2 to be a zero 3×1 vector, respectively.
Set $\mathcal{K}$ as a modified normalized truncated Gaussian function using Eq (5.22).
for $i = 1$ to N_O **do**
    Obtain the 3D convolutions $\mathcal{W} * \mathcal{K}$, $\mathcal{W}^2 * \mathcal{K}$, and $1_{\mathcal{K}}$.
    Update the centroids c_1(inside), c_2(outside) using Eq (5.25).
    for $j = 1$ to N_I **do**
        Obtain the 3D curvature κ using Eq (5.17).
        Update the 3D segmentation result ϕ using Eq (5.27).
    end for
    Update the 3D inhomogeneity field $\mathcal{W}$ using Eq (5.28).
end for
Obtain $\mathcal{I}^C$ using Eq (5.19).
return ϕ, $\mathcal{I}^C$, $\mathcal{W}$

5.4 Experimental Results

These datasets consist of 3D volumes of a rat kidney labeled with Hoechst 33342 collected using two-photon microscopy. Different from the method described in Chapter 3, this method uses entire 3D volumes to capture 3D nuclei taking consideration of different resolution across horizontal and vertical directions. Note that nuclei information of each of datasets is in the B channel so that separating the B channel from datasets is the first and necessary step as described in Chapter 3.2. As previously indicated our goal is to segment individual 3D nuclei in florescence microscopy volumes.

The performance of our method was tested on five different datasets: *WSM*, *FNa*, *FNb*, *FNa1*, and *FNb1*. The five datasets, which are volumes of rat kidney, consist of 512, 36, 41, 38, and 41 images, respectively each of size 512×512 pixels and are of 8-bit/pixel resolution. *FNa1* and *FNb1* were collected from the same specimens as *FNa* and *FNb*, respectively, but imaged in reverse directions. The following values were used for the various parameters: $\Delta t = 0.1$, $\lambda_1 = \lambda_2 = 1$, $\rho_x = 4$, $\mu = 0.001 \cdot 255^2$. Since $\rho_x = \rho_y = 4\rho_z$, ρ_y and ρ_z are automatically determined once ρ_x is set. Also, $N_I = 20$ and $N_O = 50$ so that the total number of iterations used in solving for ϕ was 1000. For the initial contours, we chose multiple spheres of radius 10 distributed evenly to cover the entire volume. The segmentation results of the proposed method and corresponding inhomogeneity corrected images taken from various depth are shown in Figure 5.2.

As can be seen from the first row of Figure 5.2, the original images suffer from significant inhomogeneous intensities. More specifically, the intensities at the center are brighter than at the boundaries of the images. In spite of this, the proposed method is able to successfully capture nuclei close to the boundaries as shown in Figure 5.2. For visualization purposes we highlight contours as red and their interiors as green. The last row of Figure 5.2 portrays the estimated inhomogeneity corrected images corresponding to the original images. These are sample images from 3D inhomogene-



(a) (b) (c)

(d) (e) (f)

(g) (h) (i)

Fig. 5.2. Segmentation results and inhomogeneity corrected images at various depth for *WSM*. Top row: Original 100^{th}, 200^{th}, 300^{th} images. Middle row: Segmentation results overlaid onto original images (red: nuclei contours, green: nuclei regions). Bottom row: Corresponding inhomogeneity corrected images based on estimated 3D inhomogeneity field

ity corrected volume (I^C) obtained by Eq (5.19). As observed in Figure 5.2, intensity inhomogeneity are corrected for all images.

The segmentation results shown in middle row of Figure 5.2 can be compared with Figure 5.1 where 2D and 3D region-based active contours used for same sample images. As indicated our proposed method successfully overcomes the inhomogeneity described in Chapter 5.2.2. Moreover, the segmentation results may be favorably compared with Figure 3.4. There are several remarkable differences. First of all, results of multiscale edge detection method is extremely sensitive to threshold value whereas segmentation results of 3D active contours with inhomogeneity correction do not have that sensitive parameters. Second, edge detection method finds edge pixels without taking into consideration of 2D or 3D topological structures. Owing to this fact some pixels which are marked as edges, it often turns to be noise and irrelevant to entire biology structure. This false positive is aggravated especially in deeper tissues due to contrast decreasing. Reversely, since active contours method uses level-set model to optimize energy function, the segmentation results using active contours naturally produce enclosed shapes. On top of that utilizing 3D information prevents segmentation regions from falling into noisy regions.

To evaluate the performance of the proposed method, images from *FNa*, *FNb*, *FNa1*, and *FNb1* were manually segmented and used as groundtruth (Figures 5.3b, 5.4b). The accuracy, Type-I error, and Type-II error metrics were obtained for our method based on the groundtruth images. Here accuracy is defined to be the ratio of the number of correctly segmented nuclei pixels (true positive) and background pixels (true negative) to the total number of pixels. Type-I error (False alarm) is the ratio of the number of background pixels wrongly detected as nuclei (false positive) to the total number of pixels. Similarly, Type-II error (Missed) is the ratio of the number of nuclei pixels wrongly detected as background (false negative) to the total number of pixels.

The proposed method's performance is provided in Table 5.1 and 5.2 where we have also included for comparison purposes the performance of five other techniques.

Table 5.1.
Comparison of the performances of proposed and other segmentation methods using *FNa*, *FNb*

	FNa			*FNb*		
Method	Accuracy	Type-I	Type-II	Accuracy	Type-I	Type-II
2Dac [46]	54.71%	43.31%	1.98%	61.89%	32.42%	10.27%
2Dlac [50]	57.62%	39.14%	3.25%	58.21%	31.52%	10.27%
2DacIC [52]	73.12%	25.09%	**1.79%**	80.35%	15.19%	**4.46%**
3Dac [21, 48]	79.76%	16.63%	3.61%	78.43%	15.14%	6.42%
3Dsquassh [63]	88.72%	8.57%	2.71%	85.32%	5.96%	8.73%
3DacIC (proposed)	**91.87%**	**5.61%**	2.53%	**89.65%**	**4.50%**	5.85%

Table 5.2.
Comparison of the performances of proposed and other segmentation methods using *FNa1*, and *FNb1*

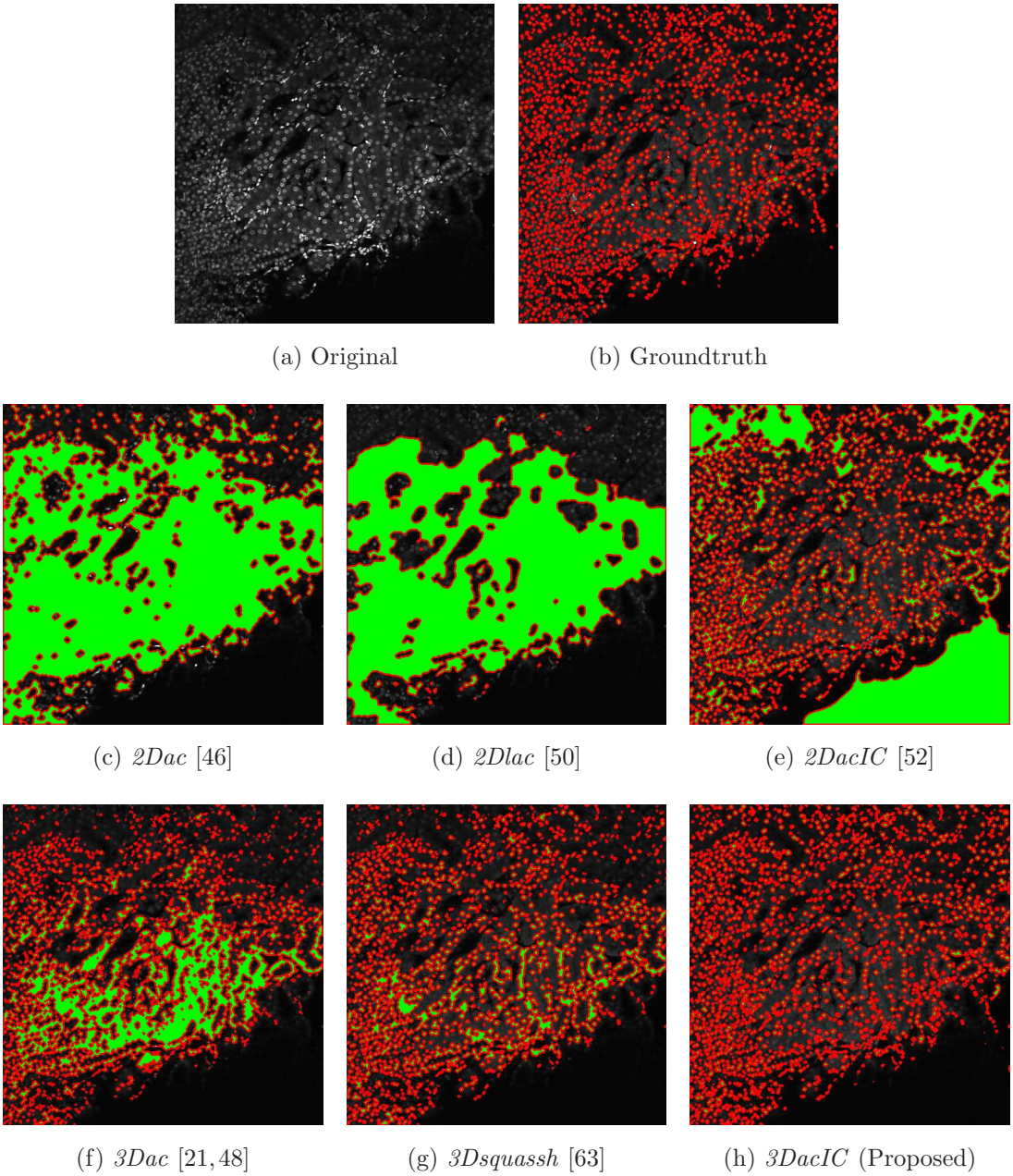
	FNa1			*FNb1*		
Method	Accuracy	Type-I	Type-II	Accuracy	Type-I	Type-II
2Dac [46]	57.39%	38.91%	3.70%	72.28%	20.44%	7.29%
2Dlac [50]	66.35%	28.13%	5.51%	63.47%	27.45%	9.08%
2DacIC [52]	86.18%	11.30%	2.52%	87.64%	8.99%	**3.38%**
3Dac [21, 48]	72.86%	24.98%	**2.16%**	81.58%	12.57%	5.86%
3Dsquassh [63]	83.35%	14.28%	2.37%	83.22%	13.01%	3.77%
3DacIC (proposed)	**87.71%**	**9.49%**	2.80%	**89.10%**	**7.00%**	3.90%

In particular, we have included the 2D region-based active contours [46] (*2Dac*), 2D region-based localized active contours [50] (*2Dlac*), 2D region-base active con-

tours with inhomogeneity correction [52] (*2DacIC*), and 3D region-based active contours [21, 48] (*3Dac*). We denote our method by *3DacIC* (3D region-based active contours with inhomogeneity correction). For consistency, the same number and same sized circles (2D)/spheres (3D) were used as initial contours in all the active contour methods. In addition, the same number of iterations were used. Finally, we also compared the performance of all these methods with a method described in [63] (*3Dsquassh*) using the default parameters setting in ImageJ.

As can be seen from Table 5.1 and 5.2, the proposed method outperformed the other methods in accuracy and Type-I error while exhibiting reasonably low Type-II error. *2DacIC* and *3Dac* had sometimes lower Type-II error than our method, but these methods suffered from high Type-I error, thus they often falsely detect nuclei. Similarly, *3Dsquassh* produced generally good accuracy and Type-II error but again suffered from high Type-I error.

For further comparison, we have included sample segmentation images of the results from *Dataset-II* and *III* produced by all the above mentioned techniques in Figure 5.3 and 5.4. The first row displays the original and groundtruth images, the second row exhibits segmentation results of *2Dac*, *2Dlac*, *2DacIC*, and the third row shows segmentation results of *3Dac*, *3Dsquassh*, and *3DacIC*, respectively. As observed, the proposed method outperforms all other methods by properly identifying nuclei, especially nuclei close to the boundary. In comparison, *2Dac* did not capture details at the center regions as well as boundary regions. *3Dac* had better performance than *2Dac* but failed to capture most nuclei located in the center. Although *3Dsquassh* was capable of capturing more nuclei at the center than *3Dac*, it still tended to group adjacent small nuclei as a single object as observed in Figure 5.3g and 5.4g. In addition, both *3Dac* and *3Dsquassh* missed many nuclei specifically at the boundary. In contrast, *2Dlac* had poorer results than the others since each localized active contour utilized local information. To be more specific, each localized contour sometimes correctly identifies nuclei as foreground but sometimes not to the extent that the segmentation results were combined with background regions



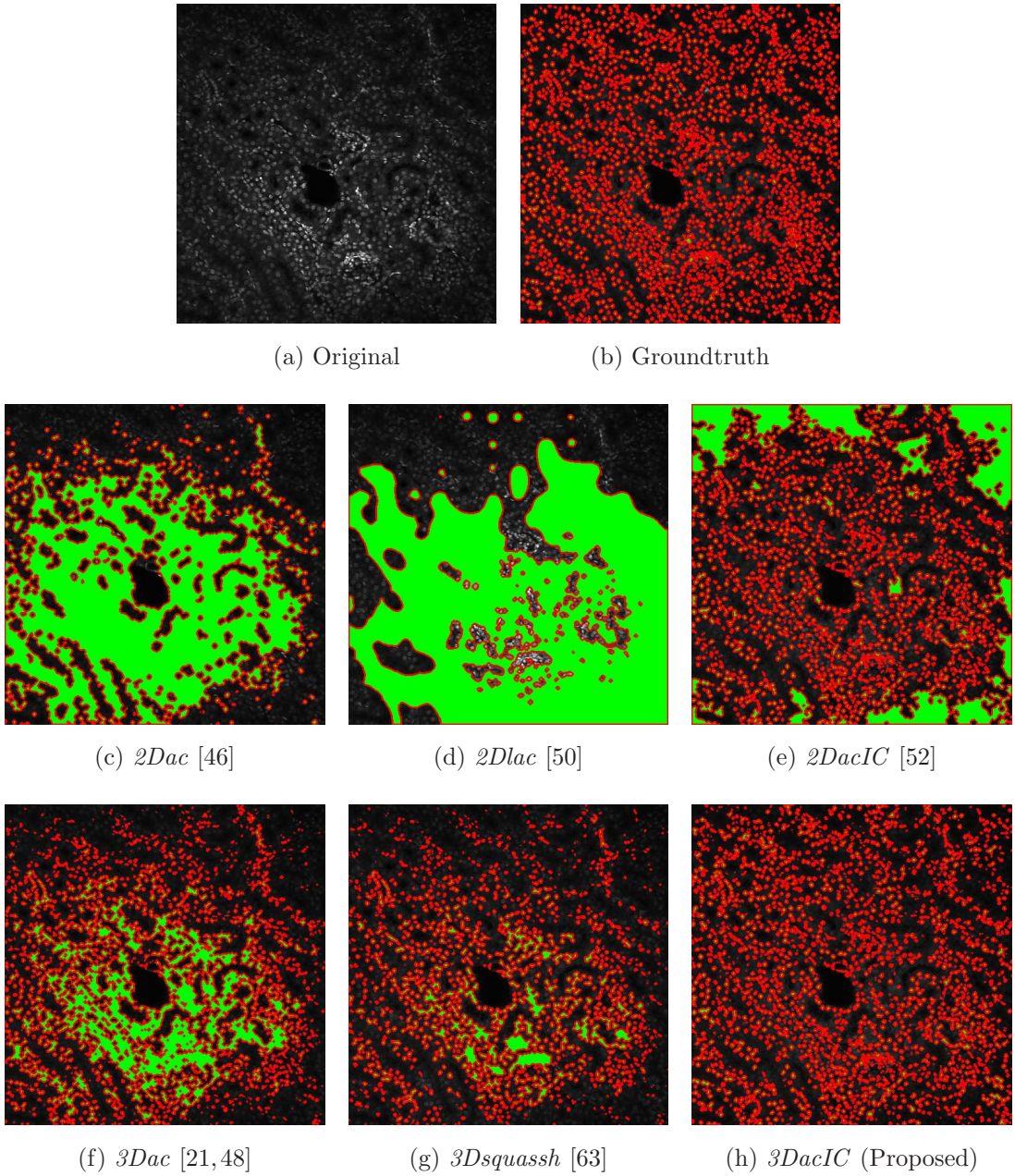
(a) Original (b) Groundtruth

(c) *2Dac* [46] (d) *2Dlac* [50] (e) *2DacIC* [52]

(f) *3Dac* [21, 48] (g) *3Dsquassh* [63] (h) *3DacIC* (Proposed)

Fig. 5.3. Comparison of segmentation results (red: nuclei contours, green: nuclei regions) of the proposed method with other methods overlaid onto original image (7^{th} image of the *FNa*)

(Figure 5.3d, 5.4d). Although *2DacIC* did produce good segmentation results at the center regions, it also incorrectly segmented background regions at the bound-



(a) Original (b) Groundtruth

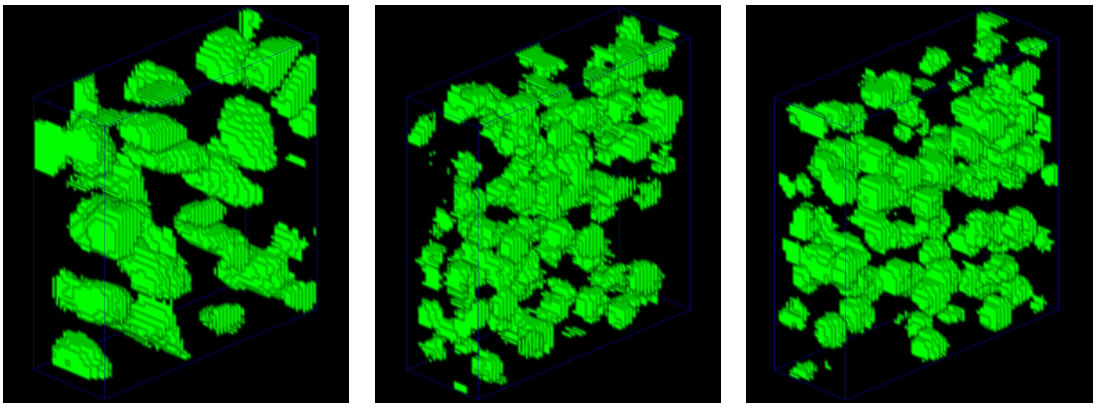
(c) *2Dac* [46] (d) *2Dlac* [50] (e) *2DacIC* [52]

(f) *3Dac* [21, 48] (g) *3Dsquassh* [63] (h) *3DacIC* (Proposed)

Fig. 5.4. Comparison of segmentation results (red: nuclei contours, green: nuclei regions) of the proposed method with other methods overlaid onto original image (16^{th} image of the *FNb*)

ary as nuclei (Figure 5.3e, 5.4e). This is a common problem for 2D based methods

since they lack 3D information and their segmentation results are often inconsistent depthwise.



(a) *WSM* (b) *FNa* (c) *FNb*

Fig. 5.5. 3D segmentation results for *WSM*, *FNa*, and *FNb* (green: nuclei regions)

Lastly, Figure 5.5 portrays 3D segmentation results of various datasets using Voxx [161], a 3D visualization tool. For visualization purposes, each segmentation result was cropped into subvolumes ($60 \times 60 \times 20$), respectively, which demonstrate that the proposed method successfully identified nuclei in 3D. Segmentation in 3D provides smooth boundaries along z direction as well as x, y directions owing to segmenting 3D entire volumes. Conversely, results from 2D based segmentation methods do not guarantee consistency in depth and the results may be unrealistic especially in the xz and yz planes [48]. This is because 2D based segmentation methods do not utilize knowledge about the entire 3D topology of the image volume simultaneously. This smoothness along z direction as well as x, y direction is critical for quantitative analysis of microscopy volume. The proposed 3D active contours with inhomogeneity correction method is a natural 3D segmentation method iteratively updating 3D segmentation results considering 3D inhomogeneity field.

6. TUBULE SYNTHESIS AND SEGMENTATION BASED ON DEEP LEARNING

Previous chapter introduced traditional methods such as wavelet, ellipse fitting, and active contours based methods to achieve nuclei and tubular structures segmentation. In this chapter we present a method for segmenting and identifying individual tubular structure based on a combination of intensity inhomogeneity correction, data augmentation, followed by a CNN architecture. Our proposed method is evaluated at object-level metrics as well as pixel-level metrics using manually annotated groundtruth images of real fluorescence microscopy data. Our datasets are comprised of images of a rat kidney labeled with a phalloidin which labels filamentous actin collected using two-photon microscopy. A typical dataset we use in our studies consists of two tissue structures, the base membrane of the tubular structures and the brush border which is generally located interior to proximal tubules. Our goal here is to segment individual tubules enclosed by their membranes[1].

In addition, we present our proposed method to generate 3D tubule synthesis using gradient noise model. In this chapter we show our preliminary results of generating binary volume of the 3D tubular structures with randomly but smoothly moving without known pattern using the Simplex noise which is an upgraded version of the Perlin noise [162]. Current outcome will be used as an input of unpaired image-to-image transfer GAN such as SpCycleGAN [116], 3-Way SpCycleGAN [118], or fully 3D SpCycleGAN [119] in the future.

[1]The tubule segmentation based on convolutional neural networks with inhomogeneity correction method can be found on our paper in [86].

6.1 Tubule Segmentation Based on Convolutional Neural Networks with Inhomogeneity Correction

In this section, we demonstrate a method for segmenting and identifying individual tubular structure using a combination of intensity inhomogeneity correction, data augmentation, followed by a CNN architecture. First of all, we use intensity inhomogeneity correction as a preprocessing step to correct inhomogeneity background. Secondly, data augmentation is performed to increase the number of paired training set. This paired training set is used as an input of our CNN architecture to learn the relationship between original image and segmentation mask. Lastly, postprocessing is done to remove small components.

6.1.1 Proposed Method

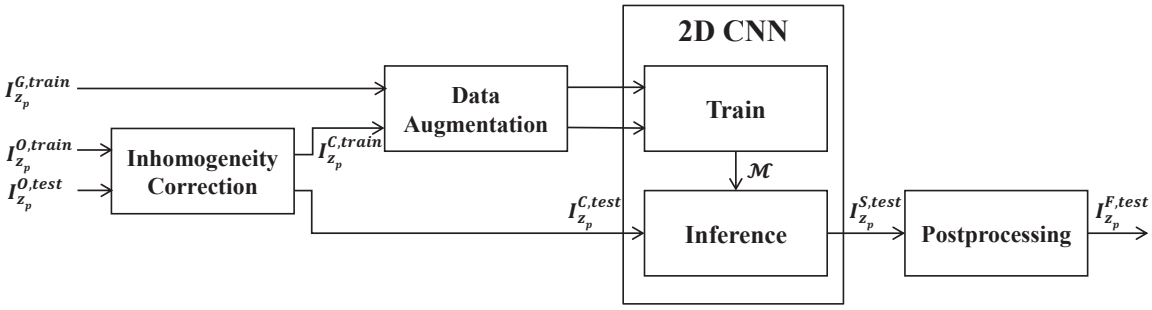


Fig. 6.1. Block diagram of the proposed segmentation method for tubule segmentation

Figure 6.1 shows a block diagram of the proposed method. Note that the notation we used in this blockdiagram is already defined in the Chapter 1.2. Additionally, we denote the original training and test images in the p^{th} focal plane by $I_{z_p}^{O,train}$ and $I_{z_p}^{O,test}$, respectively. In addition, $I_{z_p}^{G,train}$ and $I_{z_p}^{G,test}$ denote the groundtruth images that are used for training and testing that correspond to $I_{z_p}^{O,train}$ and $I_{z_p}^{O,test}$, respectively. Similarly, $I_{z_p}^{C,train}$ and $I_{z_p}^{C,test}$ denote inhomogeneity corrected training

and test images, respectively. Lastly, $I_{z_p}^{S,test}$ denotes the binary segmentation mask generated by our proposed deep learning architecture and $I_{z_p}^{F,test}$ denotes the final segmentation outcome. For example, the 100^{th} original focal plane is denoted as $I_{z_{100}}^{O}$, its corresponding groundtruth image by $I_{z_{100}}^{G}$, the inhomogeneity corrected version by $I_{z_{100}}^{C}$, the binary segmentation mask as $I_{z_{100}}^{S}$, and the final segmentation result by $I_{z_{100}}^{F}$, respectively.

As shown in Figure 6.1, our proposed network includes two stages: a training and an inference stage. During the training stage original training images ($I_{z_p}^{O,train}$) have their intensity inhomogeneities corrected ($I_{z_p}^{C,train}$) as a preprocessing step. Since fluorescence microscopy images suffer from intensity inhomogeneity due to non-uniform light attenuation, correcting intensity inhomogeneity helps improve final segmentation results. We then utilize both $I_{z_p}^{C,train}$ and $I_{z_p}^{G,train}$ as inputs to the data augmentation step to increase the number of training image pairs used for training the CNN model, $\mathcal{M}$. During the inference stage inhomogeneity correction is done on the test images ($I_{z_p}^{O,test}$) to obtain $I_{z_p}^{C,test}$. These $I_{z_p}^{C,test}$ are then used to segment tubules with the trained model $\mathcal{M}$.

Intensity Inhomogeneity Correction

Due to non-uniform intensities of fluorescence microscopy where center regions of the focal plane are generally brighter than boundary regions, simple intensity based segmentation methods failed to segment biological structures especially near image boundaries [55]. As indicated from Chapter 5, we employed a multiplicative model where the original microscopy volume is modeled as

$$I^O = W \circ I^C + N. \tag{6.1}$$

Here, W and N are a 3D weight array and a zero mean 3D Gaussian noise array, respectively, both of same size as the original microscopy volume. Specifically, W represents weight values for each voxel location that accounts for the degree of in-

tensity inhomogeneity. The $\circ$ operator denotes the Hadamard product representing voxelwise multiplication.

The main idea of the multiplicative model is that an original volume, I^O, is modeled as the product of a 3D inhomogeneity field W with a corrected volume I^C and the product corrupted by additive 3D Gaussian noise N. An iterative technique to finding W and then correcting for the intensity inhomogeneities based on this model is described in 'Method 1". Our proposed method uses this inhomogeneity correction technique as a preprocessing step for both training and inference. Examples of original and inhomogeneity corrected images are shown in Figure 6.3a and Figure 6.3b, respectively.

Data Augmentation

Our training data consists of paired images which are original microscopy images and corresponding manually annotated groundtruth images. Generating manually annotated groundtruth images is a time consuming process and thus impractical when generating large numbers of images. Data augmentation is typically used when the available training data size is relatively small to generate additional groundtruth images [87]. In this chapter we utilize an elastic deformation to generate realistic tubular structures with different shapes and orientations. This allows the network to learn various deformed tubular structures, and is particularly useful for analysis of microscopy images especially for tubular structures that appear in varying shapes and orientations [87].

We used elastic deformation by employing a grid of control points located every 64 pixels along the horizontal and vertical directions and displacing these control points randomly within 15 pixels in each direction to generate a deformation field. The deformation field is used to deform the p^{th} focal planes $I_{z_p}^{C,train}$ and $I_{z_p}^{G,train}$ by fitting 2D B-spline basis function to the grid followed by bicubic interpolation [163]. We generated 100 random deformation fields for each image pair and use them to

generate 100 deformed image pairs. Each deformed image is rotated 0°, 90°, 180°, 270° to generate four sets of rotated images while preserving the original image size. Each rotated image is then flipped left and right to generate another two sets of images. In our experiment, we manually annotated five pairs of training data during the training stage. Since the elastic deformation uses 100 deformations followed by four rotations and two flips for each deformed image, 4000 pairs of images were generated for training.

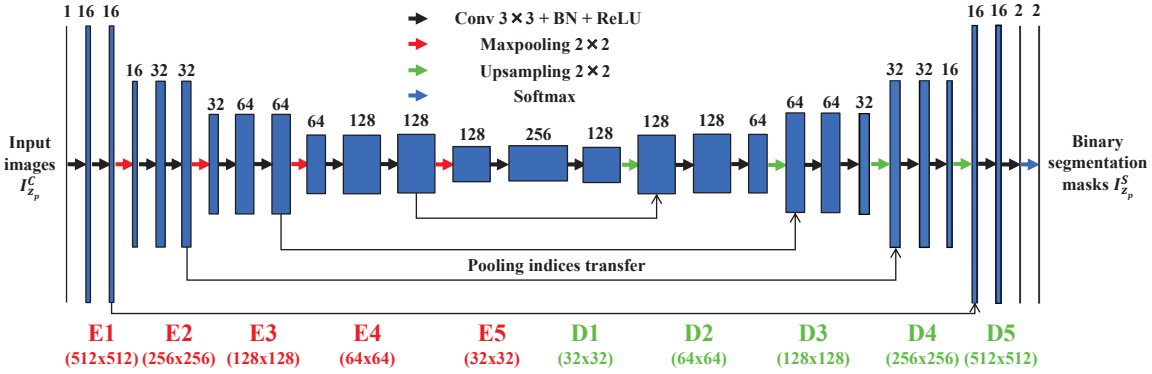


Fig. 6.2. Proposed convolutional neural network architecture

Convolutional Neural Network (CNN)

The architecture of our convolutional neural network, shown in Figure 6.2, consists of 5 encoder layers denoted as $E1$ through $E5$ and 5 decoder layers denoted as $D1$ through $D5$ that are serially connected followed by a softmax layer at the end. Each encoder layer consists of a 3×3 kernel with 1 pixel padding to maintain same image size, a batch normalization step [164] to perform image whitening, followed by a rectifier-linear unit (ReLU). The combination of convolution, batch normalization, and ReLU are performed twice at every encoder. Finally, maxpooling with a stride of 2 is used to reduce dimensionality. This encoder scheme is similar to VGGNet [84] which shrinks the input dimensions but increases the number of filters in the deeper

structures. In Figure 6.2, each encoder's input dimension is indicated in red under the encoder layers. Also note that the number shown above each layer represents the number of filters utilized for training. For example, an input image of size $512 \times 512 \times 1$ is resized to $256 \times 256 \times 16$ at the input to the E2 layer. As the image passes through the all encoder layers, its X and Y dimensions shrink to 32, respectively, but number of filters utilized increases to 256. Therefore, the input to the first decoder layer is of dimension $32 \times 32 \times 256$.

Conversely, each decoder is comprised of two 3×3 kernels with 1 pixel padding, batch normalization, and ReLU. Instead of a maxpooling layer, the decoder has an unmaxpooling layer to upsample the data to increase dimensionality. Note that this upsampling process is a reconstruction process. To achieve better upsampling maxpooling indices from each encoder layer are recorded and transferred to the corresponding same size unmaxpooling layer ($E1 \rightarrow D5, \ldots, E4 \rightarrow D2$). At the end of the encoder-decoder structure, a softmax classifier layer is utilized to determine whether each pixel location belongs to a tubule or background using a probability map. Note that the output of the softmax layer is of size $512 \times 512 \times 2$ because the final output includes two probability maps corresponding to the two classes: tubule or background. These probability maps are thresholded at 0.5 to produce binary segmentation masks.

