

Image Segmentation Techniques for Healthcare Systems

Lead Guest Editor: Orazio Gambino

Guest Editors: Cesare Valenti, Wellington P. dos Santos, Sergio Galdino,
and Vincenzo Conti





Image Segmentation Techniques for Healthcare Systems

Image Segmentation Techniques for Healthcare Systems

Lead Guest Editor: Orazio Gambino

Guest Editors: Cesare Valenti, Wellington P. dos Santos,
Sergio Galdino, and Vincenzo Conti



Copyright © 2019 Hindawi. All rights reserved.

This is a special issue published in “Journal of Healthcare Engineering.” All articles are open access articles distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Editorial Board

Saverio Affatato, Italy
Francesca Apollonio, Italy
Hasan Ayaz, USA
William Bertucci, France
Patrick Boissy, Canada
Niccolò Camarlinghi, Italy
Hongwei Chen, USA
Daniel H.K. Chow, Hong Kong
Gianluca Ciardelli, Italy
Elena De Momi, Italy
Costantino Del Gaudio, Italy
Daniel Espino, UK
Tiago H. Falk, Canada
Mostafa Fatemi, USA
Jesus Favela, Mexico
David Dagan Feng, Australia
Joseph Finkelstein, USA
Jesus Fontecha, Spain
Jean-Luc Gennisson, France
Luca Giancardo, USA
Antonio Gloria, Italy
Kheng-Lim Goh, Singapore
Pedro Gomes, Portugal
Carlos Gómez, Spain
Philippe Gorce, France
Sophia Z. Gu, Australia
Vincenzo Guarino, Italy

Valentina Hartwig, Italy
David Hewson, UK
Andreas H. Hielscher, USA
Ernesto Iadanza, Italy
Norio Iriguchi, Japan
Jingfeng Jiang, USA
Zhongwei Jiang, Japan
Rashed Karim, UK
Pasi A. Karjalainen, Finland
Chandan Karmakar, Australia
John S. Katsanis, Greece
Terry K.K. Koo, USA
Panagiotis Kosmas, UK
Michel Labrosse, Canada
Jui-Yang Lai, Taiwan
Xiang Li, USA
Feng-Huei Lin, Taiwan
Yuan-Pin Lin, Taiwan
Maria Lindén, Sweden
Dongfei Liu, Finland
Francisco Lopez-Valdes, Spain
Zufu Lu, Australia
Ilias Maglogiannis, Greece
Mehran Moazen, UK
Rafael Morales, Spain
David Moratal, Spain
Angkoon Phinyomark, Canada

Vincenzo Positano, Italy
Alessandro Ramalli, Italy
Alessandro Reali, Italy
Lei Ren, UK
Jose Joaquin Rieta, Spain
Emanuele Rizzuto, Italy
Sébastien Roth, France
Simo Saarakkala, Finland
Hélder A. Santos, Finland
Emiliano Schena, Italy
Maurizio Schmid, Italy
Jiann-Shing Shieh, Taiwan
Tiago H. Silva, Portugal
Redha Taiar, France
Jinshan Tang, USA
Vinoy Thomas, USA
Ioannis G. Tollis, Greece
Kazunori Uemura, Japan
Cesare F. Valenti, Italy
Uche Wejinya, USA
Ying Yang, UK
Haihong Zhang, Singapore
Hongbo Zhang, Finland
Ping Zhou, USA
Loredana Zollo, Italy

Contents

Image Segmentation Techniques for Healthcare Systems

Orazio Gambino , Vincenzo Conti , Sergio Galdino , Cesare Fabio Valenti ,
and Wellington Pinheiro dos Santos 
Editorial (2 pages), Article ID 2723419, Volume 2019 (2019)

Automatic Optic Disc Segmentation Based on Modified Local Image Fitting Model with Shape Prior Information

Yuan Gao , Xiaosheng Yu , Chengdong Wu , Wei Zhou , Xiaoliang Lei, and Yaoming Zhuang 
Research Article (10 pages), Article ID 2745183, Volume 2019 (2019)

Automatic Tissue Image Segmentation Based on Image Processing and Deep Learning

Zhenglun Kong , Ting Li , Junyi Luo, and Shengpu Xu 
Research Article (10 pages), Article ID 2912458, Volume 2019 (2019)

Cloud-Based Brain Magnetic Resonance Image Segmentation and Parcellation System for Individualized Prediction of Cognitive Worsening

Ryo Sakamoto , Christopher Marano, Michael I. Miller, Constantine G. Lyketsos , Yue Li ,
Susumu Mori, Kenichi Oishi , and Alzheimer's Disease Neuroimaging Initiative ADNI
Research Article (10 pages), Article ID 9507193, Volume 2019 (2019)

An Approach for Pulmonary Vascular Extraction from Chest CT Images

Wenjun Tan , Yue Yuan, Anning Chen, Lin Mao, Yuqian Ke, and Xinhui Lv
Research Article (11 pages), Article ID 9712970, Volume 2019 (2019)

Fully Convolutional DenseNet with Multiscale Context for Automated Breast Tumor Segmentation

Jinjin Hai, Kai Qiao, Jian Chen , Hongna Tan, Jingbo Xu, Lei Zeng , Dapeng Shi, and Bin Yan 
Research Article (11 pages), Article ID 8415485, Volume 2019 (2019)

Analysis of the Cluster Prominence Feature for Detecting Calcifications in Mammograms

Alejandra Cruz-Bernal, Martha M. Flores-Barranco, Dora L. Almanza-Ojeda , Sergio Ledesma ,
and Mario A. Ibarra-Manzano 
Research Article (11 pages), Article ID 2849567, Volume 2018 (2019)

Fully Automated Lumen Segmentation Method for Intracoronary Optical Coherence Tomography

Elżbieta Pociask , Krzysztof Piotr Malinowski , Magdalena Ślęzak , Joanna Jaworek-Korjakowska ,
Wojciech Wojakowski , and Tomasz Roleder 
Research Article (13 pages), Article ID 1414076, Volume 2018 (2019)

An Application for Skin Macules Characterization Based on a 3-Stage Image-Processing Algorithm for Patients with Diabetes

Cintha Lourdes Toledo Peral , Francisco José Ramos Becerril, Gabriel Vega Martínez,
Arturo Vera Hernández , Lorenzo Leija Salas, and Josefina Gutiérrez Martínez 
Research Article (13 pages), Article ID 9397105, Volume 2018 (2019)

Automated Region Extraction from Thermal Images for Peripheral Vascular Disease Monitoring

Jean Gauci , Owen Falzon , Cynthia Formosa , Alfred Gatt , Christian Ellul, Stephen Mizzi , Anabelle Mizzi , Cassandra Sturgeon Delia, Kevin Cassar, Nachiappan Chockalingam , and Kenneth P. Camilleri 

Research Article (14 pages), Article ID 5092064, Volume 2018 (2019)

Is Intensity Inhomogeneity Correction Useful for Classification of Breast Cancer in Sonograms Using Deep Neural Network?

Chia-Yen Lee , Guan-Lin Chen, Zhong-Xuan Zhang, Yi-Hong Chou, and Chih-Chung Hsu 

Research Article (10 pages), Article ID 8413403, Volume 2018 (2019)

Fully Automatic Segmentation and Three-Dimensional Reconstruction of the Liver in CT Images

ZhenZhou Wang , Cunshan Zhang, Ticao Jiao, MingLiang Gao, and Guofeng Zou

Research Article (10 pages), Article ID 6797102, Volume 2018 (2019)

A Computer-Aided Pipeline for Automatic Lung Cancer Classification on Computed Tomography Scans

Emre Dandil 

Research Article (12 pages), Article ID 9409267, Volume 2018 (2019)

An Improved Fuzzy Connectedness Method for Automatic Three-Dimensional Liver Vessel Segmentation in CT Images

Rui Zhang, Zhuhuang Zhou , Weiwei Wu , Chung-Chih Lin, Po-Hsiang Tsui , and Shuicai Wu 

Research Article (18 pages), Article ID 2376317, Volume 2018 (2019)

Pixel-Label-Based Segmentation of Cross-Sectional Brain MRI Using Simplified SegNet Architecture-Based CNN

Bijen Khagi and Goo-Rak Kwon 

Research Article (8 pages), Article ID 3640705, Volume 2018 (2019)

Rank-Two NMF Clustering for Glioblastoma Characterization

Aymen Bougacha , Ines Njeh, Jihene Boughariou, Omar Kammoun, Kheireddine Ben Mahfoudh, Mariem Dammak, Chokri Mhiri, and Ahmed Ben Hamida

Research Article (7 pages), Article ID 1048164, Volume 2018 (2019)

A Novel Method to Quantify Longitudinal Orthodontic Bone Changes with In Vivo Micro-CT Data

Chao Wang , Li Cao, Chongshi Yang , and Yubo Fan 

Research Article (8 pages), Article ID 1651097, Volume 2018 (2019)

Evaluation of Commonly Used Algorithms for Thyroid Ultrasound Images Segmentation and Improvement Using Machine Learning Approaches

Prabal Poudel , Alfredo Illanes , Debdoot Sheet, and Michael Friebe

Research Article (13 pages), Article ID 8087624, Volume 2018 (2019)

Editorial

Image Segmentation Techniques for Healthcare Systems

Orazio Gambino ¹, **Vincenzo Conti** ², **Sergio Galdino** ³, **Cesare Fabio Valenti** ⁴,
and Wellington Pinheiro dos Santos ⁵

¹Dept. of Engineering, University of Palermo, Palermo, Italy

²Faculty of Engineering and Architecture, University of Enna KORE, Enna, Italy

³Polytechnique School of Pernambuco, University of Pernambuco, Recife, Brazil

⁴Dept. of Mathematics and Informatics, University of Palermo, Palermo, Italy

⁵Dept. of Biomedical Engineering, Federal University of Pernambuco, Recife, Brazil

Correspondence should be addressed to Orazio Gambino; orazio.gambino@unipa.it

Received 16 January 2019; Accepted 16 January 2019; Published 2 April 2019

Copyright © 2019 Orazio Gambino et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The present special issue of the *Journal of Healthcare Engineering* collects articles written by researchers scattered around the world who belong to the academic and industrial environments. The papers of this special issue have been selected by a rigorous peer-reviewing process with the support of at least two reviewers per paper, along with the opinion written in the final decision by a component of the editorial staff. Different methods on biomedical image segmentation dedicated to healthcare systems have been developed regarding, for example, the fields of machine learning, deformable models, fuzzy models, and so on. Such methods have been applied on different biomedical image modalities (MRI, CT, mammograms, optical coherence tomography, and others) of various anatomical districts, such as the brain, thyroid, lung, and breast. J. Gauci et al. present an automatic approach to determine the temperature of the body's extremities by means of thermal images in diabetic patients. The approach is based on morphological operations, and geometric transformations are aimed at automatically extracting the required data from 44 predefined regions of interest. The method is tested on data from 395 participants. A correct extraction in around 90% of the images was achieved. A. Bougacha et al. investigate a novel classification method for 3D multimodal MRI glioblastoma tumor characterization. They propose the segmentation problem as a linear mixture model (LMM). First, they provide a nonnegative matrix M from every MRI slice in every segmentation process' step (matrix used as an input

for the first segmentation process to extract the edema region from T2 and FLAIR modalities); after that, they extract edema's region from T1c modality, generate the matrix M , and segment the necrosis, the enhanced tumor, and the nonenhanced tumor regions. In the segmentation process, they apply a rank-two NMF clustering. C. Wang et al. quantify the subregional alveolar bone changes during orthodontic tooth movement with a novel method. Orthodontic tooth movement (OTM) is the result of the region-specific bone modeling under a load. Quantification of this change in the alveolar bone around a tooth is a basic requirement to understand the mechanism of orthodontics. They have used 12 Sprague-Dawley (SD) rats as an orthodontic model, and one side of the first upper molar has been used to simulate OTM. The alveolar bone around the mesial root has been reconstructed from in vivo micro-CT images and separated from other parts of the alveolar bone with two semicylinder filters. The amount and rate of OTM, bone mineral density (BMD), and bone volume (BV) around the root have been calculated and compared at 5 time points. C. L. Toledo Peral et al. present an application for skin macules characterization and could be the background of a future diagnosis-assistance-tool for educational and preventive assistance technology purposes, based on a three-stage segmentation and characterization algorithm used to classify vascular, petechiae, trophic changes, and trauma macules from digital photographs of the lower limbs. First, in order to find the skin region, a logical multiplication is

performed on two skin masks obtained from color space transformations. Then, in order to locate the lesion region, illumination enhancement is performed using a chromatic model color space, followed by a principal component analysis grayscale transformation. Finally, characteristics of each type of macule are considered and classified; morphologic properties (area, axes, perimeter, and solidity), intensity properties, and a set of shade indices (red, green, blue, and brown) are proposed as a measure to obviate skin color differences among subjects. E. Pociask et al. present a method aimed at the automated lumen segmentation on optical coherence tomography images. The adopted approach is composed of a preprocessing phase (artifacts removal: speckle noise, circular rings, and guide wire) and morphological operations. The method is tested on 667 images of different patients from the Medical University of Silesia, and it has been compared with other ones by using objective measures. R. Zhang et al. propose an improved fuzzy connectedness (FC) method for three-dimensional (3D) liver vessel segmentation on computed tomography (CT) images. In particular, a novel method to define the fuzzy affinity function of FC is presented, and an improved filter based on adaptive sigmoid filtering is proposed. The proposed approach is evaluated in 40 cases of clinical CT volumetric images from public image repositories. Y. Gao et al. proposed a novel technique to perform the optic disc detection on retinal images. The elaboration pipeline involves an initial rough segmentation based on saliency detection and largest object selection to define the optic disc initial contour. Subsequently, a method based on a deformable model refines the result. The effectiveness of such an approach has been evaluated on a public retinal image repository. A. Cruz-Bernal et al. propose a method for the automated detection of malignant calcifications on mammography images. The approach is based on the analysis of the cluster prominence (cp) feature histogram. Indeed, the calcifications on the mammography are characterized by high occurrences in the histogram. Therefore, the Vandermonde interpolation is used to obtain a function which models the cp histogram, and a KNN classifier finalizes the method. Z. Kong et al. propose a method to segment MR brain images by means of a convolutional neural network. Initially, a wavelet technique is used to extract the contours of different tissues such as skull, cerebrospinal fluid (CSF), grey matter (GM), and white matter (WM). Therefore, a convolutional neural network refines the segmentation results. B. Khagi and G.-R. Kwon proposed a method to segment MR brain images using SegNet, a convolutional neural network. Such a CNN is trained by using pre-segmented MR brain images of the OASIS free dataset. Z. Z. Wang et al. described a simple, yet effective approach to segment three-dimensional livers in computed tomography imaging. This research field is of particular interest because the described task requires considerable experience by the clinician and it is subject to personal interpretation. The proposed method is based on multiple thresholds through slope difference distribution, on Gibbs energy minimization to reduce inhomogeneity and on proper mathematical morphology operations to refine the segmented components,

represented by their spline contours. This heuristic was fine tuned experimentally on a variety of real cases on public datasets. C.-Y. Lee et al. introduced a neural network approach to deal with sonograms in the case of breast cancer. Potential applications to reduce the amount of noise and intensity inhomogeneity are quite evident in this particular field. The proposed technique has been validated on two real datasets and comparing different network models. The images are passed to a stack denoised autoencoder-based classifier to enhance the contrast of the ultrasound signal. The extraction of the most promising features and the comparison with manually labeled data allowed to train the system, thus gaining a final accuracy equal to 85%. J. Hai et al. have addressed the problem of automatic breast cancer segmentation through fully convolutional networks, able to detect abstract features from the input data. Moreover, a non-standard multiresolution approach led to definition of a general technique, independent of the tumor size. Indeed, while the usual MRA analysis is efficient for compression purposes, the redundant methodology chosen by the authors retains the original resolution of the mammograms, and therefore, it is particularly suited for their automatic segmentation. Experimental results are presented on a vast proprietary dataset of real clinical cases. E. Dandil proposed a technique for the automatic detection and classification of lung cancer in computed tomography. A preliminary phase identifies candidate pulmonary nodules, which are segmented by self-organizing maps, investigated by principal component analysis and classified due to a probabilistic neural network. The computer-aided diagnostic system achieves high accuracy, sensitivity, and specificity on a publicly available dataset. W. Tan et al. have defined an interesting technique for the identification of pulmonary vascular structures in chest computed tomography. They proved that the region growing with maximum between-class variance segmentation is enough to select the regions of interest. The vascular components are therefore fully located through the fast marching method. Experimental comparisons were done against other well-known methods, and the accuracy of this technique is about 90%.

Conflicts of Interest

The editors declare that they have no conflicts of interest regarding the publication of this special issue.

Acknowledgments

The lead guest editor and his editorial staff express their gratefulness both to all the reviewers for their precious support and to the authors who decided to publish their works in this special issue.

*Orazio Gambino
Vincenzo Conti
Sergio Galdino
Cesare Fabio Valenti
Wellington Pinheiro dos Santos*

Research Article

Automatic Optic Disc Segmentation Based on Modified Local Image Fitting Model with Shape Prior Information

Yuan Gao ¹, Xiaosheng Yu ², Chengdong Wu ², Wei Zhou ¹, Xiaoliang Lei,¹
and Yaoming Zhuang ¹

¹College of Information Science and Engineering, Northeastern University, Shenyang, Liaoning 110819, China

²Faculty of Robot Science and Engineering, Northeastern University, Shenyang, Liaoning 110819, China

Correspondence should be addressed to Xiaosheng Yu; yuxiaosheng7@163.com and Chengdong Wu; wuchengdong_neu@163.com

Received 29 June 2018; Revised 18 October 2018; Accepted 15 November 2018; Published 14 March 2019

Academic Editor: Cesare Valenti

Copyright © 2019 Yuan Gao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Accurate optic disc (OD) detection is an essential yet vital step for retinal disease diagnosis. In the paper, an approach for segmenting OD boundary without manpower named full-automatic double boundary extraction is designed. There are two main advantages in it. (1) Since the performances and the computational cost produced by iterations of contour evolution of active contour models- (ACM-) based approaches greatly depend on the initialization, this paper proposes an effective and adaptive initial level set contour extraction approach using saliency detection and threshold techniques. (2) In order to handle unreliable information generated by intensity in abnormal retinal images caused by diseases, a modified LIF approach is presented by incorporating the shape prior information into LIF. We test the effectiveness of the proposed approach on a publicly available DIARETDB0 database. Experimental results demonstrate that our approach outperforms well-known approaches in terms of the average overlapping ratio and accuracy rate.

1. Introduction

Optic disc (OD) is a bright yellowish approximately circular or oval-shaped object in the retinal images [1], as shown in Figure 1.

Accurate OD localization and segmentation play an important role in retinal image analysis and eye diseases diagnosis. For instance, the localization of the OD is a crucial step for fovea detection, vessel tracking, measurement, and automated diabetic retinopathy (DR) screening [2]. Meanwhile, the segmentation of the OD can be used for diagnosing other diseases including glaucoma, papilledema, hypertensive retinopathy, and neovascularization of the disc (NVD) [3, 4]. However, in many real applications, there are some challenging problems for OD segmentation due to the complex OD appearance caused by some anomalies, such as myelinated nerve fibers, peripapillary atrophy (PPA), blood vessels covered, and poor image quality. Hence, many scholars have been proposing a series of approaches to improve the precision of

OD boundary extraction. These approaches can be divided into four categories including classification-based [5–9], template-based matching [10–17], morphology-based [18–20], active contour models- (ACM-) based approaches [15, 21–24].

Plenty of classification-based OD boundary extraction methods have been presented by Cheng et al. [5], Dutta et al. [6], Tan et al. [7], and Zhou et al. [8, 9]. They utilized image pixel-level features or superpixel-level features extracted from retinal fundus images to segment OD. However, these approaches are easy to be influenced by sample size. Namely, the segmentation results of OD have a larger bias if there is only a small amount of training data. Besides, it is also time consuming when dealing with a large amount of training data.

Template-based matching methods consider the shape prior information of the OD, i.e., the circular or elliptical shape, to match the edge maps extracted from retinal fundus images [10–17]. However, these methods always fail to detect the ODs with varied shapes.

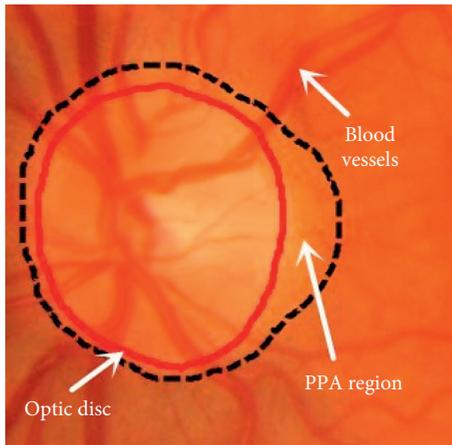


FIGURE 1: Major structures of the optic disc. Red line: the optic disc boundary.

Some morphology techniques are used to extract OD boundary, e.g., Reza et al. [18] and Welfer et al. [19]. In these approaches, the shape and bright of the OD are modeled by some morphology techniques. Nevertheless, the main disadvantage of these approaches is that bright lesions can affect their performance.

Srivastava et al. [20] applied a deep neural network composed of (unsupervised) stacked autoencoders followed by a supervised layer to distinguish OD from retinal fundus images. But it cannot deal well with the problem when the PPA is very similar to the OD.

Compared with the aforementioned approaches, ACM will obtain an excellent OD segmentation result due to the combination of the profound mathematical properties and prior knowledge of the OD. Hence, ACM-based approaches have become the most promising technique to detect the OD boundary [25]. Lee and Brady [21] firstly proposed a gradient vector flow (GVF) base active contour model for extracting the optic disc boundary with a fixed size initial contour followed by reducing the effect produced by high gradient at vessel locations. Mendels [22] presented a novel active contour approach using the gradient vector-flow-driven contour initialized manually to determine the OD boundary after preprocessing the image based on local minima detection and morphological filtering. A modified version of the conventional level-set method proposed by Wong et al. [15] is used to obtain the OD boundary with a constant scale initial contour from the red channel, and the contour is subsequently smoothed by strictly fitting an ellipse. Yu et al. [23] applied a fast, hybrid level set model, in which the deformable contour is driven by the local edge vector and the region information converging to the true optic disc boundary based on fixed size initial contour determined by experience. A variational-level set deformable model designed by Esmaili et al. [24] has higher convergence property and better computational efficiency compared with other segmentation active contour models when extracting the OD boundary with an empirical estimation initial contour around the detected OD center. These ACM-based methods can accurately segment ODs with strong boundary, but they

are always influenced by intensity inhomogeneities and blood vessels occlusion which are highly sensitive to interferences around the boundary, especially for bright lesions adjacent to the boundary of ODs, reducing their performance.

Seen from the above-mentioned OD detection methods, although the exiting ACM-based approaches can achieve better performance than classification-based approaches [5–9], template-based matching approaches [10–17] and morphology-based [18–20], most of ACM [15, 21–24] evolving the contour using the imprecise initial contour which is labeled by hand or is set based on fixed size. It not only reduces the performance for ACM but also generates the expensive computational cost. Besides, these ACM-based methods are misguided by unreliable information generated by intensity for extreme situation in abnormal retinal images caused by diseases, e.g., blurry OD boundary, bright peripapillary atrophy interference, and thick blood vessel coverage. They also need to remedy the insufficient information lost through image pre-processing which has been changed along with the different segmentation methods, making the key information lost, and have a complex operation. To address these issues, this paper proposes a novel approach by combining the local image fitting energy and shape prior information to extract OD boundary. The main contributions are as follows: (1) an automatic and robust adaptive initial level set contour extraction method by combining saliency detection and threshold techniques is designed to achieve the optimized contour evolution. (2) A novel ACM-based approach named local image fitting model with oval-shaped constraint (LIFO) is presented, which integrates the model with oval-shaped constraint into a united framework remedying the deficiency of only considering the information of intensity.

2. Methods

2.1. Optic Disc Localization. In this paper, we use our previous work [26] to locate the OD. In [26], a series of OD candidates can be firstly extracted using morphological opening by reconstruction. Then, a set of features are used to distinguish the true optic disc from the nonoptic disc candidates (for more details, refer [26]).

2.2. Optic Disc Segmentation

2.2.1. Rough Boundary Extraction of the OD. Based on the cropped region of interest around optic disc, we can further extract the optic disc boundaries. Since the contour initialization is the basic step to initialize the proposed active contour model, we propose a novel and robust contour initialization approach by combining saliency detection and threshold techniques together in this paper. The details are as follows.

Since the optic disc region is usually of a brighter pallor than the surrounding retinal areas, it can be regarded as a salient objective in retinal fundus images. Recently, inspired by saliency detection technique which aims at finding out the most important part of an image, we adopt a cellular (i.e., superpixel) automata-based saliency detection approach [27] by taking both global color and spatial distance matrices into consideration to contour initialization. First, cellular automata-based

saliency detection approach [27] is done on the tailored image. Figure 2(a) is the obtained saliency map in which the corresponding output saliency value of each superpixel is continuous between 0 and 1, as shown in Figure 2(b). Then, a mean filter is found to be a good choice [5] which is then applied on the saliency map to achieve smoothed map values, as shown in Figure 2(c). Next, the smoothed map values are then used to acquire the binary decisions for all the pixels with a threshold. In our experiment, we obtain the threshold by Otsu's thresholding and assign 1 and 0 to optic disc and nonoptic disc. After we obtain binary decisions for all the pixels, the values with 1 are regarded as object (optic disc) and 0 as background. Finally, the largest connected object (i.e., the connected region with the largest number of the pixels) can be obtained through morphological operation, as shown in Figure 2(d). And its boundary is used as the raw estimation of the optic disc, i.e., the optic disc initial contour in green, as shown in Figure 2(e).

2.2.2. Accurate Boundary Curve Extraction. Considering the intensity inhomogeneity is a frequently occurring phenomenon in the optic disc region [28]; the optic disc boundary extracted by general segmentation methods is usually inaccurate due to intensity inhomogeneity caused by imperfection of image devices or illumination variations. In order to deal with this problem, the local image fitting (LIF) model presented by Zhang et al. [28] is introduced; it defines local image fitting energy in a variational formulation which incorporates local intensity information into the active contour model. The LIF model can be described as follows:

$$E^{\text{LIF}} = \frac{1}{2} \int_{\Omega} |I - I^{\text{LFI}}|^2 dx, \quad (1)$$

where

$$I^{\text{LFI}} = m_1 H(\phi) + m_2 (1 - H(\phi)), \quad (2)$$

$$m_1 = \text{mean}(I \in (\{x \in \Omega | \phi(x) < 0\} \cap W(x))), \quad (3)$$

$$m_2 = \text{mean}(I \in (\{x \in \Omega | \phi(x) > 0\} \cap W(x))), \quad (4)$$

where I denotes an input image; I^{LFI} is a local fitted image (LFI) formulation, m_1 and m_2 are, respectively, defined as local mean near the point x described by equations (3) and (4). x is the variable to express the location information of pixel for global, Ω is the image domain, ϕ is a level set function, $H(\phi)$ is the Heaviside function, and $W(x)$ is a rectangular window function defined in [28].

Considering that the fundamental anatomical structure of the OD, e.g., it is a bright approximately circular or elliptic region, we can regard the anatomical structure as a shape prior constraint and take it into our model. In this paper, we incorporate both the smoothing item and an oval-disc prior constraint into LIF model, and the novel model named local image fitting model with oval-shaped constraint (LIFO) is proposed for OD boundary extraction. The model can remedy insufficiency of LIF, such as the LIF model will fail to extract the OD boundary with some blood vessels as shown in Figure 3(b). Seen from the result in Figure 3(c), the novel model overcomes the influence of blood vessels and intensity inhomogeneities achieving a precise OD boundary extraction of Figure 3(a).

Seen from the above results, it is necessary to introduce the smoothing item and shape prior information into LIF model aiming to acquire a whole boundary of the OD. They can be formulated as follows:

$$E^{\text{prior}} = E^{\text{smooth}} + E^{\text{ellipse}}, \quad (5)$$

$$E^{\text{smooth}} = \nu \int_{\Omega} \delta(\phi) |\nabla \phi| dx, \quad (6)$$

$$E^{\text{ellipse}} = \frac{1}{2} \alpha \int_{\Omega} (H(\phi) - H(\phi_e))^2 dx, \quad (7)$$

where

$$\phi_e = 1 - \sqrt{\frac{[(x_i - x_e) \cos \theta_e + (y_i - y_e) \sin \theta_e]^2}{a_e^2} + \frac{[-(x_i - x_e) \sin \theta_e + (y_i - y_e) \cos \theta_e]^2}{b_e^2}}, \quad (8)$$

where ∇ is the gradient operator; $\delta(\phi)$ is the smooth Dirac function; x_i and y_i are, respectively, x -coordinate and y -coordinate for global pixel information x ; x_e and y_e are oval center coordinates; θ_e is the angle of rotation; a_e denotes scaling factors of semimajor axis length; and b_e is defined as semiminor axis length. ϕ_e is the level set based on ellipse shape. Both of them are constantly changed with the curve evolution. In fact, the purpose for calculating equation (5) is to acquire the level set ϕ which is similar to ϕ_e . The novel model named LIFO can be obtained by combining equations (1) and (5) into a unified framework:

$$E^{\text{LIFO}} = (E^{\text{LIF}} + E^{\text{prior}}), \quad (9)$$

$$E = \frac{1}{2} \int_{\Omega} |I - I^{\text{LFI}}|^2 dx + \nu \int_{\Omega} \delta(\phi) |\nabla \phi| dx + \frac{1}{2} \alpha \int_{\Omega} (H(\phi) - H(\phi_e))^2 dx, \quad (10)$$

where α is the constraint coefficient for ellipse which decides the weight of elliptic constraint and ν is the coefficient of the weighted length of zero level curve of ϕ .

There are three terms in the LIFO model (equation (10)) and each of them has its unique function to deal with

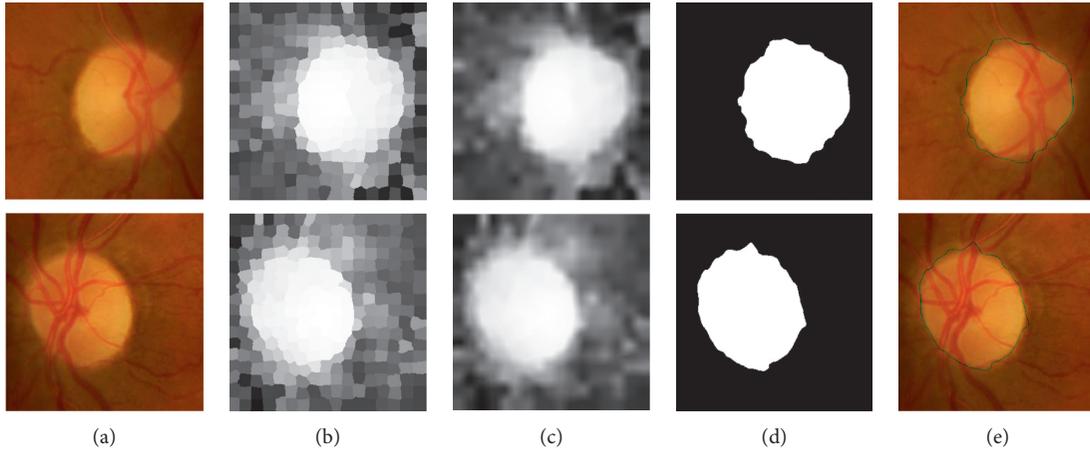


FIGURE 2: Contour initialization. (a) Cropped ROI around optic disc; (b) saliency detection result; (c) smoothed image of (b); (d) the largest connected object; (e) optic disc initial contour in green.

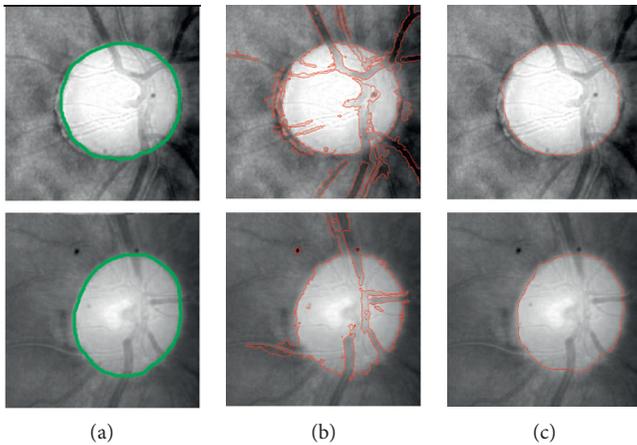


FIGURE 3: The result of OD boundary extraction obtained by LIF model and LIFO model, respectively; the ground truth is marked with a green line.

different problems in OD boundary extraction. The first term E^{LIF} is used to deal with the commonly occurred phenomenon in the optic disc regions that are always influenced by intensity inhomogeneity. The second term is the smooth item, which is used to handle drastic protuberance and sunken for evaluated contours by penalizing arc length of zero level contour of ϕ . The third term is the oval-shaped constraint term for ensuring the evaluated contour which can satisfy with the physical anatomical structure of optic disc, reducing the impact of complex environments. The LIFO model can be solved by the standard gradient descent method [28]. After a series of calculations, the solution is obtained in Appendix.

The flow diagram for segmentation of the OD is as follows:

- (1) Initialization: $v = 0.0001 \times 255 \times 255$, $\alpha = 1.0$, $x_e = \text{width}/2$, $y_e = \text{height}/2$, $\theta_e = 0$, $a_e = \sqrt{\text{width}^2 + \text{height}^2}/8$, $b_e = \sqrt{\text{width}^2 + \text{height}^2}/8$ (the width and the height are, respectively, the width and the height of the cropped region for the original image), the level set functions $\phi^l = \phi^0$, $\phi_e^l = \phi_e^0$, and l and r denote iterations.

- (2) Update m_1 and m_2 , respectively, using equations (3) and (4).
- (3) Update I^{LFI} using equation (2).
- (4) Using the standard gradient descent method, evolve the parameters of elliptical level set of the OD including x_e , y_e , θ_e , a_e , b_e according to equations (A.1)–(A.5); if x_e^r , y_e^r , θ_e^r , a_e^r , b_e^r satisfy the stationary condition, then stop; else $r = r + 1$ and return to Step 4.
- (5) Update ϕ_e^l using equation (8).
- (6) Evolve the level set functions, according to equation (A.6). If ϕ^l satisfy the stationary condition, stop; otherwise, $l = l + 1$ and return to Step 2.

3. Experimental Results

In this section, the public Standard Diabetic Retinopathy Database “Calibration Level 0” (DIARETDB0) [29] and the public dataset of retinal images namely DRISHTI-GS [30] are applied to verify the availability of our method. The DIARETDB0 and DRISHTI-GS are available and can be downloaded from the web pages <http://www.it.lut.fi/project/imageret/diaretdb0/> and <http://cvit.iit.ac.in/projects/mip/drishti-gs/mip-dataset2/Home.php>. The DIARETDB0 database is made up of 130 RGB color fundus images of which 20 are normal and 110 are abnormal (illness) with the fixed 1500×1152 resolution and 50° field of view. The ground truth is collected from two ophthalmologists. The final ground truth is acquired by averaging boundary results extracted from two ophthalmologists. The DRISHTI-GS dataset totally has 101 images of which 31 are normal and 70 are abnormal (illness). These images are produced with 30° degree field of view and have a resolution of 2896×1944 . For each image, the OD is correctly marked by four glaucoma experts. To compensate for interobserver marking variations, we also derive a majority voting manual marking as the final ground truth indicating that agreement among at least three experts [30] to qualitatively evaluate the proposed method.

Seen from Figure 4, compared with different contour evolution approaches using adaptive initial contour and

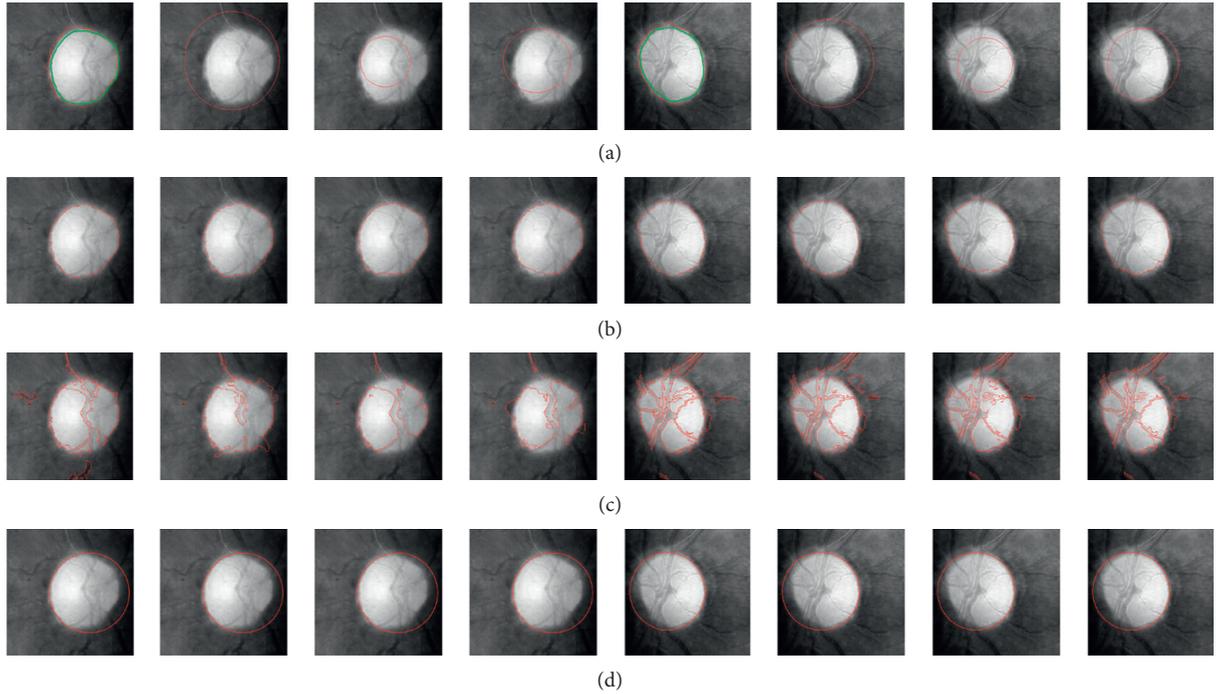


FIGURE 4: The comparisons for different segmentation models with different initial contours and Hough transform method. They, respectively, show the comparison results based on adaptive initial contour and manual initial circular contour drawing outside of the OD, inside of the OD, and intersect of the OD. The ground truth is marked with a green line. (a) Initial level set contour. (b) Presented LIFO. (c) LIF [28]. (d) Hough transform [31].

different initial circular contours based on the fundamental anatomical structure of the OD, there are some advantages for the proposed approach. First, most of ACM-based approaches are sensitive to the initialization of the contour [32]. However, the proposed initial contour can better guide the motion of the active contour since it is close to the ground truth of OD boundary. Second, the adopted initial contour which is near the OD boundary can reduce iterations of contour evolution. Therefore, it can reduce the computational cost [33, 34]. Furthermore, compared with the original LIF [28], our approach is more robust to the influence caused by the blood vessels due to the fact that the oval-shaped constraint is incorporated into our model.

The criterion is adopted to further assess the availability of LIFO model with different initial contours; it is considered that the overlapping ratio T which is computed based on the overlapping area between the true optic disc region in the ground truth and the detected optic disc region is no less than 75% for successful segmentation in terms of [11]. The accuracy ratio is the percentage ratio of successfully classified images to the total number of images. The overlapping ratio T is defined as

$$T = \frac{\text{area}(G \cap D)}{\text{area}(G \cup D)}, \quad (11)$$

where G and D are, respectively, the area of ground truth and the area extracted by the methods. Table 1 shows accuracy rate acquired by different initial contours.

Seen from Table 1, the proposed method achieves the best segmentation result with adaptive initial contour, and the accuracy rate is, respectively, 96.30% and 96.10% on the DIARETDB0 database and the DRISHTI-GS database.

TABLE 1: Performance measurement based on overlapping areas between different initial contours on the DIARETDB0 database and the DRISHTI-GS database.