During the training stage augmented training images ($I_{z_p}^{C,train}$) are randomly selected and used to train the model $\mathcal{M}$ for each iteration. The segmentation mask is compared with the corresponding groundtruth ($I_{z_p}^{G,train}$) and a loss value is obtained for each iteration. Here, we use a 2D cross entropy loss function that is minimized using stochastic gradient descent (SGD) with a fixed learning rate and a momentum. During the inference stage we use the trained model $\mathcal{M}$ with test images ($I_{z_p}^{C,test}$) to obtain binary segmentation masks ($I_{z_p}^{S,test}$). During the postprocessing step we clean up objects less than γ pixels from $I_{z_p}^{S,test}$ followed by a hole filling operation to obtain final segmentation results ($I_{z_p}^{F,test}$). Note that the hole filling operation assigns a background pixel to a tubule pixel if the background pixel's 4 neighborhood pixels are all tubule pixels.

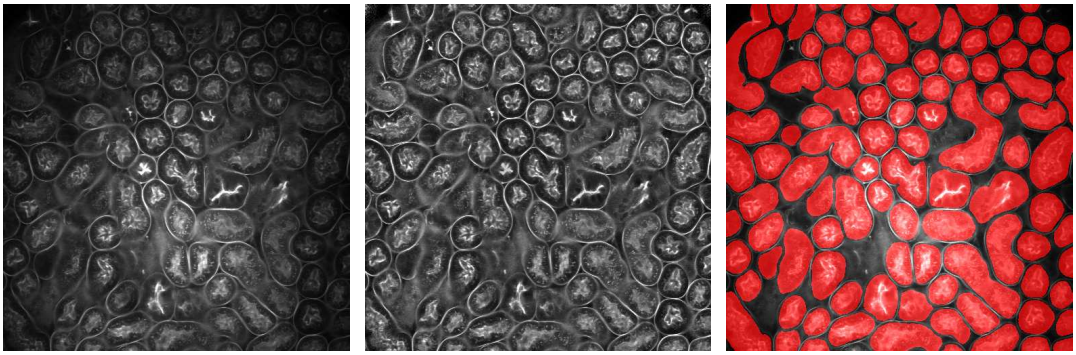
6.1.2 Experimental Results

Our datasets are comprised of images of a rat kidney labeled with a phalloidin which labels filamentous actin collected using two-photon microscopy. A typical dataset we use in our studies consists of two tissue structures, the base membrane of the tubular structures and the brush border which is generally located interior to proximal tubules. Our goal here is to segment individual tubules enclosed by their membranes.

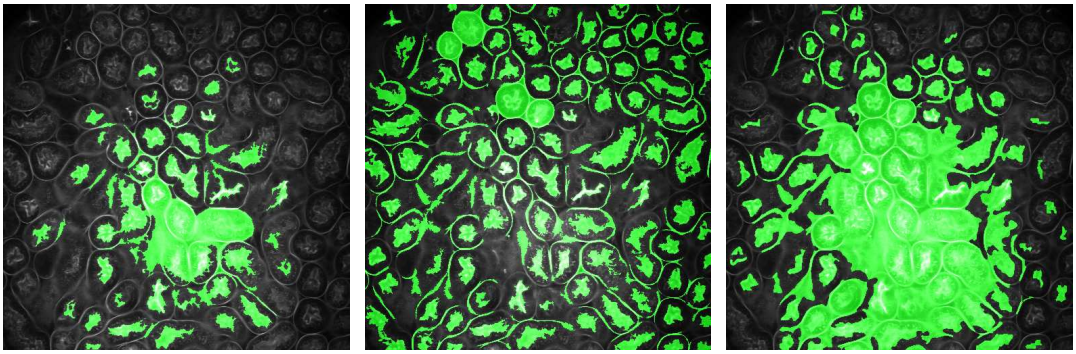
The performance of our proposed method was tested on two different datasets: *WSM* and *Lectin*. *WSM* is comprised of $Z = 512$ grayscale images, each of size $X \times Y = 512 \times 512$ pixels, whereas *Lectin* consists of $Z = 821$ grayscale images, each of size $X \times Y = 640 \times 640$ pixels. We selected five different images from *WSM* and generated corresponding manually annotated groundtruth images to train model $\mathcal{M}$. Our deep learning architecture was implemented in Torch 7 [165] using a fixed learning rate 10^{-5} and a momentum of 0.9. As indicated, 4000 pairs of images were generated by the elastic deformation, rotations, and flips using these five pairs of images. Note that each training data was used as a batch so that 4000 iterations were performed per epoch. We used 200 epochs for training our proposed network. In addition, $\gamma = 100$ was used for the removal of small objects. The performance of the proposed method was evaluated using manually annotated groundtruth images ($I_{z_p}^{G,test}$) at different depths in *WSM* that were never used during the training stage. For visual evaluation and comparison segmentation results of $I_{z_{100}}$ in *WSM* using various techniques are presented in Figure 6.3.

Qualitative Evaluation

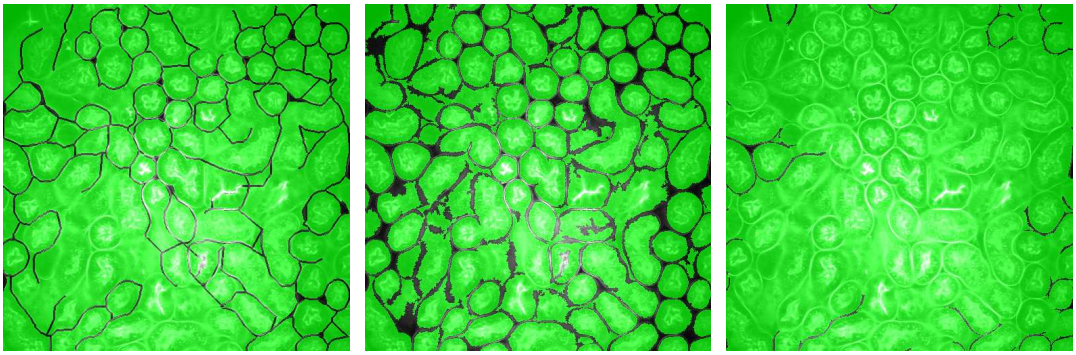
The first row in Figure 6.3 displays an original microscopy image ($I_{z_{100}}^{O}$), its inhomogeneity corrected version ($I_{z_{100}}^{C}$), and manually delineated groundtruth ($I_{z_{100}}^{G}$), respectively. For brevity we have omitted the superscript *test* in the notation. The second row shows segmentation results of various 3D methods such as 3D region-based



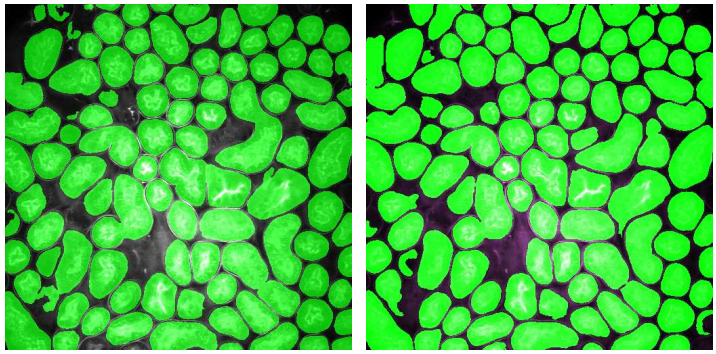
(a) $I^O_{z_{100}}$ (b) $I^C_{z_{100}}$ (c) $I^G_{z_{100}}$



(d) *3Dac* [48] (e) *3DacIC* [55] (f) *3Dsquassh* [63]

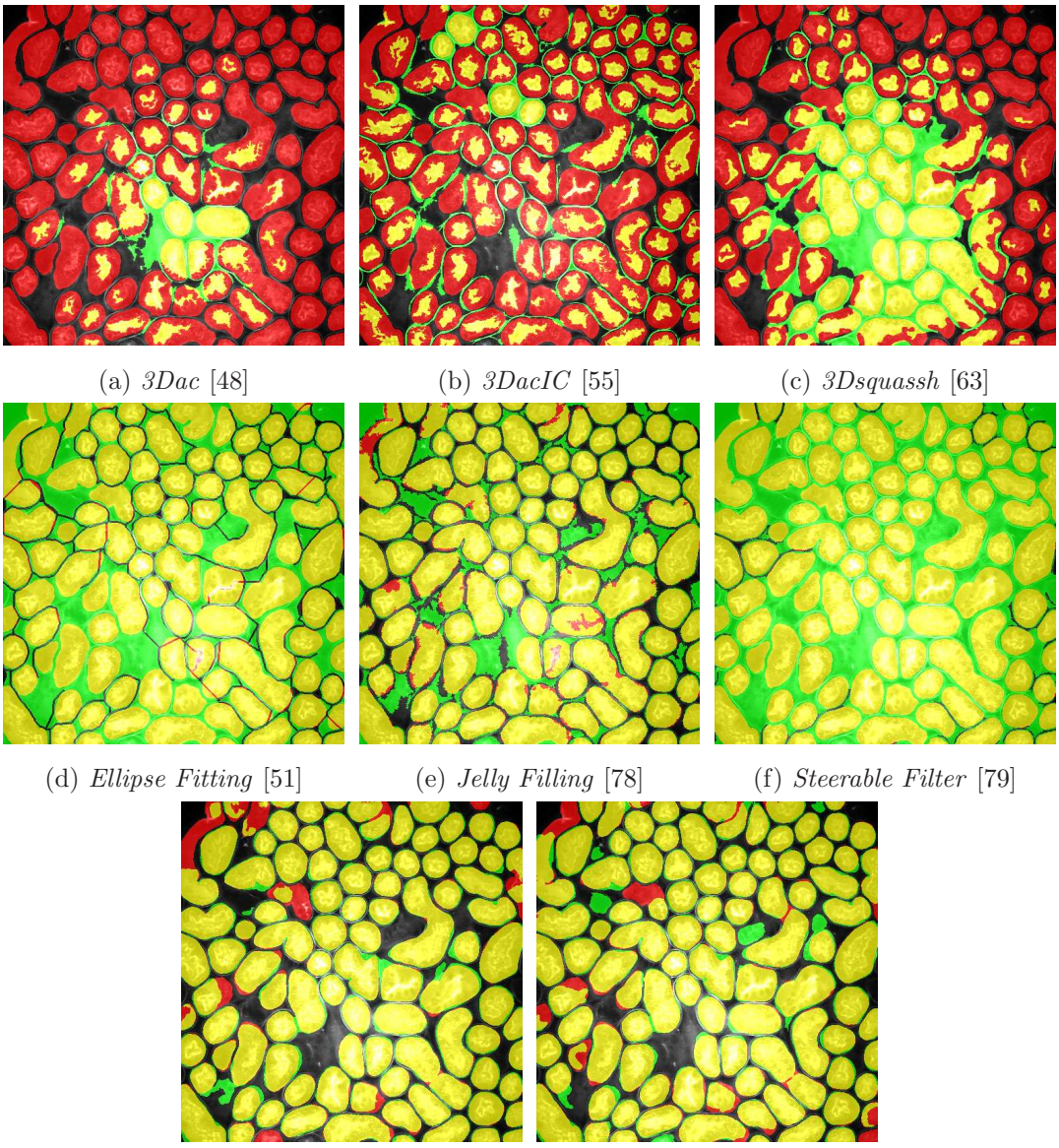


(g) *Ellipse Fitting* [51] (h) *Jelly Filling* [78] (i) *Steerable Filter* [79]



(j) *2DCNN* [85] (k) *2DCNNIC* (Proposed)

Fig. 6.3. Segmentation results obtained by the proposed method and other methods as well as the corresponding groundtruth data for the 100^{th} image ($I_{z_{100}}$) in *WSM*. Segmentation results are highlighted in green and corresponding groundtruth in red.



(a) *3Dac* [48] (b) *3DacIC* [55] (c) *3Dsquassh* [63]

(d) *Ellipse Fitting* [51] (e) *Jelly Filling* [78] (f) *Steerable Filter* [79]

(g) *2DCNN* [85] (h) *2DCNNIC* (Proposed)

Fig. 6.4. Qualitative evaluation/comparison of tubule segmentation results (shown in green) from the proposed method as well as other methods overlaid onto groundtruth image (shown in red) for $I_{z_{100}}$ belonging to *WSM*

active contours [48] (*3Dac*), 3D active contours with inhomogeneity correction [55] (*3DacIC*), and 3D Squassh presented in [63] (*3Dsquassh*). Similarly, the third row portrays various segmentation methods particularly designed for tubular structure segmentation such as ellipse fitting method presented in [51] (*Ellipse Fitting*), the Jelly filling method in [78] (*Jelly Filling*), and tubule segmentation using steerable filter [79] (*Steerable Filter*). Finally, the last row shows segmentation results of our proposed CNN architecture without inhomogeneity correction [85] (*2DCNN*) and with inhomogeneity correction (*2DCNNIC*).

For visual comparison we highlighted groundtruth regions in red, segmented tubule regions in green, and background in black. As observed in Figure 6.3, our proposed method appeared to perform better than the other six methods shown in the second and third rows by distinguishing tubules and was similar performance to *2DCNN*. Note that since some methods such as *Ellipse Fitting*, *Jelly Filling*, and *Steerable Filter* only segmented boundaries of tubule structures, tubule interiors were filled in order to perform a fair comparison using connected components with a 4-neighborhood systems. Also, based on the assumption that tubule regions should contain lumen, if a filled region contained lumen pixel, the region was identified as a tubule region. However, if a filled region did not contain any lumen pixels, the region was considered as a background region.

The segmentation results shown in the second row generally missed many tubule regions. More specifically, *3Dac* and *3Dsquassh* could not capture the tubular structures but captured some in the center regions due to the intensity inhomogeneity of microscopy images. *3DacIC* failed to segment tubular structures but captured multiple lumens inside tubules as well as some tubule boundaries. In contrast, the segmentation results displayed in the third row showed falsely detected tubules. The main reason is that these tubule segmentation methods focused only on detecting boundaries of tubular structures. In particular, due to weak/blurry edges of fluorescence microscopy images, many boundaries were not continuous causing the filling operation to overflow from one tubule to another or to the background regions. The

segmentation results using the CNN generally successfully segmented and identified each tubule region.

Figure 6.4 provides an alternative way to show the segmentation results. In particular, yellow regions correspond to true positives which are pixel locations that are identified as tubules in both the groundtruth and segmentation results. Green regions correspond to false positives which are pixel locations that are identified as background in groundtruth but tubules in segmentation results. Similarly, red pixels correspond to false negatives, namely pixel locations identified as tubules in the groundtruth but background in segmentation results, and black pixel regions correspond to true negative that are identified as background in both groundtruth and segmentation results. The green regions indicate Type-I error (false alarm) regions and the red regions represent Type-II error (miss) regions. As observed from Figure 6.4, the segmentation results in the first row contained large red regions which mean large regions of tubules were missed. Conversely, the segmentation results shown in the second row contained many green regions indicating many background regions were falsely segmented as tubule regions. In contrast, the segmentation results in the third row had reasonably small green regions and red regions which indicate that the deep learning based segmentation results had higher pixel accuracy with relatively low Type-I and Type-II errors.

Quantitative Evaluation

In addition to the qualitative evaluation, quantitative metrics for evaluating the proposed method's segmentation accuracy of objects were utilized. In particular, we used pixel-based and object-based metrics. In the pixel-based metric, the pixel accuracy (PA), Type-I error, and Type-II error of pixel segmentation were obtained based on the manually annotated groundtruth images. Here, PA, Type-I, and Type-II are defined as below:

$$PA = \frac{N^p_{tp} + N^p_{tn}}{N^p_{total}}, \quad Type - I = \frac{N^p_{fp}}{N^p_{total}}, \quad Type - II = \frac{N^p_{fn}}{N^p_{total}} \tag{6.2}$$

where N^p_{tp}, N^p_{tn}, N^p_{fp}, and N^p_{fn} are defined to be the number of segmented pixels that were labeled as true positives, true negatives, false positives, false negatives, respectively. N^p_{total} denotes the total number of pixels in a image. These three pixel-based metrics obtained for 8 different segmentation results are provided in Table 6.1. As shown in Figure 6.4, Type-II errors of the first three methods (*3Dac*, *3DacIC*, *3Dsquassh*) were much higher compared to other methods. Similarly, Type-I errors of next three methods (*Ellipse Fitting*, *Jelly Filling*, *Steerable Filter*) were much higher than those of the other methods. However, *2DCNN* and *2DCNNIC* had high PA and relatively low Type-I and Type-II errors.

Table 6.1.
Pixel-based evaluation of the proposed method and other known methods in terms of Pixel Accuracy (PA), Type-I error, and Type-II error

	$I_{z_{100}}$ of the *WSM*			$I_{z_{200}}$ of the *WSM*		
Method	PA	Type-I	Type-II	PA	Type-I	Type-II
3Dac [48]	37.74%	**3.31%**	58.95%	38.98%	**2.72%**	58.30%
3DacIC [55]	42.92%	8.06%	49.02%	44.58%	4.84%	50.59%
3Dsquassh [63]	47.02%	11.80%	41.18%	48.37%	9.55%	42.09%
Ellipse Fitting [51]	76.17%	22.79%	1.04%	76.11%	22.98%	0.91%
Jelly Filling [78]	83.91%	13.36%	2.73%	81.76%	15.38%	2.86%
Steerable Filter [79]	70.98%	28.98%	**0.04%**	71.00%	28.97%	**0.03%**
2DCNN [85]	**90.57%**	5.25%	4.17%	86.92%	5.64%	7.44%
2DCNNIC (Proposed)	90.04%	6.44%	3.52%	**88.66%**	5.77%	5.57%

In addition, our segmentation methods were evaluated using object-based criteria described in the *2015 MICCAI Grand Segmentation Challenge* [88, 166] namely: the F1 score metric, the Dice Index, and the Hausdorff Distance.

Table 6.2.
Object-based evaluation of the proposed method and other known methods in terms of F1 score, the Dice Index (OD), and the Hausdorff Distance (OH)

	$I_{z_{100}}$ of the *WSM*			$I_{z_{200}}$ of the *WSM*		
Method	F1	OD	OH	F1	OD	OH
3Dac [48]	0.90%	20.86%	95.72	1.83%	20.74%	125.69
3DacIC [55]	0.86%	36.45%	35.07	0.00%	38.02%	30.41
3Dsquassh [63]	1.83%	11.64%	223.34	1.94%	14.34%	181.58
Ellipse Fitting [51]	61.15%	47.10%	144.28	48.48%	29.34%	303.34
Jelly Filling [78]	81.82%	71.58%	52.93	74.53%	60.73%	76.26
Steerable Filter [79]	9.90%	5.32%	455.83	4.12%	4.00%	521.83
2DCNN [85]	91.49%	90.09%	13.28	86.96%	87.10%	16.80
2DCNNIC (Proposed)	**92.63%**	**90.12%**	**11.95**	**90.61%**	**89.65%**	**11.76**

The F1 score metric is a measure of the segmentation/detection accuracy of individual objects. The evaluation of the F1 score metric is based on two metrics, precision P and recall R. Denoting the number of tubules correctly identified by N^o_{tp}, the number of objects that are non-tubules but identified as tubules by N^o_{fp}, and the number of tubules that are not correctly identified as tubules by N^o_{fn}, respectively, then precision P and recall R are obtained as [88]

$$P = \frac{N^o_{tp}}{N^o_{tp} + N^o_{fp}} \quad \text{and} \quad R = \frac{N^o_{tp}}{N^o_{tp} + N^o_{fn}}. \tag{6.3}$$

Given the values of P and recall R, the F1 is found using

$$F1 = \frac{2PR}{P + R}. \tag{6.4}$$

It is to be noted that a tubule segmented by the proposed method (or any other method for that matter) that overlaps at least 50% with its corresponding manually annotated tubule is labeled as a true positive and added to the count of the true positives (N^o_{tp}), otherwise it is considered as a false positive and added to the count of the false positives (N^o_{fp}). Similarly, a manually annotated tubule that has no corresponding segmented tubule or overlaps less than 50% with segmented tubular regions is considered to be a false negative and added to the count of the false negatives (N^o_{fn}).

As observed in Table 6.2, F1 score of proposed method outperformed other competed methods and had some margin compared to *2DCNN*. In particular, *3Dac*, *3DacIC*, *3Dsquassh*, and *Steerable Filter* were extremely low F1 score since these segmentation methods had huge Type-I or Type-II error and it resulted in high P with low R or low P with high R. Since F1 score is the harmonic mean of P and R, both P and R should be high to achieve high F1 score.

As mentioned above a second metric used to evaluate segmentation accuracy is the Dice Index (OD). The Dice Index [167] is a measure of similarity between two sets of samples. In our case, the two sets of samples are the sets of voxels belonging

to a manually annotated tubule denoted by G, and the set of voxels belonging to a segmented tubule denoted by S. The Dice Index between G and S is defined as

$$D(G,S) = \frac{2|G \cap S|}{|G| + |S|} \tag{6.5}$$

where $|\cdot|$ denotes set cardinality which in this case will be the number of voxels belonging to an object. A higher value of the Dice Index indicates better segmentation match/results relative to the groundtruth data. A practical way of evaluating the Dice Index for segmented objects is described in [166] and is given by

$$D(G,S) = \frac{1}{2}\left[\sum_{i=1}^{n_S} w_i D(G_i, S_i) + \sum_{j=1}^{n_G} \tilde{w}_j D(\tilde{G}_j, \tilde{S}_j)\right] \tag{6.6}$$

where

$$w_i = |S_i|/|\sum_{p=1}^{n_S} |S_p|, \quad \tilde{w}_j = |\tilde{G}_j|/|\sum_{q=1}^{n_G} |\tilde{G}_q|. \tag{6.7}$$

In Eq (6.6), S_i denotes the i^{th} tubule ($i \in \{1, \ldots, n_S\}$) obtained by a segmentation method and G_i denotes a manually annotated tubule that is maximally matched with S_i. Similarly, $\tilde{G}_j$ denotes the j^{th} tubule ($j \in \{1, \ldots, n_G\}$) identified in the groundtruth data and $\tilde{S}_j$ denotes a segmented tubule that is maximally matched with $\tilde{G}_j$. Finally, n_S and n_G denote the total number of segmented and manually annotated tubules, respectively. The first summation term in Eq (6.6) represents how well each groundtruth tubule overlaps with its segmented counterpart, whereas the second summation term represents how well each segmented tubule overlaps with its manually annotated counterpart. The terms w_i and $\tilde{w}_j$ which are used to weight the summation terms represent the fraction of the space that each tubule S_i and $\tilde{G}_j$ occupies within the entire tubule region, respectively.

While the Dice Index measures segmentation accuracy, a third metric, the Hausdorff Distance (OH), is needed to evaluate shape similarity. The Hausdorff Distance [168], $H(G,S)$, between a segmented tubule S and its manually annotated counterpart G, is defined to be

$$H(G,S) = \max\{\sup_{x \in G} \inf_{y \in S} ||x - y||_2, \ \sup_{y \in S} \inf_{x \in G} ||x - y||_2\}. \tag{6.8}$$

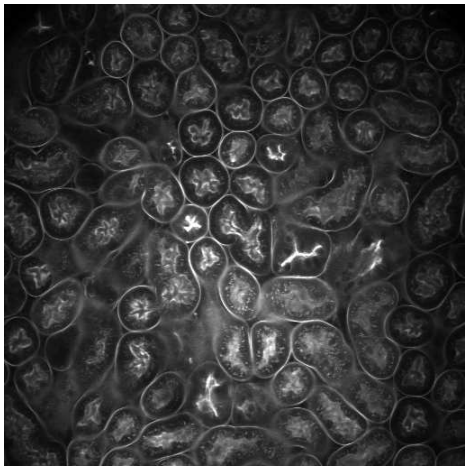
Here, $||x-y||_2$ denotes the Euclidean distance between a pair of pixels x and y. Based on Eq (6.8), the Hausdorff Distance obtains the maximum distance among all pairs of voxels on the boundaries of S and G. Therefore, a smaller value of the Hausdorff Distance indicates a higher similarity in shape between the boundaries of S and G. As done above (see Eq (6.6)), a practical way of finding the Hausdorff Distance between a segmented tubule S and its manually annotated counterpart G is given by [166]:

$$H(G,S) = \frac{1}{2}\left[\sum_{i=1}^{n_S} w_i H(G_i, S_i) + \sum_{j=1}^{n_G} \tilde{w}_j H(\tilde{G}_j, \tilde{S}_j)\right] \tag{6.9}$$

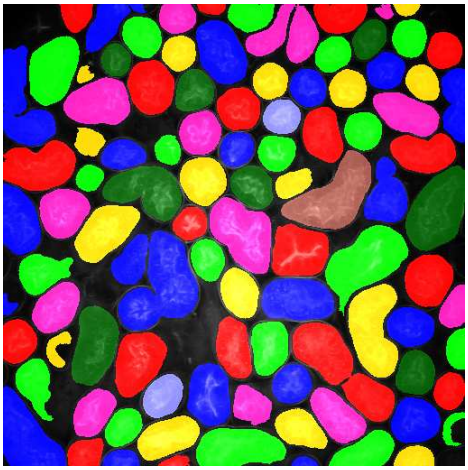
where the parameters w_i and $\tilde{w}_j$ are defined in Eq (6.7).

The performance of the proposed method and other methods based on the F1 score, OD, and OH metrics were obtained and tabulated in Table 6.2. As mentioned above higher values of F1 and OD are considered to be indicators of better segmentation results. In contrast, lower values of OH indicate better segmentation result. As can be seen in Table 6.2, our proposed method outperformed all the other segmentation methods against which the proposed method is being evaluated. In particular, *3Dac*, *3DacIC*, *3Dsquassh*, and *Steerable Filter* had low F1 scores since these segmentation methods had large Type-I or Type-II errors. Similarly, all of the methods except for *2DCNN* suffered from low OD and high OH values. In particular, since the segmentation results of *3Dsquassh*, *Ellipse Fitting*, and *Steerable Filter* failed to distinguish most of the individual tubules, they exhibited low OD and high OH values. Note that *3DacIC* had relatively low OH *and* low OD values since it segmented some tubule boundaries as well as some partial regions (lumen) inside the tubules. Lastly, the use of intensity inhomogeneity correction in the proposed method improved its performance relative to that of *2DCNN*.

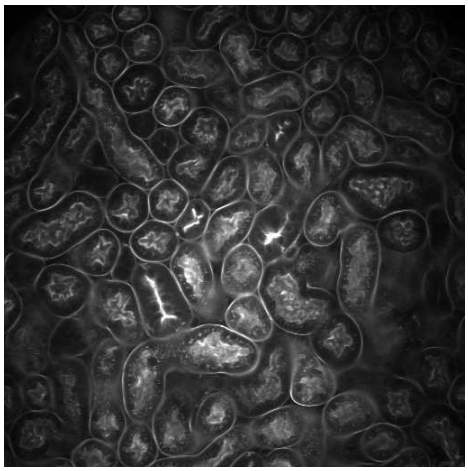
For visual evaluation we provide the segmentation results of the proposed method using two different datasets: *WSM* and *Lectin*, sampled at different depths within the volumes. The first row shows original microscopy images $I^O_{z_{100}}$, $I^O_{z_{150}}$, and $I^O_{z_{200}}$ from *WSM* and the second row displays the segmentation results corresponding to the first row. To better visualize the segmentation results, we highlighted individual



(a) $I^O_{z_{100}}$ of *WSM*

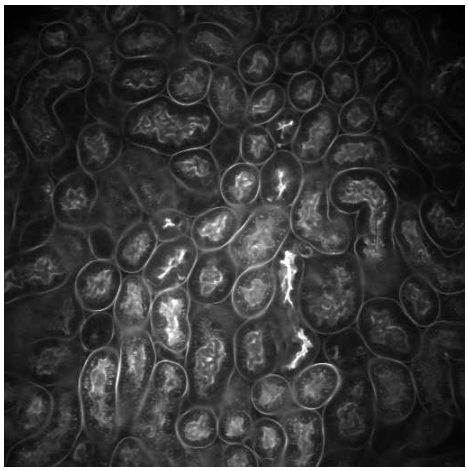


(b) $I^F_{z_{100}}$ of *WSM*



(c) $I^O_{z_{150}}$ of *WSM*

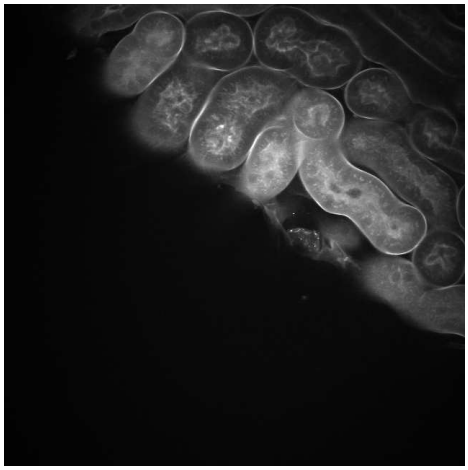
(d) $I^F_{z_{150}}$ of *WSM*



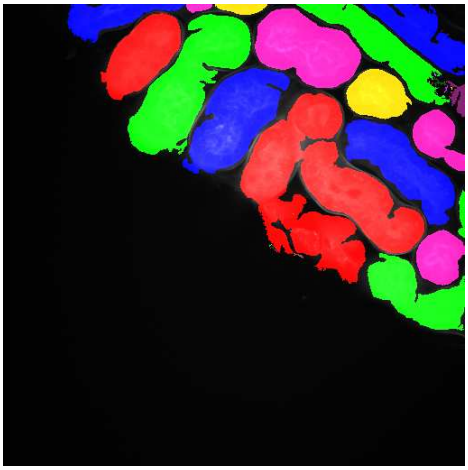
(e) $I^O_{z_{200}}$ of *WSM*

(f) $I^F_{z_{200}}$ of *WSM*

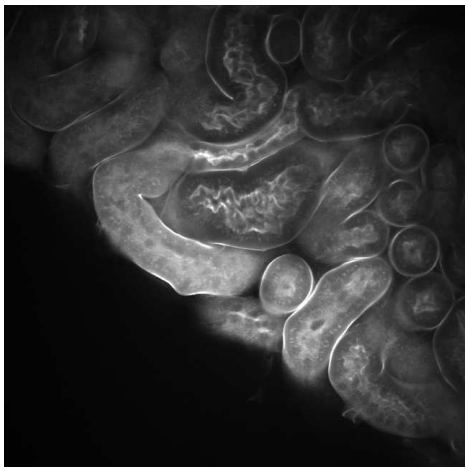
Fig. 6.5. Original and color coded segmentation results of the proposed method on different depth of *WSM* using model $\mathcal{M}$



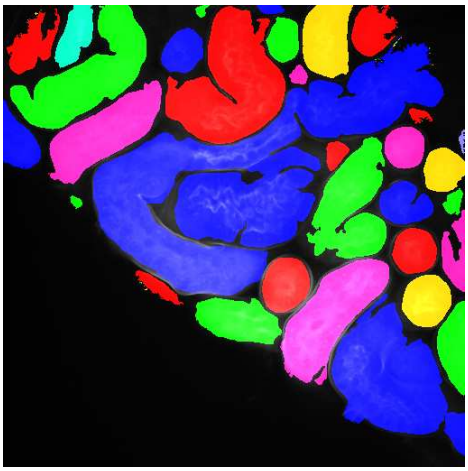
(a) $I^{O}_{z_{50}}$ of *Lectin*



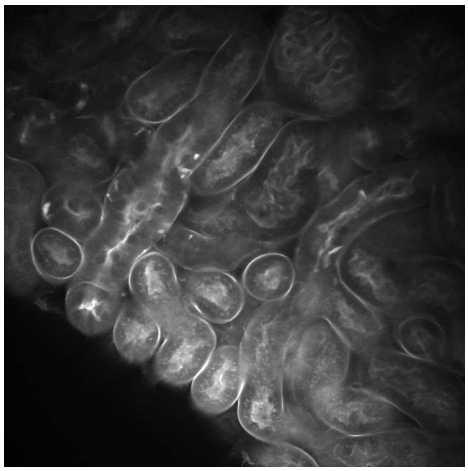
(b) $I^{F}_{z_{50}}$ of *Lectin*



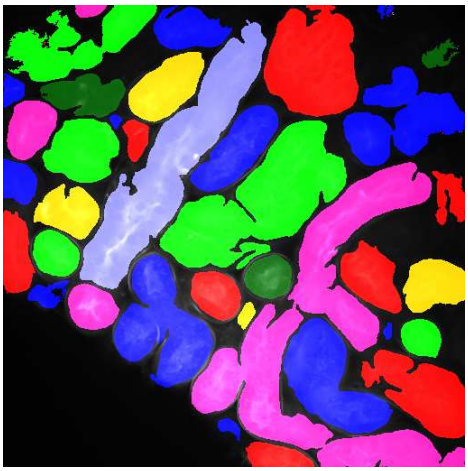
(c) $I^{O}_{z_{150}}$ of *Lectin*



(d) $I^{F}_{z_{150}}$ of *Lectin*



(e) $I^{O}_{z_{250}}$ of *Lectin*



(f) $I^{F}_{z_{250}}$ of *Lectin*

Fig. 6.6. Original and color coded segmentation results of the proposed method on different depth of *Lectin* using same trained model $\mathcal{M}$

tubules with different colors and overlaid them onto the original microscopy images. Similarly, the third row exhibits original microscopy images I^O_{z50}, I^O_{z150}, and I^O_{z250} from *Lectin*. Their corresponding segmentation results are shown in the fourth row. Note that the model $\mathcal{M}$ which was trained on *WSM* was used for *Lectin* during the inference stage. Although the shape, size, and orientation of tubular structures presented in *Lectin* are all different from *WSM*, the proposed method can still successfully segment and identify individual tubules presented in *Lectin* as well as individual tubules in *WSM*.