Initial contour	Accuracy rate (DIARETDB0) (%)	Accuracy rate (DRISHTI-GS) (%)
Contour intersecting the OD	94.50	94.10
Contour within the OD	94.80	94.50
Contour outside the OD	95.10	95.30
Adaptive contour	96.30	96.10

In order to better verify the effectiveness of the proposed method, we compare our method with some related and newest approaches for segmentation in medical image processing area such as Hough transform method [31], modified radial symmetry method (MRS) [35], GVF method [36], Chan-Vese (CV) ACM [37], LIF ACM [28], and LSACM ACM [38]. The different segmentation results obtained by all five methods from retinal images are given in Figure 5, in which the green line denotes the ground truth obtained from the experts' marking and the red line represents some segmentation results extracted by different approaches. The examples of the OD having peripapillary atrophy are shown in the first three columns, and the OD with irregular shape and high gradient variations is shown in the fourth column. The Hough transform and the GVF model fail to extract the whole OD boundary due to the fact that they are sensitive to the varying of local gradient. Although MRS can achieve more accurate result than Hough transform, it ignores that the OD is an approximately circular or elliptic region rather than rigid circular

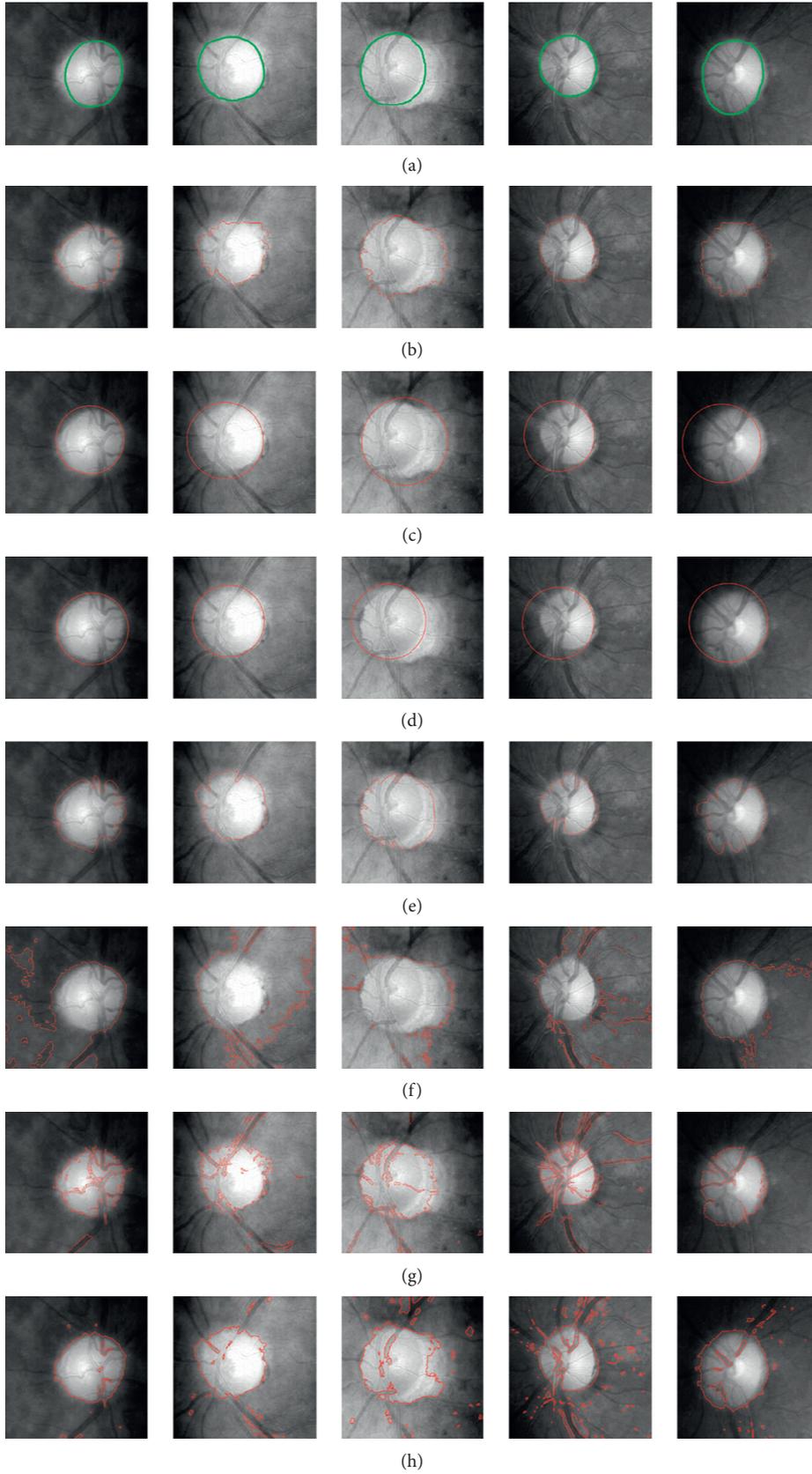


FIGURE 5: Continued.

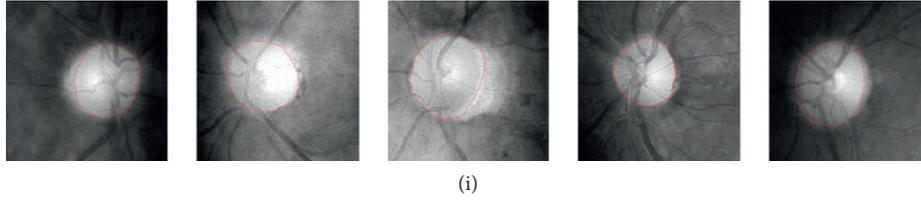


FIGURE 5: OD segmentation results: (a) original image with the ground truth; (b) adaptive initialized contour; (c) Hough transform results [31]; (d) MRS results [35]; (e) GVF model results [36]; (f) CV model results [37]; (g) LIF model results [28]; (h) LSACM model results [38]; (i) proposed LIFO model results. Green color indicates boundary marked by the expert and red color indicates achieved boundary by a method.

region. However, the CV model models image as piecewise constant function which fails to handle intensity inhomogeneity in retinal image, and thereby achieves unsatisfactory segmentation result. Although the LIF model can deal with these local gradient variations well compared to GVF and Hough transform and reduce the influence of intensity inhomogeneity because of considering local intensity information; it is severely influenced by blood vessel covering the OD surface. The LSACM model also can handle the intensity inhomogeneity and achieve more integrated OD boundary compared to the LIF model because it models the objects as Gaussian distributions of different means and variances; however, it is defeated by blood vessels and PPA obtaining a deficient segmentation result. Seen from the aforementioned methods, our method performs better and captures the whole OD boundary, which overcomes the influence caused by intensity inhomogeneity, PPA, and blood vessels. The fifth column shows a successful result segmented by LIFO model in blurry OD region with smooth transition boundary. This is mainly due to the fact that the prior shape information in some regions is a stronger cue than the intensity information. Therefore, combining the prior information and intensity information together can obtain the smooth and precise OD boundary.

Table 2, respectively, shows the average overlapping ratio and accuracy rate acquired by different models.

As seen from Table 2, we can clearly see that our method can get a better performance from DIARETDB0/DRISHTI-GS compared with other methods in terms of average overlapping ratio 66.59%/65.61% and accuracy rate 96.30%/96.10% for successful segmentation in retinal images including normal and abnormal (illness). The average overlapping ratio of segmentation obtained by proposed method in retinal image for normal/abnormal is 67.33% and 66.25%/65.53% and 64.87%; the accuracy rate of segmentation obtained by the proposed method in retinal image for normal/abnormal is 98.40% and 98.90%/95.90% and 94.90% on the DIARETDB0 and the DRISHTI-GS, respectively.

Besides, we also use an important evaluation metric F -score (F) which is the harmonic mean of precision and recall between the achieved boundary by the method and ground truth to test the performance of the proposed model. The pixelwise precision and recall values are, respectively, defined as

$$\begin{aligned} \text{precision} &= \frac{tp}{tp + fp}, \\ \text{recall} &= \frac{tp}{tp + fn}, \end{aligned} \quad (12)$$

where true positive (tp) indicates the number of pixels in the coverage areas between the ground truth and achieved segmented area by the methods; false positive (fp) expresses the number of pixels in the area where the pixel is classified only in the segmented area by the methods and is excluded belonging to the ground truth; false negative (fn) is the number of pixels in the area where a pixel is classified only in the ground truth and is excluded belonging to the segmented area by the methods. Then, the single performance measure, namely, F -score (F) is computed and defined as

$$F = 2 \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}} \quad (13)$$

The value of F -score always lies between 0 and 1 and will be high if the performance of method is good.

Table 3 depicts the quantitative assessment for segmentation results in terms of the F -score. The best and the worst achieved by the proposed method are, respectively, the best case and the worst case for fundamental results of the optical disc from the DIARETDB0 and the DRISHTI-GS. Seen from Table 3, it can be inferred that our method has a significant improvement in the segmentation results compared to others methods.

4. Conclusions

In this paper, we design a strategy to accurately segment OD boundary without manpower. First, an automatic and robust adaptive initial level set contour extraction method consisting of saliency detection and threshold techniques is presented for making the contour evolution. Then, in order to remedy the deficiency that only considers the intensity and ignores the prior information for OD shape, an excellent local image fitting model with oval-shaped constraint (LIFO) is presented to extract the whole and precise OD boundary. Comparing with the original LIF model only based on intensity information, the LIFO model uses both of the intensity information and shape information which has the following advantages. First, the original model is easily influenced by PPA, blood vessels, and noise due to only considering the intensity information. On the contrary, the proposed model can overcome these issues by using both of the intensity information and the shape prior information without any preprocessing. Second, the proposed model introduces the shape prior information based on the physical anatomical structure of the optic disc, and it can extract the whole boundary of the optic disc especially for the irregular shape of the optic disc. The experimental results

TABLE 2: Performance measurement based on overlapping areas between the proposed approach and other segmentation approaches on the DIARETDB0 database and DRISHTI-GS database.

	Average overlapping ratio (DIARETDB0) (%)	Accuracy rate (DIARETDB0) (%)	Average overlapping ratio (DRISHTI-GS) (%)	Accuracy rate (DRISHTI-GS) (%)
Hough [31]	61.42	89.60	60.55	88.10
MRS [35]	61.96	90.80	60.81	88.60
GVF [36]	63.66	92.80	61.86	91.30
CV [37]	55.15	86.10	55.15	85.30
LIF [28]	63.89	93.10	63.02	91.70
LSACM [38]	64.24	93.90	63.91	93.50
Ours (LIFO)	66.59	96.30	65.61	96.10
Normal	67.33	98.40	66.25	98.90
Abnormal	65.53	95.90	64.87	94.90

TABLE 3: Performance measurement based on F -score between the proposed approach and other segmentation approaches on the DIARETDB0 database and DRISHTI-GS database.

Methods	F -score (average) (DIARETDB0)	F -score (average) (DRISHTI-GS)
Hough [31]	0.853	0.841
MRS [35]	0.865	0.859
GVF [36]	0.885	0.882
CV [37]	0.792	0.786
LIF [28]	0.915	0.908
LSACM [38]	0.937	0.919
Ours(LIFO)	0.951	0.946
Best	0.986	0.990
Worst	0.658	0.646

demonstrate the availability of the proposed method. Now, the deep learning has attracted attention and achieves a good performance when the number of training samples is enough. However, it is hard to collect enough data in medical field such as the retinal fundus images, which will greatly reduce the performance of model. That is the main reason why we did not employ the deep learning technique to segment the optic disc and optic cup. In the future, we will try to use the deep learning approaches on the larger database.

Appendix

The LIFO model can be solved by the standard gradient descent method [28]. After a series of calculations, the solution is obtained as follows:

$$\begin{aligned} \frac{dx_e}{dt} = & \alpha \int_{\Omega} (H(\phi) - H(\phi_e)) \delta(\phi_e) \\ & \times \left(\frac{[(x_i - x_e) \cos \theta_e + (y_i - y_e) \sin \theta_e]^2}{a_e^2} \right. \\ & \left. + \frac{[-(x_i - x_e) \sin \theta_e + (y_i - y_e) \cos \theta_e]^2}{b_e^2} \right)^{-1/2} \\ & \cdot \left(\frac{[(x_i - x_e) \cos \theta_e + (y_i - y_e) \sin \theta_e] \cos \theta_e}{a_e^2} \right. \\ & \left. + \frac{[-(x_i - x_e) \sin \theta_e + (y_i - y_e) \cos \theta_e] \sin \theta_e}{b_e^2} \right) dx, \end{aligned} \quad (\text{A.1})$$

$$\begin{aligned} \frac{dy_e}{dt} = & \alpha \int_{\Omega} (H(\phi) - H(\phi_e)) \delta(\phi_e) \\ & \times \left(\frac{[(x_i - x_e) \cos \theta_e + (y_i - y_e) \sin \theta_e]^2}{a_e^2} \right. \\ & \left. + \frac{[-(x_i - x_e) \sin \theta_e + (y_i - y_e) \cos \theta_e]^2}{b_e^2} \right)^{-1/2} \\ & \cdot \left(\frac{[(x_i - x_e) \cos \theta_e + (y_i - y_e) \sin \theta_e] \sin \theta_e}{a_e^2} \right. \\ & \left. + \frac{[-(x_i - x_e) \sin \theta_e + (y_i - y_e) \cos \theta_e] \cos \theta_e}{b_e^2} \right) dx, \end{aligned} \quad (\text{A.2})$$

$$\begin{aligned} \frac{d\theta_e}{dt} = & \alpha \int_{\Omega} (H(\phi) - H(\phi_e)) \delta(\phi_e) \times \left(\frac{[(x_i - x_e) \cos \theta_e + (y_i - y_e) \sin \theta_e]^2}{a_e^2} + \frac{[-(x_i - x_e) \sin \theta_e + (y_i - y_e) \cos \theta_e]^2}{b_e^2} \right)^{-1/2} \\ & \cdot \left(\frac{-[(x_i - x_e) \cos \theta_e + (y_i - y_e) \sin \theta_e] [- (x_i - x_e) \sin \theta_e + (y_i - y_e) \cos \theta_e]}{a_e^2} \right. \\ & \left. + \frac{[(x_i - x_e) \cos \theta_e + (y_i - y_e) \sin \theta_e] [- (x_i - x_e) \sin \theta_e + (y_i - y_e) \cos \theta_e]}{b_e^2} \right) dx, \end{aligned} \quad (\text{A.3})$$

$$\begin{aligned} \frac{da_e}{dt} = \alpha \int_{\Omega} & (H(\phi) - H(\phi_e))\delta(\phi_e) \\ & \times \left(\frac{[(x_i - x_e)\cos\theta_e + (y_i - y_e)\sin\theta_e]^2}{a_e^2} \right. \\ & \left. + \frac{[-(x_i - x_e)\sin\theta_e + (y_i - y_e)\cos\theta_e]^2}{b_e^2} \right)^{-(1/2)} \\ & \cdot \left(\frac{[(x_i - x_e)\cos\theta_e + (y_i - y_e)\sin\theta_e]^2}{a_e^3} \right) dx, \end{aligned} \quad (\text{A.4})$$

$$\begin{aligned} \frac{db_e}{dt} = \alpha \int_{\Omega} & (H(\phi) - H(\phi_e))\delta(\phi_e) \\ & \times \left(\frac{[(x_i - x_e)\cos\theta_e + (y_i - y_e)\sin\theta_e]^2}{a_e^2} \right. \\ & \left. + \frac{[-(x_i - x_e)\sin\theta_e + (y_i - y_e)\cos\theta_e]^2}{b_e^2} \right)^{-(1/2)} \\ & \cdot \left(\frac{[-(x_i - x_e)\sin\theta_e + (y_i - y_e)\cos\theta_e]^2}{b_e^3} \right) dx, \end{aligned} \quad (\text{A.5})$$

$$\begin{aligned} \frac{\partial\phi}{\partial t} = \nu\delta(\phi)\text{div}\left(\frac{\nabla\phi}{|\nabla\phi|}\right) & + \delta(\phi)(I - I^{\text{LFI}})(m_1 - m_2) \\ & - \alpha\delta(\phi)[H(\phi) - H(\phi_e)], \end{aligned} \quad (\text{A.6})$$

where x_e , y_e , θ_e , a_e , b_e continually vary along with the changing of information in image and t is the time step of the experiment.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China under Grant nos. 61701101, U1713216, 61803077, and 61603080, the National Key Robot Project 2017YFB1300900, and the Fundamental Research Fund for the Central Universities of China N172603001 and N172604004.

References

- [1] K. Akita and H. Kuga, "A computer method of understanding ocular fundus images," *Pattern Recognition*, vol. 15, no. 6, pp. 431–443, 1982.
- [2] Y. Wang, G. Ji, P. Lin, and E. Trucco, "Retinal vessel segmentation using multiwavelet kernels and multiscale hierarchical decomposition," *Pattern Recognition*, vol. 46, no. 8, pp. 2117–2133, 2013.
- [3] J. Xu, O. Chutatape, E. Sung, C. Zheng, and P. C. T. Kuan, "Optic disk feature extraction via modified deformable model technique for glaucoma analysis," *Pattern Recognition*, vol. 40, no. 7, pp. 2063–2076, 2007.
- [4] B. Dashtbozorg, A. M. Mendonça, and A. Campilho, "Optic disc segmentation using the sliding band filter," *Computers in Biology and Medicine*, vol. 56, pp. 1–12, 2015.
- [5] J. Cheng, J. Liu, Y. Xu et al., "Superpixel classification based optic disc and optic cup segmentation for glaucoma screening," *IEEE Transactions on Medical Imaging*, vol. 32, no. 6, pp. 1019–1032, 2013.
- [6] M. K. Dutta, A. K. Mourya, A. Singh, M. Parthasarathi, R. Burget, and K. Riha, "Glaucoma detection by segmenting the super pixels from fundus colour retinal images," in *Proceedings of 2014 International Conference on Medical Imaging, m-Health and Emerging Communication Systems (Med Com)*, pp. 86–90, Noida, India, 2014.
- [7] N.-M. Tan, Y. Xu, W. B. Goh, and J. Liu, "Robust multi-scale superpixel classification for optic cup localization," *Computerized Medical Imaging and Graphics*, vol. 40, pp. 182–193, 2015.
- [8] W. Zhou, C. D. Wu, D. L. Chen, Z. Z. Wang, Y. G. Yi, and W. Y. Du, "Automatic microaneurysm detection using the sparse principal component analysis based unsupervised classification method," *IEEE Access*, vol. 5, no. 1, pp. 2169–3536, 2017.
- [9] W. Zhou, C. Wu, Y. Yi, and W. Du, "Automatic detection of exudates in digital color fundus images using superpixel multi-feature classification," *IEEE Access*, vol. 5, no. 1, pp. 17077–17088, 2017.
- [10] A. Pinz, S. Bernogger, P. Datlinger, and A. Kruger, "Mapping the human retina," *IEEE Transactions on Medical Imaging*, vol. 17, no. 4, pp. 606–619, 1988.
- [11] M. Lalonde, M. Beaulieu, and L. Gagnon, "Fast and robust optic disc detection using pyramidal decomposition and Hausdorff-based template matching," *IEEE Transactions on Medical Imaging*, vol. 20, no. 11, pp. 1193–1200, 2001.
- [12] S. Sekhar, W. Al-Nuaimy, and A. K. Nandi, "Automated localisation of retinal optic disk using Hough transform," in *Proceedings of 2008 5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pp. 1577–1580, Paris, France, May 2008.
- [13] A. Bhuiyan, R. Kawasaki, T. Y. Wong, and R. Kotagiri, "A new and efficient method for automatic optic disc detection using geometrical features," in *Proceedings of World Congress on Medical Physics and Biomedical Engineering*, pp. 1131–1134, Munich, Germany, September 2009.
- [14] A. Aquino, M. E. Gegúndez-Arias, and D. Marín, "Detecting the optic disc boundary in digital fundus images using morphological, edge detection, and feature extraction techniques," *IEEE Transactions on Medical Imaging*, vol. 29, no. 11, pp. 1860–1869, 2010.
- [15] D. Wong, J. Liu, J. Lim et al., "Level-set based automatic cup-to-disc ratio determination using retinal fundus images in ARGALI," in *Proceedings of 2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, pp. 2266–2269, Vancouver, BC, Canada, August 2008.
- [16] S. Morales, V. Naranjo, J. Angulo, and M. Alcaniz, "Automatic detection of optic disc based on PCA and mathematical

- morphology," *IEEE Transactions on Medical Imaging*, vol. 32, no. 4, pp. 786–796, 2013.
- [17] S. Roychowdhury, D. Koozekanani, S. Kuchinka, and K. Parhi, "Optic disc boundary and vessel origin segmentation of fundus images," *IEEE Journal of Biomedical and Health Informatics*, vol. 20, no. 6, pp. 1562–1574, 2017.
- [18] A. W. Reza, C. Eswaran, and S. Hati, "Automatic tracing of optic disc and exudates from color fundus images using fixed and variable thresholds," *Journal of Medical Systems*, vol. 33, no. 1, pp. 73–80, 2008.
- [19] D. Welfer, J. Scharcanski, and D. R. Marinho, "A morphologic two-stage approach for automated optic disk detection in color eye fundus images," *Pattern Recognition Letters*, vol. 34, no. 5, pp. 476–485, 2013.
- [20] R. Srivastava, J. Cheng, D. W. K. Wong, and J. Liu, "Using deep learning for robustness to parapapillary atrophy in optic disc segmentation," in *Proceedings of 2015 IEEE 12th International Symposium on Biomedical Imaging*, pp. 768–771, Brooklyn, NY, USA, April 2015.
- [21] S. Lee and M. Brady, *Optic Disk Boundary Detection*, Springer London, London, UK, 1991.
- [22] F. Mendels, "Identification of the optic disk boundary in retinal images using active contours," in *Proceedings of Irish Machine Vision and Image Processing Conference*, pp. 103–115, Dublin, Ireland, September 1999.
- [23] H. Yu, E. S. Barriga, C. Agurto et al., "Fast localization and segmentation of optic disk in retinal images using directional matched filtering and level sets," *IEEE Transactions on Information Technology in Biomedicine*, vol. 16, no. 4, pp. 644–657, 2012.
- [24] M. Esmaeili, H. Rabbani, and A. M. Dehnavi, "Automatic optic disk boundary extraction by the use of curvelet transform and deformable variational level set model," *Pattern Recognition*, vol. 45, no. 7, pp. 2832–2842, 2012.
- [25] M. C. V. S. Mary, E. B. Rajsingh, J. K. K. Jacob, D. Anandhi, U. Amato, and S. E. Selvan, "An empirical study on optic disc segmentation using an active contour model," *Biomedical Signal Processing and Control*, vol. 18, pp. 19–29, 2015.
- [26] W. Zhou, C. Wu, Y. Gao, and X. Yu, "Automatic optic disc boundary extraction based on saliency object detection and modified local intensity clustering model in retinal images," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. E100-A, no. 9, pp. 2069–2072, 2017.
- [27] Y. Qin, H. Lu, Y. Xu, and H. Wang, "Saliency detection via cellular automata," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 110–119, Boston, MA, USA, June 2015.
- [28] K. Zhang, H. Song, and L. Zhang, "Active contours driven by local image fitting energy," *Pattern Recognition*, vol. 43, no. 4, pp. 1199–1206, 2010.
- [29] T. Kauppi, V. Kalesnykiene, J. K. Kamarainen et al., "DIA-RETDB0: evaluation database and methodology for diabetic retinopathy algorithms," Technical Report, Lappeenranta University of Technology, Lappeenranta, Finland, 2008.
- [30] J. Sivaswamy, S. R. Krishnadas, G. D. Joshi, M. Jain, and A. U. S. Tabish, "Drishti-G. S.: retinal image dataset for optic nerve head (ONH) segmentation," in *Proceedings of 2014 IEEE 11th International Symposium on Biomedical Imaging*, pp. 53–56, Beijing, China, April-May 2014.
- [31] R. Chrástek, M. Wolf, K. Donath, G. Michelson, and H. Niemann, "Optic disc segmentation in retinal images," in *Proceedings of Bildverarbeitung für die Medizin*, pp. 263–266, Leipzig, Germany, March 2002.
- [32] C. Li, R. Huang, Z. Ding, J. 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, 2011.
- [33] L. Wang, C. Li, Q. Sun, D. Xia, and C.-Y. Kao, "Active contours driven by local and global intensity fitting energy with application to brain MR image segmentation," *Computerized Medical Imaging and Graphics*, vol. 33, no. 7, pp. 520–531, 2009.
- [34] C. Li, C. Y. Kao, J. C. Gore, and Z. Ding, "Implicit active contours driven by local binary fitting energy," in *Proceedings of 2007 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–7, Minneapolis, MN, USA, June 2007.
- [35] G. Sciortino, E. Orlandi, C. Valenti, and D. Tegolo, "Wavelet analysis and neural network classifiers to detect mid-sagittal sections for nuchal translucency measurement," *Image Analysis & Stereology*, vol. 35, no. 2, pp. 105–115, 2016.
- [36] A. Osareh, M. Mirmehdi, B. Thomas, and R. Markham, "Colour morphology and snakes for optic disc localization," in *Proceedings of the 6th Medical Image Understanding and Analysis Conference*, pp. 21–24, Portsmouth, UK, July 2002.
- [37] G. D. Joshi, J. Sivaswamy, K. Karan, and R. Krishnadas, "Optic disk and cup boundary detection using regional information," in *Proceedings of IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pp. 948–951, Rotterdam, The Netherlands, April 2010.
- [38] K. Zhang, L. Zhang, K.-M. Lam, and D. Zhang, "A level set approach to image segmentation with intensity inhomogeneity," *IEEE Transactions on Cybernetics*, vol. 46, no. 2, pp. 546–557, 2016.

Research Article

Automatic Tissue Image Segmentation Based on Image Processing and Deep Learning

Zhenglun Kong ¹, Ting Li ², Junyi Luo,³ and Shengpu Xu ²

¹Northeastern University, Boston, MA, USA

²Institute of Biomedical Engineering, Chinese Academy of Medical Science and Peking Union, Tianjin 300192, China

³University of Electronic Science and Technology of China, Chengdu, China

Correspondence should be addressed to Shengpu Xu; xushengpu@163.com

Received 4 May 2018; Revised 17 September 2018; Accepted 20 December 2018; Published 31 January 2019

Academic Editor: Cesare Valenti

Copyright © 2019 Zhenglun Kong et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Image segmentation plays an important role in multimodality imaging, especially in fusion structural images offered by CT, MRI with functional images collected by optical technologies, or other novel imaging technologies. In addition, image segmentation also provides detailed structural description for quantitative visualization of treating light distribution in the human body when incorporated with 3D light transport simulation methods. Here, we first use some preprocessing methods such as wavelet denoising to extract the accurate contours of different tissues such as skull, cerebrospinal fluid (CSF), grey matter (GM), and white matter (WM) on 5 MRI head image datasets. We then realize automatic image segmentation with deep learning by using convolutional neural network. We also introduce parallel computing. Such approaches greatly reduced the processing time compared to manual and semiautomatic segmentation and are of great importance in improving the speed and accuracy as more and more samples are being learned. The segmented data of grey and white matter are counted by computer in volume, which indicates the potential of this segmentation technology in diagnosing cerebral atrophy quantitatively. We demonstrate the great potential of such image processing and deep learning-combined automatic tissue image segmentation in neurology medicine.

1. Introduction

Nuclear magnetic resonance imaging gives a clear and high-resolution image of brain tissues [1]. It is a common method for clinical examination of brain diseases. The human brain structure is very complicated. Important tissues include grey matter, white matter, and cerebrospinal fluid [2] (Figure 1(a)). These tissues play a key role in memory, cognition, awareness, and language [3]. Cerebral atrophy/expansion [4, 5] and leukodystrophy [6] are serious brain dysfunction diseases that have a high incidence in infants and elderly people [7]. However, crucial tissues such as cerebrospinal fluid, grey matter, and white matter are hard to differentiate due to blurry boundaries, especially in the cross-sectional images that do not show the center of the brain, as shown in (Figure 1(b)). As a result, it is hard for doctors to analyze them separately and find the location of the disease [8]. With the popularization of image-aided

medical diagnosis, computer-aided doctors can improve the efficiency of segmenting [9] the grey matter and white matter of the brain MRI. In MR imaging, different signal intensities and weighted images (T1 weighted and T2 weighted) can make the image display at different grey levels. Since the T1 brain magnetic image shows that the soft tissue is better [10], the experiment selects the brain magnetic resonance T1-W image as the experimental sample.

Many approaches have been made to segment the brain image automatically. Segmentation algorithms based on regional, texture, and histogram thresholds [11, 12] are simple but lack accuracy. Threshold is a simple but effective way to segment images. However, there are some limits regarding only using this method for segmentation. First, the grey scale of tissues may not be restricted in one range. This means that, if we simply use threshold to locate the tissues, it may fail to separate all the parts. Secondly, the threshold usually does not consider the spatial properties of an image.

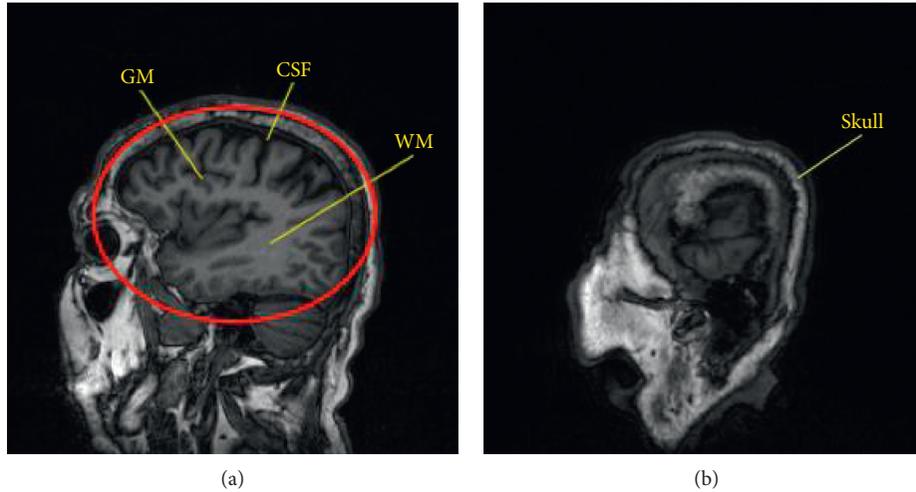


FIGURE 1: MRI images: (a) location of the 3 tissues. WM is the area where the color is light; GM is the gray boundary around the WM; CSF is the black parts inside the skull. We focus on segmenting the tissues in the red circle. (b) Cross section of a side of the brain; the tissues are hard to distinguish by eyes.

For example, the skull is a round structure that covers the other tissues. This can help us to determine the location of tissues and get more accurate segmentation images. As a result, threshold determination is often considered as an early stage sequential image process. Later, methods related to fuzzy c-means (FCM) [13, 14] and machine learning are introduced. The atlas-based method is also widely used for brain image segmentation [15]. It has a relatively complete system framework. However, explicit information such as intensity and spatial features is required in order to get accurate results [16]. Spatial and intensity features could be avoided by using convolutional neural networks (CNNs) [16]. Convolution neural network proposed by LeCun et al. [17] is a deep supervised learning method [18]. It has been applied in many fields and has made great success in image recognition [19, 20], speech recognition [21, 22], natural language processing, and so on. CNNs obtain the convolution weight by means of cyclic convolution and samples with a supervised training mode. The final realization is directly extracted from the original input, which is conducive to the classification features. The features in the image recognition are texture, shape, and structure.

MICCAI is a conference held every year focusing on medical image computing and computer-assisted intervention [23]. Recently, many methods related to deep learning were presented in the conference. Zhang et al. presented a 2D patch-wise convolutional neural networks (CNNs) approach to segment tissues from multimodal MR images of infants [24]. An $N \times N$ size picture block was extracted from a given image, and the model is trained with these blocks. Then, the label was given to the correct identification class. In order to improve the performance of block training framework, multiscale CNNs used a variety of ways with different patch sizes. The outputs of these approaches were combined with the neural network, and the model was trained to give the correct label. This method in this paper did not include a pooling layer or consider the

relation between the patches. Yang et al. [25] used a deep active learning framework to reduce the annotation effort. It was combined with fully convolutional network and active learning. Man et al. [26] proposed combining MRI multimodal information to extend CNNs to 3D, which was composed of multiple modes that formed 3D raw data.

In our paper, we use image enhancement, operators, and morphometry methods to extract the accurate contours of different tissues on 5 MRI head image datasets. After that, we utilize convolutional neural network to realize automatic segmentation of images with deep learning. Such approaches greatly reduced the processing time compared to the other methods. We also introduce parallel computing to further speed up the processing speed. Our work has a great potential in the medical field for diagnosing brain disease.

The rest of the paper is organized as follows. In Section 2, we describe our dataset, model, and training method. Our experiments and comparison with other methods are discussed in Section 3. Section 4 concludes the paper.

2. Materials and Methods

2.1. Dataset. Our dataset includes 5 patient's brain MRI T1-W images. For every patient, we have 160 images, with a total of 800 images. The size of the images is 256×256 pixels. Every pixel value in the matrix is an integer between 0 and 255. Figure 2 shows some typical MRI of a human brain. The MRI data used to support the findings of this study are available from the corresponding author upon request.

2.2. Preprocessing. Due to the complicity of the brain structure, there exist many overlapping regions in each MR image. Image preprocessing can improve both the efficiency of the algorithm and the reliability of the segmentation results. Image noise reduction [27] and enhancement can make the image more conforming for viewing. By removing

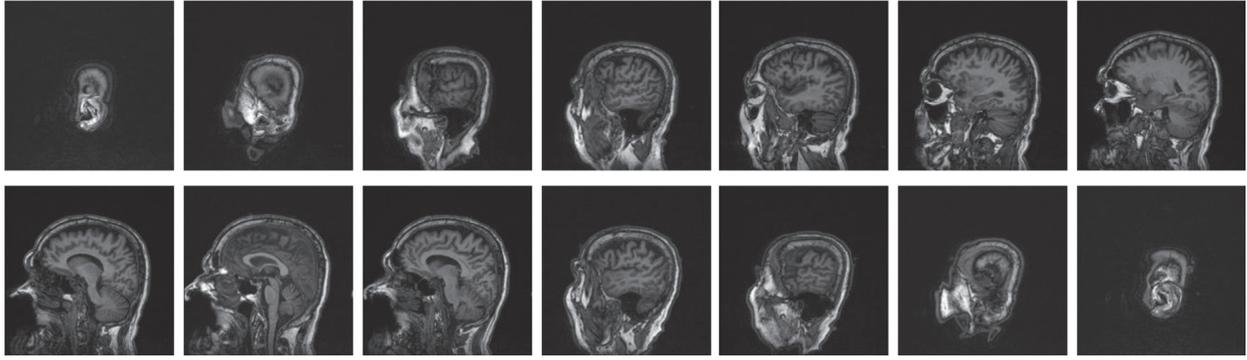


FIGURE 2: Samples of a patient's brain cross-sectional image.

the bright skull, we can avoid it from affecting the accuracy of the segmentation of the brain. Image noise reduction directly affects the result of segmentation.

Wavelet domain denoising is used to transform noisy signals from time domain to wavelet domain [10] by using multiscale transformation. We removed the wavelet coefficients of noise from all scales to obtain the wavelet coefficients of signals. Finally, the signals are reconstructed by a wavelet transform. The image after noise reduction preserves the details of the original image, and the visual effect becomes clearer. The histogram equalization method is used to enhance the image of the cerebrospinal fluid, grey matter, and white matter.

We collected the grey level of all 800 MRI images and generated a histogram that contains all the points, as shown in Figure 3. The result shows 4 peaks, each stands for a kind of tissue [28]. The background grey value is not shown in the figure, which is smaller than 35. From the histogram, we can remove pixels that are not in the grey level range of the GM, WM, CSF, and skull. We convert them to level 0 to reduce the noise. The histogram shows four thresholds, which stands for the four tissues. It seems that we can segment the image by only using this result. However, there are some limits. Threshold usually does not consider spatial properties of an image. For example, the shape of the skull is round and located around the other tissues. Also, the grey level of a tissue may not be restricted around one region. The grey level of GM may be in the CSF region depending on the location. Thus, the result may not be accurate by only using the threshold as an analysis measuring method.

2.3. Convolutional Neural Network. Convolutional neural networks (CNNs) have recently enjoyed a great success in image recognition and segmentation. The basic structure of CNNs consists of two layers. One is the feature extraction layer (C1, C3). The input of each neuron is connected to the local receptive domain of the previous layer, extracting the local feature. Once the local feature is extracted, its positional relationship between the others can be determined. The other layer is the feature mapping network layer (S2, S4). Each computing layer is composed of multiple feature maps [29]. The feature map is a flat plane; all neuron weights are equal. The feature mapping structure uses the sigmoid

function [30] as the activation function of the convolution network. In addition, the number of free parameters of the network is reduced because of the weights shared by a neuron on a mapping surface. Each convolutional layer in the CNN closely follows a computing layer for local average and second extraction. This unique extraction structure reduces the feature resolution.

The C1 layer (Figure 4) is a convolutional layer with six feature maps. Each neuron in the feature map is connected to the 5×5 input. The size of the feature map is 28×28 . S2 is a pooling layer with six 14×14 features. Each unit in the feature map is connected to the 2×2 neighborhood of the corresponding feature map in the C1. The four inputs in each unit are added in S2 and multiplied by a trainable parameter, along with a trainable offset. The 2×2 receptive field of each unit does not overlap, so the size of each feature map in S2 is $1/4$ of the size as in C1.

The C3 layer is also a convolutional layer which uses a kernel of 5×5 to convolute the layer S2. The feature map has only 10×10 neurons but with 16 different convolution kernels. Hence, there are 16 feature maps. Each map in C3 consists of all 6 or several feature maps in S2. The reason why we do not connect each feature map of the S2 to C3 is that the incomplete connection mechanisms keep the number of connections within a reasonable range. Moreover, it destroys the symmetry of the network. Since different feature maps have different inputs, it forces them to extract different features. The S4 layer is a pooling layer that consists of sixteen 5×5 size feature maps. Each unit in the feature map is connected to the 2×2 neighborhood of the corresponding feature map in the C3, same as the C1 and S2. The F6 layer has 84 units and is fully connected to the C5 layer. Finally, the output layer is composed of a Euclidean radial basis function unit, each of which has a unit with 84 inputs.

The output of the convolution layer is the sum of the convolution kernel and the output of the upper layer:

$$x_j^l = f \left(\sum_{i \in M_j} x_i^{l-1} * k_{ij}^l + b_j^l \right), \quad (1)$$

where l is the layer number, k is the convolution kernel, b is the bias, x is the input, and M_j is the chosen feature map.

The number of input and output in the pooling layer are the same, while the dimension of the maps is reduced:

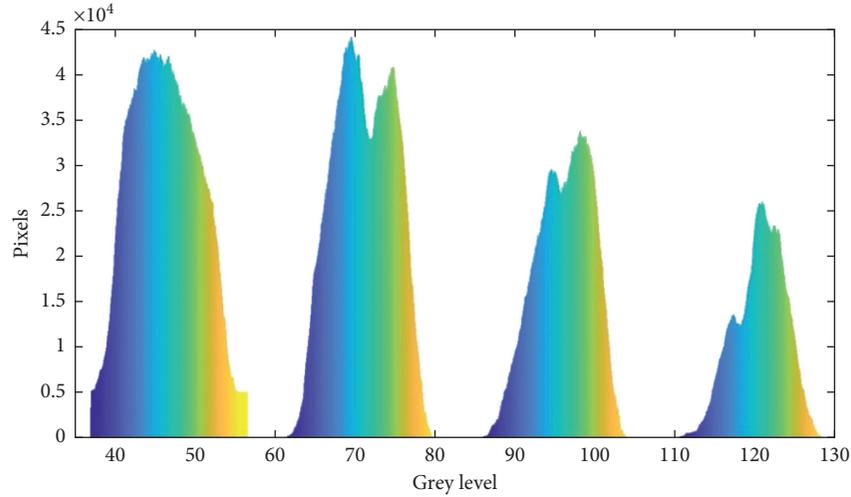


FIGURE 3: There are four peaks in the histogram. From left to right, the peaks stand for the following: cerebrospinal fluid (40–57), grey matter (61–79.8), white matter (86–110), and skull (110–130).

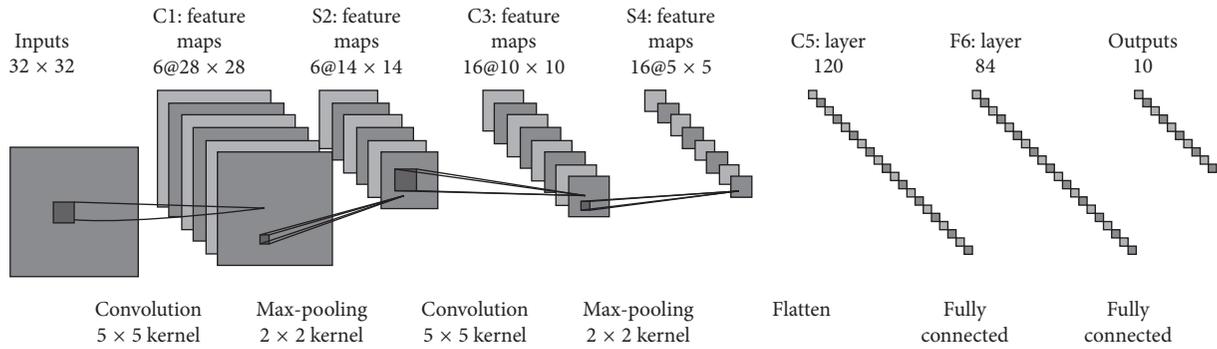


FIGURE 4: General structure of CNN. The input layer is 32×32 . The input is convoluted to six feature maps in the C1 layer by 5×5 convolution kernel. S2 is a pooling layer with six 14×14 features. Each unit in the feature map is connected to the 2×2 neighborhood of the corresponding feature map in the C1 layer. The C3 layer is also a convolutional layer that uses a kernel of 5×5 to convolute the layer S2. The S4 layer is a pooling layer that consists of sixteen 5×5 size feature maps. The C5 layer is a convolutional layer with 120 feature maps. The F6 layer has 84 units and is fully connected to the C5 layer. The output layer has a unit with 84 inputs.

$$x_j^l = f(\beta_j^l \text{down}(x_i^{l-1}) + b_j^l). \quad (2)$$

The parameters of the CNN are set as follows (Figure 5). The neural network is divided into three layers: the input layer is many 4×4 pixel units, the second layer is the convolution layer with 6 kernel functions of 3×3 , and the third layer is the down-sampling layer with six 2×2 pixel units. Finally, the parameters of the stable network after training are obtained. In each layer, there are many 2D planar elements, and each 2D planar element has many independent neurons. The output has six pixels, thus extracting the deep feature data. We used Adam algorithm for the learning method. We set the initial learning rate to 0.001 and the momentum to 0.5. We used the cross entropy as the loss function. Our CNN was trained for 50 epochs, each consisting of 20 subepochs. Our batch size is 5. Training samples are randomly selected from the total 800 images: 600 images are for training, 100 images are for validation, and 100 images are for testing.