6.2 Three Dimensional Tubule Synthesis and Segmentation

In this section, we demonstrate a method for generating 3D binary synthetic tubules. This will be the first step toward to generate realistic 3D tubular structures using SpCycleGAN or any of unpaired GAN. Our previous approach for nuclei synthesis and segmentation was demonstrated in [116]. For the 3D nuclei synthesis, we have an assumption that each nucleus can be modeled as an 3D ellipsoidal shape with random orientation and rotation angle with translation. We also allow a few voxels overlapping with adjacent nuclei.

However, the shape of the tubular structure is impossible to take simple model like nuclei since the orientation, shape, and length, and size of tubule is varied without known specific patterns. There are some existed methods that try to generate synthetic shape of biological structures. The method called VascuSynth which proposed simulating vascular trees for generating volumetric image data was demonstrated in [169,170]. However, their synthetic vascular tree structures is quite different from our tubule shape. Moreover, a synthetic model of the spatial tumor including many structures was proposed in [171] for generating different grades of the tissues in histopathological image domain. Among many shapes generated in [171], the shape of crypt composed of a single layer of epithelium glandular structures [171] may look similar as our tubular structure which assumed 2D ellipse with varying major and mi-

nor axes. Also, tubular structure segmentation method was presented in [73] where a tubular shape was assumed as the envelop of a set of disks/spheres with continuously changing center points and radii. Inspired by the idea presented in [73], we assume our tubules can be modeled as an union of the multiple disk (2D) or spheres (3D).

6.2.1 Proposed Method

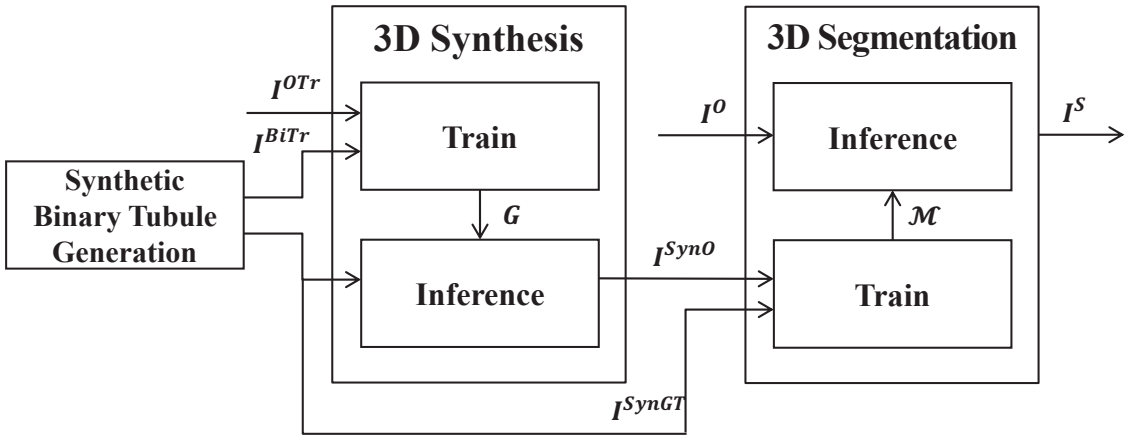
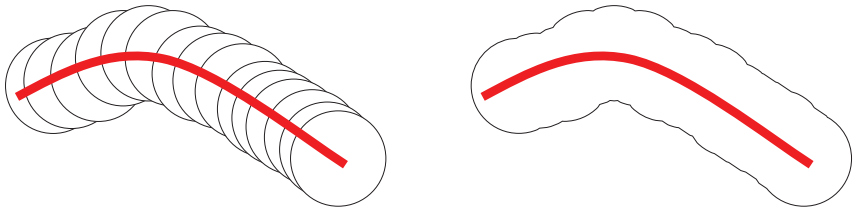


Fig. 6.7. Block diagram of the our proposed 3D tubule synthesis method

Figure 6.7 shows a block diagram of the proposed method. As observed, our proposed method is a two stage method: 3D synthesis and 3D segmentation. Again, the notation we used in this block diagram is based on the definition in the Chapter 1.2. Additionally, an original volume used for synthesis training is denoted as I^{OTr} and an original volume to be segmented as I^{O}. Also, synthetically generated binary volume used for synthesis training is denoted as I^{BiTr}. Similarly, I^{SynO} and I^{SynGT} are synthetic microscopy volume and corresponding synthetic ground truth volume, respectively. Once the synthetic binary tubule generation is done, I^{OTr} and I^{BiTr} are used for 3D synthesis training and obtain a generative model $\mathcal{G}$. Then, from the 3D synthesis inference, I^{SynO} can be generated based on I^{SynGT}. Here, I^{SynGT} is again generated by the synthetic binary tubule generation block but different set

with I^{BiTr}. Once this I^{SynO} and I^{SynGT} are obtained, the 3D segmentation training become possible since we have paired dataset to use. The model M is obtained after the 3D segmentation training and model $\mathcal{M}$ is used for the 3D segmentation inference to get final 3D segmentation results (I^S).

Synthetic Binary Tubule Generation

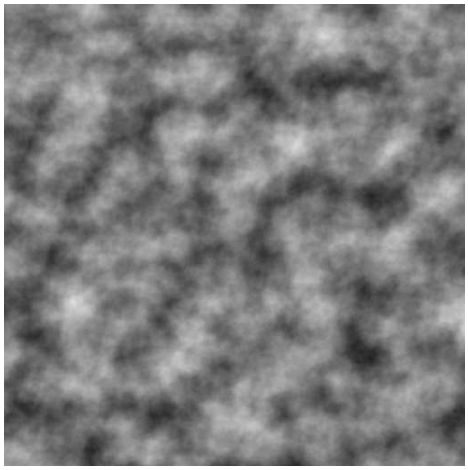


(a) Before union of the multiple disks (2D)

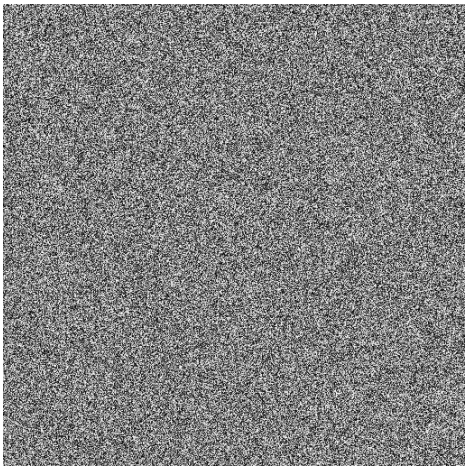
(b) After union of the multiple disks (2D)

Fig. 6.8. The union of the synthetic tubule boundary (Black) with corresponding centerline (Red)

As aforementioned we assume our tubules can be modeled as an union of the multiple disk (2D) or spheres (3D). The example of the 2D cases are shown in Figure 6.8. As observed multiple disk can be formed as smooth curves so that their union can form as a tubule boundary. Note that the centerline corresponding to each disk is highlighted in red. Here, the centerline is defined as a collection of center location of the disks used as the union of the disks. Thus, the centerline generation is an important to task which should be smooth curves but with random orientation so that the centerline can resemble the realistic microscopy tubular structures. Also, our goal is extending this to 3D, the centerline should eventually travel on the 3D space. One of the example that has similar idea of randomly moving is a random walk [172] easily extend to the 3D space. However, the random walk does not provide the smooth curve in 2D or 3D. Rather, it moves with sharp turn which could not represent the tubular structures well enough.



(a) Example of the Simplex noise (512 × 512)



(b) Example of the White noise (512 × 512)

Fig. 6.9. Comparison of different type of noises

Instead of the random walk, the Simplex noise which is an improved version of the Perlin noise [162] with flow field is used for generating a centerline for the tubule candidate. Note that the Perlin noise is a type of gradient noise. The comparison of gradient noise and white noise are shown in Figure 6.9. We generate both noises size of 512×512 resolution. As seen from the Figure 6.9 the gradient noise values are not independent with their locations. One of the properties of the white noise is their independent regardless of the locations. These kinds of noises are typically used for the computer graphics that looks more realistic but still needs random shape or orientation. For example, clouds on the sky should be realistically generated by modern computer graphics but it contains randomness of the shape. Another example is topographic maps. Again, this map should look realistic but needs randomness to generate different scenes.

To introduce random movement in 2D space, we firstly generate 3D Simplex noise denoted as $n^{3D}(x, y, t)$. Here, the first two of the 3D coordinates corresponds to the 2D space location (x, y) and third coordinate represents time (t) stamp. Then, 2D

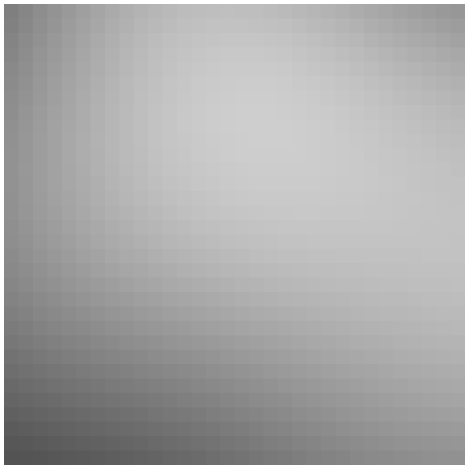
vector (v_x, v_y) is generated in each grid location at every time point as a function of the noise level such that

$$v_x = \cos(2\pi n^{3D} c)$$
$$v_y = \sin(2\pi n^{3D} c) \tag{6.10}$$

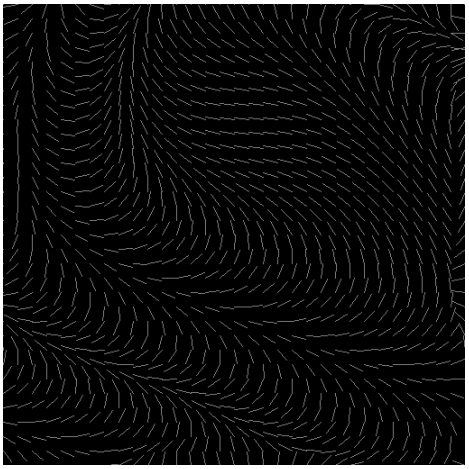
where c is a coefficient to control vector angle changes within time stamp. Note that the Simplex noise is a type of gradient noise so that their corresponding flow fields generated by 3D Simplex noise is temporally correlated as well as spatially correlated. This characteristics results in generating smooth changing vectors in adjacent locations and close time stamp. Once this 3D flow field is obtained, the particle is randomly located in the 2D space in the first time frame and move along with flow field in time. The trajectory of this particle is used as a centerline of the synthetic tubules.

The example of the Simplex noise, corresponding flow fields, and particle located in 2D with and without flow fields are exhibited in Figure 6.10. To generate the full size of flow field the Simplex noise is generated with size of 32×32. Later, when flow fields are generated, 16 pixel interval in between each grid point so that the flow fields are occupied in the size of 512×512. Then, the particle is randomly located and move along with the flow fields as shown in Figure 6.10d.

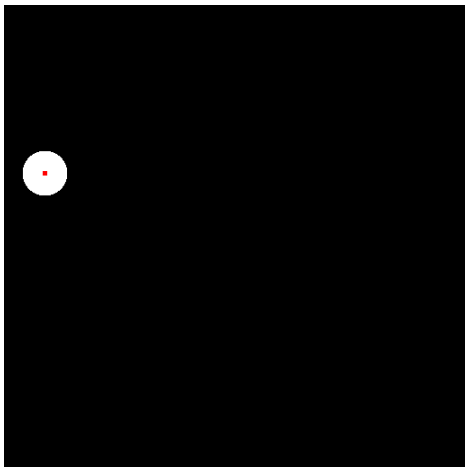
To extend this approach to 3D, 4D Simplex noise was generated and denoted as $n^{4D}(x, y, z, t)$. Extending from 3D Simplex noise to 4D Simplex noise, first three coordinates represent the volumetric locations (x,y,z) and the fourth coordinate represents the time (t) stamp. Similar to previous approach, we need to have a 3D flow field to generate 3D flow field so that the particle can randomly move guided by the 3D flow fields. However, using one 4D noise value could not represent three independent parameters in 3D location. To circumvent this problem, we generate two set of the 4D Simplex noise denote as n_1^{4D} and n_2^{4D}. Similar to Eq (6.10) 3D vector (v_x, v_y, v_z) is generated in each grid location at every time point as a function of the two noise levels such that



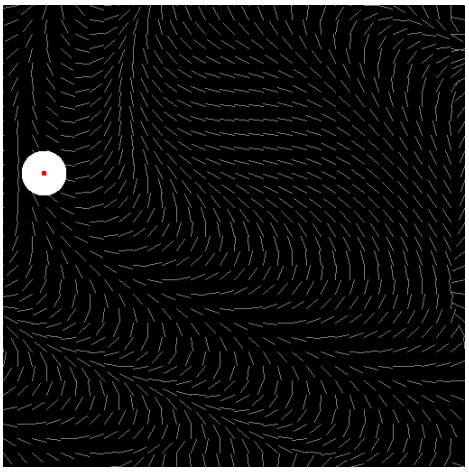
(a) 2D Simplex noise (32×32)



(b) 2D Flow fields (512×512) with interval 16



(c) Particle (Center: Red, Disk: White)



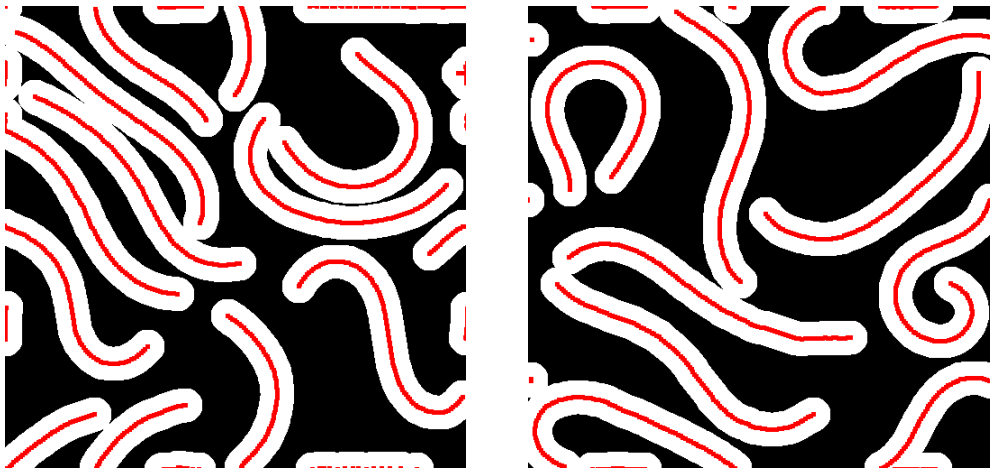
(d) 2D Flow fields with particles

Fig. 6.10. The Example of the Simplex noise, flow fields, and particle in 2D

$$
\begin{aligned}
v_x &= \cos(2\pi n_1^{4D} c)\cos(2\pi n_2^{4D} c) \\
v_y &= \sin(2\pi n_2^{4D} c)\cos(2\pi n_2^{4D} c) \\
v_z &= \sin(2\pi n_2^{4D} c)
\end{aligned} \tag{6.11}
$$

where c is a coefficient to control vector angle changes within time stamp same as in Eq (6.10). Then, v_x, v_y, and v_z are all independently generated so that the 3D flow fields are independent on their 3D location as well as time. This helps to simulate and generate randomly moving particle in 3D space.

6.2.2 Experimental Results



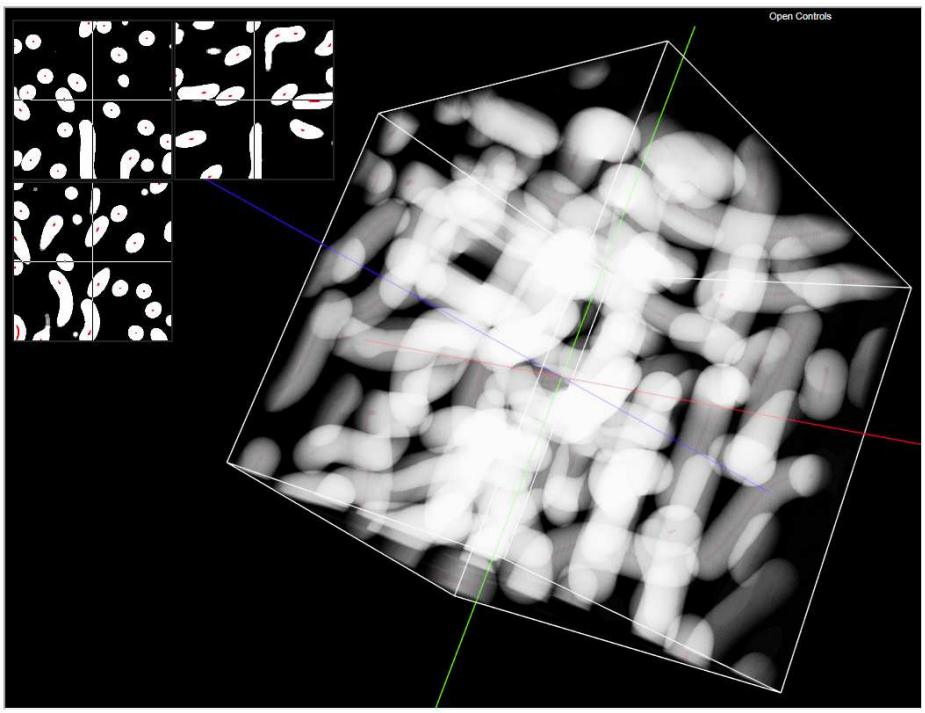
(a) (b)

Fig. 6.11. 2D synthetic binary tubule generation examples (Tubule region: White, Centerline: Red) by proposed method

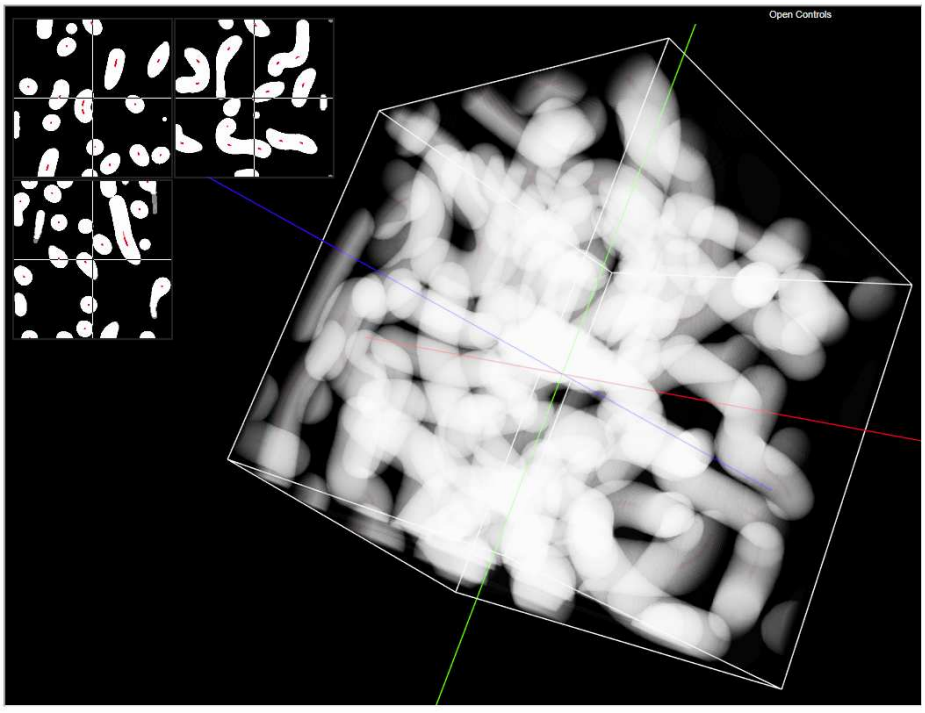
Two examples of the synthetically generated 2D binary tubules are displayed in Figure 6.11. The image size of two examples are 512×512 and each radius of the disk is in between 18 pixels and 22 pixels. Also, each tubule length contains between 90 to 110 disks to form a single tubule. Here, red lines represent the centerlines

of the disk trajectories and white regions are corresponding union of all the disks follows the trajectory. As observed our proposed synthetic tubule generation methods successfully generate synthetic binary tubules in 2D space.

Similarly, two examples of the synthetically generated 3D binary tubules are illustrated in Figure 6.12. To visualize these 3D volumes, we developed an interactive 3D volume visualizer. Also, we added a sections (slices) viewer at the left corner to visualize each xy, xz, and yz section. More detail of this visualization technique will be introduced in Chapter 8. The volume size of two examples are $512 \times 512 \times 512$ and each radius of the sphere is in between 23 voxels and 27 voxels. Also, each tubule length contains between 60 to 80 spheres to form a single 3D tubule. Again, the red lines represent the centerlines of the sphere trajectories in 3D and white regions are corresponding union of all the spheres followed the trajectory. As observed our proposed synthetic tubule generation methods successfully generate synthetic binary tubules in the 3D space. Thus, this will be used as a I^{BiTr} and input of the 3D synthesis training to generate I^{SynO} for the 3D synthesis stage.



(a)



(b)

Fig. 6.12. 3D synthetic binary tubule generation examples (Tubule region: White, Centerline: Red) by proposed method

7. THREE DIMENSIONAL BLIND IMAGE DECONVOLUTION USING GENERATIVE ADVERSARIAL NETWORKS

Previous chapters introduced various segmentation methods for fluorescence microscopy image (volume) data. In this chapter we describe image/volume deconvolution which is inversion process of the convolution. In particular, we present a new approach to restore various biological structures in 3D microscopy images in deeper tissue without knowing the 3D PSF using a spatially constrained CycleGAN (SpCycleGAN) [116]. We train and inference the SpCycleGAN in three directions along with xy, yz, and xz sections (3-Way SpCycleGAN) to incorporate 3D information inspired by [173, 174]. These restored 3-way microscopy images are then averaged and evaluated with three different image quality metrics. Our datasets consist of Hoechst 33342 labeled nuclei and a phalloidin labeled filamentous actin collected from a rat kidney using two-photon microscopy. The goal is to restore blurred and noisy 3D microscopy images to the level of well-defined images so that the deeper depth tissues can be used for biological study[1].

7.1 3-Way Spatially Constrained Cycle-Consistent Adversarial Networks for Blind Deconvolution [2]

Figure 7.1 shows a block diagram of the proposed 3D images deconvolution method. Note that the notation we used here is already defined in the Chapter 1.2. As shown in Figure 7.1, we divide an original florescence microscopy volume denoted as I^O into two subvolumes such as an out-of-focus and noisy subvolume and a well-defined

[1]The 3D blind image deconvolution method using generative adversarial networks presented in this chapter is based on the paper in [118].

[2]The work demonstrated in this section was jointly done with Shuo Han of Purdue University.

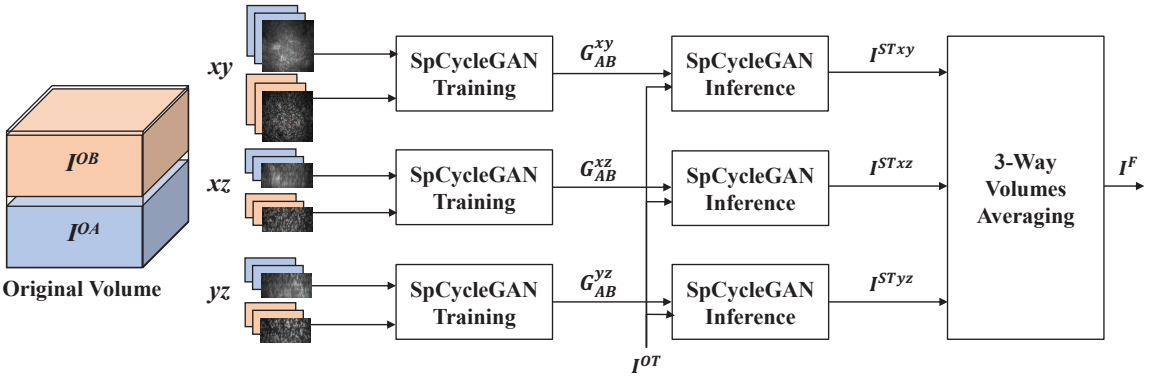


Fig. 7.1. Blockdiagram of the proposed deconvolution method using a 3-Way SpCycleGAN

subvolume denoted as I^{OA} and I^{OB}, respectively. In particular, we choose I^{OA} from deep sections and I^{OB} from shallow sections since shallow sections of fluorescence microscopy volumes typically have a better image quality than deep sections. Due to diffraction limit of fluorescence microscopy, shallow sections of fluorescence microscopy volumes typically have a better image quality than deep sections. These two volumes are sliced in the z-, y-, and x-direction to form the xy, xz, and yz sections of the images. Then, the xy sections from I^{OA} and I^{OB} are used for the training of the SpCycleGAN [116] to obtain the trained generative network denoted as G_{AB}^{xy}. Similarly doing this with the xz sections and the yz sections, trained generative networks G_{AB}^{xz} and G_{AB}^{yz} are obtained. These generative networks are used for inference with a test volume denoted as I^{OT} in the xy, xz, and yz sections. Next, these synthetically generated results by the SpCycleGAN inference are stacked with z-, y-, and x-direction to form 3D volumes denoted as I^{STxy}, I^{STxz}, and I^{STyz}, respectively. Finally, we obtain the final volume I^{F} by voxelwise weighted averaging of these volumes.

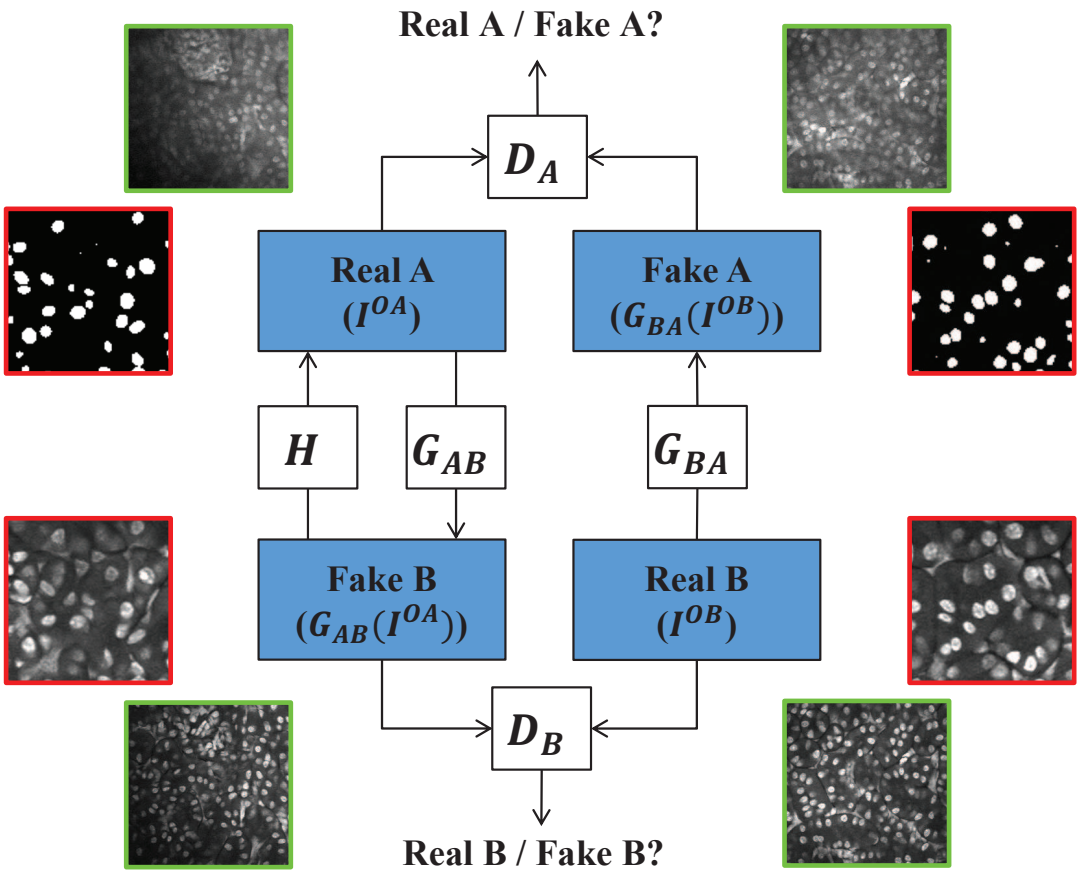


Fig. 7.2. Blockdiagram of the SpCycleGAN

7.1.1 Spatially Constrained CycleGAN (SpCycleGAN)

The SpCycleGAN was firstly introduced in our previous work [116] which extended the CycleGAN [113] by adding one more term to the loss function and introducing an additional generative network. As mentioned in Chapter 2.1, the CycleGAN can be utilized as a style transfer without having paired images to generate synthetic data. One problem with the CycleGAN is that the generated images sometimes are misaligned with the input images [116]. To address this issue, the SpCycleGAN introduced new loss function (spatial loss) with new generative model H. These spatial loss and new generative model help to prevent each nucleus from misalignment between

two different styles. For the segmentation purpose, the synthetic groundtruth images are denoted as a style A and the real microscopy synthetic images are denoted as a style B. Then, main goal for the training is training generative network (G_{AB}) which generates from the style A to the style B. Similarly, for the deconvolution purpose, the out-of-focus images are denoted as a style A and the well-defined images are denoted as a style B. Again, main goal of the training here is training generative network to generate from the style A to the style B. These segmentation and deconvolution examples are shown in Figure 7.2. Each image highlighted in red box is the example of the segmentation method presented in [116]. Also, each image highlighted in green box is the example of the deconvolution method presented in this chapter [118]. More specifically, "Real A" and "Real B" are the set of given input images which represents the style A and the style B and using them to train generative models and discriminative models for the training images. Then, the images next to "Fake A" and "Fake B" are the example images synthetically generated by "Real B" and G_{BA}, and "Real A" and G_{AB}, respectively. As observed these synthetically generated images ("Fake A" and "Fake B") are very similar to the original images ("Real A" and "Real B"). These "Real A" and "Fake A" are set of paired images of the style A whereas these "Real B" and "Fake B" are set of paired images of the style B.