2.4. Parallel Computing. MPI is at the core of many supercomputing software frameworks [31]. We used a Caffe framework that adopts the MPI. The MPI enables the cluster version to optimize the data parallel to Caffe. It supports command line, Python, and MATLAB interfaces and various programming methods. The CPU information is Intel(R) Xeon(R) CPU E5-2670 0 @ 2.60 GHz. We convert the 800 matrices into one large matrix, which has the size 800×65536 . In addition, we added a row for placing the image number. Each row is an image. The first number in each row is the index of the image. The image data are put together in one matrix. Thus, the processing involves all the images at one time. We adopt the master-slave mode [32]. It includes two sets of processes: the master processor is in charge of processing the work orders [33]. The slaves execute the work that the master processor assigns.

In our work, one node acts as the master node, which is responsible for data partition and allocation. The other nodes complete the calculation of local data and return the

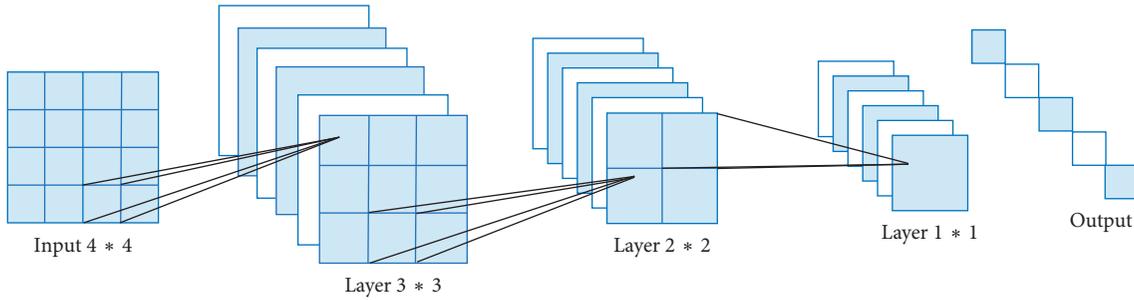


FIGURE 5: Our CNN structure.

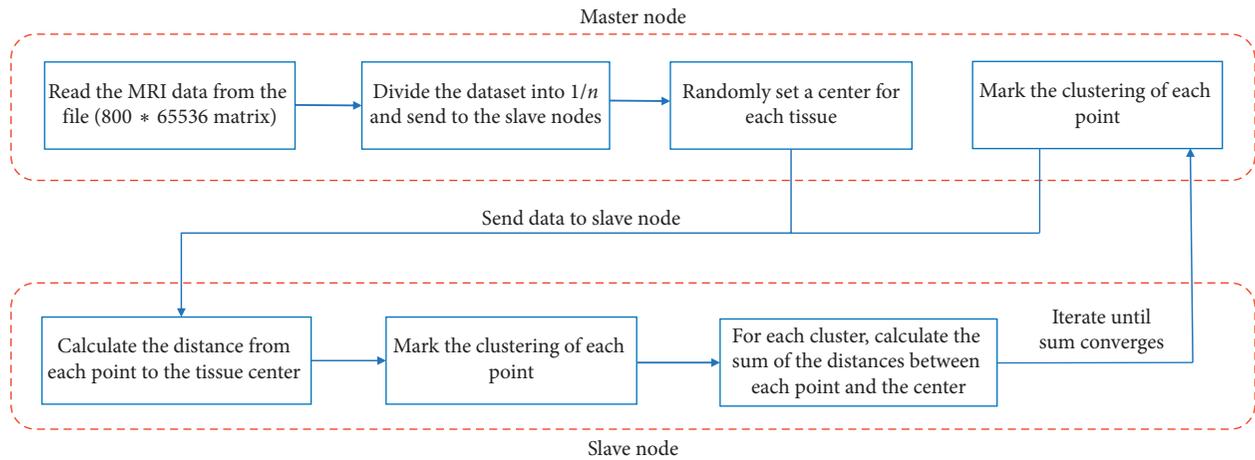


FIGURE 6: Parallel computing master-slave method. The master processor is in charge of processing the work orders. The slaves execute the work that the master processor assigns. The process repeats until the sum of the distances of all the clusters is constant.

result to the master node. As shown in Figure 6, the master node first reads the data and assigns them to the other nodes and then selects the center of each cluster. The slaves calculate the distance from each point to the center of the data block, then mark the clustering of each point, calculate the sum of the distances between all the points of each cluster to the center of the cluster, and finally return these results to the master. The cluster centers stand for the tissues in the GM, WM, CSF, and skull, which are spatial coordinates. We use the Euclidean distance to find the center of the tissue feature clusters, respectively, and set the parameter $\partial = 0.5$, which takes out 50% feature points that are nearest to the feature center points to accurately characterize the quantitative characteristics of the different types of data. The master node will calculate the new center point, send to other processes, and calculate the other process from the clustering of all points to the center of the sum of the distance. The process will continue until the sum of the distances of all the clusters is constant.

3. Results

3.1. Tissue Segmentation. Our work shows some satisfying results (Figure 7(b)). The images in the left column are the original images. They show a full MRI image of the tissues of the brain and other parts such as the facial skull, muscle, and ears. After removing the other parts of the head apart from

the brain that are considered noise, we successfully segment one brain image into 4 images. In each image, we set the gray value of the tissue as 255 and the background as 0.

The result of the skull shows a curved shape located on the frontier of the brain. The cerebrospinal fluid is the segment between the skull and the grey matter/white matter. The grey matter and white matter are also accurately segmented.

To test the efficiency of our method, we also did segmentation without preprocessing. As shown in Figure 7(a), the results include a large amount of noise including parts that are not brain such as the nose, eyes, and other facial structures.

3.2. Comparison with Visible Chinese Human (VCH). Our result was compared pixel by pixel with the images segmented by an expert operator. To quantize our result, we calculated the percentage of each tissue in the human brain and then compared it to the visible Chinese human head (VCH) model [34]. The VCH model is a very good presentation for the anatomical structure. It is mostly used in modeling light propagation [35]. It can help calculate the volume of brain tissues because the intensity of light changes while propagating. The VCH model is developed from high-resolution cryosectional color photographs of a reference adult male [36, 37]. It includes various types of tissues from a

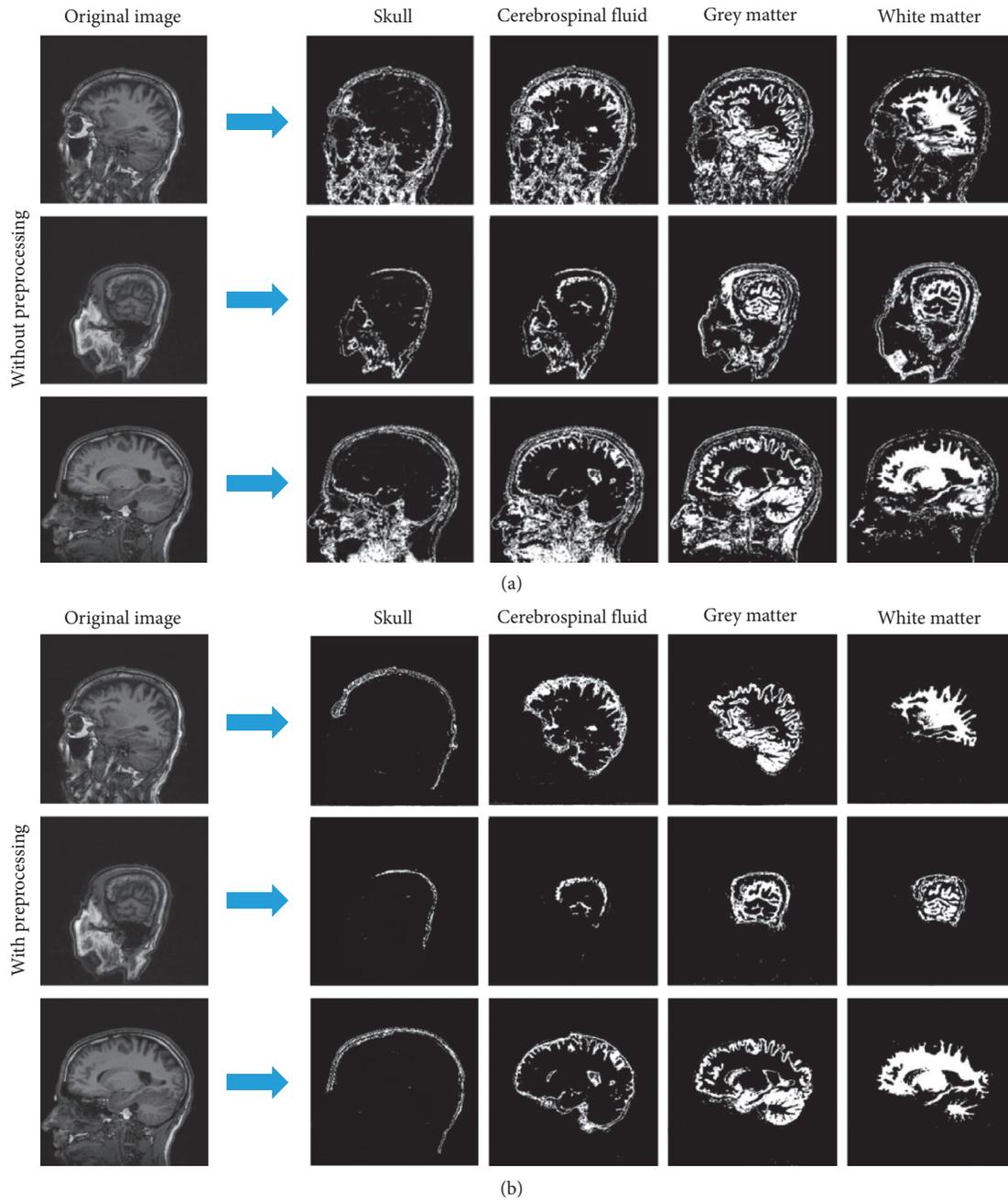


FIGURE 7: (a) Images generated without preprocessing. The first column shows the skull results; the second column shows the cerebrospinal fluid results; the third column shows the grey matter results; the fourth column shows the white matter results. (b) Images generated with preprocessing.

standing frozen man body. The section precision is horizontally 0.02 cm interval, and the digital color image has a resolution of 0.01 cm per pixel, which is higher than CT and MRI [38]. Thus, it is one of the most realistic head models that contains precise cerebral cortex folding geometry [39].

The data are shown in Table 1. Overall, our result is quite satisfying. We manage to locate the boundary and segment the tissues accurately (Figure 8(a)). We then calculated the ratio between the grey matter and white matter and compared it to the VCH result, giving an accuracy of 95%

(Figure 8(b)). This can be used in diagnosing disease such as cerebral atrophy, which is caused by grey matter or white matter reduction. There are still some deviations between our result and the ground truth. This is caused by the remaining noise and the insufficiency of the algorithm. The MRI dataset contains slit images of the brain instead of crosscut, which shows a full head rather than only the brain. As a result, more noise will be generated from the other part of the head. Taking the GM percentage, for example, some parts that are apart from the brain have the same grey value

TABLE 1: A comparison of the percentage of each tissue in the brain.

Type	VCH (%)	MRI (%)
CSF	37.65	38.54 \pm 2.21
GM	27.08	27.33 \pm 1.47
WM	12.35	12.95 \pm 0.94
Skull	12.81	11.32 \pm 1.33

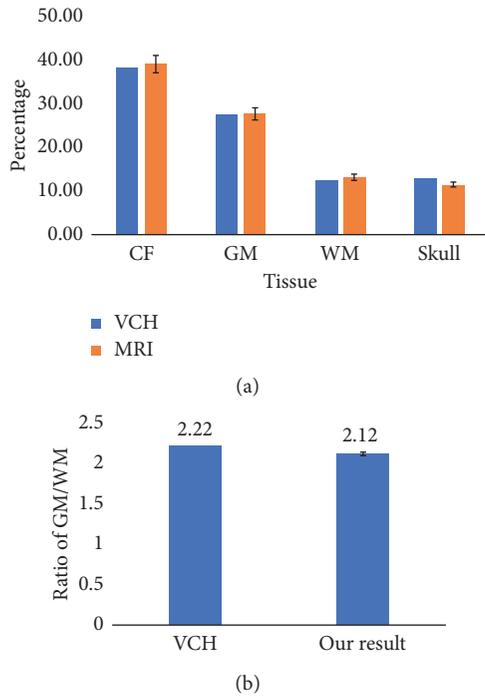


FIGURE 8: (a) A comparison of the percentage of each tissue in the brain between our result and the VCH result. (b) Ratio of grey matter and white matter.

as the GM, which adds to the total percentage of the GM. For further research, we will use the dataset from BrainWeb, an online interface that provides a 3D MRI-simulated brain database [40]. It also provides a fuzzy model for users to estimate the partial volume and is a good way to verify the accuracy of our method.

For the ratio between the grey matter and white matter, we compared our result with the VCH model, along with some studies done by others. Bartlett et al. introduced an interactive segmentation (IS) method to get the volume of the GM and WM from MRI images [41]. Ge et al. investigated the effects of age and sex on the GM and WM volumes by using volumetric MR imaging in healthy adults [42]. The average results are also shown in Table 2.

We also calculated the Jaccard index for each tissue, as shown in Table 3.

Our ground truth is the visible Chinese human. We calculated the coefficient between our result and the ground truth. Our method performs well on segmenting CF, GM, and WM, outperforming the methods from Hasanzadeh et al. [44] and Luo et al. [26].

TABLE 2: The ratio between the grey matter and white matter.

Type	GM/WM
MRI	2.12
VCH	2.22
IS	2.03
Ge	1.5

TABLE 3: Jaccard index of four tissues.

Type	Jaccard index
CSF	0.9431
GM	0.9020
WM	0.9142
Skull	0.8799

3.3. Comparison with FreeSurfer. For further experiment, we compared our result with FreeSurfer. FreeSurfer is the software built for analyzing and visualizing the structural and functional neuroimaging data from cross-sectional or longitudinal studies [43]. The results are shown in Figure 9. Our method has many advantages. FreeSurfer can only locate the WM and GM from an MRI image. Also, it cannot show separate results, except the WM. The image of the WM has many defects: it has many miss-labeling and error-labeling, such as white spots in the image. Our method performs better than this software; we can segment every tissue clearly and also display the results separately.

3.4. Runtime. By introducing parallel computing, we managed to reduce the runtime. We adopt the master-slave mode. One node acts as the master node, which is responsible for data partition and allocation. The other nodes complete the calculation of local data and return the result to the master node. This is important when facing large data. Our result is not significant due to the limitation of the data size. The larger the dataset, the better the result obtained. Figure 10 shows a visualized result of the tendency of the runtime.

4. Discussion and Conclusion

We used image enhancement, operators, and morphometry methods to extract the accurate contours of four tissues: the skull, cerebrospinal fluid (CSF), grey matter (GM), and white matter (WM) on 5 MRI head image datasets. Then, we realize automatic image segmentation with deep learning by using the convolutional neural network. Approaches such as regional, texture, and histogram threshold algorithms and fuzzy c-means (FCM) have limitations in processing time, accuracy, and datasets [16]. In our paper, the percentage of each tissue is calculated, which can be used as a criterion when diagnosing diseases such as cerebral atrophy, often caused by the grey matter or white matter reduction [8]. We also used parallel computing to reduce the runtime.

In our method, the preprocessing step improved the efficiency of the algorithm and the reliability of the segmentation result. Wavelet domain denoising is used to transform noisy signal from time domain to wavelet domain

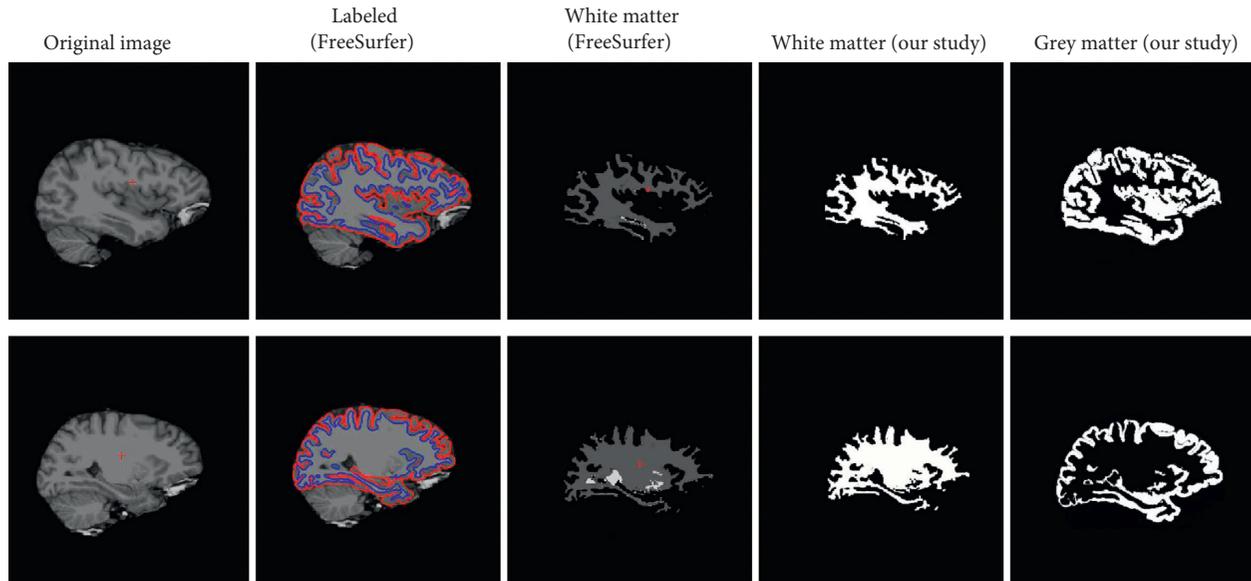


FIGURE 9: Comparison between our method and FreeSurfer.

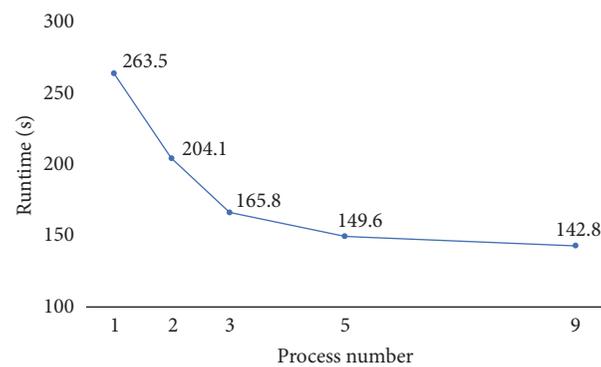


FIGURE 10: Runtime under different numbers of processes.

[10] by using multiscale transformation. We removed the wavelet coefficients of noise from all scales to obtain the wavelet coefficients of signals. This way, the efficiency of the algorithm and the reliability of the segmentation result were improved.

We introduced convolutional neural networks to our work. The CNN consists of two types of layers, a convolution layer and a pooling layer. Multiple feature maps are generated from the convolution layer after convolution. After that, the pixels of each group in the feature map are modified by adding weighted values and offset, along with a sigmoid function to get the feature map in the pooling layer. With multiple convolution and pooling layers, we were able to get more accurate results in less amount of time compared to manual and semiautomatic segmentation. We also used parallel computing to further reduce the runtime of the process. We adopted the master-slave mode by setting one node as the master node, which is responsible for data partition and allocation, along with other nodes to complete the calculation of local data and return the result to the master node.

We compared our results with the data from the visible Chinese human (VCH) head model [34]. The VCH model gives a very good presentation of the anatomical structure and is mainly used in modeling light propagation [35]. The data are collected from high-resolution cryosectional color photographs of a reference adult male. It includes various types of tissues from a standing frozen man body, which includes precise cerebral cortex folding geometry [37]. It gave us an average percentage of each tissue (skull, cerebrospinal fluid, grey matter, and white matter) in the brain. The deviation of our result was less than 2.21%. Another important index is the ratio of GM to WM, which helps in evaluating certain pathological changes of brain. Our result is also very satisfying, with an accuracy rate up to 95%. Our dataset includes 5 human brains with 160 images each. Therefore, the results are convincing. Our work concentrates on the total percentage between the tissues; we did not compare the accuracy of our boundary with other studies. For further research, we will focus more on the comparison between the boundary and use dataset from BrainWeb, which is an online interface that provides 3D MRI-simulated

brain database. It provides a fuzzy model for users to estimate the partial volume and is a good way to verify the accuracy of our method.

Our research has some limitations. First of all, due to limited time, our dataset is not very large. We will increase the quantity and also the variety of the samples, including different races such as black, white, and yellow people in the future. We also need to collect samples from different ages, ranging from the infant, the juvenile, to the elderly. Currently, our dataset includes only adults. With a bigger dataset, we can classify the samples into age, gender, race, and more. We hope to set up a criterion of judgment for medical diagnosing. Researchers and doctors can compare the brain data of a patient with our data and confirm the abnormal proportion of tissues in the brain and further diagnose what disease the patient has.

To conclude, we presented a method to successfully segment the brain tissues from MRI using the convolutional neural network. The percentage results are very close to the average human brain data generated by the VCH model. This is a breakthrough since artificial intelligence and machine learning have become more and more widely used in research. By introducing deep learning into the therapeutic field, the speed and accuracy can be improved. This is because machines can automatically analyze the data, which can be much faster and accurate than the manual and semiautomatic analysis. For future work, we can visualize the contours of the borders of different tissues in 3D so that it can be integrated with optical simulation software such as MCVM for low-level light therapy. Our work has a great potential in the medical field, and we hope that our technique can be a criterion of judgment for diagnosing.

Data Availability

The MRI data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This study was supported by the National Natural Science Fund Project (no. 61675039), the National Key R&D Program of China (nos. 2017YFB1302305 and 2017YFB1300301), the Tianjin Key Project Grant for Natural Science (no. 18JJCZDJC32700), and the CAMS Innovation Fund for Medical Sciences (no. 2016-I2M-3-023).

References

- [1] J. Dolz, N. Betrouni, M. Quidet et al., "Stacking denoising auto-encoders in a deep network to segment the brainstem on MRI in brain cancer patients: a clinical study," *Computerized Medical Imaging and Graphics*, vol. 52, no. 52, pp. 8–18, 2016.
- [2] P. Thirumurugan and P. Shanthakumar, "Brain tumor detection and diagnosis using ANFIS classifier," *International Journal of Imaging Systems and Technology*, vol. 26, no. 2, pp. 157–162, 2016.
- [3] E. R. Kandel, J. H. Schwartz, and T. M. Jessell, *Principles of Neural Science*, United State of America: McGraw-Hill, New York, PA, USA, 4th edition, 2000.
- [4] N. A. Losseff, L. Wang, H. M. Lai et al., "Progressive cerebral atrophy in multiple sclerosis A serial MRI study," *Brain*, vol. 119, no. 6, pp. 2009–2019, 1996.
- [5] N. C. Fox, R. S. Black, S. Gilman et al., "Effects of A immunization (AN1792) on MRI measures of cerebral volume in Alzheimer disease," *Neurology*, vol. 64, no. 9, pp. 1563–1572, 2005.
- [6] P. Demaeral, C. Faubert, G. Wilms, P. Casaer, U. Piepgras, and A. L. Baert, "MR findings in leukodystrophy," *Neuroradiology*, vol. 33, pp. 368–371, 1991.
- [7] T. Kurokawa, Y. Chen, M. Nagata, K. Hasuo, T. Kobayashi, and T. Kitaguchi, "Late infantile krabbe leukodystrophy: MRI and evoked potentials in a Japanese girl," *Neuropediatrics*, vol. 18, no. 3, pp. 182–183, 2008.
- [8] Z. Kong, T. Li, J. Luo, and S. Xu, "Automatic tissue image segmentation based on image processing and deep learning," *Neural Imaging and Sensing*, vol. 63, 2018.
- [9] M. I. Razzak, S. Naz, and A. Zaib, *Deep Learning for Medical Image Processing: Overview, Challenges and Future*, 2017, arXiv preprint arXiv:1704.06825.
- [10] J. Cheng and Y. Zhao, "Brain NMR image segmentation algorithm based on deep learning," *Light Industry Science and Technology*, vol. 8, pp. 93–97, 2017.
- [11] C.-C. Lai, "A novel image segmentation approach based on particle swarm optimization," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. E89-A, pp. 324–327, 2006.
- [12] C. Chu and J. K. Aggarwal, "The integration of image segmentation maps using region and edge information," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 15, pp. 1241–1252, 1993.
- [13] X. Zhou, D. Chen, and S. Yuan, "Medical brain MRI images segmentation by improved fuzzy c-means clustering analysis," *Journal of Jilin University, Engineering and Technology Edition*, 2009.
- [14] B. Miao and Y. Hou, "Medical Image segmentation based on improved fuzzy c-means clustering algorithm," *Laser Journal*, vol. 36, no. 1, 2015.
- [15] V. Srhoj-Egekher, M. J. Benders, M. A. Viergever, and I. Isgum, "Automatic neonatal brain tissue segmentation with MRI," in *Proceedings of SPIE Medical Imaging*, International Society for Optics and Photonics, Bellingham, WA, USA, 2013.
- [16] P. Moeskops, M. A. Viergever, A. M. Mendrik, L. S. de Vries, M. J. N. L. Benders, and I. Isgum, "Automatic segmentation of MR brain images with a convolutional neural network," *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1252–1261, 2016.
- [17] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," in *Proceedings of the IEEE*, vol. 86, 11, pp. 2278–2324, 1998.
- [18] G. E. Hinton, S. Osindero, and Y.-W. Teh, "A fast learning algorithm for deep belief nets," *Neural Computation*, vol. 18, pp. 1527–1554, 2006.
- [19] C. Farabet, C. Couprie, L. Najman, and Y. LeCun, "Learning hierarchical features for scene labeling," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 8, pp. 1915–1929, 2013.
- [20] A. Krizhevsky, I. Sutskever, and G. Hinton, *ImageNet Classification with Deep Convolutional Neural Networks*, NIPS, 2012.

- [21] A. Mohamed, T. N. Sainath, G. E. Dahl et al., "Deep belief networks using discriminative features for phone recognition," in *Proceedings of 2011 IEEE International Conference on IEEE Acoustics, Speech and Signal Processing (ICASSP)*, pp. 5060–5063, Prague, Czech Republic, 2011.
- [22] G. E. Dahl, Y. Dong, D. Li et al., "Large vocabulary continuous speech recognition with context-dependent DBN-HMMS," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 4688–4691, Prague, Czech Republic, 2011.
- [23] Z. Akkus, A. Galimzianova, A. Hoogi, D. L. Rubin, and B. J. Erickson, "Deep learning for brain MRI segmentation: state of the art and future directions," *Journal of Digital Imaging*, vol. 1–11, 2017.
- [24] W. Zhang, R. Li, H. Deng et al., "Deep convolutional neural networks for multi-modality iso-intense infant brain image segmentation," *NeuroImage*, vol. 108, pp. 214–224, 2015.
- [25] L. Yang, Y. Zhang, J. Chen, S. Zhang, and Z. Danny, "Chen suggestive annotation: a deep active learning framework for biomedical image segmentation," 2017, MICCAI arXiv: 1706.04737.
- [26] L. Man, H. Jing, and Y. Feng, "Multimodal 3D convolutional neural networks features for brain tumor segmentation," *Science Technology and Engineering*, vol. 31, 2014.
- [27] C. Huang, "Research of image denoising method about wavelet transform with neighborhood average," *Advanced Materials Research*, vol. 989-994, pp. 4054–4057, 2014.
- [28] S. Wu, Y. Xue, G. Cheng et al., "Research on K-means clustering segmentation method for MRI brain image based on selecting multi-peaks in grey histogram," *Journal of Biomedical Engineering*, vol. 30, no. 6, pp. 1164–1170, 2013.
- [29] M. Havaei, A. Davy, D. Warde-Farley et al., "Brain tumor segmentation with deep neural networks," *Medical Image Analysis*, vol. 35, pp. 18–31, 2017.
- [30] T. Liu, S. Fang, Y. Zhao, P. Wang, and J. Zhang, *Implementation of Training Convolutional Neural Networks*, 2015, arXiv:1506.01195.
- [31] W. Gropp, E. Lusk, N. Doss, and A. Skjellum, "A high-performance, portable implementation of the MPI message passing interface standard," *Parallel Computing*, vol. 22, no. 6, pp. 789–828, 1996.
- [32] S. Sahni and G. Vairaktarakis, "The master-slave paradigm in parallel computer and industrial settings," *Journal of Global Optimization*, vol. 9, no. 3-4, pp. 357–377, 1996.
- [33] I. Sokolinskaya and L. B. Sokolinsky, "Scalability evaluation of NSLP algorithm for solving non-stationary linear programming problems on cluster computing systems," in *Supercomputing. RuSCDays, Communications in Computer and Information Science*, V. Voevodin and S. Sobolev, Eds., vol. 793, Springer, Cham, Switzerland, 2017.
- [34] T. Li, Y. Zhao, Y. Sun, and M. Duan, "Effects of wavelength, beam type and size on cerebral low-level laser therapy by a Monte Carlo study on visible Chinese human," *Journal of Innovative Optical Health Sciences*, vol. 8, no. 1, article 154002, 2014.
- [35] T. Li, Y. Li, Y. Sun, M. Duan, and L. Peng, "Effect of head model on Monte Carlo modeling of spatial sensitivity distribution for functional near-infrared spectroscopy," *Journal of Innovative Optical Health Sciences*, vol. 8, no. 5, article 1550024, 2015.
- [36] T. Li, H. Gong, and Q. Luo, "Visualization of light propagation in visible Chinese human head for functional near-infrared spectroscopy," *Journal of Biomedical Optics*, vol. 16, article 045001, 2011.
- [37] L. Wu, Y. Lin, and T. Li, "Effect of human brain edema on light propagation: a Monte Carlo modeling based on the visible Chinese human dataset," *IEEE Photonics Journal*, vol. 9, no. 5, pp. 1–10, 2017.
- [38] T. Li, C. Xue, P. Wang, and L. Wu, "Photon penetration depth in human brain for light stimulation and treatment: a Realistic Monte Carlo Simulation Study," *Journal of Innovative Optical Health Sciences*, vol. 10, no. 5, article 1743002, 2017.
- [39] T. Li, H. Gong, and Q. Luo, "Visualization of light propagation in visible Chinese human head for functional near-infrared spectroscopy," *Journal of Biomedical Optics*, vol. 16, no. 4, article 045001, 2011.
- [40] C. A. Cocosco, V. Kollokian, R. K.-S. Kwan, and A. C. Evans, "BrainWeb: online interface to a 3D MRI simulated brain Database," in *Proceedings of 3rd International Conference on Functional Mapping of the Human Brain*, Copenhagen, Denmark, May 1997.
- [41] Y. Ge, R. I. Grossman, J. S. Babb, M. L. Rabin, L. J. Mannon, and D. L. Kolson, "Age-related total grey matter and white matter changes in normal adult brain. Part I: volumetric MR imaging analysis," *American Journal of Neuroradiology*, vol. 23, no. 8, pp. 1327–1333, September 2002.
- [42] T. Q. Bartlett, M. W. Vannier, D. W. McKeel Jr., M. Gado, C. F. Hildebolt, and R. Walkup, "Interactive segmentation of cerebral grey matter, white matter, and CSF: photographic and MR images," *Computerized Medical Imaging and Graphics*, vol. 18, no. 6, pp. 449–460, 1994.
- [43] FreeSurfer Wiki, <http://surfer.nmr.mgh.harvard.edu/fswiki>.
- [44] M. Hasanzadeh and S. Kasaei, "Multispectral brain MRI segmentation using genetic fuzzy systems," in *Proceedings of 9th International Symposium on Signal Processing and Its Applications*, pp. 1–4, Sharjah, UAE, February 2007.

Research Article

Cloud-Based Brain Magnetic Resonance Image Segmentation and Parcellation System for Individualized Prediction of Cognitive Worsening

Ryo Sakamoto ^{1,2}, Christopher Marano,^{3,4} Michael I. Miller,^{5,6} Constantine G. Lyketsos ³, Yue Li ⁶, Susumu Mori,^{1,6,7} Kenichi Oishi ¹, and Alzheimer's Disease Neuroimaging Initiative ADNI⁸

¹Department of Radiology, Johns Hopkins University School of Medicine, Baltimore, MD, USA

²Department of Radiology, Kyoto University School of Medicine, Kyoto, Japan

³Department of Psychiatry and Behavioral Sciences, Johns Hopkins Bayview and Johns Hopkins University, Baltimore, MD, USA

⁴Division of Geriatric Psychiatry, University of Maryland School of Medicine, Baltimore, MD, USA

⁵Center for Imaging Science, School of Engineering, Johns Hopkins University, Baltimore, MD, USA

⁶AnatomyWorks, LLC, Baltimore, MD, USA

⁷F. M. Kirby Research Center for Functional Brain Imaging, Kennedy Krieger Institute, Baltimore, MD, USA

⁸Alzheimer's Disease Neuroimaging Initiative Study, USA

Correspondence should be addressed to Kenichi Oishi; koishi2@jhmi.edu

Received 8 August 2018; Revised 24 October 2018; Accepted 5 November 2018; Published 29 January 2019

Academic Editor: Norio Iriguchi

Copyright © 2019 Ryo Sakamoto et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

For patients with cognitive disorders and dementia, accurate prognosis of cognitive worsening is critical to their ability to prepare for the future, in collaboration with health-care providers. Despite multiple efforts to apply computational brain magnetic resonance image (MRI) analysis in predicting cognitive worsening, with several successes, brain MRI is not routinely quantified in clinical settings to guide prognosis and clinical decision-making. To encourage the clinical use of a cutting-edge image segmentation method, we developed a prediction model as part of an established web-based cloud platform, MRICloud. The model was built in a *training dataset* from Alzheimer's Disease Neuroimaging Initiative (ADNI) where baseline MRI scans were combined with clinical data over time. Each MRI was parcellated into 265 anatomical units based on the MRICloud fully automated image segmentation function, to measure the volume of each parcel. The Mini Mental State Examination (MMSE) was used as a measure of cognitive function. The normalized volume of 265 parcels, combined with baseline MMSE score, age, and sex were input variables for a Least Absolute Shrinkage and Selection Operator (LASSO) regression analysis, with MMSE change in the subsequent two years as the target for prediction. A leave-one-out analysis performed on the training dataset estimated a correlation coefficient of 0.64 between true and predicted MMSE change. A receiver operating characteristic (ROC) analysis estimated a sensitivity of 0.88 and a specificity of 0.76 in predicting substantial cognitive worsening after two years, defined as MMSE decline of ≥ 4 points. This MRICloud prediction model was then applied to a *test dataset* of clinically acquired MRIs from the Johns Hopkins Memory and Alzheimer's Treatment Center (MATC), a clinical care setting. In the latter setting, the model had both sensitivity and specificity of 1.0 in predicting substantial cognitive worsening. While the MRICloud prediction model demonstrated promise as a platform on which computational MRI findings can easily be extended to clinical use, further study with a larger number of patients is needed for validation.

1. Introduction

Cognitive disorders and dementia, heterogeneous conditions that include various brain diseases, are common in old age. Regardless of the diagnosis, one of the greatest stressors

for dementia patients and caregivers is future uncertainty surrounding change progression of their condition. In clinical care settings, including memory clinics, medical providers make the best possible clinical diagnosis to inform the patient and caregivers about future progression, the type

of care needed, problems that might occur in the future, and how to prevent or ameliorate these problems [1]. Accurate, individualized ways to predict cognitive progression, ones that could easily be applied in typical clinical settings, would be a great advance with huge clinical benefit for patients and caregivers.

Many neuroanatomical predictors of cognitive decline have been identified in previous studies of cognitively normal or cognitively impaired elders. These best imaging predictors involve atrophy in selected brain structures, such as areas of the mesial and lateral temporal lobes, the posterior cingulate, the orbitofrontal gyri, and white matter hyperintensity [2–4]. Non-neuro-anatomical predictors [5], such as age—with younger having worse prognosis [6], baseline cognitive function [7–9], and vascular risk factors [10], are well established. However, individualized prediction (*precision medicine*) of future decline based on individual variables is difficult because each factor only weakly correlates with the outcome, and often overlaps and interacts with other factors. Integration of these variables is, thus, essential for accurate prediction [11–14]. For such integration, important features that relate to disease progression must be extracted and recorded in a standardized quantitative manner. However, a widely available platform that could enable such quantification and integration has not been established.

In an initial attempt to develop clinically useful individualized prediction, we used a web-accessible, cloud-based platform MRICloud (<https://mricloud.org/>) [15] to achieve image standardization, quantification, and cross-variable integration. An integrative cloud platform is the critical enabling technology for the proposed prediction as it requires a large amount of atlas resources and intensive computation. While previous research has succeeded in predicting cognitive outcomes for research populations [3, 13, 16, 17], one of the greatest barriers to clinical application is fragmentation of analysis pipelines, which require aggregation of various tools with different capabilities on different platforms to extract a prediction value from a single image. MRICloud provides for seamless integration of whole-brain segmentation and subsequent prediction. Moreover, it enables users to develop their own applications using its programming interfaces. For example, users can implement their own data processing and analysis pipeline in MRICloud. This flexibility is particularly important in the era of machine-learning and artificial intelligence where rapid improvement, advancement, and innovation are expected.

Heterogeneity in symptoms and comorbidities encountered in clinical practice, compared to research populations with strict inclusion and exclusion criteria, is one of the major causes for which research discovery has not made its way to clinical application, as pointed out by [14]. In the era of “big data” science, the application of a trained algorithm to *real-world data* is necessary to validate its usefulness in day to day clinical practice. The MRICloud platform is promising for such a validation study since it enables the collection of raw MRIs from all over the world through its web interface. To test this concept, we applied a prediction model, developed on a research cohort, to a real clinical population of cognitively impaired patients to

investigate the applicability of the prediction model to a heterogeneous clinical setting.

2. Methods

2.1. Training Dataset. Data used in the preparation of this article were obtained from Alzheimer’s Disease Neuroimaging Initiative (ADNI) database (<https://adni.loni.usc.edu>). ADNI was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner, MD. The primary goal of ADNI has been to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early Alzheimer’s disease (AD). The data were analyzed anonymously, using publicly available secondary data. Therefore, no specific ethics approval was required for this work.

A total of 402 individuals available from the ADNI-1 database, with corresponding baseline MRIs and two years of follow-up data, were analyzed (AD = 75, MCI = 176, and cognitively normal individuals = 151). Structural MRIs were acquired from 1.5T scanners with a protocol individualized for each scanner, as defined in <http://adni.loni.usc.edu/>. The MRIs were downloaded from <https://ida.loni.usc.edu/> in NiFTI formats, with geometry distortion correction and B1 correction. The demographics and characteristics of the study population are in Table 1.

2.2. Test Dataset. Imaging and clinical data acquired as part of clinical care in the Johns Hopkins Memory and Alzheimer’s Treatment Center (MATC) located at the Bayview Medical Center were used as the test dataset. The MATC dataset consists of patients with memory problems, self-referred or referred by other physicians, evaluated in an outpatient memory disorders clinic. The creation and use of the database occurred under oversight by the Johns Hopkins Institutional Review Board, which provided waiver of consent, as the data were all collected in clinical care [18]. This patient cohort was heterogeneous, with various etiologies and levels of severity, and represented diverse people with memory problems. A total of 17 patients with baseline MRIs and two years of clinical follow-up data were analyzed. MRI scans, acquired for clinical care only, followed the ADNI protocol: a three-dimension (3D), magnetization-prepared, rapid gradient-echo sequence, with a repetition time of 2300 ms, an echo time of 2.98 ms, and a voxel resolution of $1 \times 1 \times 1$ mm, scanned on a 3T scanner (SIEMENS Vario). The demographics and characteristics of the selected population are given in Table 2.

2.3. Image Processing. A multiatlas label fusion method in which an entire brain is automatically parcellated into 265 anatomical units [19] was applied to the MRIs. This is a fully automated method that is open to public use through our website (<https://mricloud.org/>). All MRIs were bias-corrected and linearly aligned to the JHU-MNI atlas [20]

TABLE 1: Demographics and characteristics of the training dataset (ADNI1).

Diagnosis at baseline	<i>N</i>	Age	Sex (men/women)	MMSE at baseline	MMSE after 2 years	MMSE decline in 2 years	Number of patients with substantial worsening*
CN	151	75.5 ± 5.0	80/71	29.2 ± 1.0	29.0 ± 1.2	-0.2 ± 1.3	1
MCI	176	73.4 ± 7.1	113/63	27.2 ± 1.7	25.4 ± 3.9	-1.8 ± 3.3	46
AD	75	73.8 ± 7.4	34/41	23.3 ± 2.0	19.0 ± 5.6	-4.3 ± 5.3	36
Total	402	74.9 ± 6.5	227/175	27.2 ± 2.6	25.6 ± 5.1	-1.6 ± 3.6	83

*Substantial worsening: MMSE declines ≤ -4 within two years.

TABLE 2: Demographics and characteristics of the test dataset (MATC database).

Suspected diagnosis at baseline	<i>N</i>	Age	Sex (men/women)	MMSE at baseline	MMSE after 2 years	MMSE decline in 2 years	Number of patients with substantial worsening*
AD	7	73.7 ± 11.0	4/1	21.9 ± 4.5	16.9 ± 8.0	-5.0 ± 5.8	4
MCI	3	82.5 ± 2.9	1/2	29.0 ± 0.8	26.3 ± 3.9	-2.7 ± 4.6	1
Others	7	69.8 ± 10.1	3/4	23.1 ± 7.0	20.6 ± 10.1	-2.6 ± 3.8	2
Total	17	73.7 ± 10.7	8/9	23.6 ± 6.0	20.0 ± 9.1	-3.6 ± 5.0	7

Others: mixed dementia, 2; vascular dementia, 1; frontotemporal dementia, 2; nonspecific cognitive disorder with depression, 2. *Substantial worsening: MMSE declines ≤ -4 within two years.

space. Atlases were warped to the linearly aligned subject image using Large Deformation Diffeomorphic Metric Mapping (LDDMM) [21], followed by application of the multiatlas fusion algorithm [19]. The JHU T1 Geriatric Multi-Atlas Inventory V5 [22, 23], designed for older patient populations with potential brain atrophy, was used as a set of atlases. The volume of each parcel was measured and normalized based on the whole-brain volume.