As discussed above the SpCycleGAN introduced a spatial loss and new generative model. More precisely, we added a spatial constrained term ($\mathcal{L}_{spatial}$) to the loss function and minimize the loss function together with the two original GAN losses ($\mathcal{L}_{\text{GAN}}$) and the cycle consistent loss ($\mathcal{L}_{cyc}$) as:

$$\begin{aligned}\mathcal{L}(G_{AB}, G_{BA}, H, D_A, D_B) = \mathcal{L}_{\text{GAN}}(G_{AB}, D_B, I^{OA}, I^{OB}) \\ + \mathcal{L}_{\text{GAN}}(G_{BA}, D_A, I^{OB}, I^{OA}) \\ + \lambda_1 \mathcal{L}_{cyc}(G_{AB}, G_{BA}, I^{OA}, I^{OB}) \\ + \lambda_2 \mathcal{L}_{spatial}(G_{AB}, H, I^{OA}, I^{OB}) \end{aligned} \tag{7.1}$$

where

$$\mathcal{L}_{\text{GAN}}(G_{AB}, D_B, I^{OA}, I^{OB}) = \mathbb{E}_{I^{OB}}[\log(D_B(I^{OB}))] + \mathbb{E}_{I^{OA}}[\log(1 - D_B(G_{AB}(I^{OA})))]$$

$$\mathcal{L}_{\text{GAN}}(G_{BA}, D_A, I^{OB}, I^{OA}) = \mathbb{E}_{I^{OA}}[\log(D_A(I^{OA}))] + \mathbb{E}_{I^{OB}}[\log(1 - D_A(G_{BA}(I^{OB})))]$$

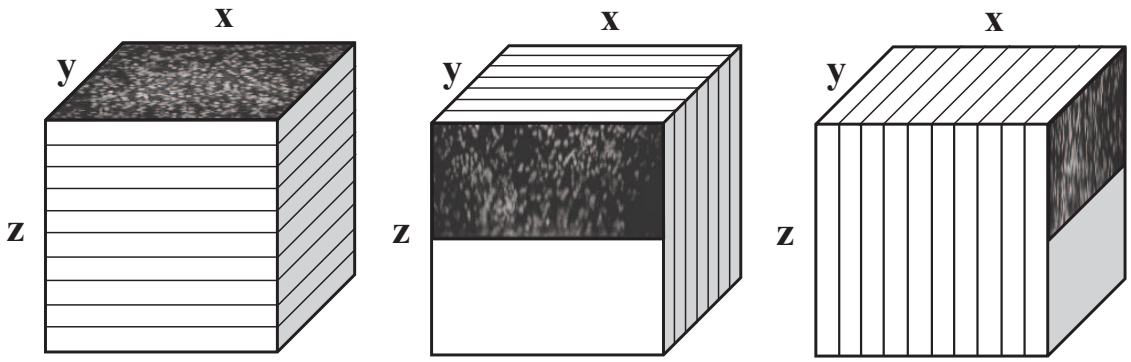
$$\mathcal{L}_{cyc}(G_{AB}, G_{BA}, I^{OA}, I^{OB}) = \mathbb{E}_{I^{OA}}[||G_{BA}(G_{AB}(I^{OA})) - I^{OA}||_1] + \mathbb{E}_{I^{OB}}[||G_{AB}(G_{BA}(I^{OB})) - I^{OB}||_1]$$

$$\mathcal{L}_{spatial}(G_{AB}, H, I^{OA}, I^{OB}) = \mathbb{E}_{I^{OA}}[||H(G_{AB}(I^{OA})) - I^{OA}||_2].$$

Note that λ_1 and λ_2 are the controllable coefficients for $\mathcal{L}_{cyc}$ and $\mathcal{L}_{spatial}$. Also, $||\cdot||_1$ and $||\cdot||_2$ represent L_1 and L_2 norms, respectively. The generative model G_{AB} transfers I^{OA} to I^{OB} and the generative model G_{BA} transfers I^{OB} to I^{OA}. Similarly, the discriminative model D_A and D_B distinguish between I^{OA} and $G_{BA}(I^{OB})$ and between I^{OB} and $G_{AB}(I^{OA})$. In particular, $G_{AB}(\cdot)$ is a transfer function using model G_{AB} and $G_{BA}(\cdot)$ is another transfer function using model G_{BA}. For example, $G_{AB}(I^{OA})$ is a synthetically restored volume generated by model G_{AB} using a blurred and noisy volume. Also, $G_{BA}(I^{OB})$ is a synthetically generating blurred and noisy volume by model G_{BA} using a well-defined volume. Additionally, another generative model H takes $G_{AB}(I^{OA})$ as an input to generate a synthetically blurred and noisy volume $H(G_{AB}(I^{OA}))$ using synthetically restored volume. This generative model H minimizes $\mathcal{L}_2$ loss between I_{OA} and $H(G_{AB}(I^{OA}))$.

7.1.2 3-Way SpCycleGAN and Volumes Averaging

One drawback of the SpCycleGAN is that it works only in 2D. Since our fluorescence microscopy data is a 3D volume, we form the 3D volume by stacking 2D



(a) I^{STxy} (b) I^{STxz} (c) I^{STyz}

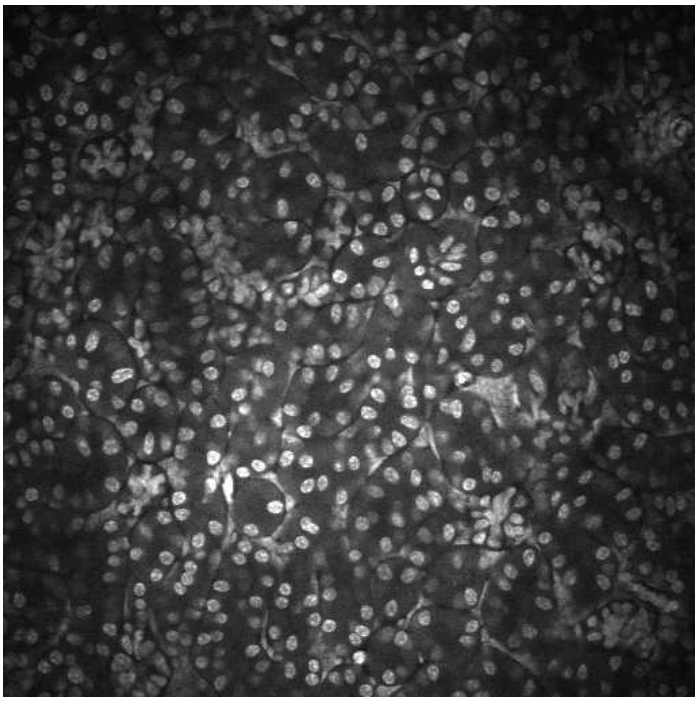
Fig. 7.3. The 3D volume formation using xy, xz, and yz sections

images obtained from different focal planes during data acquisition [125]. Therefore, we employ 3-Way SpCycleGAN training which uses the SpCycleGAN in the xy, xz, and yz sections independently and obtain the generative models per each sections. We use three generative models (G_{AB}^{xy}, G_{AB}^{xz}, and G_{AB}^{yz}) for the inference using I^{OT} which transfer noisy and out-of-focus images to well-defined and focused images in the xy, xz, and yz sections. More specifically, the test volume is sliced into three sets of sectional images and each image is used as an input of inference to generate synthetic well-defined and focused image. Then, these synthetically generated images are stacked in the z-, y-, and x-direction to form I^{STxy}, I^{STxz}, and I^{STyz}, respectively. Figure 7.3 displays the 3D volumes formation using xy, xz, and yz sections. In general, the number of the xy, xz, and yz sections are different from each other, we use zero padding to make the dimension of three volumes identical. Lastly, the final volume (I^F) is obtained as

$$I^F = w_1 I^{STxy} + w_2 I^{STxz} + w_3 I^{STyz} \tag{7.2}$$

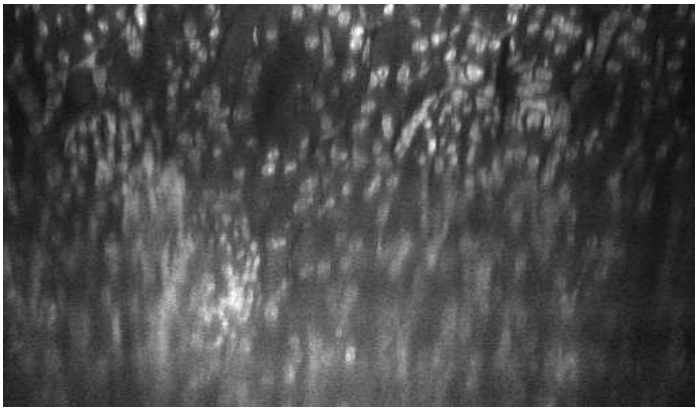
where w_1, w_2, and w_3 are weight coefficients of I^{STxy}, I^{STxz}, and I^{STyz}, respectively.

7.2 Experimental Results



(a) xy section (I^{OT}_{z1})

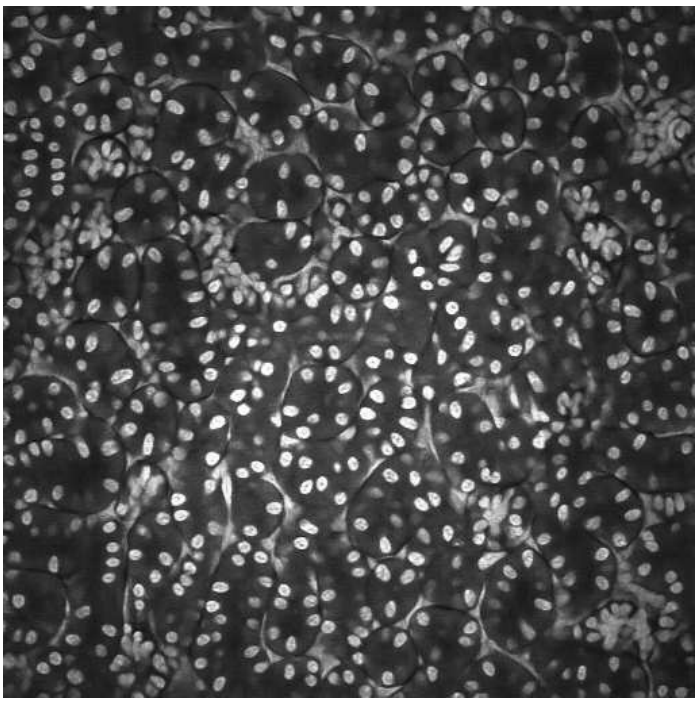
(b) yz section (I^{OT}_{x256})



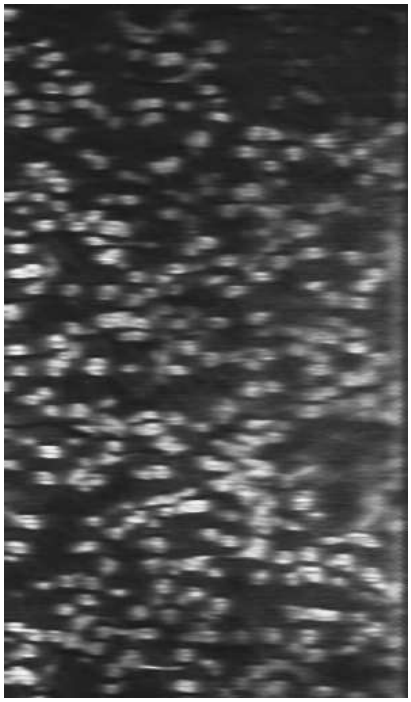
(c) xz section (I^{OT}_{y256})

Fig. 7.4. Original test volume (I^{OT}) of *WSM* blue channel displayed in three orthogonal sections

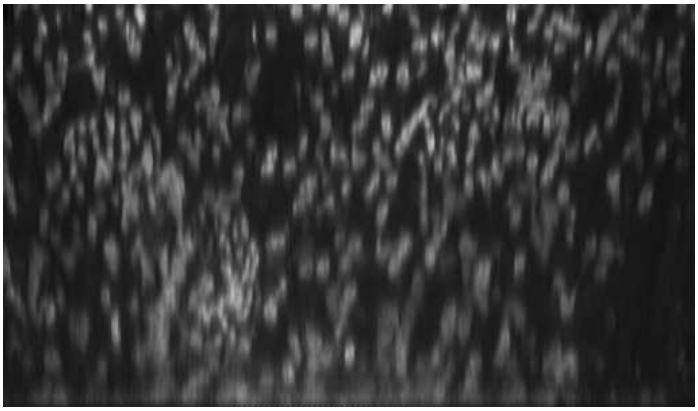
The performance of our proposed deconvolution method was tested on two different datasets: *WSM* blue channel and *WSM* red channel. *WSM* are originally



(a) xy section (I^{ST}_{z1})



(b) yz section (I^{ST}_{x256})



(c) xz section (I^{ST}_{y256})

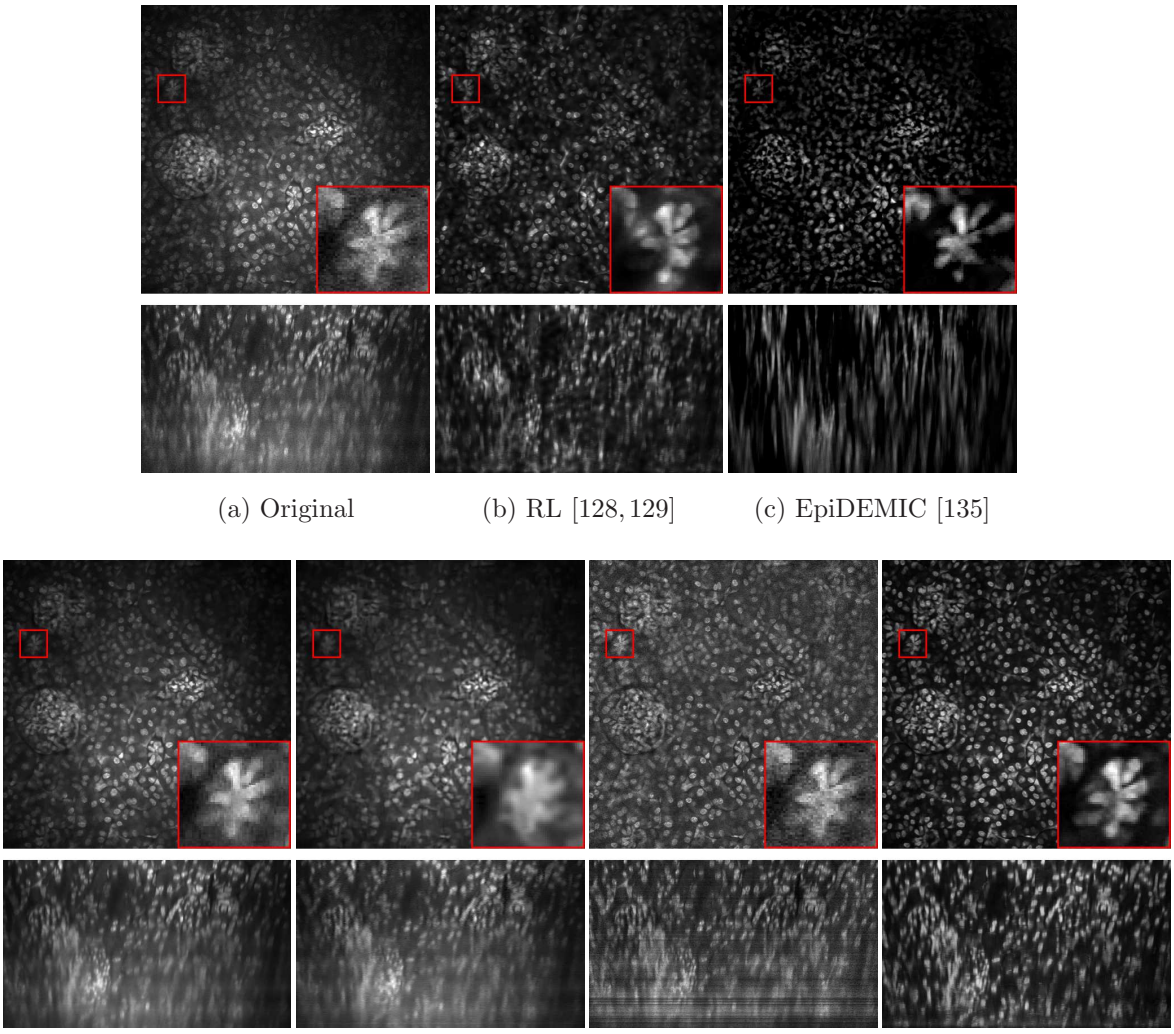
Fig. 7.5. Synthetically generated volume (I^{ST}) of *WSM* blue channel displayed in three orthogonal sections

obtained at the same time with different fluorophores such as blue dying and red dying to delineate different biological structures. Both datasets are comprised of $Z = 512$ grayscale images, each of size $X \times Y = 512 \times 512$ pixels. We selected a blurred and noisy subvolume (I^{OA}) from last 200 images of given fluorescence microscopy volume as $I^{O}_{(1:512,1:512,313:512)}$ with a size of $512 \times 512 \times 200$ for both datasets. Also, a good quality subvolume (I^{OB}) was selected based on a biologist's opinion as $I^{O}_{(1:512,1:512,15:214)}$ with size of $512 \times 512 \times 200$. The test volume ($I^{OT}$) for each dataset was selected at deeper tissue depth than I^{OB} as $I^{O}_{(1:512,1:512,215:512)}$ with size of $512 \times 512 \times 298$.

Our 3-Way SpCycleGAN is implemented in PyTorch using the Adam optimizer [175] with constant learning rate 0.0002 for the first 100 epochs and gradually decreased to 0 for the next 100 epochs. Also, we use the ResNet 9 blocks [92] for all generative models (G_{AB}, G_{BA}, and H) with 64 feature maps at the first layer. For the corresponding discriminative models (D_A and D_B), same discriminative models are used in the CycleGAN [113]. We randomly select patches size of 256×256 from 512×512 for the xy sections and 200×200 from 512×200 for the xz and yz sections for the SpCycleGAN training, respectively. We choose larger resolution for the xy sections since xy sections is a finer resolution than those xz and yz sections. Also, we set the coefficients $\lambda_1 = \lambda_2 = 10$ for all 3-Way SpCycleGAN training for both *WSM* blue channel and red channel. Lastly, the weights for 3-way volume averaging is set as $w_1 = w_2 = w_3 = 1/3$ so that each sectional results equally contribute the final volume.

Figure 7.4 and 7.5 exhibits the orthogonal sections of original test volume (I^{OT}) of *WSM* blue channel and corresponding synthetically generated volume (I^{ST}). Here, the orthogonal section view is one of the popular technique to visualize 3D volumes. As observed from Figure 7.4 and 7.5, each section of I^{ST} shows better defined nuclei than blurred shown in I^{OT}. Also, the inhomogeneity problem shown in Figure 7.4c is significantly improved in Figure 7.5c. On top of that, the noisy bright lines portrayed

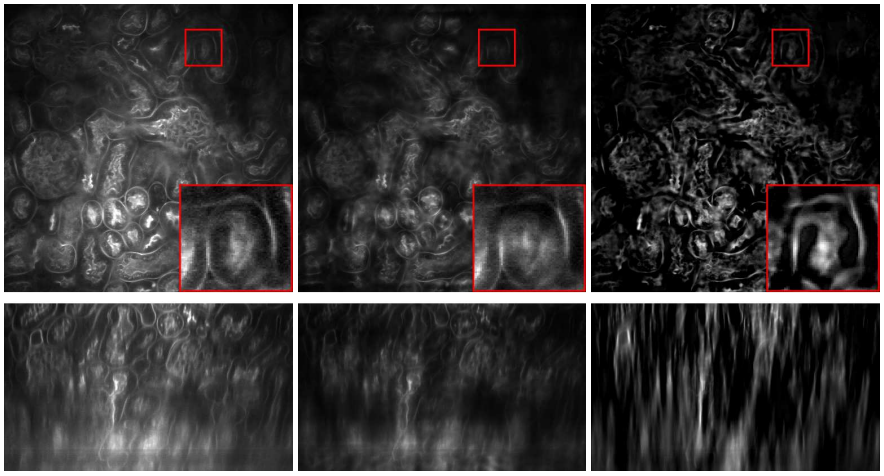
in Figure 7.4b and 7.4c are disappeared and the quality of images become better in Figure 7.5b and 7.5c, respectively.



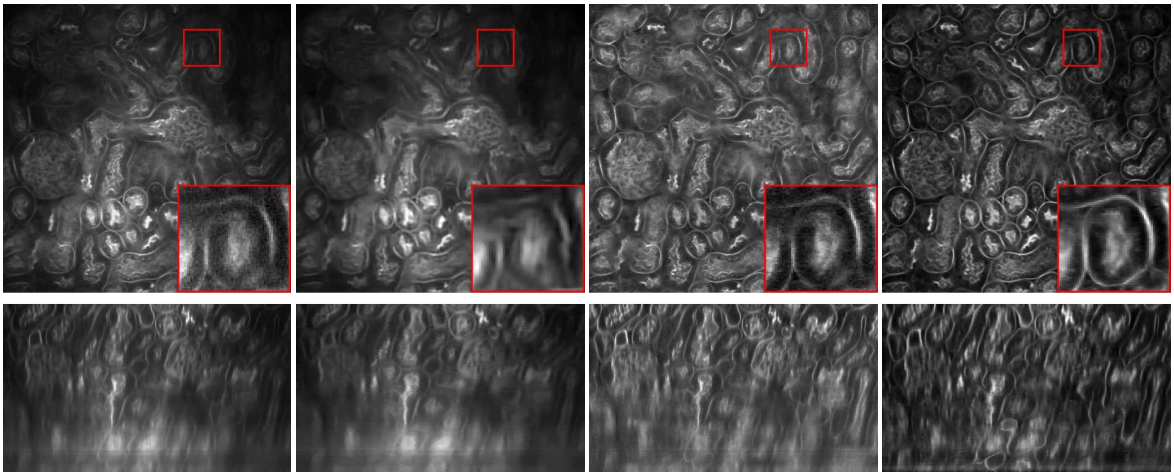
(a) Original (b) RL [128, 129] (c) EpiDEMIC [135]

(d) PureDenoise [136] (e) iterVSTpoisson-Deb [138] (f) 3DacIC [55] (g) 3-Way SpCycleGAN (Proposed)

Fig. 7.6. Comparison of the original volume and 3D restored volume results with the xy and xz sections of *WSM* blue channel using various methods. We use the original test volume of the xy section (I^{OT}_{z126}) and the yz section (I^{OT}_{y256})



(a) Original (b) RL [128, 129] (c) EpiDEMIC [135]



(d) PureDenoise [136] (e) iterVSTpoisson-Deb [138] (f) 3DacIC [55] (g) 3-Way SpCycleGAN (Proposed)

Fig. 7.7. Comparison of the original volume and 3D restored volume results with the xy and xz sections of *WSM* red channel using various methods. Similar to the blue channel images, we use the original test volume of the xy section (I^{OT}_{z126}) and the yz section (I^{OT}_{y256})

Moreover, our proposed deconvolution results were visually compared with five different techniques including RL [128, 129], EpiDEMIC [135], PureDenoise [136], iterVSTPoissonDeb [138], and 3DacIC [55] shown in Figure 7.6 and 7.7. Note that we used default settings for the methods RL and PureDenoise in ImageJ [176] plugins, EpiDEMIC in Icy [177] plugin, and iterVSTpoissonDeb.

As shown in Figure 7.6 and 7.7, first column displays a sample xy section (I^{OT}_{z126}) and xz section (I^{OT}_{y256}) of original test volumes in *WSM* blue channel and *WSM* red channel, respectively. The original test volumes suffer from significant intensity inhomogeneity, blur, and noise. Also, this degradation gets worse at deeper depth as shown in the xz section. As observed, our proposed method showed the best performance among presented methods in terms of inhomogeneity correction, clarity of the shape of nuclei and tubules/glomeruli structure, and noise level. More specifically, two deconvolution methods (RL and EpiDEMIC) successfully reduced blur but the original shapes of the biological structures were lost. Also, EpiDEMIC's xz section deconvolution results were all connected each other since EpiDEMIC learned 2D features as a prior to enhance 3D. Similarly, two denoising methods (PureDenoise and iterVSTpoissonDeb) successfully suppressed Poisson noise but these denoising results were still suffered from intensity inhomogeneity and blur. In fact, the denoising results added more blur than original test volume. Meanwhile, 3DacIC method successfully corrected inhomogeneity but this method amplified background noise level and aggravated image quality. Moreover, 3DacIC exacerbated line shape noise shown in the xz sections.

In addition to the visual comparison, three image quality metrics were utilized for evaluating volume quality of restored volumes of proposed and other presented methods. Since our microscopy volumes do not have reference volumes to compare, we need to use no reference image quality assessment (NR-IQA) [146] instead of traditional PSNR, SSIM, and FSIM. One problem is that there is no gold standard image quality metric for 3D fluorescence microscopy. We employed the blind/referenceless image spatial quality evaluator (BRISQUE) [146], the oriented-gradient image qual-

Table 7.1.
Comparison of the performance of proposed and other restoration methods with three image quality metrics using *WSM* blue channel

	$I^{OT}_{(1:512,1:512,149:298)}$ of the *WSM* blue channel		
Method	3-Way BRISQUE [146]	3-Way OG-IQA [147]	3-Way Microscopy IFQ [149]
$I^{OT}_{(1:512,1:512,149:298)}$	35.50	−0.34	1.95
RL [128, 129]	41.19	−0.80	0.67
EpiDEMIC [135]	58.96	−0.75	0.62
PureDenoise [136]	39.90	−0.67	2.04
iterVSTpoissonDeb [138]	35.01	−0.44	2.84
3DacIC [55]	37.96	−0.26	0.66
3-Way SpCycleGAN (Proposed)	**34.05**	**−0.88**	**0.52**

Table 7.2.
Comparison of the performance of proposed and other restoration methods with three image quality metrics using *WSM* red channel

	$I^{OT}_{(1:512,1:512,149:298)}$ of the *WSM* red channel		
Method	3-Way BRISQUE [146]	3-Way OG-IQA [147]	3-Way Microscopy IFQ [149]
$I^{OT}_{(1:512,1:512,149:298)}$	**15.68**	−0.64	3.07
RL [128, 129]	23.97	−0.49	3.98
EpiDEMIC [135]	50.97	−0.29	0.96
PureDenoise [136]	24.40	−0.47	3.34
iterVSTpoissonDeb [138]	32.38	−0.36	4.05
3DacIC [55]	19.68	−0.61	1.64
3-Way SpCycleGAN (Proposed)	31.14	**−0.82**	**0.94**

ity assessment (OG-IQA) [147], and the microscopy image focus quality assessment (Microscopy IFQ) [149] for evaluating quality of restored microscopy volumes. In particular, BRISQUE model is a regression model learned from local statistics of natural scenes in the spatial domain to measure image quality where the quality value range is from 0 to 100. OG-IQA model is a gradient feature based model that maps from image features to image quality via an adaboosting back propagation neural network where the quality value is from -1 to 1. Lastly, Microscopy IFQ measures discrete defocus level from 0 to 10 using 84×84 local patches using a CNN. Instead of using discrete defocus level, we got a probability ($p(l)$) for each corresponding defocus level (l) before the softmax layer and used these probabilities and corresponding defocus levels to obtain expected value which defined as

$$\text{Microscopy IFQ} = \sum_{l=0}^{10} l \cdot p(l). \tag{7.3}$$

Since this Microscopy IFQ value was obtained for each individual 84×84 local patches, we resized our images to be nearest integer multiple of local patch size and took an average through entire image. Note that the smaller values of all three image quality assessments indicate the better image quality. In addition, since these three image quality assessments can only measure the quality of 2D images, we again utilized 3-way idea to obtain the image quality of the xy, xz, and yz sections and took an average of them. The end result was a single representative value for volume quality per each volume.

We used these three image quality metrics to test seven different volumes including the original test volume. This is provided in Table 7.1 and 7.2. Note that we selected the 150 most blurred and noisy image volumes from test volume as $I^{OT}_{(1:512,1:512,149:298)}$ for the evaluation purpose. As mentioned above the smaller values are considered to be indicators of the better image volume quality. As observed in Table 7.1 and 7.2, our proposed method outperformed the other methods and original volume except from 3-Way BRISQUE in *WSM* red channel. This is because BRISQUE measures the quality from natural image statistics and this model is a favor of blurred volume. Therefore, RL and EpiDEMIC had higher values in BRISQUE image quality metric.

Similarly, OG-IQA is a gradient based measurement so edge preserved restoration volume can get smaller image quality values. Also, Microscopy IFQ is a defocus level measurement. Hence, PureDenoise and iterVSTpoissonDeb had sometimes higher OG-IQA and Microscopy IFQ. 3DacIC produced reasonably lower values for the xy sections but the quality of the xz and yz sections were poor so that the entire volume quality was inferior than proposed method's volume quality.

Figure 7.8 portrays the visual comparison between proposed 3-Way SpCycleGAN and SpCycleGAN using the xy sections only ($w_1 = 1$, $w_2 = w_3 = 0$). Without having z-direction information, SpCycleGAN using the xy sections cannot correctly restore the glomerulus displayed in the red box. In particular, the glomerulus restored by only the xy sections of SpCycleGAN looks like group of nuclei presented rather than glomerulus shown in first row of Figure 7.8. Similarly, inside region of glomerulus shown in *WSM* red channel may not be correctly restored if using only xy sections of SpCycleGAN. Also, the z-direction images are frequently discontinued as shown in the xz section in the green box. Compared to that, proposed 3-Way SpCycleGAN can successfully restore glomerulus and connect smoothly in z-direction.

Lastly, segmentation method described in [116] was used to various restored results. To be more specific, these various restored volumes were used as preprocessing steps for the segmentation. We selected subvolume region which contains the glomerulus shown in Figure 7.8 inside the red box using *WSM* blue channel. In particular, this subvolume was sampled as $I^{OT}_{(18:81,243:306,94:157)}$ with size of $64 \times 64 \times 64$. We provided two view angles using Volume Viewer [178] in ImageJ [176] plugin as shown in Figure 7.9 and 7.10. As observed from Figure 7.9 and 7.10, our proposed method can capture more nuclei compared to other presented methods. For example, our proposed method can segment more nuclei in the lower left regions shown in Figure 7.9 and perform better segmentation results for z-direction as can be seen in Figure 7.10.