2.4. Neuroimage Features. One of the challenges of atlas-based image analysis is the granularity of the anatomical parcellation map used to quantify brain MRI [24–26]. The statistical power to characterize anatomical features related to cognitive decline is maximized when the size and shape of the parcel exactly follow the pathological locations that determine the prognosis. If there is an *a priori* hypothesis about the distribution of the pathologic tissues, predefinition of the size and shape of parcels can follow the hypothesis. However, impaired anatomical structures and spatial distribution depend on the disease severity seen in the brain MRI, which are unknown prior to the analysis. To account for the heterogeneity of this patient population, we applied a tool that can flexibly change the granularity level based on the hierarchical relationships of 254 structures defined in our atlas [27], in which the 254 structures were assigned a hierarchical relationship based on their ontological relationship [28, 29]. This relationship consists of five hierarchical levels that were named: Level 1 (11 parcels); Level 2 (17 parcels); Level 3 (36 parcels); Level 4 (54 parcels); and Level 5 (254 parcels) (Figure 1). The volumes of all structures in each level were obtained.

2.5. Non-Image Features. Many non-image predictors of cognitive decline have been identified in previous studies of cognitively normal or cognitively impaired elders. These can be summarized into seven major categories [5]:

sociodemographics; clinical characteristics; cognitive or neuropsychological features; behavioral or psychological factors; cardiovascular risk factors; genetics; and biological markers. Although inclusion of all these factors into the prediction model might maximize the prediction accuracy, obtaining and recording such information comprehensively in a structured way is not a feasible standard for the vast majority of clinical care settings. We attempted to balance this by focusing on (preselecting) predictor variables available in pretty much all clinical care setting: age, sex, and MMSE score at the time of the MRI scan.

2.6. Prediction Algorithm. A Least Absolute Shrinkage and Selection Operator (LASSO) regression analysis [30, 31] was estimated to predict worsening in cognition two years after the baseline. The LASSO was chosen to address the multicollinearity problem and to identify important predictors. The MMSE measured cognitive function, and change in MMSE was set as the target for prediction. Among various cognitive measures, the MMSE was chosen because it is most commonly measured in patients who visit memory clinics and can easily be also measured in primary care settings. The LASSO regularization weight parameter λ was selected in order to minimize a mean squared prediction error between the measured and the predicted Δ MMSE obtained by leave-one-out cross validation. The normalized volume of anatomical units, baseline MMSE, age, and sex were used as a set of variables.

2.7. Validation. To validate the prediction algorithm within the ADNI population, a leave-one-out analysis was performed to investigate the correlation between the measured and the predicted MMSE change. The analyses were performed based on level of granularity (Levels 1–5) and all levels combined. The predictor model that provided the best

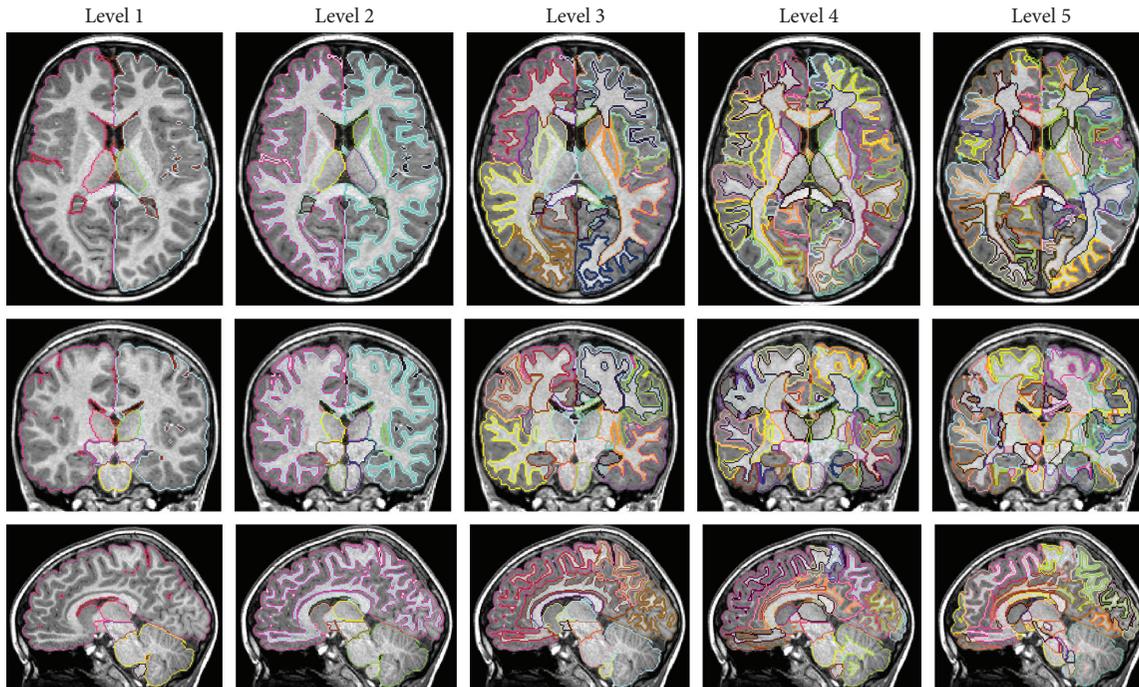


FIGURE 1: Hierarchical relationships of 254 structures defined in the MRICloud.

correlation with cognitive outcome in ADNI data was then applied to the clinical population.

There is considerable debate about the disease-specific speed of cognitive worsening [32, 33], although a meta-analysis suggested a similar pace of cognitive decline in two of the most common diagnoses of dementias, AD and dementia with Lewy bodies [34]. Faster cognitive decline is seen in autopsy-confirmed frontotemporal dementia (FTD) compared to AD [33]. Within the AD population, there are at least two distinctive types of disease progression, slow and rapid [7, 35–37]. Since we aimed to identify rapid progressors in the clinical cohort of mild cognitive impairment or mild dementia ($\text{MMSE} \geq 23$) at baseline, and majority of the ADNI training data fall into this category (Table 1), we used MMSE decline of ≥ 4 points after two years to define substantial cognitive worsening, according to the reported average decline in MMSE score over two years (3.3 points decline in MMSE score) in the mild AD population with a baseline MMSE score of 23 [38]. To investigate the accuracy to predict substantial cognitive worsening, ROC analysis was performed. The cutoff of the predicted MMSE change that maximized the Youden index [39] was adopted as the threshold to predict substantial cognitive worsening (true MMSE decline of ≥ 4 points). The glmnet package [31] implemented in R software [<http://cran.r-project.org>, R Core Team, version 3.2.3] was used for the LASSO regression analysis and the ROC curve analysis.

2.8. Implementation of the Prediction Model into the MRICloud. The MRICloud provides a cloud-based architecture for neuroimage analysis tools through the web. It has three components: storage; computation; and applications. It also provides visualization applications and enables users

to develop their own application with the application programming interfaces (API). It provides low-barrier access to the algorithms and tools and accommodates high throughput, as well as parallel computation, to render intensive computations tractable. The prediction model developed was integrated into the MRICloud using the API, which allows users to upload their own image and non-image data to predict the MMSE change.

2.9. Application of the Prediction Algorithm to the Clinical Population. The prediction model implemented in MRICloud was applied to the clinical cohort for individualized prediction of the MMSE change over the two years after the initial evaluation. A correlation between true and predicted MMSE change, as well as sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) to predict substantial cognitive worsening were calculated.

3. Results

3.1. Generation of Prediction Model. A leave-one-out analysis of the training dataset estimated a larger correlation coefficient between true and predicted MMSE change at the highest granularity level (level 5, $R = 0.635$) and at all levels combined ($R = 0.636$) (Figure 2(a)) compared to the lower granularity levels (Level 1, $R = 0.464$, Level 2 = 0.489, Level 3 = 0.527, and Level 4 = 0.599). The ROC analysis (Figure 2(b)) using an applied regression model estimated areas under the curve (AUC) of 0.898 (Level 5, 95% CI: 0.862–0.935) or 0.899 (all levels combined, 95% CI: 0.862–0.936) for the prediction of substantial cognitive worsening. Therefore, the prediction model using local brain volumes of all granularity levels combined was used in later

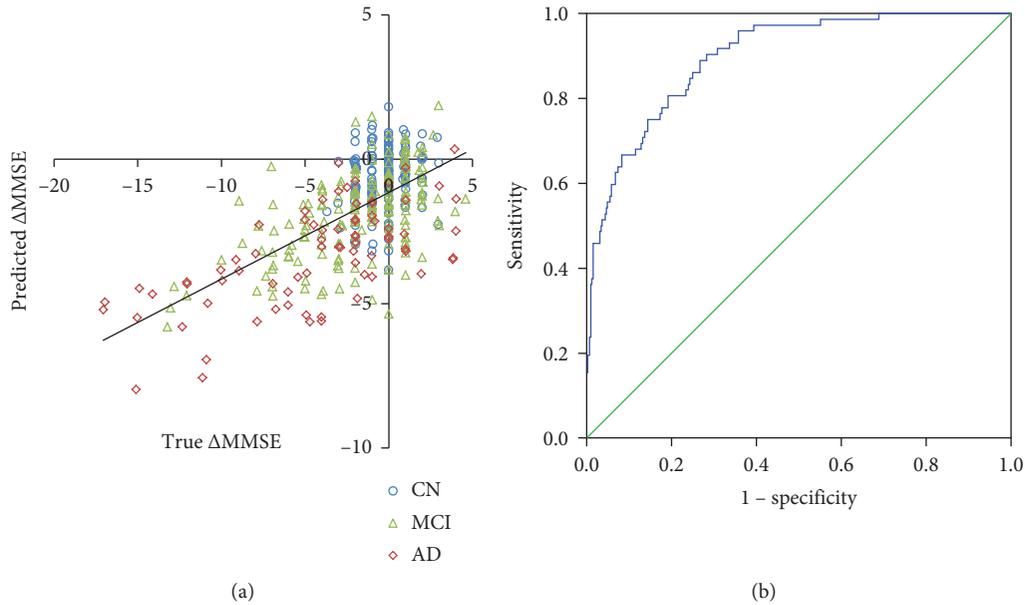


FIGURE 2: (a) Scatter plot showing the relationship between true and predicted MMSE change based on the training dataset. (b) ROC analysis showing the relationship between sensitivity and specificity to predict substantial cognitive worsening. CN: cognitively normal, MCI: mild cognitive impairment, and AD: Alzheimer’s disease.

analyses. The optimal cut of value for the prediction of substantial cognitive decline, calculated from the ROC curve, was a predicted MMSE decline of -1.9 , which demonstrated sensitivity of 0.847 , specificity of 0.779 , PPV of 0.455 , and NPV of 0.959 for substantial cognitive worsening. The mean absolute difference between the true and predicted MMSE change was 2.0 ($SD = 2.0$).

3.2. Contribution of Each Factor to the Prediction. Among 402 ADNI participants, 83 (21%) were rapid progressors (Table 1). Thirty-one factors (volumes of 29 anatomical areas, baseline MMSE score, and age, but no sex) were selected through the LASSO regression analysis using all granularity levels combined and built into the final prediction model. The standardized regression coefficients for each of the image and non-image factors are in Table 3. Regression coefficients for selected anatomical structures are color-coded and overlaid on the JHU-MNI atlas (Figure 3). Atrophy in the bilateral middle temporal gyri, the claustrum complex, the superior parietal white matter, expansion of the Sylvian fissure, and the lower baseline MMSE score and younger age predicted greater likelihood being a rapid progressor by MMSE change score.

3.3. Implementation of the Model into the MRICloud. The model was then implemented into the MRICloud. This enables external users to obtain access to the prediction model for additional validation. Users can visit the website (<https://www.mricloud.org>) to log in to the “BrainGPS” module and then select the “T1 segmentation” tool listed in the upper row. This allows users to upload their own ADNI-compatible, high-resolution, 3D-MPRAGE images, MMSE

score, and age at scan to obtain the predicted Δ MMSE in two years (Figure 4).

3.4. Application to the Clinical Population. Among 17 MATC patients, 7 (41%) were rapid progressors (Table 2). A graph showing the correlation ($R = 0.69$) between the actual and predicted MMSE change is shown in Figure 5. All patients with substantial cognitive worsening were accurately predicted when a predicted MMSE decline of -1.9 was applied. The mean absolute difference between the true and predicted MMSE change was 3.4 (standard deviation = 2.5).

4. Discussion

4.1. Generation of the Prediction Model. In current clinical practice for patients with cognitive disorders, the main role of structural MRI is to exclude causes such as neoplasms, hydrocephalus, trauma, or ischemic disease that are clearly visible on a scan. In clinical research, MRI is used as one of the measures of neurodegeneration caused by AD pathology [40]. Our results suggest a substantial contribution of local neuro-anatomical changes to predict cognitive worsening in addition to non-image features. The prediction model trained by LASSO indicated younger age and lower baseline cognitive function were related to faster deterioration in cognitive function, consistent with prior studies [7, 41]. The final model did not include sex as a predictor, suggesting that sex has little effect on rate of cognitive decline, as expected [42]. Since our goal was to investigate a robust index for individualized prediction from information available through clinical practice not necessarily obtained by specialists, the inclusion of only two simple non-image factors—age and MMSE—has practical significance.

TABLE 3: List of the standardized regression coefficients for each of the image and non-image factors, obtained from the LASSO regression analysis.

Factors	Standardized regression coefficients
Age	0.606014
Middle temporal gyrus, left (Level 5)	0.555413
Clastrum, right (Level 5)	0.33835
Baseline MMSE score	0.221676
External capsule, right (Level 5)	0.203446
Temporal lobe, left (Level 4)	0.184202
Angular gyrus, right (Level 4)	0.179487
Superior parietal white matter, left (Level 5)	0.147153
Angular gyrus, left (Level 4)	0.146566
Fimbria, left (Level 5)	0.131843
Middle temporal gyrus, right (Level 5)	0.119239
Inferior occipital gyrus, left (Level 4)	0.09018
Middle frontal gyrus, left (Level 5)	0.083102
Posterior cingulate cortex, right (Level 5)	0.080805
Posterior cingulate cortex, left (Level 5)	0.074983
Superior parietal white matter, right (Level 5)	0.068838
Middle occipital gyrus, left (Level 4)	0.060132
Superior frontal gyrus, left (Level 5)	0.052995
Fimbria, right (Level 5)	0.045444
Inferior deep parietal white matter, left (Level 4)	0.014723
Peripheral frontal white matter, right (Level 4)	$-3.00E - 05$
Inferior frontal white matter, right (Level 5)	-0.01381
Lateral frontoorbital gyrus white matter (Level 5)	-0.03149
Lingual gyrus white matter, right (Level 5)	-0.04243
Postcentral gyrus white matter, left (Level 5)	-0.07299
Postcentral gyrus, right (Level 4)	-0.08223
Fornix, right (Level 4)	-0.08814
Dorsal anterior cingulate cortex, right (Level 5)	-0.10213
Gyrus rectus white matter, left (Level 5)	-0.16791
Mammillary body, right (Level 5)	-0.25573
Sylvian fissure and temporal sulcus, right (Level 4)	-0.29769

4.2. *Brain Areas Important for the Prediction.* Three anatomical areas—middle temporal gyrus, peri-Sylvian, and parietal areas—were the areas that contributed to prediction of cognitive worsening. These areas did not necessarily include anatomical structures that are known to be most affected in early AD, such as medial temporal area, but rather, overlapped with local brain areas whose volumes correlated with MMSE score, such as the temporal, the middle frontal, and the left angular and supramarginal gyri [43–45]. Previously reported correlations between enlargement of the Sylvian fissure and MMSE score in AD and MCI [46] support our findings. These results suggest that the mesial

temporal atrophy, while correlated with the MMSE score at the baseline [43, 47–49] is not necessarily the best predictor of cognitive worsening in coming years.

4.3. *Clinical Validation Study Performed on the MRICloud Platform.* Although the MATC population consisted of patients with various clinical diagnoses or with mixed causes of cognitive decline, the sensitivity and specificity to predict rapid cognitive decline was comparable to that obtained from an ADNI cohort with strict inclusion and exclusion criteria. This was unexpected because we initially assumed that the prediction model trained from the ADNI dataset would be applicable only to patients with a clinical diagnosis of probable AD. One possible explanation is the existence of a spatial pattern in local brain atrophy related to an increased risk of cognitive worsening, which is less specific to AD, but rather common to various types of dementias. The applicability of the model to nonAD dementias, such as frontotemporal dementia, dementia with Lewy bodies, or vascular dementia, needs to be further investigated.

We also noted overestimated prediction of MMSE decline for patients with a low baseline MMSE score ($MMSE < 23$). This was probably due to the relative imprecision of MMSE in measuring cognitive decline in demented patients [50]. This explains the results of the ROC curve analysis, in which the cutoff predicted MMSE change of -1.9 was higher than the true actual change in MMSE of -4 . Therefore, care should be taken when the predicted change is less than -1.9 ; in this case, the practical interpretation should be a “MMSE change less than -4 can be predicted.”

Other limitations include the small clinical sample size, which may have involved selection bias or confounding factors, such as age. Indeed, this limitation was one of our motivations to develop an open-access platform for future validation study, in which MRIs from other institutes can be easily submitted to increase the number of test MRIs, as detailed in the next section. The method used to select the LASSO regularization parameter could also be an issue, since inappropriate selection leads to overfitting of training data; therefore, the performance measures for the training data (ADNI data) were possibly upward-biased. Application of the nested cross validation is one possibility to remain unbiased in the cross validation-based evaluation, where the parameters are selected by cross validation [17].

4.4. *MRICloud as a Platform for Algorithm Development and Clinical Validation Studies.* External and/or clinical validation of prediction models developed in research is not easy. Such models are typically published in print with only the theoretical aspects, such as functions and variables, elaborated. However, practical implementation of the models depends on the computational environment, the operating systems used, software versions, methods to extract variables, and coding of the functions, all of which have an impact on the reproducibility of the research findings. There have been efforts to distribute the open-source scripts

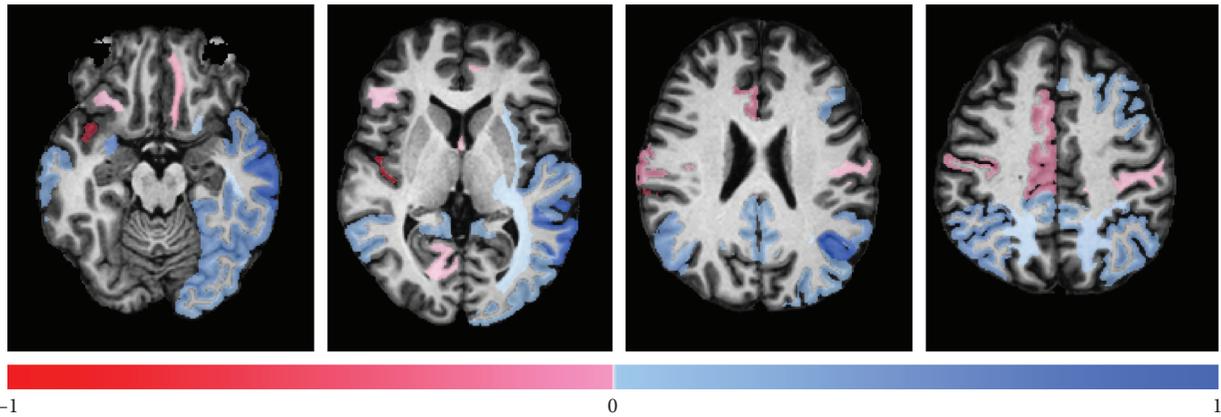


FIGURE 3: Regression coefficients of the selected anatomical structures are color-coded (blue: positive regression coefficient and red: negative regression coefficient) and overlaid on the JHU-MNI atlas.