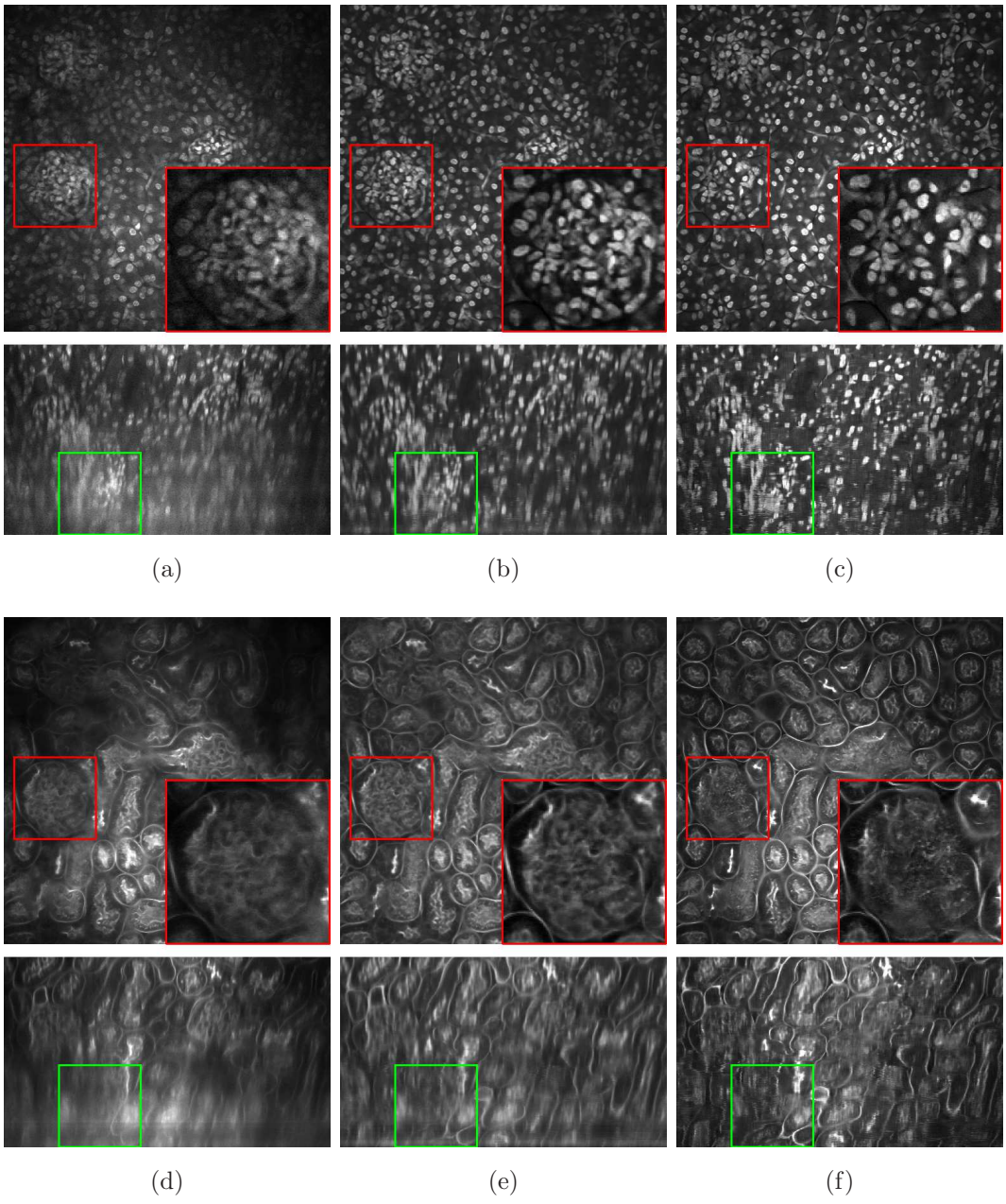
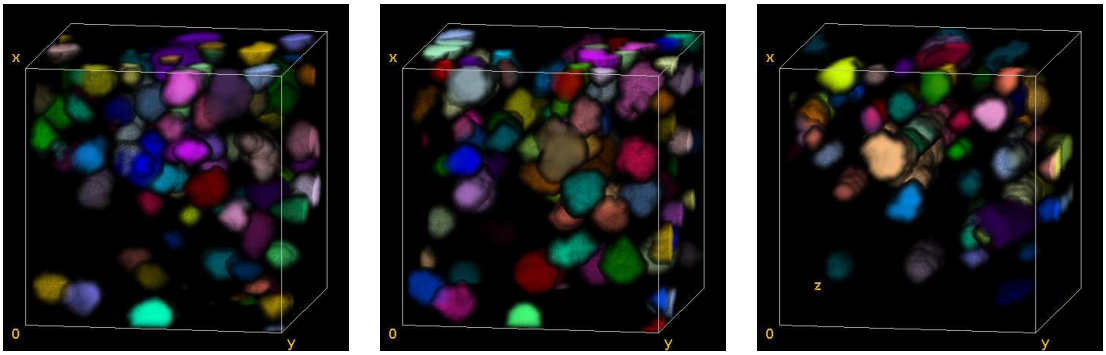
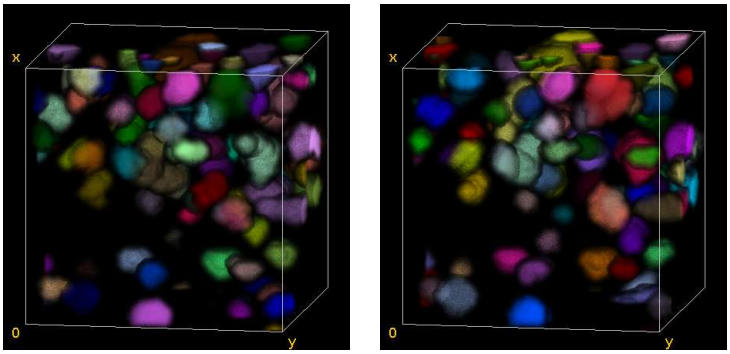


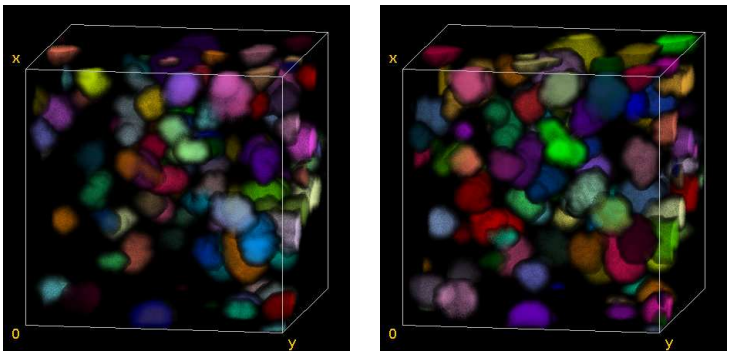
Fig. 7.8. Comparison of original test volume, restored volume using proposed 3-Way SpCycleGAN, and restored volume using the *xy* sections of SpCycleGAN using *WSM* blue channel (up) and *WSM* red channel (down)



(a) Original (b) RL [128, 129] (c) EpiDEMIC [135]

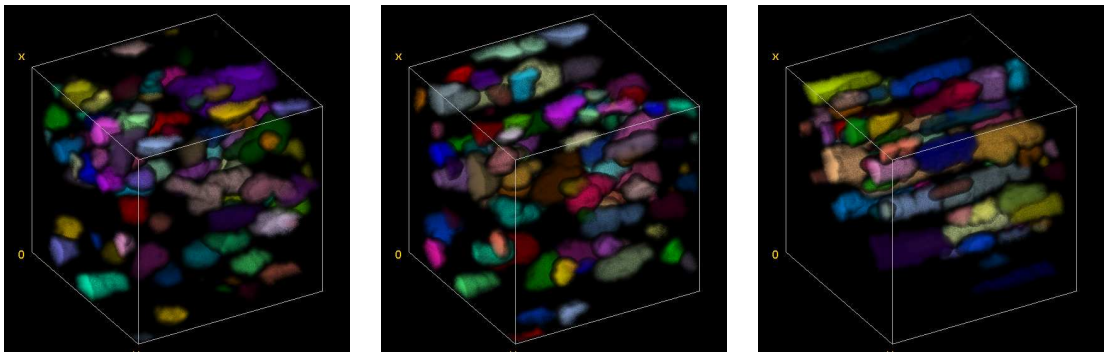


(d) PureDenoise [136] (e) iterVSTpoissonDeb [138]

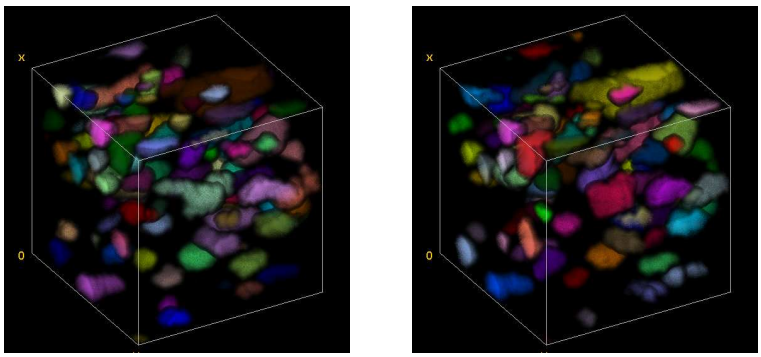


(f) 3DacIC [55] (g) 3-Way SpCycleGAN (Proposed)

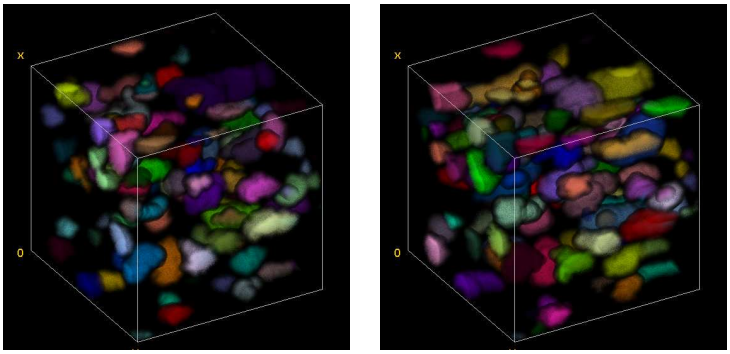
Fig. 7.9. 3D segmentation results comparison of the original volume and 3D restored volume results using *WSM* blue channel using various methods



(a) Original (b) RL [128, 129] (c) EpiDEMIC [135]



(d) PureDenoise [136] (e) iterVSTpoissonDeb [138]



(f) 3DacIC [55] (g) 3-Way SpCycleGAN (Proposed)

Fig. 7.10. 3D segmentation results comparison of the original volume and 3D restored volume results with different angle using *WSM* blue channel using various methods

8. MICROSCOPY VOLUME VISUALIZATION

8.1 Overall System [1]

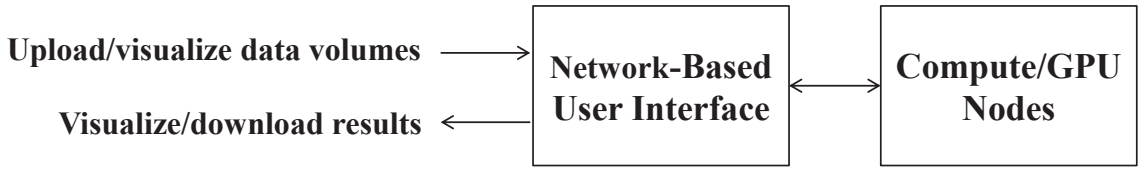


Fig. 8.1. The overview of the proposed DINAVID system

Image/volume processing/analysis and 2D/3D interactive visualization tool are fundamental needs for the most biologists to study and understand the complex 3D biological structures. For that reason, many biologists use the open source image processing package that can support plug-ins such as ImageJ [176], Icy [177], CellProfiler [179, 180] and so on. However, neither of these supports most recent deep learning technique with 3D interactive visualization tools. ImageJ [176] is the most popular free software and it recently supports the U-Net [87] but the client needs to provide the set of groundtruth so that the network can learn the relationship between original images and groundtruth images. Also, the fully 3D segmentation in U-Net which is 3D U-Net [90] is not fully functional in ImageJ yet. CellProfiler [180] also recently added a function called U-Net to support segmentation with deep learning, but it is more difficult to use compared to ImageJ [176]. There is a need for the system which can do both image/volume analysis and 2D/3D visualization with easily handle data to upload and download. Therefore, we designed and developed a

[1]The work demonstrated in this section was jointly done with Shuo Han and Chichen Fu of Purdue University.

web-based microscopy analysis and visualization system called the Distributed and Networked Analysis of Volumetric Image Data (DINAVID).

Figure 8.1 shows an overview of the proposed DINAVID system. First of all, the client upload their volumes to the network based user interface. Then, the high performance computing/Graphics Processing Unit (GPU) nodes are supported in the back to run deep learning inference. Once this inference is done, the segmentation results are given to the network based user interface and can be downloaded and visualized. We also support the visualization of their input data volume so that the client can observe their data volumes.

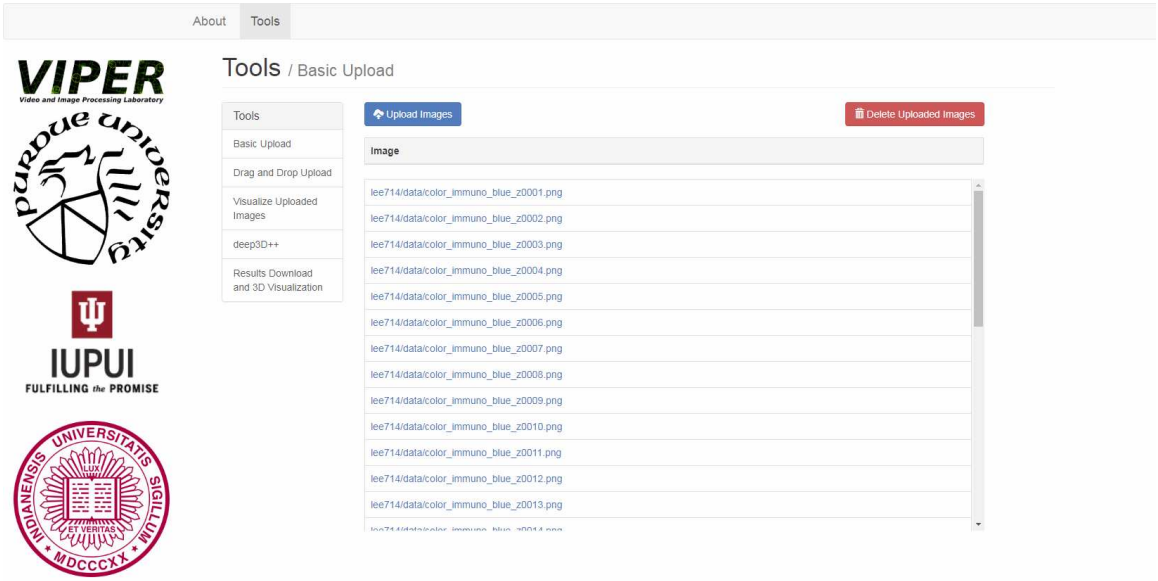


Fig. 8.2. The DINAVID webpage example

The DINAVID webpage example is shown in Figure 8.2. As you can see from the left side bar you have multiple tabs such as upload, drag and drop upload, visualize uploaded images, deep 3D++, and results download and 3D visualization. First two are the upload tabs. Third one is the 3D interactive visualization of uploaded data. Fourth one is the deep learning based nuclei segmentation method presented in [116]. Here, we support five pretrained segmentation model to run inference. The client can

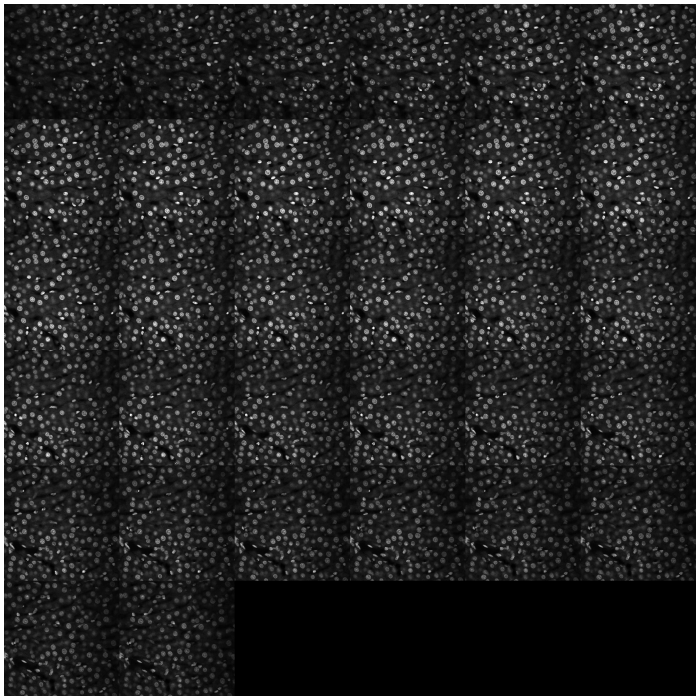
choose the most similar pretrained models to their uploaded images. Also, there is an option to add color labeling and small component removal so that segmentation results are more distinctive to adjacent nuclei and clearer. This color labeling method will be discussed later this chapter.

8.2 Microscopy Volume Visualization [2]

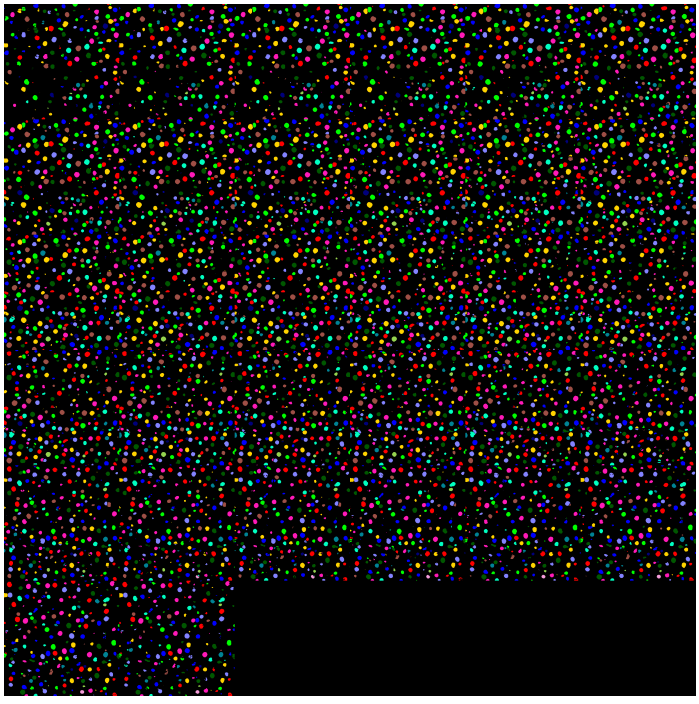
In this chapter, we present our 3D interactive visualization tool. For stacked images to form a 3D volume with different focal planes, there are a variety of ways such as composite or projection view, three-dimension views, or transverse xz or yz cross sectional views [1]. Composite or projection view is like maximum intensity projection (MIP) so that an image stack is projected into a single image. Three-dimension views are displaying the stacked images as a volume and perform 3D rendering to show three dimensional objects. Lastly, transverse xz or yz cross sectional views are the technique to move back and forth to check each 2D cross sectional images inside of the 3D volume. The Voxx [161] is an example of the 3D views software that render entire images as a single volume and show to the client. Also, the volume viewer [178], a plug-in in ImageJ [176], also supports volume rendering. One of the drawback of fully volume viewer is that observing inside of the volume is difficult since it is blocked by other surfaces. Another viewer provided by ImageJ plug-in is called the 3D viewer [181] is not fully 3D render viewer but it contains 3D viewer aspect and the cross section viewers aspect. Note that ImageJ [176] also provides embedded the cross sectional viewer called orthogonal views. One drawback of the cross sectional viewer is that it could not show the 3D views so that biologist needs to go back and forth of the sections and imagine how it looks like in the 3D space.

Consequently, we determine to show the 3D volume rendering and cross section viewers at the same time. To achieve this goal, we investigate the Open Graphics Library (OpenGL) functionality for the 3D volume rendering [182]. However, the

[2]The work demonstrated in this section was jointly done with Shuo Han of Purdue University.



(a) Mosaic image of original volume



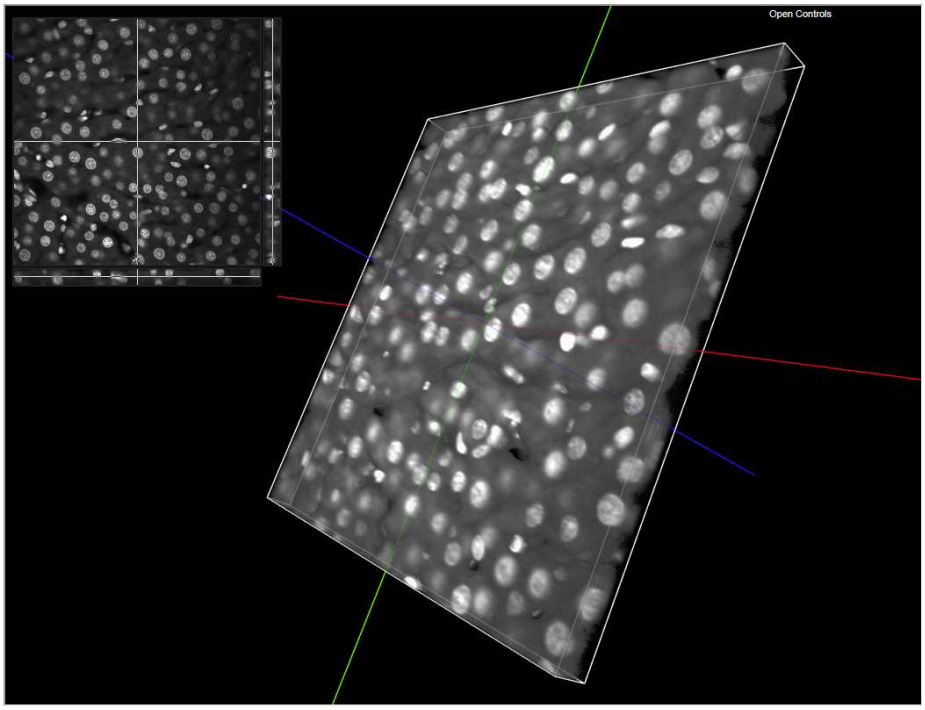
(b) Mosaic image of color labeled segmentation results

Fig. 8.3. Examples of the mosaic images of original and color labeled of the *immuno* dataset

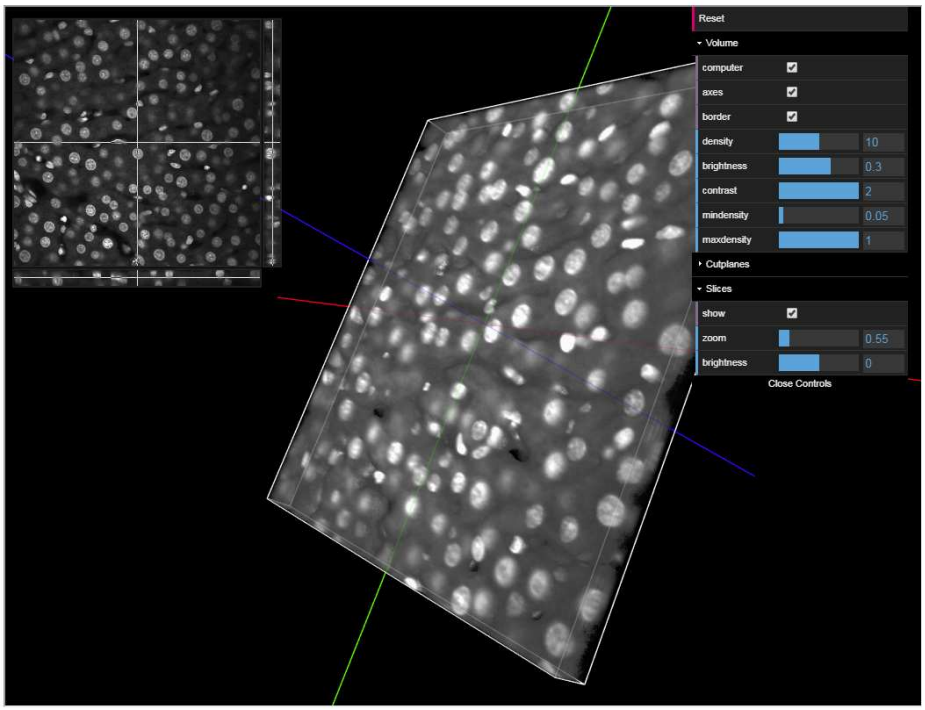
OpenGL is not supported in the web-based system but utilized more with desktop or server side graphical user interface or gaming development. Compared to the OpenGL, the Web Graphics Library (WebGL) derived from the OpenGL Embedded Systems 2.0 (OpenGL ES 2.0) [183] which less functionalities but is lighter load then OpenGL [184, 185]. In fact, this WebGL is intended to design for the mobile devices and even in a browser. Along with recent emerging HTML5, the more functionalities of the WebGL have become supported in a web-based system [186]. Thus, we determine to use for the WebGL as a main framework for the DINAVID visualization tool.

To achieve developing the 3D volume rendering, we use a framework similar to lightweight volume viewer in WebGL called "ShareVol" [187]. For the dependencies, we need to import "glMatrix" [188] which is Javascript matrix and vector library for real time 3D graphics via WebGL and "dat.GUI" [189] which provides a lightweight graphical user interface for changing variables in JavaScript. Having the glMatrix, the client is able to interactively handle the 3D volume with rotating (mouse left click) and panning (mouse right click). Since the glMatrix is designed for the real time application, this 3D volume rotating and panning motion are instantaneously and therefore real time 3D interactive visualization become possible. Also, the dat.GUI supports the parameters changes in the windows while the visualization is running. This means the client can interactively adjust parameters until the 3D volumes and 2D cross sections are displayed in the best setting. Having these two dependencies, the DINAVID can offer 3D volume visualization and 2D cross sections view at the same time.

One drawback of the WebGL is that the WebGL does not support 3D texture due to lightweight version of OpenGL. Therefore, simply loading the stack of 2D images to generate a 3D texture type input volume as typically OpenGL did is impossible from the WebGL [183]. To address this problem, the stack of 2D images are saved as a 2D texture with a mosaic configuration as described in [190]. Figure 8.3 portrays

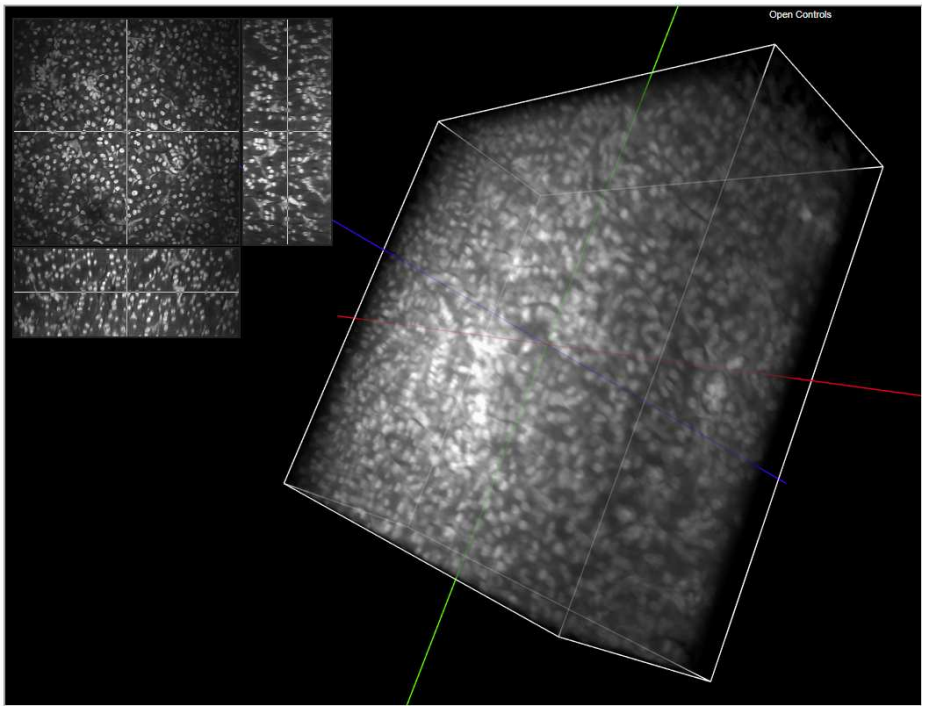


(a) Original volume visualization collapsing light weighted GUI

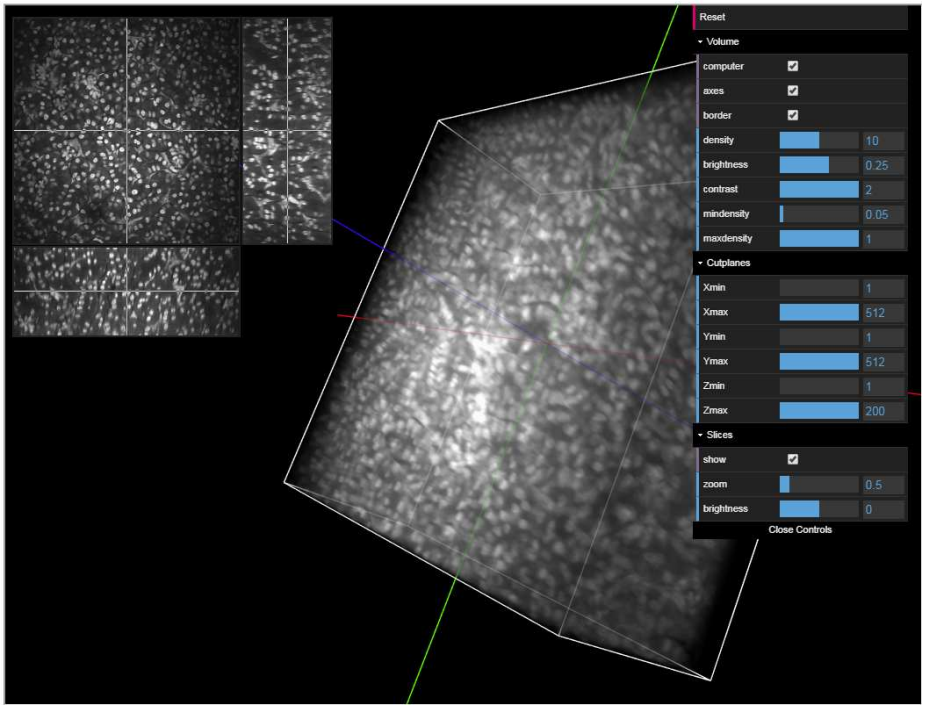


(b) Original volume visualization with light weighted GUI

Fig. 8.4. 3D visualization using the original volume of the *immuno* dataset



(a) Original volume visualization collapsing light weighted GUI



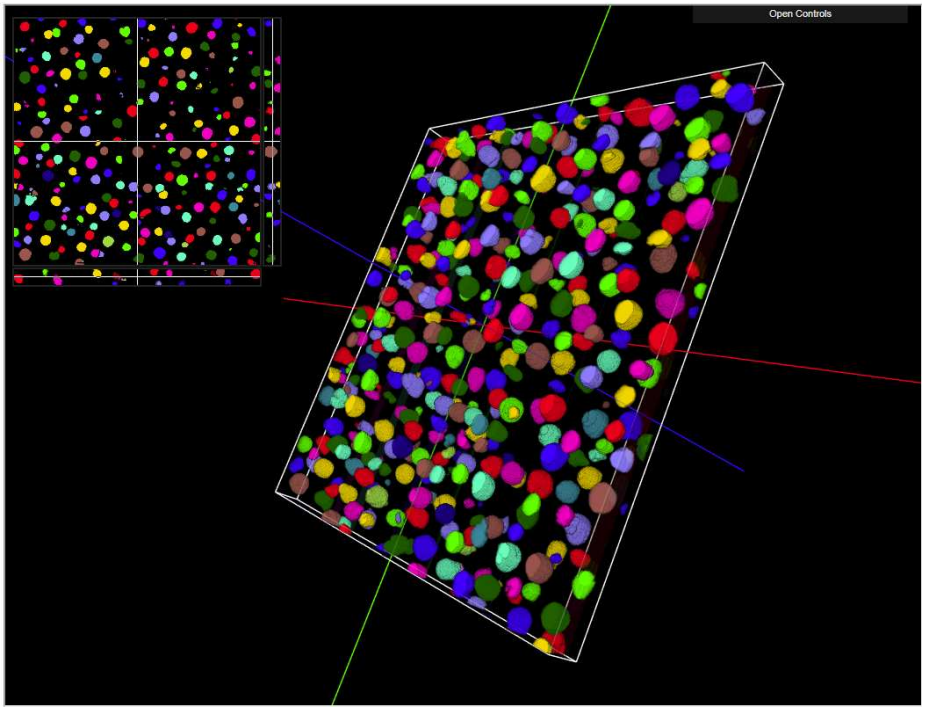
(b) Original volume visualization with light weighted GUI

Fig. 8.5. 3D visualization using the original volume of the *WSM* dataset

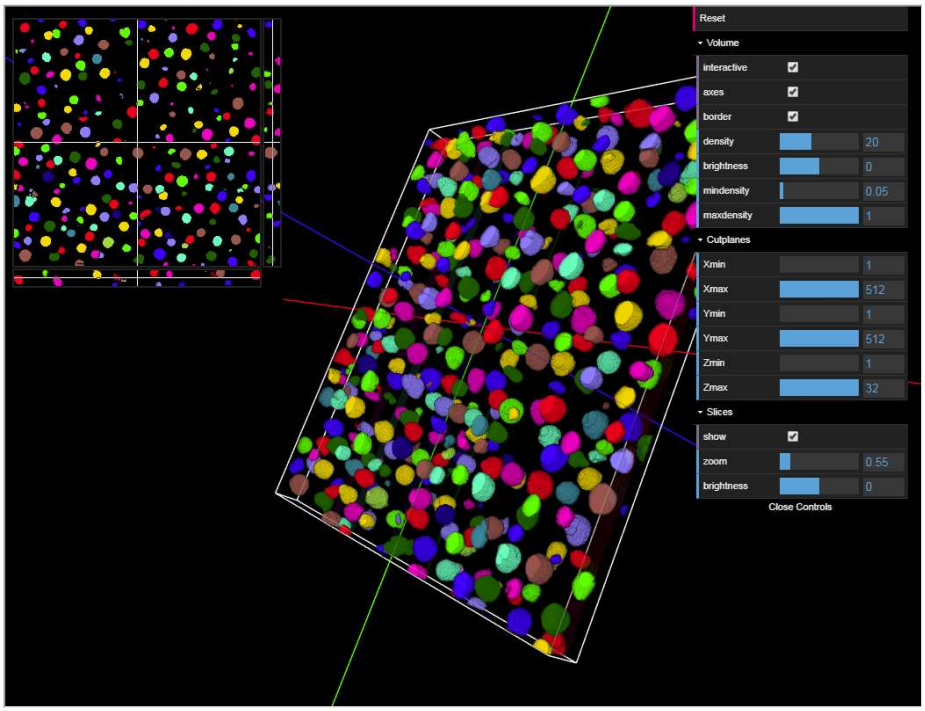
the an example of the mosaic images of original and color labeled of the *immuno* dataset.

As mentioned in Chapter 8.1 we provide two different version of the 3D interactive visualization: One for the purpose of input data visualization and the other for the purpose of segmented results visualization along with color labeling. These visualization can be observed in the third and fifth tab in the DINAVID system as shown in Figure 8.2. The examples of the 3D visualization of the original volume of the *immuno* and *WSM* datasets are exhibited in Figure 8.4 and 8.5, respectively. As observed each visualization window contains three things: 3D volume visualization (center), 2D cross sections visualization (upper left), and the light weighted GUI (right) to control parameters to get better visualization adjusted for the client. Note that 3D volume visualization and 2D cross sections can be selectively removed in the display from the GUI. Also, the parameters shown in GUI can be collapsed so that the client can utilize large space of the windows. In addition, in order to connect between 3D volume and 2D cross sections, there are red, green, and blue axes located at the volume to indicate which sections are displayed. We provide set of the default parameters for the visualization and these parameters can be adjustable by client to have a better visualization. Also, in side of the window, we can adjust parameters to display certain range of the xy sections ($z_{\min}$, $z_{\max}$), xz sections ($y_{\min}$, $y_{\max}$), and yz sections ($x_{\min}$, $x_{\max}$) to selectively visualize the 3D subvolume.

Similarly, the visualization after segmentation using the method proposed in [116] also provided in the DINAVID system. Note that a segmentation output of the method in [116] is a binary mask. We perform our proposed color labeling method (which will describe later in this chapter) to assign colors to the output of the binary mask. The examples of the 3D visualization of the segmentation results after color labeling of the *immuno* and *WSM* datasets are exhibited in Figure 8.6 and 8.7, respectively. Note that the default parameters for the color labeled segmentation results are different from the original volume visualization since the grayscale images and color labeled images look different. Same as previous original volume visualiza-

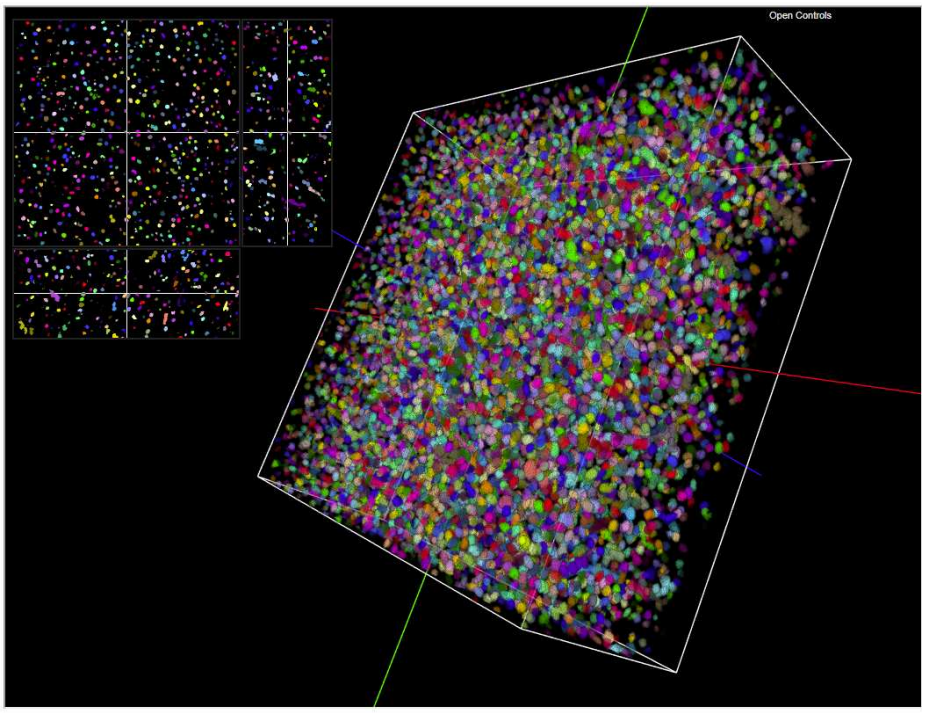


(a) Color labeled segmentation result visualization collapsing light weighted GUI

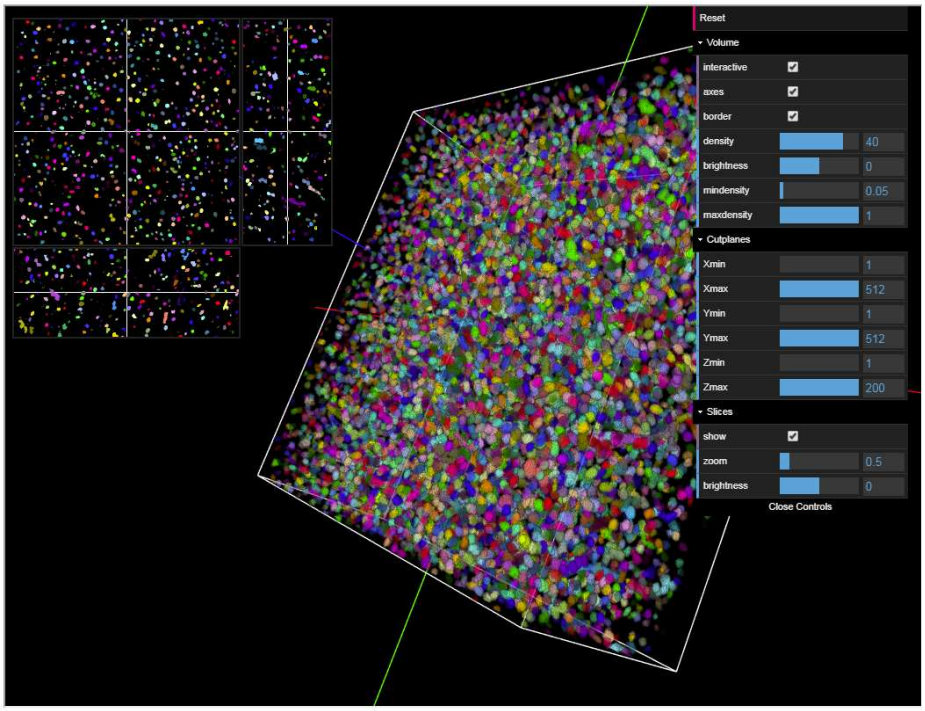


(b) Color labeled segmentation result visualization with light weighted GUI

Fig. 8.6. 3D visualization using the color labeled segmentation results of the *immuno* dataset



(a) Color labeled segmentation result visualization collapsing light weighted GUI



(b) Color labeled segmentation result visualization with light weighted GUI

Fig. 8.7. 3D visualization using the color labeled segmentation results of the *WSM* dataset

tion, all parameters are adjustable so that the client can change parameters for the better observation.

8.3 Color Labeling [3]

The objective of color labeling is observing segmentation results of structure of interest. In general, the final output of segmentation results are a binary mask which each pixel belongs to background region or foreground region. The color labeling assigned to this binary mask make biologists be capable of determining the segmentation results of individual structures and visually evaluate the performance of segmentation. Typical color mapping provided by commercial software does not use distinctive color for the adjacent segmented objects. Human eyes could not distinguish separate but adjacent objects especially with dense microscopy nuclei objects in 3D if their colors are same or similar. In this chapter, we demonstrate our proposed color labeling method we developed.

8.3.1 Proposed Color Labeling Method

Figure 8.8 is a block diagram of our proposed color labeling method. The input is a binary image/mask after segmentation and we denote this binary volume results as I^{BW}. Also, the output of this color labeling method is a color labeled volume denoted as I^{CL}. We are particularly interested in color labeling to the nuclei segmentation since nuclei is small but densely located in 3D. However, our proposed color labeling method can also be used for other segmented objects to help human distinguish the objects of interest as shown in Figure 6.5 and 6.6 in Chapter 6. First of all, colormap is generated. Then, 3D connected component labeling, morphological operations are performed followed by small component removal. After these labeling and clean-up processes are done, 3D windows is selected to assign colors based on the generated

[3]The work described in this section was jointly done with David Ho of Purdue University.

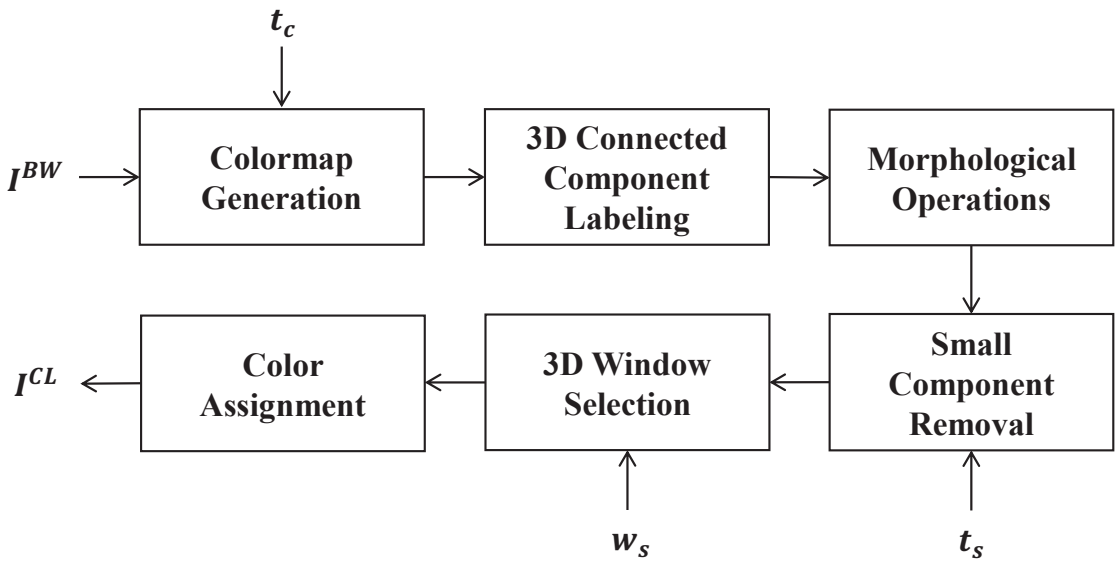


Fig. 8.8. Block diagram of our color labeling method

colormap. So, the output should be the same size of volume as I^{BW} but individual segmented nucleus has a distinctive color compared to its adjacent nuclei.

Colormap Generation

Fig. 8.9. The sample first 100 colors of our colormap

Appropriate colormap generation is a key step of being successfully distinct each adjacent nuclei. To address this problem we generate a colormap which has a distinct

each other so that we can assign color based on order of this colormap. Here, an RGB color coordinate for the i^{th} color denoted as c_i are determined from the *Lab* color space [191]. Note that the *Lab* color space is usually chosen because the *Lab* color space is perceptually uniform color space with human vision system. More specifically, L value is obtained from luminance mainly weighed from G channel values, a reflects nonlinear transform obtained by G and R channel values from the RGB space, b reflects nonlinear transform obtained by B and Y channel values. This *Lab* color space helps human vision system that the amount of color changes in these values are same amount that visually perceived changes from human visual system.

Our first color (c_1) for the proposed colormap is chosen as blue which the RGB color space value is $c_1 = (0, 0, 255)$. After that, we choose red ($c_2 = (255, 0, 0)$) and green ($c_3 = (0, 255, 0)$) sequentially since each B, R, and G are located at the end of orthogonal axes in the color space so that they can be maximized color the sum of distance each other. For $i \geq 4$, c_i are determined by maximizing the distance between c_i and already existed color values (c_j where $j = 1, \ldots, i$) such that

$$\underset{c_i}{\operatorname{argmax}} \sum_{j=1}^{i-1} d_{Lab}(c_i, c_j), \text{ when } i \geq 4 \tag{8.1}$$

where $d_{Lab}(c_i, c_j)$ is the Euclidean distance between two color value locations in the *Lab* color space [191]. Thus, each color value depends on previous colors set which already existed in the colormap. We generate 9000 colors with sequential order as an initial colormap by Eq (8.1). Note that this colormap is distinct color set used for highlighting the foreground of segmentation results. However, we use a background color as black $(0, 0, 0)$ in color space so there is a need to modify given initial colormap so that biologist is able to distinguish between assigned color for the foreground and black background. To achieve this, we introduce color threshold denoted as t_c that reject color candidate which is similar color as background from an initial colormap. In particular, we use $t_c = 50$ and compared with each R, G, and B channel values. If all of the channel values are less than t_c, we drop that color candidate. However, if at least one channel value are greater than t_c, we keep that color candidate. Using this

thresholding step, we obtain final colormap which contains 8863 colors. Figure 8.9 is the first 100 colors of our colormap. As observed in Figure 8.9 neighboring colors are distinguishable by human eyes as well as black especially in first 30 colors. In general, we do not use more than 30 colors since we reuse the color from the c_1 if we do not find any objects adjacent to the target object.

Labeling

Once we generate colormap as we discussed in previous Chapter 8.3.1, we need to assign labels to each individual objects to segment from the I^{BW}. Here, our proposed approach is performing 3D connected component followed by morphological operations [34] in order to assign different labels to the adjacent/touched segmented objects. More specifically, we firstly perform a 3D morphological erosion with a structuring element of the sphere of radius 1. This 3D erosion helps slightly touched objects to separate two or multiple objects. Once this erosion is done, we utilize 3D connected component with 6-neighborhood and assign different label to each component. Then, using 3D morphological dilation with a structuring element of the sphere of radius 1 to recover the eroded part. Note that this is different from simple erosion followed by dilation since we already assign different labels to touched objects. Therefore, although the touched objects are getting together, the assigned labels are different. Finally, we use threshold denoted as t_s for cleaning small components. This t_s is adjustable parameter and the default value of the small component threshold value as 20. If each labeled component is smaller than 20 voxels, make the component belong to background region. Otherwise, leave it as it is.

Color Assignment

The last step of the proposed color labeling is a color assignment step which assigns color to each labeled component. Since we need to assign distinct colors to neighbor objects of interest, we firstly use a 3D bounding box which is the smallest size that

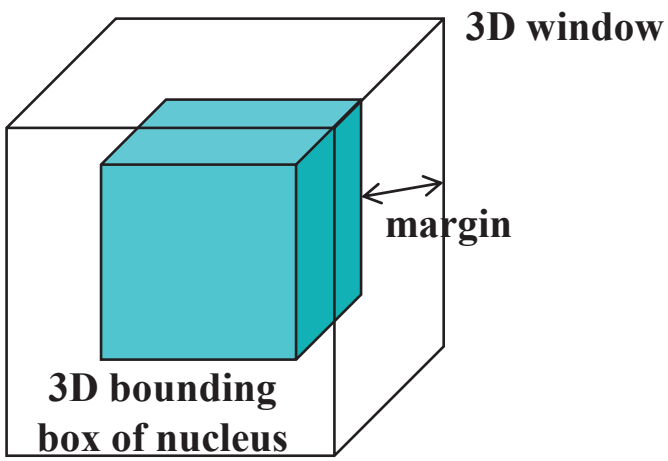
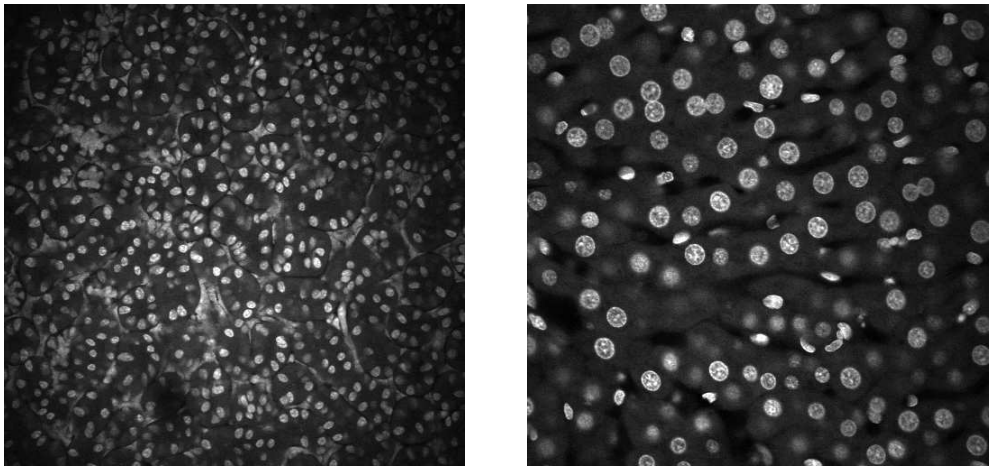


Fig. 8.10. The 3D bounding box to be considered as a 3D neighborhood for the color assignment

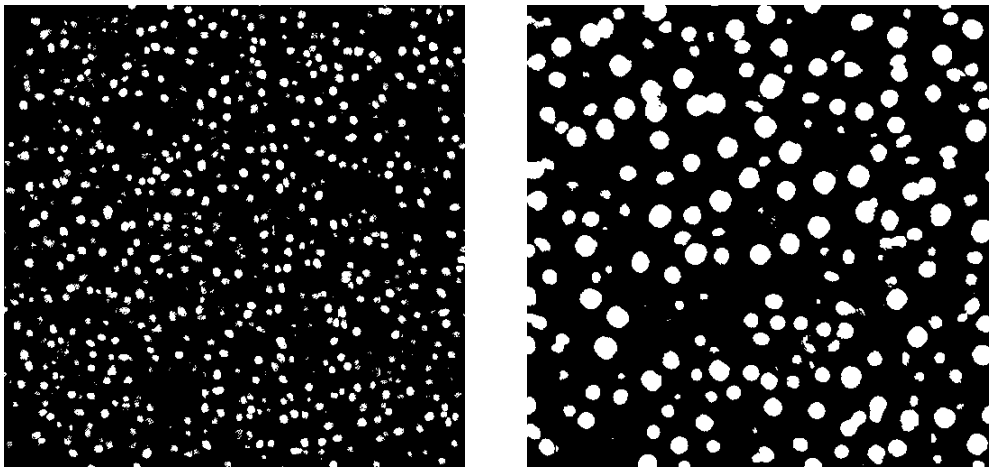
contains the segmented objects. Then we can set a 3D window which locates the 3D bounding box at the center with a margin denoted as m portrayed in Figure 8.10. For example, if the 3D bounding box is $I_{(x_1:x_2,y_1:y_2,z_1:z_2)}$, then the corresponding 3D window becomes $I_{(x_1-m:x_2+m,y_1-m:y_2+m,z_1-m:z_2+m)}$. Here, we set m to be 40 voxels. Then, we assigned color c_k where k is the smallest index that not taken by other segmented objects inside of the 3D window. This procedure is repeated until all the segmented objects are assigned colors. In this ways, we ensure that any of the segmented objects within the certain margin distance cannot have same colors. Also, since our colormap generation is based on maximum distance in the *Lab* color space, the assigned colors are distinctive. The proposed color labeling method, however, could not assign all different colors to the entire 3D volume space. To be more specific, same color value can be used with different segmented objects like c_1. Despite the fact that proposed colormap did not assign different colors to entire volume, it guarantees the adjacent segmented objects are assigned with distinctive colors.

8.3.2 Experimental Results

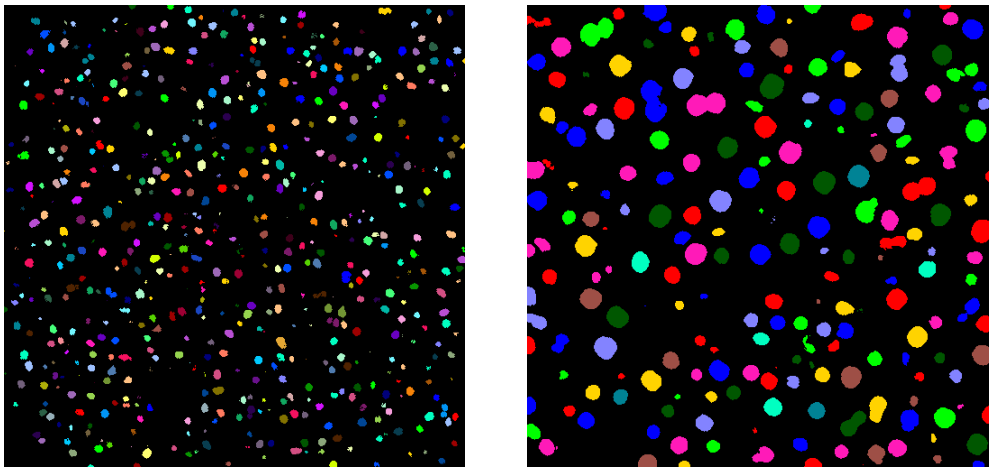
Figure 8.11 exhibits our color labeling result. Notify that this is the sample 2D image out of 3D color labeling results obtained from our DINAVID system that used for segmentation and visualization. Here, the binary segmentation image is obtained using the SpCycleGAN described in [116]. As observed in Figure 8.11 our proposed color labeling method successfully assign distinctive colors to adjacent nuclei. Consequently, the segmented results are visually distinguishable and much clearer than conventional color labeling method.



(a) $I^{O}_{z_{100}}$ of *WSM* (b) $I^{O}_{z_{8}}$ of *immuno*



(c) $I^{BW}_{z_{100}}$ of *WSM* (d) $I^{BW}_{z_{8}}$ of *immuno*



(e) $I^{CL}_{z_{100}}$ of *WSM* (f) $I^{CL}_{z_{8}}$ of *immuno*

Fig. 8.11. Original, binary, and proposed color labeled image of *WSM* blue channel (left column) and *immuno* blue channel (right column)

9. CONCLUSIONS

9.1 Summary

This thesis introduced various segmentation methods to understand and quantify biological structures observed in fluorescence microscopy volume. Our datasets are obtained from rat kidney labeled with Hoechst 33342 (nuclei) and phalloidin (tubules) collected using two-photon microscopy. Since nuclei and tubules have different characteristics, we developed various segmentation methods depending on different biological structures. Also, this thesis demonstrated deep learning based three dimensional image deconvolution method to enhance microscopy image quality. This deconvolution method could be used as a preprocessing step to improve segmentation results. In addition, we developed tools for visualizing 3D segmentation results so that biologists can evaluate segmentation results. This tool is used along with pretrained deep learning network model. The main contributions of this thesis are:

- Wavelet Based Multiscale Edge Detection

  We reviewed current edge detection techniques using wavelet and contourlet transform. We introduced our multiscale edge detection method using dyadic wavelet transform. To be more specific, the proposed method utilizes a combination of highpass components in x and y directions passed from 1D dyadic filter, pointwise multiplications across two adjacent dyadic wavelet transform, the modulus and angle computation followed by thresholding, and nonmaximum suppression in 8 directions to get edge maps. This method was tested on our two-photon microscopy volume acquired from a rat kidney labeled with Hoechst 33342. The edge detection results from different adjacent scales were compared with Canny edge detector side by side. Experimental results showed ability to capture nuclei edges despite inhomogeneity background.

- Boundary Fitting Based Segmentation

  We conducted a literature review focusing on tubular shapes segmentation schemes. We demonstrated boundary fitting based segmentation technique for segmenting boundaries of tubular structures. In particular, this technique segments 2D cross-sections of tubules in microscopy images using a combination of adaptive and global thresholding, potentials, z direction refinement, branch pruning, end point matching, and curve fitting. The aim of this technique is to be able to segment the 2D cross-sections at various depth as an initial step to 3D segmentation of these objects. This technique was tested on our datasets that are 3D fluorescence microscopy volume of rat kidney labeled with fluorescent phalloidin dyes using two photon fluorescence excitation microscopy. The experimental results were compared to that of an active contours based method and showed efficacy of our proposed boundary fitting based segmentation technique.

- Three Dimensional Active Contours with Inhomogeneity Correction

  We conducted a literature review regarding edge-based active contours, 2D region-based active contours, 3D region-based active contours. We extended previous 3D region-based active contours method by utilizing 3D inhomogeneity correction. Prior work that incorporated 3D snakes did not account for intensity inhomogeneities. More particularly, our proposed method modified energy function of 3D snakes and add inhomogeneity field terms so that active contours find optimum 3D segmentation as well as 3D inhomogeneity field information. The goal of this method is to account for intensity inhomogeneity while incorporating 3D information so as to segment 3D nuclei successfully. This method was tested on five different datasets that contain 3D nuclei structures obtained from rat kidney labeled with Hoechst 33342 collected using two-photon microscopy. The segmentation results of proposed method were compared with five different methods visually and numerically. Also, we presented 3D segmentation results

using a 3D visualization tool. Experimental results demonstrate that the proposed method achieves better performance than other reported methods.

- Tubule Synthesis and Segmentation Based on Deep Learning

  We reviewed recent deep learning schemes for image segmentation. We also reviewed multiple schemes for the image synthesis. We presented tubular structure segmentation method using convolutional neural networks with data augmentation and inhomogeneity correction using 3D region-based active contours method as a preprocessing to correct inhomogeneity background. After that data augmentation was performed to increase size of paired training set. This paired training set becomes convolutional neural networks input to train model. This trained model was then used for inference for the test set. Lastly, postprocessing was done for filling holes and small components removal. The goal of this method is identifying and segmenting individual tubules enclosed by their membranes. In order to identify individual tubule regions, we used 3D connected component labeling to assign a color to each tubule region. This method was tested on two different datasets that contain 3D tubular structures imaged from rat kidney labeled with phalloidin which labels filamentous actin collected using two-photon microscopy. The segmentation results of the proposed methods were evaluated qualitatively and quantitatively compared with other microscopy segmentation methods. Also, we introduced object-based metrics as well as pixel-based metric to evaluate segmentation correctness. Experimental results showed that our proposed method outperforms with identifying multiple tubules than other compared methods. Moreover, we presented a 3D tubular structure synthesis method using a gradient noise model by generating random but smooth centerlines. The results demonstrated that our proposed method can successfully generate 3D synthetic binary tubular structures.

- Three Dimensional Blind Image Deconvolution Using Generative Adversarial Networks

We reviewed various methods from conventional methods to recent deep learning methods for image/volume deconvolution. We presented a blind image deconvolution method for fluorescence microscopy volumes using the 3-Way SpCycleGAN. In particular, we presented a new approach to restore various biological structures in 3D microscopy images in deeper tissue without knowing the 3D point spread function using a spatially constrained CycleGAN (SpCycleGAN). We performed train and inference the SpCycleGAN in three directions along with xy, yz, and xz sections (3-Way SpCycleGAN). These restored 3-way microscopy volumes were averaged to incorporate 3D information. The restored volumes of proposed deconvolution method and other well-known deconvolution methods, denoising methods, and an inhomogeneity correction method were visually and numerically evaluated. To produce numerical evaluations, three image quality metrics were used. We tested on our proposed and compared methods to two datasets which consist of Hoechst 33342 labeled nuclei and phalloidin labeled filament actin collected from a rat kidney, respectively. Moreover, we used this restored volume as a preprocessing step to perform SpCycleGAN segmentation and visualize the 3D segmentation results. Experimental results indicated that our proposed blind deconvolution method can restore and improve the quality of blurred and noisy deep depth microscopy image visually and quantitatively.

- Microscopy Volume Visualization

  We developed the Distributed and Networked Analysis of Volumetric Image Data (DINAVID) system that provides image analysis and visualization tools to biologists. In particular, this DINAVID system provides pretrained deep learning model for the nuclei segmentation and biologist can choose area to perform deep learning inference. We also developed this color labeling method such that biologist can distinguish adjacent nuclei as much as possible. Moreover, we developed 3D interactive visualization tools to help biologist observe the original volume as well as segmentation results. Along with color labeling,

the 3D interactive visualization could provide better insight for biologists to evaluate segmentation results.

9.2 Future Work

Our proposed methods can be improved and extended as follows:

- Wavelet Based Multiscale Edge Detection

  We reviewed current edge detection techniques and proposed our multiscale edge detection technique using dyadic wavelet transform, pointwise multiplications, the modulus and angle computation, thresholding followed by nonmaximum suppression. However, our multiscale edge detection method used thresholding technique and the edge detection results heavily rely on thresholding values. This thresholds are empirically selected in each image and therefore it is impractical. Also, our future work after segmentation will be quantification of biological structures. In this perspective, closed boundaries detection of each nuclei is expected. Due to this fact, we abandoned this technique and proposed active contours based method.

- Boundary Fitting Based Segmentation

  We reviewed tubular shapes segmentation schemes and demonstrated a 2D tubule segmentation scheme that segments tubular structures in microscopy images based on adaptive and global thresholding, potentials, z direction refinement, branch pruning, end point matching, and boundary fitting. This boundary fitting sometimes failed if end point matching is not correctly done. Also, blurred edges of tubules sometimes yield false end points so that the false end points may match wrongly with existed end points. In future, we will extend the current techniques to 3D so that we can get more information from adjacent image slices. Also, we will extend this method to use region merging to combine small regions, utilizing additional shape information to aid in the sep-

aration of tubular structures that were joined together during the segmentation process.

- Three Dimensional Active Contours with Inhomogeneity Correction
  We reviewed traditional active contours including edge-based and region-based method and extended 3D active contours with incorporating 3D inhomogeneity field correction. One of the limitation of this method is that this proposed method cannot separate overlapped nuclei since this method only provides segmentation mask but not individual nuclei labels. Also, proposed method sometimes over-corrected inhomogeneity background so that the nuclei which should not be shown in the specific slices may appear. Therefore, future work will include quantifying the segmented individual nuclei by separating multiple overlapped nuclei and develop automatic ways to determine appropriate level of inhomogeneity correction.

- Tubule Synthesis and Segmentation Based on Deep Learning
  We reviewed recent deep learning schemes used for image segmentation and introduced our proposed tubule segmentation method using a combination of intensity inhomogeneity correction, data augmentation, followed by a CNN architecture. Currently, our proposed method should need manually annotated groundtruth but manual annotation is time consuming and intractable. In future, we plan to utilize 3D CNN architecture to fully utilize 3D information to improve segmentation results. Also, we plan to generate realistic 3D synthetic tubules using generative adversarial network to reduce manual annotation work. We already generated 3D synthetic binary tubules and need to use them for generating realistic 3D tubules. Also, current 3D synthetic tubules are not packed in the volume but sparsely located each other. Hence, we plan to generate more tubules densely packed to given 3D volume boundary box.

- Three Dimensional Blind Image Deconvolution For Fluorescence Microscopy Using Generative Adversarial Networks

We reviewed various methods from conventional methods to recent deep learning methods for image/volume deconvolution. We demonstrated our proposed three dimensional blind image deconvolution method using 3-Way SpCycleGAN. Currently, our proposed method is not fully 3D method but utilizing 3 directions of 2D method since the SpCycleGAN works only in 2D. Therefore, our future work should include fully 3D SpCycleGAN for achieving blind 3D volume deconvolution. Moreover, we plan to investigate the image/volume quality metric for fluorescence microscopy since there is no gold standard image/volume quality metric yet. Lastly, we will also need to develop 3D tubule structure segmentation so that we can compare tubule segmentation results with various restored method as a preprocessing step.

- Microscopy Volume Visualization

  We developed the DINAVID system which can offer image analysis and 3D interactive visualization tools. Our current version of the DINAVID system is smaller version of the system and currently testing with biologists who works with us. In the future, we plan to update system for the larger group of people can use. Also, current 3D interactive visualization system is locally rendered which means it cannot use the power of server side GPU. Thus, our future goal could be utilizing server side GPU and send information to clients so that there is no burden from client side for the volume rendering.

9.3 Publications Resulting from This Thesis

Journal Papers

1. C. Fu, S. Han, **S. Lee**, D. J. Ho, P. Salama, K. W. Dunn, and E. J. Delp, "Three dimensional nuclei synthesis and instance segmentation," *To be submitted to the IEEE Transactions on Medical Imaging.*

2. **S. Lee**, S. Han, C. Fu, P. Salama, K. W. Dunn, and E. J. Delp, "Three dimensional tubule synthesis and segmentation for fluorescence microscopy using generative adversarial networks," *To be submitted to the IEEE Transactions on Medical Imaging.*

3. S. Han, **S. Lee***, C. Fu*, D. J. Ho, P. Salama, K. W. Dunn, and E. J. Delp, "A high-performance distributed and networked image analysis system for volumetric image data (DINAVID)," *To be submitted to the IEEE Transactions on Medical Imaging.*

4. K. W. Dunn, C. Fu, D. J. Ho, **S. Lee**, S. Han, P. Salama, and E. J. Delp, "DeepSynth: Three-dimensional nuclear segmentation of biological images using neural networks trained with synthetic data," *Submitted to Scientific Reports.*

Conference Papers

1. **S. Lee**, S. Han, P. Salama, K. W. Dunn, and E. J. Delp, "Three dimensional blind image deconvolution for fluorescence microscopy using generative adversarial networks," *Proceedings of the IEEE International Symposium on Biomedical Imaging*, pp. 538-542, April 2019, Venice, Italy.

2. S. Han, **S. Lee**, C. Fu, P. Salama, K. W. Dunn, and E. J. Delp, "Nuclei counting in microscopy images with three dimensional generative adversarial networks," *Proceedings of the SPIE Conference on Medical Imaging*, pp. 109492Y-1-11, February 2019, San Diego, CA.

3. C. Fu, **S. Lee**, D. J. Ho, S. Han, P. Salama, K. W. Dunn, and E. J. Delp, "Fluorescence microscopy image segmentation using convolutional neural network with generative adversarial networks," *Proceedings of the IEEE International Conferences on Computer Vision and Pattern Recognition Workshop*, pp. 2302-2310, June 2018, Salt Lake City, UT.

4. **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*, vol. 2018, no. 15, pp. 199-1-199-8, January 2018, Burlingame, CA.

5. **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.

6. **S. Lee**, P. Salama, K. W. Dunn, and E. J. Delp, "Boundary fitting based segmentation of fluorescence microscopy images," *Proceedings of the IS&T / SPIE International Symposium on Electronic Imaging*, pp. 940805-1-10, February 2015, San Francisco, CA.

REFERENCES

REFERENCES

[1] D. B. Murphy and M. W. Davidson, *Fundamentals of light microscopy and electronic imaging*, 2nd ed. Hoboken, NJ: Wiley-Blackwell, 2012. [Online]. Available: https://doi.org/10.1002/9781118382905

[2] A. Koehler, "New method of illumination for phomicrographical purposes," *Journal of the Royal Microscopical Society*, vol. 14, pp. 261–262, 1894.

[3] M. Minsky, "Memoir on inventing the confocal scanning microscope," *Scanning*, vol. 10, no. 4, pp. 128–138, 1988. [Online]. Available: https://doi.org/10.1002/sca.4950100403

[4] M. G. L. Gustafsson, D. A. Agard, and J. W. Sedat, "I^{5}M: 3D widefield light microscopy with better than 100 nm axial resolution," *Journal of Microscopy*, vol. 195, no. 1, pp. 10–16, July 1999. [Online]. Available: https://doi.org/10.1046/j.1365-2818.1999.00576.x

[5] K. A. Stankov, "A mirror with an intensity-dependent reflection coefficient," *Applied Physics B*, vol. 45, no. 3, pp. 191–195, March 1988. [Online]. Available: https://doi.org/10.1007/BF00695290

[6] A. Jablonski, "Efficiency of anti-stokes fluorescence in dyes," *Nature*, vol. 131, pp. 839–840, June 1933. [Online]. Available: https://doi.org/10.1038/131839b0

[7] J. R. Lakowicz, *Introduction to fluorescence*, 3rd ed. Boston, MA: Springer US, 2006. [Online]. Available: https://doi.org/10.1007/978-0-387-46312-4_1

[8] H. Fujisaki and J. E. Straub, "Vibrational energy relaxation in proteins," *Proceedings of the National Academy of Sciences*, vol. 102, no. 19, pp. 6726–6731, May 2005. [Online]. Available: https://doi.org/10.1073/pnas.0409083102

[9] R. K. P. Benninger, M. Hao, and D. W. Piston, "Multi-photon excitation imaging of dynamic processes in living cells and tissues," *Reviews of Physiology Biochemistry and Pharmacology*, vol. 160, pp. 71–92, April 2008. [Online]. Available: https://doi.org/10.1007/112_2008_801

[10] D. W. Piston, "Imaging living cells and tissues by two-photon excitation microscopy," *Trends in Cell Biology*, vol. 9, no. 2, pp. 66–69, February 1999. [Online]. Available: https://doi.org/10.1016/S0962-8924(98)01432-9

[11] W. Denk, J. H. Strickler, and W. W. Webb, "Two-photon laser scanning fluorescence microscopy," *Science*, vol. 248, no. 4951, pp. 73–76, April 1990. [Online]. Available: https://doi.org/10.1126/science.2321027

[12] K. Svoboda and R. Yasuda, "Principles of two-photon excitation microscopy and its applications to neuroscience," *Neuron*, vol. 50, no. 6, pp. 823–839, June 2006. [Online]. Available: https://doi.org/10.1016/j.neuron.2006.05.019

[13] S. G. Clendenon, P. A. Young, M. Ferkowicz, C. Phillips, and K. W. Dunn, "Deep tissue fluorescent imaging in scattering specimens using confocal microscopy," *Microscopy and Microanalysis*, vol. 17, no. 4, pp. 614–617, August 2011. [Online]. Available: https://doi.org/10.1017/S1431927611000535

[14] W. Denk and K. Svoboda, "Photon upmanship: Why multiphoton imaging is more than a gimmick," *Neuron*, vol. 18, no. 3, pp. 351–357, 1997. [Online]. Available: https://doi.org/10.1016/S0896-6273(00)81237-4

[15] A. T. Young, "Rayleigh scattering," *Applied Optics*, vol. 20, no. 4, pp. 533–535, February 1981. [Online]. Available: https://doi.org/10.1364/AO.20.000533

[16] D. W. Piston, "The coming of age of two-photon excitation imaging for intravital microscopy," *Advanced Drug Delivery Reviews*, vol. 58, no. 7, pp. 770–772, 2006. [Online]. Available: https://doi.org/10.1016/j.addr.2006.07.003

[17] F. Helmchen and W. Denk, "Deep tissue two-photon microscopy," *Nature Methods*, vol. 2, no. 12, pp. 932–940, December 2005. [Online]. Available: https://doi.org/10.1038/nmeth818

[18] W. R. Zipfel, R. M. Williams, and W. W. Webb, "Nonlinear magic: Multiphoton microscopy in the biosciences," *Nature Biotechnology*, vol. 21, no. 11, pp. 1369–1377, October 2003. [Online]. Available: https://doi.org/10.1038/nbt899

[19] J. Mertz, "Nonlinear microscopy: New techniques and applications," *Current Opinion in Neurobiology*, vol. 14, no. 5, pp. 610–616, October 2004. [Online]. Available: https://doi.org/10.1016/j.conb.2004.08.013

[20] R. G. Presson Jr, M. B. Brown, A. J. Fisher, R. M. Sandoval, K. W. Dunn, K. S. Lorenz, E. J. Delp, P. Salama, B. A. Molitoris, and I. Petrache, "Two-photon imaging within the murine thorax without respiratory and cardiac motion artifact," *The American Journal of Pathology*, vol. 179, no. 1, pp. 75–82, July 2011. [Online]. Available: https://doi.org/10.1016/j.ajpath.2011.03.048

[21] A. Dufour, V. Shinin, S. Tajbakhsh, N. Guillen-Aghion, J. C. Olivo-Marin, and C. Zimmer, "Segmenting and tracking fluorescent cells in dynamic 3-D microscopy with coupled active surfaces," *IEEE Transactions on Image Processing*, vol. 14, no. 9, pp. 1396–1410, September 2005. [Online]. Available: https://doi.org/10.1109/TIP.2005.852790

[22] J. A. Rosado-Toro and J. J. Rodriguez, "Cell splitting using dynamic programming," *Proceedings of the IEEE Southwest Symposium on Image Analysis and Interpretation*, pp. 33–36, April 2012, Santa Fe, NM. [Online]. Available: https://doi.org/10.1109/SSIAI.2012.6202446

[23] C. L. Phillips, L. J. Arend, A. J. Filson, D. J. Kojetin, J. L. Clendenon, S. Fang, and K. W. Dunn, "Three-dimensional imaging of embryonic mouse kidney by two-photon microscopy," *The American Journal of Pathology*, vol. 158, no. 1, pp. 49–55, January 2001. [Online]. Available: https://doi.org/10.1016/S0002-9440(10)63943-0

[24] M. J. Kyan, L. Guan, M. R. Arnison, and C. J. Cogswell, "Feature extraction of chromosomes from 3-D confocal microscope images," *IEEE Transactions on Biomedical Engineering*, vol. 48, no. 11, pp. 1306–1318, November 2001. [Online]. Available: https://doi.org/10.1109/10.959326

[25] L. G. Brown, "A survey of image registration techniques," *ACM Computing Survey*, vol. 24, no. 4, pp. 325–376, December 1992. [Online]. Available: https://doi.org/10.1145/146370.146374

[26] P. T. C. So, C. Y. Dong, B. R. Masters, and K. M. Berland, "Two-photon excitation fluorescence microscopy," *Annual Review of Biomedical Engineering*, vol. 2, no. 1, pp. 399–429, August 2000. [Online]. Available: https://doi.org/10.1146/annurev.bioeng.2.1.399

[27] K. W. Dunn, R. M. Sandoval, K. J. Kelly, P. C. Dagher, G. A. Tanner, S. J. Atkinson, R. L. Bacallao, and B. A. Molitoris, "Functional studies of the kidney of living animals using multicolor two-photon microscopy," *American Journal of Physiology-Cell Physiology*, vol. 283, no. 3, pp. C905–C916, September 2002. [Online]. Available: https://doi.org/10.1152/ajpcell.00159.2002

[28] L. Vincent and P. Soille, "Watershed in digital spaces: An efficient algorithm based on immersion simulations," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, no. 6, pp. 583–598, June 1991. [Online]. Available: https://doi.org/10.1109/34.87344

[29] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *International Journal of Computer Vision*, vol. 1, no. 4, pp. 321–331, January 1988. [Online]. Available: https://doi.org/10.1007/BF00133570

[30] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62–66, 1979. [Online]. Available: https://doi.org/10.1109/TSMC.1979.4310076

[31] W. Niblack, *An introduction to digital image processing.* Englewood Cliffs, NJ: Prentice-Hall, 1986.

[32] J. Sauvola and M. Pietikainen, "Adaptive document image binarization," *Pattern Recognition*, vol. 33, no. 2, pp. 225–236, February 2000. [Online]. Available: https://doi.org/10.1016/S0031-3203(99)00055-2

[33] G. Li, T. Liu, A. Tarokh, J. Nie, L. Guo, A. Mara, S. Holley, and S. T. C. Wong, "3D cell nuclei segmentation based on gradient flow tracking," *BMC Cell Biology*, vol. 8, no. 1, pp. 40–1–10, September 2007. [Online]. Available: https://doi.org/10.1186/1471-2121-8-40

[34] R. C. Gonzalez and R. E. Woods, *Digital image processing*, 2nd ed. Upper Saddle River, NJ: Pearson, 2002. [Online]. Available: https://www.pearson.com/us/higher-education/program/Gonzalez-Digital-Image-Processing-4th-Edition/PGM241219.html

[35] X. Yang, H. Li, and X. Zhou, "Nuclei segmentation using marker-controlled watershed, tracking using mean-shift, and Kalman filter in time-lapse microscopy," *IEEE Transactions on Circuits and Systems*, vol. 53, no. 11, pp. 2405–2414, November 2006. [Online]. Available: https://doi.org/10.1109/TCSI.2006.884469

[36] X. Chen, X. Zhou, and S. T. C. Wong, "Automated segmentation, classification, and tracking of cancer cell nuclei in time-lapse microscopy," *IEEE Transactions on Biomedical Engineering*, vol. 53, no. 4, pp. 762–766, April 2006. [Online]. Available: https://doi.org/10.1109/TBME.2006.870201

[37] A. Krtolica, C. O. Solorzano, S. Lockett, and J. Campisi, "Quantification of epithelial cells in coculture with fibroblasts by fluorescence image analysis," *Cytometry*, vol. 49, no. 2, pp. 73–82, October 2002. [Online]. Available: https://doi.org/10.1002/cyto.10149

[38] G. Lin, U. Adiga, K. Olson, J. Guzowski, C. Barnes, and B. Roysam, "A hybrid 3-D watershed algorithm incorporating gradient cues and object models for automatic segmentation of nuclei in confocal image stacks," *Cytometry*, vol. 56, no. 1, pp. 23–36, November 2003. [Online]. Available: https://doi.org/10.1002/cyto.a.10079

[39] K. S. Lorenz, F. Serrano, P. Salama, and E. J. Delp, "Segmentation and registration based analysis of microscopy images," *Proceedings of the IEEE International Conference on Image Processing*, pp. 4213–4216, November 2009, Cairo, Egypt. [Online]. Available: https://doi.org/10.1109/ICIP.2009.5413531

[40] R. Delgado-Gonzalo, V. Uhlmann, D. Schmitter, and M. Unser, "Snakes on a plane: A perfect snap for bioimage analysis," *IEEE Signal Processing Magazine*, vol. 32, no. 1, pp. 41–48, January 2015. [Online]. Available: https://doi.org/10.1109/MSP.2014.2344552

[41] V. Caselles, R. Kimmel, and G. Sapiro, "Geodesic active contours," *International Journal of Computer Vision*, vol. 22, no. 1, pp. 61–79, February 1997. [Online]. Available: https://doi.org/10.1023/A:1007979827043

[42] Y. He, N. Khanna, C. J. Boushey, and E. J. Delp, "Snakes assisted food image segmentation," *Proceedings of the IEEE International Workshop on Multimedia Signal Processing*, pp. 181–185, September 2012, Banff, Canada. [Online]. Available: https://doi.org/10.1109/MMSP.2012.6343437

[43] C. Xu and J. L. Prince, "Snake, shapes, and gradient vector flow," *IEEE Transactions on Image Processing*, vol. 7, no. 3, pp. 359–369, March 1998. [Online]. Available: https://doi.org/10.1109/83.661186

[44] B. Li and S. T. Acton, "Active contour external force using vector field convolution for image segmentation," *IEEE Transactions on Image Processing*, vol. 16, no. 8, pp. 2096–2106, August 2007. [Online]. Available: https://doi.org/10.1109/TIP.2007.899601

[45] B. Li and S. T. Acton, "Automatic active model initialization via Poisson inverse gradient," *IEEE Transactions on Image Processing*, vol. 17, no. 8, pp. 1406–1420, August 2008. [Online]. Available: https://doi.org/10.1109/TIP.2008.925375

[46] T. F. Chan and L. A. Vese, "Active contours without edges," *IEEE Transactions on Image Processing*, vol. 10, no. 2, pp. 266–277, February 2001. [Online]. Available: https://doi.org/10.1109/83.902291

[47] L. A. Vese and T. F. Chan, "A multiphase level set framework for image segmentation using the Mumford and Shah model," *International Journal of Computer Vision*, vol. 50, no. 3, pp. 271–293, December 2002. [Online]. Available: https://doi.org/10.1023/A:1020874308076

[48] K. S. Lorenz, P. Salama, K. W. Dunn, and E. J. Delp, "Three dimensional segmentation of fluorescence microscopy images using active surfaces," *Proceedings of the IEEE International Conference on Image Processing*, pp. 1153–1157, September 2013, Melbourne, Australia. [Online]. Available: https://doi.org/10.1109/ICIP.2013.6738238

[49] O. Dzyubachyk, W. A. van Cappellen, J. Essers, W. J. Niessen, and E. Meijering, "Advanced level-set-based cell tracking in time-lapse fluorescence microscopy," *IEEE Transactions on Medical Imaging*, vol. 29, no. 3, pp. 852–867, March 2010. [Online]. Available: https://doi.org/10.1109/TMI.2009.2038693

[50] S. Lankton and A. Tannenbaum, "Localizing region-based active contours," *IEEE Transactions on Image Processing*, vol. 17, no. 11, pp. 2029–2039, November 2008. [Online]. Available: https://doi.org/10.1109/TIP.2008.2004611

[51] S. Lee, P. Salama, K. W. Dunn, and E. J. Delp, "Boundary fitting based segmentation of fluorescence microscopy images," *Proceedings of the IS&T/SPIE International Symposium on Electronic Imaging*, pp. 940 805–1–10, February 2015, San Francisco, CA. [Online]. Available: https://doi.org/10.1117/12.2085417

[52] C. Li, R. Huang, Z. Ding, C. Gatenby, D. N. Metaxas, and J. C. Gore, "A level set method for image segmentation in the presence of intensity inhomogeneities with application to MRI," *IEEE Transactions on Image Processing*, vol. 20, no. 7, pp. 2007–2016, July 2011. [Online]. Available: https://doi.org/10.1109/TIP.2011.2146190

[53] D. Schmitter, C. Gaudet-Blavignac, D. Piccini, and M. Unser, "New parametric 3D snake for medical segmentation of structures with cylindrical topology," *Proceedings of the IEEE International Conference on Image Processing*, pp. 276–280, September 2015, Quebec City, Canada. [Online]. Available: https://doi.org/10.1109/ICIP.2015.7350803

[54] A. Badoual, D. Schmitter, and M. Unser, "Locally refinable parametric snakes," *Proceedings of the IEEE International Conference on Image Processing*, pp. 354–358, September 2015, Quebec City, Canada. [Online]. Available: https://doi.org/10.1109/ICIP.2015.7350819

[55] 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. [Online]. Available: https://doi.org/10.1109/ISBI.2017.7950618

[56] S. Mallat and S.Zhong, "Characterization of signals from multiscale edges," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 7, pp. 710–732, July 1992. [Online]. Available: https://doi.org/10.1109/34.142909

[57] M. N. Do and M. Vetterli, "The contourlet transform: An efficient directional multiresolution image representation," *IEEE Transactions on Image Processing*, vol. 14, no. 12, pp. 2091–2106, December 2005. [Online]. Available: https://doi.org/10.1109/TIP.2005.859376

[58] S. Mallat and W. L. Hwang, "Singularity detection and processing with wavelets," *IEEE Transactions on Information Theory*, vol. 38, no. 2, pp. 617–643, March 1992. [Online]. Available: https://doi.org/10.1109/18.119727

[59] L. Zhang and P. Bao, "Edge detection by scale multiplication in wavelet domain," *Pattern Recognition Letters*, vol. 23, no. 14, pp. 1771–1784, December 2002. [Online]. Available: https://doi.org/10.1016/S0167-8655(02)00151-4

[60] Z. Zhang, S. Ma, H. Liu, and Y. Gong, "An edge detection approach based on directional wavelet transform," *Computers and Mathematics with Applications*, vol. 57, no. 8, pp. 1265–1271, April 2009. [Online]. Available: https://doi.org/10.1016/j.camwa.2008.11.013

[61] V. Velisavljevic, P. L. Dragotti, and M. Vetterli, "Directional wavelet transforms and frames," *Proceedings of the IEEE International Conference on Image Processing*, pp. 589–592, June 2002, Rochester, NY. [Online]. Available: https://doi.org/10.1109/ICIP.2002.1039039

[62] S. Ma, G. Zheng, L. Jin, S. Han, and R. Zhang, "Directional multiscale edge detection using the contourlet transform," *Proceedings of the IEEE International Conference on Advanced Computer Control*, pp. 58–62, March 2010, Shenyang, China. [Online]. Available: https://doi.org/10.1109/ICACC.2010.5487180

[63] G. Paul, J. Cardinale, and I. F. Sbalzarini, "Coupling image restoration and segmentation: A generalized linear model/Bregman perspective," *International Journal of Computer Vision*, vol. 104, no. 1, pp. 69–93, March 2013. [Online]. Available: https://doi.org/10.1007/s11263-013-0615-2

[64] S. Arslan, T. Ersahin, R. Cetin-Atalay, and C. Gunduz-Demir, "Attributed relational graphs for cell nucleus segmentation in fluorescence microscopy images," *IEEE Transactions on Medical Imaging*, vol. 32, no. 6, pp. 1121–1131, June 2013. [Online]. Available: https://doi.org/10.1109/TMI.2013.2255309

[65] G. Srinivasa, M. C. Fickus, Y. Guo, A. D. Linstedt, and J. Kovacevic, "Active mask segmentation of fluorescence microscope images," *IEEE Transactions on Image Processing*, vol. 18, no. 8, pp. 1817–1829, August 2009. [Online]. Available: https://doi.org/10.1109/TIP.2009.2021081

[66] N. Gadgil, P. Salama, K. W. Dunn, and E. J. Delp, "Nuclei segmentation of fluorescence microscopy images based on midpoint analysis and marked point process," *Proceedings of the IEEE Southwest Symposium on Image Analysis and Interpretation*, pp. 37–40, March 2016, Santa Fe, NM. [Online]. Available: https://doi.org/10.1109/SSIAI.2016.7459169

[67] Q. Wen, H. Chang, and B. Parvin, "A Delaunay triangulation approach for segmenting clumps of nuclei," *Proceedings of the IEEE International Symposium on Biomedical Imaging*, pp. 9–12, June 2009, Boston, MA. [Online]. Available: https://doi.org/10.1109/ISBI.2009.5192970

[68] H. Chang, Q. Wen, and B. Parvin, "Coupled segmentation of nuclear and membrane-bound macromolecules through voting and multiphase level set," *Pattern Recognition*, vol. 48, no. 3, pp. 882–893, March 2015. [Online]. Available: https://doi.org/10.1016/j.patcog.2014.10.005

[69] S. Kothari, Q. Chaudry, and M. D. Wang, "Automated cell counting and cluster segmentation using concavity detection and ellipse fitting techniques," *Proceedings of the IEEE International Symposium on Biomedical Imaging*, pp. 795–798, June 2009, Boston, MA. [Online]. Available: https://doi.org/10.1109/ISBI.2009.5193169

[70] X. Bai, C. Sun, and F. Zhou, "Splitting touching cells based on concave points and ellipse fitting," *Pattern Recognition*, vol. 42, no. 11, pp. 2434–2446, November 2009. [Online]. Available: https://doi.org/10.1016/j.patcog.2009.04.003

[71] Y. He, Y. Meng, H. Gong, S. Chen, B. Zhang, W. Ding, Q. Luo, and A. Li, "An automated three-dimensional detection and segmentation method for touching cells by integrating concave points clustering and random walker algorithm," *PLOS ONE*, vol. 9, no. 8, pp. e104 437–1–15, August 2014. [Online]. Available: https://doi.org/10.1371/journal.pone.0104437

[72] Y. He, H. Gong, B. Xiong, X. Xu, A. Li, T. Jiang, Q. Sun, S. Wang, Q. Luo, and S. Chen, "iCut: An integrative cut algorithm enables accurate segmentation of touching cells," *Scientific Reports*, vol. 5, pp. 12 089–1–17, July 2015. [Online]. Available: https://doi.org/10.1038/srep12089

[73] F. Benmansour and L. D. Cohen, "Tubular structure segmentation based on minimal path method and anisotropic enhancement," *International Journal of Computer Vision*, vol. 92, no. 2, pp. 192–210, March 2010. [Online]. Available: https://doi.org/10.1007/s11263-010-0331-0

[74] H. Li and A. Yezzi, "Vessels as 4-D curves: Global minimal 4-D paths to extract 3-D tubular surfaces and centerlines," *IEEE Transactions on Medical Imaging*, vol. 26, no. 9, pp. 1213–1223, September 2007. [Online]. Available: https://doi.org/10.1109/TMI.2007.903696

[75] S. Worz and K. Rohr, "A new 3D parametric intensity model for accurate segmentation and quantification of human vessels," *Proceedings of the International Conference on Medical Image Computing and Computer Assisted Intervention*, pp. 491–499, September 2004, Saint-Malo, France. [Online]. Available: https://doi.org/10.1007/978-3-540-30135-6_60

[76] A. Fakhrzadeh, E. Sporndly-Nees, L. Holm, and C. L. L. Hendriks, "Analyzing tubular tissue in histopathological thin sections," *Proceedings of the IEEE International Conference on Digital Image Computing Techniques and Applications*, pp. 1–6, December 2012, Fremantle, WA. [Online]. Available: https://doi.org/10.1109/DICTA.2012.6411735

[77] L. M. Lorigo, O. Faugeras, W. E. L. Grimson, R. Keriven, R. Kikinis, A. Nabavi, and C.-F. Westin, "Codimension-two geodesic active contours for the segmentation of tubular structures," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, vol. 1, pp. 444–451, June 2000, Hilton Head Island, SC. [Online]. Available: https://doi.org/10.1109/CVPR.2000.855853

[78] N. Gadgil, P. Salama, K. W. Dunn, and E. J. Delp, "Jelly filling segmentation of fluorescence microscopy images containing incomplete labeling," *Proceedings of the IEEE International Symposium on Biomedical Imaging*, pp. 531–535, April 2016, Prague, Czech Republic. [Online]. Available: https://doi.org/10.1109/ISBI.2016.7493324

[79] D. J. Ho, P. Salama, K. W. Dunn, and E. J. Delp, "Boundary segmentation for fluorescence microscopy using steerable filters," *Proceedings of the SPIE Conference on Medical Imaging*, pp. 10 133–1–11, February 2017, Orlando, FL. [Online]. Available: https://doi.org/10.1117/12.2254627

[80] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015. [Online]. Available: https://doi.org/10.1038/nature14539

[81] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. van der Laak, B. van Ginneken, and C. I. Sanchez, "A survey on deep learning in medical image analysis," *Medical Image Analysis*, vol. 42, no. 1, pp. 60–88, December 2017. [Online]. Available: https://doi.org/10.1016/j.media.2017.07.005

[82] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3431–3440, June 2015, Boston, MA. [Online]. Available: https://doi.org/10.1109/CVPR.2015.7298965

[83] V. Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet: A deep convolutional encoder-decoder architecture for image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 12, pp. 2481–2495, January 2017. [Online]. Available: https://doi.org/10.1109/TPAMI.2016.2644615

[84] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, pp. 1–14, April 2015. [Online]. Available: https://arxiv.org/abs/1409.1556

[85] 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. [Online]. Available: https://doi.org/10.1109/ISBI.2017.7950617

[86] 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*, vol. 2018, no. 15, pp. 199–1–199–8, January 2018, Burlingame, CA. [Online]. Available: https://doi.org/10.2352/ISSN.2470-1173.2018.15.COIMG-199

[87] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," *Proceedings of the International Conference on Medical Image Computing and Computer Assisted Intervention*, vol. 9351, pp. 234–241, October 2015, Munich, Germany. [Online]. Available: https://doi.org/10.1007/978-3-319-24574-4_28

[88] H. Chen, X. Qi, L. Yu, and P. A. Heng, "DCAN: Deep contour-aware networks for accurate gland segmentation," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2487–2496, June 2016, Las Vegas, NV. [Online]. Available: https://doi.org/10.1109/CVPR.2016.273

[89] S. E. A. Raza, L. Cheung, D. Epstein, S. Pelengaris, M. Khan, and N. M. Rajpoot, "MIMO-NET: A multi-input multi-output convolutional neural network for cell segmentation in fluorescence microscopy images," *Proceedings of the IEEE International Symposium on Biomedical Imaging*, pp. 337–340, April 2017, Melbourne, Australia. [Online]. Available: https://doi.org/10.1109/ISBI.2017.7950532

[90] O. Cicek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger, "3D U-Net: Learning dense volumetric segmentation from sparse annotation," *Proceedings of the International Conference on Medical Image Computing and Computer Assisted Intervention*, vol. 9901, pp. 424–432, October 2016, Athens, Greece. [Online]. Available: https://doi.org/10.1007/978-3-319-46723-8_49

[91] F. Milletari, N. Navab, and S. A. Ahmadi, "V-Net: Fully convolutional neural networks for volumetric medical image segmentation," *Proceedings of the IEEE 2016 Fourth International Conference on 3D Vision*, pp. 565–571, October 2016, Stanford, CA. [Online]. Available: https://doi.org/10.1109/3DV.2016.79

[92] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778, June 2016, Las Vegas, NV. [Online]. Available: https://doi.org/10.1109/CVPR.2016.90

[93] H. Chen, Q. Dou, L. Yu, J. Qin, and P. A. Heng, "VoxResNet: Deep voxelwise residual networks for brain segmentation from 3D MR images," *Neuroimage*, vol. 170, no. 1, pp. 446–455, April 2018. [Online]. Available: https://doi.org/10.1016/j.neuroimage.2017.04.041

[94] X. Zhang, Y. Fu, A. Zang, L. Sigal, and G. Agam, "Learning classifiers from synthetic data using a multichannel autoencoder," *arXiv preprint arXiv:1503.03163*, pp. 1–11, March 2015. [Online]. Available: https://arxiv.org/abs/1503.03163

[95] I. B. Barbosa, M. Cristani, B. Caputo, A. Rognhaugen, and T. Theoharis, "Looking beyond appearances: Synthetic training data for deep CNNs in re-identification," *Computer Vision and Image Understanding*, vol. 167, no. 1, pp. 50–62, February 2018. [Online]. Available: https://doi.org/10.1016/j.cviu.2017.12.002

[96] 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. [Online]. Available: https://doi.org/10.1109/CVPRW.2017.116

[97] L. A. Gatys, A. S. Ecker, and M. Bethge, "Texture synthesis using convolutional neural networks," *Proceedings of the Advances in Neural Information Processing Systems*, pp. 262–270, December 2015, Montreal, Canada. [Online]. Available: http://papers.nips.cc/paper/5633-texture-synthesis-using-convolutional-neural-networks

[98] L. A. Gatys, A. S. Ecker, and M. Bethge, "Image style transfer using convolutional neural networks," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2414–2423, June 2016, Las Vegas, NV. [Online]. Available: https://doi.org/10.1109/CVPR.2016.265

[99] J. Johnson, A. Alahi, and L. Fei-Fei, "Perceptual losses for real-time style transfer and super-resolution," *Proceedings of the European Conference on Computer Vision*, pp. 694–711, October 2016, Amsterdam, Netherlands. [Online]. Available: https://doi.org/10.1007/978-3-319-46475-6_43

[100] D. Ulyanov, V. Lebedev, A. Vedaldi, and V. S. Lempitsky, "Texture networks: Feed-forward synthesis of textures and stylized images," *Proceedings of the International Conference on Machine Learning*, vol. 48, pp. 1349–1357, June 2016, New York, NY. [Online]. Available: http://proceedings.mlr.press/v48/ulyanov16.html

[101] D. Ulyanov, A. Vedaldi, and V. S. Lempitsky, "Improved texture networks: Maximizing quality and diversity in feed-forward stylization and texture synthesis," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4105–4113, July 2017, Honolulu, HI. [Online]. Available: https://doi.org/10.1109/CVPR.2017.437

[102] 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. [Online]. Available: https://papers.nips.cc/paper/5423-generative-adversarial-nets

[103] A. Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," *International Conference on Learning Representations*, pp. 1–16, May 2016. [Online]. Available: https://arxiv.org/abs/1511.06434

[104] M. Arjovsky, S. Chintala, and L. Bottou, "Wasserstein GAN," *arXiv preprint arXiv:1701.07875*, pp. 1–32, December 2017. [Online]. Available: https://arxiv.org/abs/1701.07875

[105] D. Berthelot, T. Schumm, and L. Metz, "BEGAN: Boundary equilibrium generative adversarial networks," *arXiv preprint arXiv:1703.10717*, pp. 1–10, May 2017. [Online]. Available: https://arxiv.org/abs/1703.10717

[106] Y. Bengio, "Learning deep architectures for ai," *Foundations and trends® in Machine Learning*, vol. 2, no. 1, pp. 1–127, 2009. [Online]. Available: https://doi.org/10.1561/2200000006

[107] V. Dumoulin, I. Belghazi, B. Poole, O. Mastropietro, A. Lamb, M. Arjovsky, and A. Courville, "Adversarially learned inference," *International Conference on Learning Representations*, pp. 1–18, April 2017. [Online]. Available: https://openreview.net/forum?id=B1ElR4cgg

[108] J. Donahue, P. Krahenbuhl, and T. Darrell, "Adversarial feature learning," *International Conference on Learning Representations*, pp. 1–18, April 2017. [Online]. Available: https://openreview.net/forum?id=BJtNZAFgg

[109] A. B. L. Larsen, S. K. Sonderby, H. Larochelle, and O. Winther, "Autoencoding beyond pixels using a learned similarity metric," *Proceedings of the International Conference on Machine Learning*, vol. 48, pp. 1558–1566, June 2016, New York NY. [Online]. Available: http://proceedings.mlr.press/v48/larsen16.html

[110] P. Isola, J. Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 5967–5976, July 2017, Honolulu, HI. [Online]. Available: https://doi.org/10.1109/CVPR.2017.632

[111] M. Y. Liu and O. Tuzel, "Coupled generative adversarial networks," *Proceedings of the Advances in Neural Information Processing Systems*, pp. 469–477, December 2016, Barcelona, Spain. [Online]. Available: https://papers.nips.cc/paper/6544-coupled-generative-adversarial-networks

[112] A. Shrivastava, T. Pfister, O. Tuzel, J. Susskind, W. Wang, and R. Webb, "Learning from simulated and unsupervised images through adversarial training," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2242–2251, July 2017, Honolulu, HI. [Online]. Available: https://doi.org/10.1109/CVPR.2017.241

[113] 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.

[114] T. Kim, M. Cha, H. Kim, J. K. Lee, and J. Kim, "Learning to discover cross-domain relations with generative adversarial networks," *Proceedings of the International Conference on Machine Learning*, vol. 70, pp. 1857–1865, August 2017, Sydney, Australia. [Online]. Available: http://proceedings.mlr.press/v70/kim17a.html

[115] Y. Huo, Z. Xu, S. Bao, A. Assad, R. G. Abramson, and B. A. Landman, "Adversarial synthesis learning enables segmentation without target modality ground truth," *Proceedings of the IEEE International Symposium on Biomedical Imaging*, pp. 1217–1220, April 2018, Washington, DC. [Online]. Available: https://doi.org/10.1109/ISBI.2018.8363790

[116] 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 Conference on Computer Vision and Pattern Recognition Workshop*, pp. 2302–2310, June 2018, Salt Lake City, UT. [Online]. Available: https://doi.org/10.1109/CVPRW.2018.00298

[117] S. Han, S. Lee, C. Fu, P. Salama, K. W. Dunn, and E. J. Delp, "Nuclei counting in microscopy images with three dimensional generative adversarial networks," *Proceedings of the SPIE Conference on Medical Imaging*, pp. 109 492Y–1–11, February 2019, San Diego, CA. [Online]. Available: https://doi.org/10.1117/12.2512591

[118] S. Lee, S. Han, P. Salama, K. W. Dunn, and E. J. Delp, "Three dimensional blind image deconvolution for fluorescence microscopy using generative adversarial networks," *Proceedings of the IEEE International Symposium on Biomedical Imaging*, pp. 538–542, April 2019, Venice, Italy.

[119] Z. Zhang, L. Yang, and Y. Zheng, "Translating and segmenting multimodal medical volumes with cycle- and shape-consistency generative adversarial network," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 9242–9251, June 2018, Salt Lake City, UT. [Online]. Available: https://doi.org/10.1109/CVPR.2018.00963

[120] A. Bansal, S. Ma, D. Ramanan, and Y. Sheikh, "Recycle-gan: Unsupervised video retargeting," *Proceedings of the European Conference on Computer Vision*, pp. 122–138, September 2018, Munich, Germany. [Online]. Available: https://doi.org/10.1007/978-3-030-01228-1_8

[121] T. C. Wang, M. Y. Liu, J. Y. Zhu, G. Liu, A. Tao, J. Kautz, and B. Catanzaro, "Video-to-video synthesis," *Proceedings of the Advances in Neural Information Processing Systems*, pp. 1144–1156, December 2018, Washington, DC. [Online]. Available: https://papers.nips.cc/paper/7391-video-to-video-synthesis

[122] C. Vonesch, F. Aguet, J. Vonesch, and M. Unser, "The colored revolution of bioimaging," *IEEE Signal Processing Magazine*, vol. 23, no. 3, pp. 20–31, May 2006. [Online]. Available: https://doi.org/10.1109/MSP.2006.1628875

[123] S. Koho, E. Fazeli, J. E. Eriksson, and P. E. Hänninen, "Image quality ranking method for microscopy," *Scientific Reports*, vol. 6, no. 28962, pp. 1–14, July 2016. [Online]. Available: https://doi.org/10.1038/srep28962

[124] S. Yang and B. U. Lee, "Poisson-Gaussian noise reduction using the hidden Markov model in contourlet domain for fluorescence microscopy images," *PLOS ONE*, vol. 10, no. 9, pp. 1–19, September 2015. [Online]. Available: https://doi.org/10.1371/journal.pone.0136964

[125] D. Sage, L. Donati, F. Soulez, D. Fortun, G. Schmit, A. Seitz, R. Guiet, C. Vonesch, and M. Unser, "DeconvolutionLab2: An open-source software for deconvolution microscopy," *Methods*, vol. 115, no. 1, pp. 28–41, February 2017. [Online]. Available: https://doi.org/10.1016/j.ymeth.2016.12.015

[126] P. Sarder and A. Nehorai, "Deconvolution methods for 3-D fluorescence microscopy images," *IEEE Signal Processing Magazine*, vol. 23, no. 3, pp. 32–45, May 2006. [Online]. Available: https://doi.org/10.1109/MSP.2006.1628876

[127] D. Agard and J. W. Sedat, "Three-dimensional architecture of a polytene nucleus," *Nature*, vol. 302, no. 5910, pp. 676–681, April 1983. [Online]. Available: https://doi.org/10.1038/302676a0

[128] W. H. Richardson, "Bayesian-based iterative method of image restoration," *Journal of the Optical Society of America*, vol. 62, no. 1, pp. 55–59, January 1972. [Online]. Available: https://doi.org/10.1364/JOSA.62.000055

[129] L. B. Lucy, "An iterative technique for the rectification of observed distributions," *The Astronomical Journal*, vol. 79, no. 6, pp. 745–754, June 1974. [Online]. Available: https://doi.org/10.1086/111605

[130] N. Dey, L. Blanc-Feraud, C. Zimmer, P. Roux, Z. Kam, J.-C. Olivo-Marin, and J. Zerubia, "Richardson-Lucy algorithm with total variation regularization for 3D confocal microscope deconvolution," *Microscopy Research and Technique*, vol. 69, no. 4, pp. 260–266, April 2006. [Online]. Available: https://doi.org/10.1002/jemt.20294

[131] A. Shajkofci and M. Liebling, "Semi-blind spatially-variant deconvolution in optical microscopy with local point spread function estimation by use of convolutional neural networks," *Proceedings of the IEEE International Conference on Image Processing*, pp. 3818–3822, October 2018, Athens, Greece. [Online]. Available: https://doi.org/10.1109/ICIP.2018.8451736

[132] D. A. Fish, A. M. Brinicombe, E. R. Pike, and J. G. Walker, "Blind deconvolution by means of the Richardson-Lucy algorithm," *Journal of the Optical Society of America A*, vol. 12, no. 1, pp. 58–65, January 1995. [Online]. Available: https://doi.org/10.1364/JOSAA.12.000058

[133] F. Soulez, L. Denis, Y. Tourneur, and E. Thiebaut, "Blind deconvolution of 3D data in wide field fluorescence microscopy," *Proceedings of the IEEE International Symposium on Biomedical Imaging*, pp. 1735–1738, May 2012, Barcelona, Spain. [Online]. Available: https://doi.org/10.1109/ISBI.2012.6235915

[134] T. Kenig, Z. Kam, and A. Feuer, "Blind image deconvolution using machine learning for three-dimensional microscopy," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 12, pp. 2191–2204, December 2010. [Online]. Available: https://doi.org/10.1109/TPAMI.2010.45

[135] F. Soulez, "A learn 2D, apply 3D method for 3D deconvolution microscopy," *Proceedings of the IEEE International Symposium on Biomedical Imaging*, pp. 1075–1078, April 2014, Beijing, China. [Online]. Available: https://doi.org/10.1109/ISBI.2014.6868060

[136] F. Luisier, C. Vonesch, T. Blu, and M. Unser, "Fast interscale wavelet denoising of Poisson-corrupted images," *Signal Processing*, vol. 90, no. 2, pp. 415–427, February 2010. [Online]. Available: https://doi.org/10.1016/j.sigpro.2009.07.009

[137] J. Li, F. Luisier, and T. Blu, "PURE-LET deconvolution of 3D fluorescence microscopy images," *Proceedings of the IEEE International Symposium on Biomedical Imaging*, pp. 723–727, April 2017, Melbourne, Australia. [Online]. Available: https://doi.org/10.1109/ISBI.2017.7950621

[138] L. Azzari and A. Foi, "Variance stabilization in Poisson image deblurring," *Proceedings of the IEEE International Symposium on Biomedical Imaging*, pp. 728–731, April 2017, Melbourne, Australia. [Online]. Available: https://doi.org/10.1109/ISBI.2017.7950622

[139] J. Xie, L. Xu, and E. Chen, "Image denoising and inpainting with deep neural networks," *Proceedings of the Advances in Neural Information Processing Systems*, pp. 341–349, December 2012, Lake Tahoe, NV. [Online]. Available: https://papers.nips.cc/paper/4686-image-denoising-and-inpainting-with-deep-neural-networks

[140] L. Xu, J. S. Ren, C. Liu, and J. Jia, "Deep convolutional neural network for image deconvolution," *Proceedings of the Advances in Neural Information Processing Systems*, pp. 1790–1798, December 2014, Montreal, Canada. [Online]. Available: https://papers.nips.cc/paper/5485-deep-convolutional-neural-network-for-image-deconvolution

[141] R. Yan and L. Shao, "Blind image blur estimation via deep learning," *IEEE Transactions on Image Processing*, vol. 25, no. 4, pp. 1910–1921, April 2016. [Online]. Available: https://doi.org/10.1109/TIP.2016.2535273

[142] M. Weigert, L. Royer, F. Jug, and G. Myers, "Isotropic reconstruction of 3D fluorescence microscopy images using convolutional neural networks," *Proceedings of the International Conference on Medical Image Computing and Computer Assisted Intervention*, pp. 126–134, September 2017, Quebec, Canada. [Online]. Available: https://doi.org/10.1007/978-3-319-66185-8_15

[143] X. Yi, E. Walia, and P. Babyn, "Generative adversarial network in medical imaging: A review," *arXiv preprint arXiv:1809.07294*, pp. 1–20, September 2018. [Online]. Available: https://arxiv.org/abs/1809.07294

[144] E. Kang, H. J. Koo, D. H. Yang, J. B. Seo, and J. C. Ye, "Cycle consistent adversarial denoising network for multiphase coronary CT angiography," *arXiv preprint arXiv:1806.09748*, pp. 1–9, June 2018. [Online]. Available: https://arxiv.org/abs/1806.09748

[145] R. Zhang, P. Isola, A. A. Efros, E. Shechtman, and O. Wang, "The unreasonable effectiveness of deep features as a perceptual metric," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 586–595, June 2018, Salt Lake City, UT. [Online]. Available: https://doi.org/10.1109/CVPR.2018.00068

[146] A. Mittal, A. K. Moorthy, and A. C. Bovik, "No-reference image quality assessment in the spatial domain," *IEEE Transactions on Image Processing*, vol. 21, no. 12, pp. 4695–4708, December 2012. [Online]. Available: https://doi.org/10.1109/TIP.2012.2214050

[147] L. Liu, Y. Hua, Q. Zhao, H. Huang, and A. C. Bovik, "Blind image quality assessment by relative gradient statistics and adaboosting neural network," *Signal Processing: Image Communication*, vol. 40, no. 1, pp. 1–15, January 2016. [Online]. Available: https://doi.org/10.1016/j.image.2015.10.005

[148] L. S. Chow and H. Rajagopal, "Modified-BRISQUE as no reference image quality assessment for structural MR images," *Magnetic Resonance Imaging*, vol. 43, no. 1, pp. 74–87, November 2017. [Online]. Available: https://doi.org/10.1016/j.mri.2017.07.016

[149] S. J. Yang, M. Berndl, D. M. Ando, M. Barch, A. Narayanaswamy, E. Christiansen, S. Hoyer, C. Roat, J. Hung, C. T. Rueden, A. Shankar, S. Finkbeiner, and P. Nelson, "Assessing microscope image focus quality with deep learning," *BMC Bioinformatics*, vol. 19, no. 1, pp. 77–1–9, March 2018. [Online]. Available: https://doi.org/10.1186/s12859-018-2087-4

[150] J. Canny, "A computational approach to edge detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8, no. 6, pp. 679–698, November 1986. [Online]. Available: https://doi.org/10.1109/TPAMI.1986.4767851

[151] I. Daubechies, *Ten lectures on wavelets*. Philadelphia, PA: Society for Industrial and Applied Mathematics, 1992. [Online]. Available: https://doi.org/10.1137/1.9781611970104

[152] L. Lam, S.-W. Lee, and C. Y. Suen, "Thinning methodologies-a comprehensive survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 9, pp. 869–885, September 1992. [Online]. Available: https://doi.org/10.1109/34.161346

[153] C. Lantuejoul and F. Maisonneuve, "Geodesic methods in quantitative image analysis," *Pattern Recognition*, vol. 17, no. 2, pp. 177–187, 1984. [Online]. Available: https://doi.org/10.1016/0031-3203(84)90057-8

[154] A. Fitzgibbon, M. Pilu, and R. B. Fisher, "Direct least square fitting of ellipses," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no. 5, pp. 476–480, May 1999. [Online]. Available: https://doi.org/10.1109/34.765658

[155] R. Malladi, J. A. Sethian, and B. C. Vemuri, "Shape modeling with front propagation: A level set approach," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17, no. 2, pp. 158–175, February 1995. [Online]. Available: https://doi.org/10.1109/34.368173

[156] S. C. Zhu and A. Yuille, "Region competition: Unifying snakes, region growing, and Bayes/MDL for multi-band image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 18, no. 9, pp. 884–900, September 1996. [Online]. Available: https://doi.org/10.1109/ICCV.1995.466909

[157] H.-K. Zhao, T. Chan, B. Merriman, and S. Osher, "A variational level set approach to multiphase motion," *Journal of Computational Physics*, vol. 127, no. 1, pp. 179–195, August 1996. [Online]. Available: https://doi.org/10.1006/jcph.1996.0167

[158] Y. Zhang, B. J. Matuszewski, L. Shark, and C. J. Moore, "Medical image segmentation using new hybrid level-set method," *Proceedings of the IEEE International Conference on BioMedical Visualization*, pp. 71–76, July 2008, London, UK. [Online]. Available: https://doi.org/10.1109/MediVis.2008.12

[159] A. du Chene, C. Min, and F. Gibou, "Second-order accurate computation of curvatures in a level set framework using novel high-order reinitialization schemes," *Journal of Scientific Computing*, vol. 35, no. 2, pp. 114–131, June 2008. [Online]. Available: https://doi.org/10.1007/s10915-007-9177-1

[160] C. Li, C. Kao, J. C. Gore, and Z. Ding, "Minimization of region-scalable fitting energy for image segmentation," *IEEE Transactions on Image Processing*, vol. 17, no. 10, pp. 1940–1949, October 2008. [Online]. Available: https://doi.org/10.1109/TIP.2008.2002304

[161] J. L. Clendenon, C. L. Phillips, R. M. Sandoval, S. Fang, and K. W. Dunn, "Voxx: A PC-based, near real-time volume rendering system for biological microscopy," *American Journal of Physiology-Cell Physiology*, vol. 282, no. 1, pp. C213–C218, January 2002. [Online]. Available: https://doi.org/10.1152/ajpcell.2002.282.1.C213

[162] K. Perlin, "Improving noise," *ACM Transactions on Graphics*, vol. 21, no. 3, pp. 681–682, July 2002. [Online]. Available: http://doi.acm.org/10.1145/566654.566636

[163] K. S. Lorenz, P. Salama, K. W. Dunn, and E. J. Delp, "Digital correction of motion artefacts in microscopy image sequences collected from living animals using rigid and nonrigid registration," *Journal of Microscopy*, vol. 245, no. 2, pp. 148–160, February 2012. [Online]. Available: https://doi.org/10.1111/j.1365-2818.2011.03557.x

[164] 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. [Online]. Available: http://proceedings.mlr.press/v37/ioffe15.html

[165] R. Collobert, K. Kavukcuoglu, and C. Farabet, "Torch7: A Matlab-like environment for machine learning," *Proceedings of the BigLearn workshop at the Neural Information Processing Systems*, pp. 1–6, December 2011, Granada, Spain. [Online]. Available: https://citeseerx.ist.psu.edu/viewdoc/similar?doi=10.1.1.231.4195&type=sc

[166] K. Sirinukunwattana, J. P. W. Pluim, H. Chen, X. Qi, P. A. Heng, Y. B. Guo, L. Y. Wang, B. J. Matuszewski, E. Bruni, U. Sanchez, A. Bohm, O. Ronneberger, B. B. Cheikh, D. Racoceanu, P. Kainz, M. Pfeiffer, M. Urschler, D. R. J. Snead, and N. M. Rajpoot, "Gland segmentation in colon histology images: The glas challenge contest," *Medical Image Analysis*, vol. 35, no. 1, pp. 489–502, January 2017. [Online]. Available: https://doi.org/10.1016/j.media.2016.08.008

[167] L. R. Dice, "Measures of the amount of ecologic association between species," *Ecology*, vol. 26, no. 3, pp. 297–302, July 1945. [Online]. Available: https://doi.org/10.2307/1932409

[168] D. P. Huttenlocher, G. A. Klanderman, and W. J. Rucklidge, "Comparing images using the Hausdorff distance," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 15, no. 9, pp. 850–863, September 1993. [Online]. Available: https://doi.org/10.1109/34.232073

[169] G. Hamarneh and P. Jassi, "VascuSynth: Simulating vascular trees for generating volumetric image data with ground truth segmentation and tree analysis," *Computerized Medical Imaging and Graphics*, vol. 34, no. 8, pp. 605–616, December 2010. [Online]. Available: https://doi.org/10.1016/j.compmedimag.2010.06.002

[170] P. Jassi and G. Hamarneh, "VascuSynth: Vascular tree synthesis software," *Insight Journal*, pp. 1–12, April 2011. [Online]. Available: https://doi.org/10380/3260

[171] V. N. Kovacheva, D. Snead, and N. M. Rajpoot, "A model of the spatial tumour heterogeneity in colorectal adenocarcinoma tissue," *BMC Bioinformatics*, vol. 17, no. 1, pp. 255–1–16, June 2016. [Online]. Available: https://doi.org/10.1186/s12859-016-1126-2

[172] K. Pearson, "The problem of the random walk," *Nature*, vol. 72, no. 1867, p. 342, August 1905. [Online]. Available: https://doi.org/10.1038/072342a0

[173] H. R. Roth, L. Lu, A. Seff, K. M. Cherry, J. Hoffman, S. Wang, J. Liu, E. Turkbey, and R. M. Summers, "A new 2.5D representation for lymph node detection using random sets of deep convolutional neural network observations," *Proceedings of the International Conference on Medical Image Computing and Computer Assisted Intervention*, pp. 520–527, September 2014, Boston, MA. [Online]. Available: https://doi.org/10.1007/978-3-319-10404-1_65

[174] S. Lee and D. Kim, "Background subtraction using the factored 3-way restricted Boltzmann machines," *arXiv preprint arXiv:1802.01522*, pp. 1–10, February 2018. [Online]. Available: https://arxiv.org/abs/1802.01522

[175] D. P. Kingma and J. L. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, pp. 1–15, January 2017. [Online]. Available: https://arxiv.org/abs/1412.6980

[176] J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J.-Y. Tinevez, D. J. White, V. Hartenstein, K. Eliceiri, P. Tomancak, and A. Cardona, "Fiji: An open-source platform for biological-image analysis," *Nature Methods*, vol. 9, no. 7, pp. 676–682, July 2012. [Online]. Available: https://doi.org/10.1038/nmeth.2019

[177] F. de Chaumont, S. Dallongeville, N. Chenouard, N. Herve, S. Pop, T. Provoost, V. Meas-Yedid, P. Pankajakshan, T. Lecomte, Y. L. Montagner, T. Lagache, A. Dufour, and J.-. C. Olivo-Marin, "Icy: An open bioimage informatics platform for extended reproducible research," *Nature Methods*, vol. 9, no. 7, pp. 690–696, July 2012. [Online]. Available: https://doi.org/10.1038/nmeth.2075

[178] K. U. Barthel, "3D-data representation with ImageJ," *ImageJ User and Developer Conference*, pp. 1–4, May 2006.

[179] A. E. Carpenter, T. R. Jones, M. R. Lamprecht, C. Clarke, I. H. Kang, O. Friman, D. A. Guertin, J. H. Chang, R. A. Lindquist, J. Moffat, P. Golland, and D. M. Sabatini, "Cellprofiler: Image analysis software for identifying and quantifying cell phenotypes," *Genome Biology*, vol. 7, no. 10, pp. R100–1–11, October 2006. [Online]. Available: https://doi.org/10.1186/gb-2006-7-10-r100

[180] C. McQuin, A. Goodman, V. Chernyshev, L. Kamentsky, B. A. Cimini, K. W. Karhohs, M. Doan, L. Ding, S. M. Rafelski, D. Thirstrup, W. Wiegraebe, S. Singh, T. Becker, J. C. Caicedo, and A. E. Carpenter, "Cellprofiler 3.0: Next-generation image processing for biology," *PLOS Biology*, vol. 16, no. 7, pp. 1–17, July 2018. [Online]. Available: https://doi.org/10.1371/journal.pbio.2005970

[181] B. Schmid, J. Schindelin, A. Cardona, M. Longair, and M. Heisenberg, "A high-level 3D visualization API for Java and ImageJ," *BMC Bioinformatics*, vol. 11, no. 274, pp. 1–7, May 2010. [Online]. Available: https://doi.org/10.1186/1471-2105-11-274

[182] D. Shreiner, G. Sellers, J. M. Kessenich, and B. M. Licea-Kane, *OpenGL programming guide: The official guide to learning OpenGL, Version 4.3*, 8th ed. Upper Saddle River, NJ: Addison-Wesley Professional, 2013. [Online]. Available: https://www.pearson.com/us/higher-education/product/Shreiner-Open-GL-Programming-Guide-The-Official-Guide-to-Learning-Open-GL-Version-4-9780321773036.html

[183] J. M. Noguera and J. R. Jiménez, "Visualization of very large 3D volumes on mobile devices and WebGL," *International Conference on Computer Graphics, Visualization and Computer Vision*, pp. 105–112, June 2012, Pilsen, Czech Republic. [Online]. Available: https://gggj.ujaen.es/publications/details/62

[184] D. Cantor and B. Jones, *WebGL beginner's guide*, 1st ed. Birmingham, UK: Packt Publishing, 2012. [Online]. Available: https://www.packtpub.com/game-development/webgl-beginners-guide

[185] T. Parisi, *WebGL: Up and running*, 1st ed. Sebastopol, CA: O'Reilly Media, 2012. [Online]. Available: http://shop.oreilly.com/product/0636920024729.do

[186] T. Parisi, *Programming 3D applications with HTML5 and WebGL*, 1st ed. Sebastopol, CA: O'Reilly Media, 2014. [Online]. Available: http://shop.oreilly.com/product/0636920029205.do

[187] D. Vohl, D. G. Barnes, C. J. Fluke, G. Poudel, N. Georgiou-Karistianis, A. H. Hassan, Y. Benovitski, T. H. Wong, O. L. Kaluza, T. D. Nguyen, and C. P. Bonnington, "Large-scale comparative visualization of sets of multidimensional data," *PeerJ Computer Science*, vol. 2, no. 88, pp. 1–36, October 2016. [Online]. Available: https://doi.org/10.7717/peerj-cs.88

[188] "glMatrix." [Online]. Available: http://glmatrix.net

[189] "dat.GUI." [Online]. Available: https://github.com/dataarts/dat.gui

[190] J. Congote, A. Segura, L. Kabongo, A. Moreno, J. Posada, and O. Ruiz, "Interactive visualization of volumetric data with WebGL in real-time," *Proceedings of the ACM International Conference on 3D Web Technology*, pp. 137–146, June 2011, Paris, France. [Online]. Available: http://doi.acm.org/10.1145/2010425.2010449

[191] C. Glasbey, G. van der Heijden, V. F. K. Toh, and A. Gray, "Colour displays for categorical images," *Color Research & Application*, vol. 32, no. 4, pp. 304–309, June 2007. [Online]. Available: https://doi.org/10.1002/col.20327

VITA

VITA

Soonam Lee was born in Seoul, Korea. He received the B.S. degree in Electrical and Computer Engineering (Summa Cum Laude) and another B.S. degree in Information Technology Management from Hanyang University, Seoul, Korea in 2008. He received the M.S. degree in Electrical Engineering:Systems and another M.S. degree in Mathematics in University of Michigan, Ann Arbor, Michigan in 2012.

Mr. Lee joined the Ph.D. program at the School of Electrical and Computer Engineering, Purdue University, West Lafayette, Indiana in August 2012. He has worked as a Research Assistant and a Teaching Assistant in the Video and Image Processing Laboratory (VIPER) under the supervision of Professor Edward J. Delp since January 2013. While in the graduate program, he worked on projects sponsored by the National Institutes of Health (NIH).

He was a research intern at the vision algorithm research team in Delphi, Agoura Hills, CA and at the machine learning team of the new business initiative (NBI) group in Intel Corporation, Hillsboro, OR in the summer of 2014 and 2015, respectively. He was also a research intern at the flat panel display team in Apple Inc., Cupertino, CA in the summer of 2016. His research interests includes image processing, computer vision, deep learning, and medical imaging. He is a student member of the IEEE, the IEEE Signal Processing Society, the IS&T, and the Society for Industrial and Applied Mathematics (SIAM).