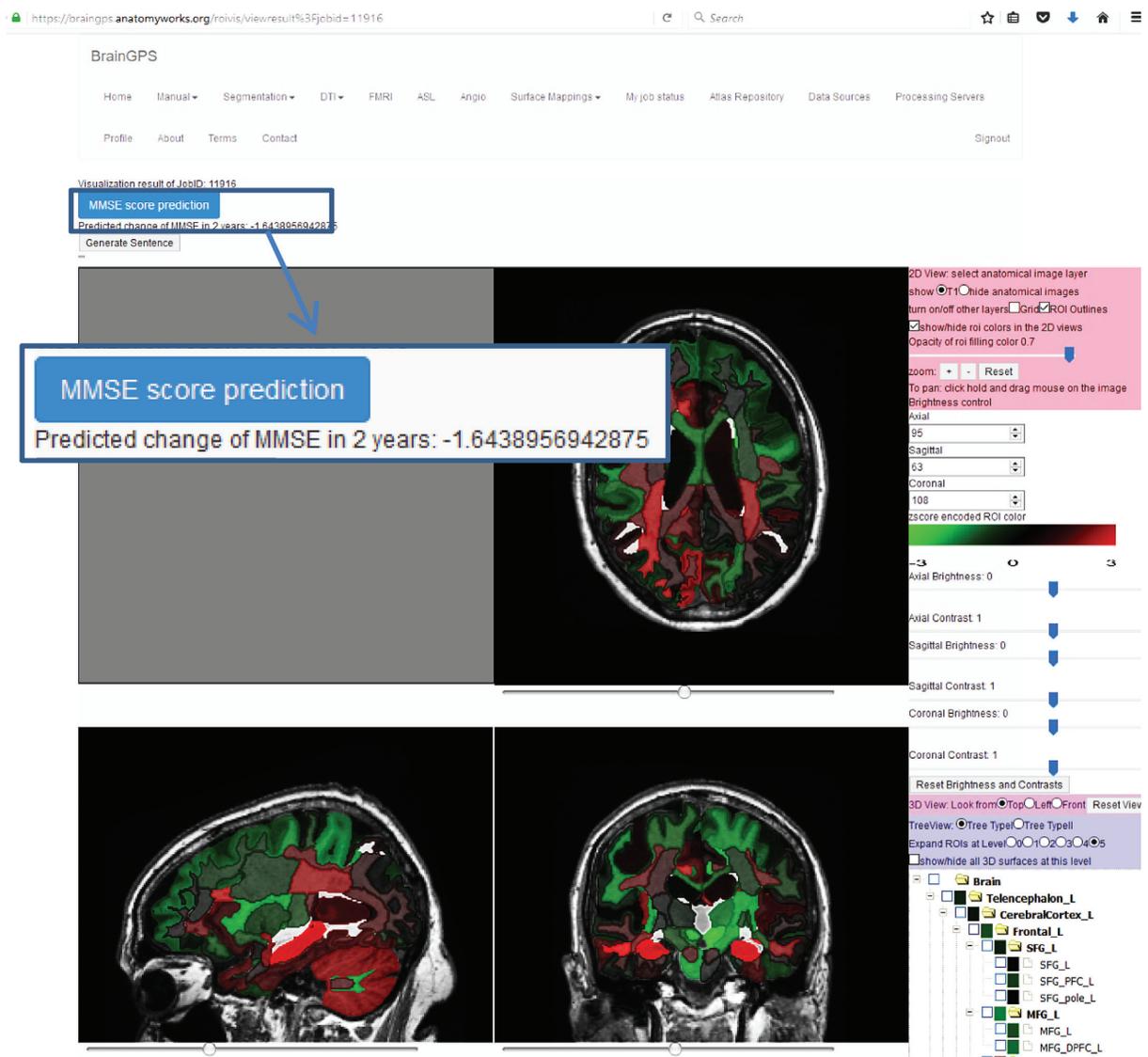


FIGURE 4: Screenshot of the MMSE score prediction function implemented in the BrainGPS module of the MRICloud. The module allows users to submit their own high-resolution, 3D, T1-weighted images with the age and MMSE score at scan. The module provides a color-coded z-score map of the local volume (lower left) as well as the predicted Δ MMSE (magnified view in the blue rectangle).

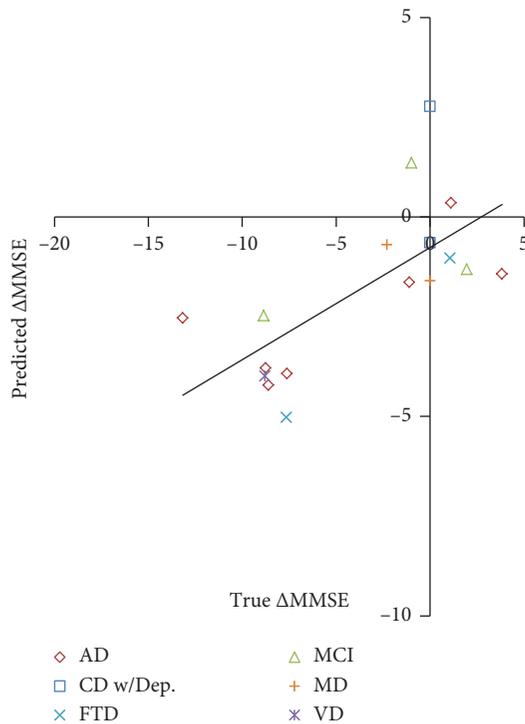


FIGURE 5: Results of the fully automated MMSE change prediction using the BrainGPS module of the MRICloud, compared to the true MMSE change. AD: Alzheimer's disease, MCI: mild cognitive impairment, CD w/Dep: nonspecific cognitive disorder with depression, FTD: frontotemporal dementia, MD: mixed dementia, and VD: vascular dementia.

or software packages as the solutions for effective deployment of new models. While successful, there are several limitations in this approach. First, it places a considerable amount of the burden on developers, which includes program development for cross platforms and reengineering after a version change of the operating system, efforts to redistribute after a version-update, management of users with different versions, and inquiries from other developers about the content of the source codes. The cloud-based software-as-a-service (SaaS) model has emerged as a solution to these barriers for cross-program communications, platform independence, and efficient computation strategies. The field of machine-learning and artificial intelligence is developing rapidly, and we expect algorithms that will replace LASSO, in the near future. Currently available algorithms, such as elastic net, are already demonstrating excellent performance in MRI feature selection [51]. The SaaS through API, which allows users to implement their novel image analysis algorithm to be shared and tested with users, seems to be one of the best solutions computer science can offer at this moment. The MRICloud provided a user-friendly environment to share the prediction model with external users for rigorous validation.

5. Conclusion

Our results indicated the potential to apply results from a study population to clinical practice, at least in a limited

venue, such as the MATC, but further study with a larger number of patients is needed to characterize the features of cognitive worsening. The MRICloud provided a user-friendly environment suitable for multi-institutional clinical validation studies to predict future cognitive worsening from image and non-image data.

Data Availability

The ADNI dataset is downloadable through the website <http://adni.loni.usc.edu/>. MRICloud is available for registered users (<https://mricloud.org/>). Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of the ADNI and/or provided data but did not participate in the data analysis or in the writing of this report. A complete listing of ADNI investigators can be found at http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf.

Disclosure

Parts of the preliminary results of this study were presented at the 22nd Annual Meeting of the Organization for Human Brain Mapping. The contents of this paper are solely the responsibility of the authors and do not necessarily represent the official view of NIH or the Department of Radiology of the Johns Hopkins University School of Medicine.

Conflicts of Interest

Dr. Oishi is a consultant for AnatomyWorks. Dr. Mori is a Chief Executive Officer of AnatomyWorks. The terms of these arrangements are being managed by the Johns Hopkins University in accordance with its conflicts of interest policies. Dr. Lyketos has received support from the following organizations: Associated Jewish Federation of Baltimore, Weinberg Foundation, Forest, Glaxo-Smith-Kline, Eisai, Pfizer, Astra-Zeneca, Lilly, Ortho-McNeil, Bristol-Myers, and Novartis. CGL has served as a consultant/advisor for Astra-Zeneca, Glaxo-Smith-Kline, Eisai, Novartis, Forest, Supernus, Adlyfe, Takeda, Wyeth, Lundbeck, Merz, Lilly, and Genentech. CGL has received honorarium or travel support from Pfizer, Forest, Glaxo-Smith-Kline, and Health Monitor. This arrangement has been approved by the Johns Hopkins University in accordance with its conflicts of interest policies.

Acknowledgments

We would like to acknowledge support for the statistical analysis from Dr. Martin A. Lindquist and thank Ms. Mary McAllister for help with manuscript editing. Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI was funded by the National Institute on Aging, by the National Institute of Biomedical Imaging and

Bioengineering, and through generous contributions from AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd. and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research provided funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (<https://www.fnih.org>). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California. This publication was made possible by grants R01HD065955, R44 NS078917, and P50AG005146 from the National Institutes of Health; the Pilot Project Discovery Program from the Johns Hopkins Individualized Health Initiative (inHealth); and the Fakhri Rad BriteStar award from the Department of Radiology Johns Hopkins University School of Medicine.

References

- [1] P. Rabins, C. Lyketsos, and C. Steele, *Practical Dementia Care*, Oxford University Press, New York, NY, USA, 3rd edition, 2016.
- [2] C. Fennema-Notestine, L. K. McEvoy, D. J. Hagler Jr., M. W. Jacobson, A. M. Dale, and The Alzheimer's Disease Neuroimaging, "Structural neuroimaging in the detection and prognosis of pre-clinical and early AD," *Behavioural Neurology*, vol. 21, no. 1-2, pp. 3-12, 2009.
- [3] L. K. McEvoy, C. Fennema-Notestine, J. C. Roddey et al., "Alzheimer disease: quantitative structural neuroimaging for detection and prediction of clinical and structural changes in mild cognitive impairment," *Radiology*, vol. 251, no. 1, pp. 195-205, 2009.
- [4] M. Mortamais, S. Artero, and K. Ritchie, "Cerebral white matter hyperintensities in the prediction of cognitive decline and incident dementia," *International Review of Psychiatry*, vol. 25, no. 6, pp. 686-698, 2013.
- [5] M. E. Soto, S. Andrieu, C. Arbus et al., "Rapid cognitive decline in Alzheimer's disease. Consensus paper," *The Journal of Nutrition, Health & Aging*, vol. 12, no. 10, pp. 703-713, 2008.
- [6] D. Peng, Z. Shi, J. Xu et al., "Demographic and clinical characteristics related to cognitive decline in Alzheimer disease in China: a multicenter survey from 2011 to 2014," *Medicine*, vol. 95, no. 26, p. e3727, 2016.
- [7] R. S. Doody, P. Massman, and J. K. Dunn, "A method for estimating progression rates in Alzheimer disease," *Archives of Neurology*, vol. 58, no. 3, pp. 449-454, 2001.
- [8] J. A. Pillai, A. Bonner-Jackson, E. Walker, L. Mourany, and J. L. Cummings, "Higher working memory predicts slower functional decline in autopsy-confirmed Alzheimer's disease," *Dementia and Geriatric Cognitive Disorders*, vol. 38, no. 3-4, pp. 224-233, 2014.
- [9] S. Marquis, M. M. Moore, D. B. Howieson et al., "Independent predictors of cognitive decline in healthy elderly persons," *Archives of Neurology*, vol. 59, no. 4, pp. 601-606, 2002.
- [10] F. F. Oliveira, E. S. Chen, M. C. Smith, and P. H. Bertolucci, "Predictors of cognitive and functional decline in patients with Alzheimer disease dementia from Brazil," *Alzheimer Disease & Associated Disorders*, vol. 30, no. 3, pp. 243-250, 2016.
- [11] E. Moradi, A. Pepe, C. Gaser, H. Huttunen, and J. Tohka, "Machine learning framework for early MRI-based Alzheimer's conversion prediction in MCI subjects," *NeuroImage*, vol. 104, pp. 398-412, 2015.
- [12] C. Salvatore, A. Cerasa, and I. Castiglioni, "MRI characterizes the progressive course of AD and predicts conversion to Alzheimer's dementia 24 Months before probable diagnosis," *Frontiers in Aging Neuroscience*, vol. 10, p. 135, 2018.
- [13] A. Hall, M. Munoz-Ruiz, J. Mattila et al., "Generalizability of the disease state index prediction model for identifying patients progressing from mild cognitive impairment to Alzheimer's disease," *Journal of Alzheimer's Disease*, vol. 44, no. 1, pp. 79-92, 2015.
- [14] S. Kloppel, J. Peter, A. Ludl et al., "Applying automated MR-based diagnostic methods to the memory clinic: a prospective study," *Journal of Alzheimer's Disease*, vol. 47, no. 4, pp. 939-954, 2015.
- [15] S. Mori, D. Wu, C. Ceritoglu et al., "MRICloud: delivering high-throughput MRI neuroinformatics as cloud-based software as a service," *Computing in Science & Engineering*, vol. 18, no. 5, p. 15, 2016.
- [16] S. F. Eskildsen, P. Coupé, D. García-Lorenzo et al., "Prediction of Alzheimer's disease in subjects with mild cognitive impairment from the ADNI cohort using patterns of cortical thinning," *NeuroImage*, vol. 65, pp. 511-521, 2013.
- [17] S. F. Eskildsen, P. Coupé, V. S. Fonov, J. C. Pruessner, D. L. Collins, and A. S. D. N. Initiative, "Structural imaging biomarkers of Alzheimer's disease: predicting disease progression," *Neurobiology of Aging*, vol. 36, no. 1, pp. S23-S31, 2015.
- [18] A. Alzheimer's, "Research consent for cognitively impaired adults: recommendations for institutional review boards and investigators," *Alzheimer Disease & Associated Disorders*, vol. 18, no. 3, pp. 171-175, 2004.
- [19] X. Tang, K. Oishi, A. V. Faria et al., "Bayesian parameter estimation and segmentation in the multi-atlas random orbit model," *PLoS One*, vol. 8, no. 6, Article ID e65591, 2013.
- [20] K. Oishi, A. Faria, H. Jiang et al., "Atlas-based whole brain white matter analysis using large deformation diffeomorphic metric mapping: application to normal elderly and Alzheimer's disease participantstlas," *NeuroImage*, vol. 46, no. 2, pp. 486-499, 2009.
- [21] C. Ceritoglu, K. Oishi, X. Li et al., "Multi-contrast large deformation diffeomorphic metric mapping for diffusion tensor imaging," *NeuroImage*, vol. 47, no. 2, pp. 618-627, 2009.
- [22] A. Djamanakova, A. V. Faria, J. Hsu et al., "Diffeomorphic brain mapping based on T1-weighted images: improvement of registration accuracy by multichannel mapping," *Journal of Magnetic Resonance Imaging*, vol. 37, no. 1, pp. 76-84, 2012.
- [23] D. Wu, T. Ma, C. Ceritoglu et al., "Resource atlases for multi-atlas brain segmentations with multiple ontology levels based

- on T1-weighted MRI," *NeuroImage*, vol. 125, pp. 120–130, 2016.
- [24] M. I. Miller, A. V. Faria, K. Oishi, and S. Mori, "High-throughput neuro-imaging informatics," *Frontiers in Neuroinformatics*, vol. 7, p. 31, 2013.
- [25] S. Mori, K. Oishi, A. V. Faria, and M. I. Miller, "Atlas-based neuroinformatics via MRI: harnessing information from past clinical cases and quantitative image analysis for patient care," *Annual Review of Biomedical Engineering*, vol. 15, no. 1, pp. 71–92, 2013.
- [26] K. Oishi, M. M. Mielke, M. Albert, C. G. Lyketsos, and S. Mori, "DTI analyses and clinical applications in Alzheimer's disease," *Journal of Alzheimer's Disease*, vol. 26, no. 3, pp. 287–296, 2011.
- [27] A. Djamanakova, X. Tang, X. Li et al., "Tools for multiple granularity analysis of brain MRI data for individualized image analysis," *NeuroImage*, vol. 101, pp. 168–176, 2014.
- [28] J. Mai, G. Paxinos, and T. Voss, *Atlas of Human Brain*, Academic Press, San Diego, CA, USA, 2007.
- [29] L. Puelles, M. Harrison, G. Paxinos, and C. Watson, "A developmental ontology for the mammalian brain based on the prosomeric model," *Trends in Neurosciences*, vol. 36, no. 10, pp. 570–578, 2013.
- [30] R. Tibshirani, "Regression shrinkage and selection via the Lasso," *Journal of the Royal Statistical Society Series B-Methodological*, vol. 58, no. 1, pp. 267–288, 1996.
- [31] J. Friedman, T. Hastie, and R. Tibshirani, "Regularization paths for generalized linear models via coordinate descent," *Journal of Statistical Software*, vol. 33, no. 1, p. 1, 2010.
- [32] J. M. Olichney, D. Galasko, D. P. Salmon et al., "Cognitive decline is faster in Lewy body variant than in Alzheimer's disease," *Neurology*, vol. 51, no. 2, pp. 351–357, 1998.
- [33] K. Rascofsky, D. P. Salmon, A. M. Lipton et al., "Rate of progression differs in frontotemporal dementia and Alzheimer disease," *Neurology*, vol. 65, no. 3, pp. 397–403, 2005.
- [34] M. H. Breitve, L. J. Chwiszczuk, M. J. Hynninen et al., "A systematic review of cognitive decline in dementia with Lewy bodies versus Alzheimer's disease," *Alzheimers Research Therapy*, vol. 6, no. 5, p. 53, 2014.
- [35] C. J. Thalhauser and N. L. Komarova, "Alzheimer's disease: rapid and slow progression," *Journal of the Royal Society Interface*, vol. 9, no. 66, pp. 119–126, 2012.
- [36] D. Bhargava, M. F. Weiner, L. S. Hynan, R. Diaz-Arrastia, and A. M. Lipton, "Vascular disease and risk factors, rate of progression, and survival in Alzheimer's disease," *Journal of Geriatric Psychiatry and Neurology*, vol. 19, no. 2, pp. 78–82, 2006.
- [37] Y. Nagahama, H. Nabatame, T. Okina et al., "Cerebral correlates of the progression rate of the cognitive decline in probable Alzheimer's disease," *European Neurology*, vol. 50, no. 1, pp. 1–9, 2003.
- [38] M. S. Mendiondo, J. W. Ashford, R. J. Kryscio, and F. A. Schmitt, "Modelling mini mental state examination changes in Alzheimer's disease," *Statistics in Medicine*, vol. 19, no. 11–12, pp. 1607–1616, 2000.
- [39] W. J. Youden, "Index for rating diagnostic tests," *Cancer*, vol. 3, no. 1, pp. 32–35, 1950.
- [40] M. S. Albert, S. T. DeKosky, D. Dickson et al., "The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease," *Alzheimer's & Dementia*, vol. 7, no. 3, pp. 270–279, 2011.
- [41] U. Lucca, M. Comelli, M. Tettamanti, P. Tiraboschi, and A. Spagnoli, "Rate of progression and prognostic factors in Alzheimer's disease: a prospective study," *Journal of the American Geriatrics Society*, vol. 41, no. 1, pp. 45–49, 1993.
- [42] D. M. Lipnicki, P. S. Sachdev, J. Crawford et al., "Risk factors for late-life cognitive decline and variation with age and sex in the Sydney Memory and Ageing Study," *PLoS One*, vol. 8, no. 6, Article ID e65841, 2013.
- [43] L. G. Apostolova, P. H. Lu, S. Rogers et al., "3D mapping of mini-mental state examination performance in clinical and preclinical Alzheimer disease," *Alzheimer Disease & Associated Disorders*, vol. 20, no. 4, pp. 224–231, 2006.
- [44] L. C. Baxter, D. L. Sparks, S. C. Johnson et al., "Relationship of cognitive measures and gray and white matter in Alzheimer's disease," *Journal of Alzheimer's Disease*, vol. 9, no. 3, pp. 253–260, 2006.
- [45] Z. Yao, B. Hu, C. Liang, L. Zhao, and M. Jackson, "A longitudinal study of atrophy in amnesic mild cognitive impairment and normal aging revealed by cortical thickness," *PLoS One*, vol. 7, no. 11, Article ID e48973, 2012.
- [46] T. Liu, D. M. Lipnicki, W. Zhu et al., "Cortical gyrification and sulcal spans in early stage Alzheimer's disease," *PLoS One*, vol. 7, no. 2, Article ID e31083, 2012.
- [47] Y. Li, S. M. Shea, C. H. Lorenz, H. Jiang, M. C. Chou, and S. Mori, "Image corruption detection in diffusion tensor imaging for post-processing and real-time monitoring," *PLoS One*, vol. 8, no. 10, Article ID e49764, 2013.
- [48] C. R. Jack Jr., D. W. Dickson, J. E. Parisi et al., "Antemortem MRI findings correlate with hippocampal neuropathology in typical aging and dementia," *Neurology*, vol. 58, no. 5, pp. 750–757, 2002.
- [49] T. Nickl-Jockschat, A. Kleiman, J. B. Schulz et al., "Neuro-anatomic changes and their association with cognitive decline in mild cognitive impairment: a meta-analysis," *Brain Structure & Function*, vol. 217, no. 1, pp. 115–125, 2012.
- [50] L. E. Harrell, D. Marson, A. Chatterjee, and J. A. Parrish, "The Severe Mini-Mental State Examination: a new neuropsychologic instrument for the bedside assessment of severely impaired patients with Alzheimer disease," *Alzheimer Disease and Associated Disorders*, vol. 14, no. 3, pp. 168–175, 2000.
- [51] J. Tohka, E. Moradi, H. Huttunen, and Alzheimer's Disease Neuroimaging, "Comparison of feature selection techniques in machine learning for anatomical brain MRI in dementia," *Neuroinformatics*, vol. 14, no. 3, pp. 279–296, 2016.

Research Article

An Approach for Pulmonary Vascular Extraction from Chest CT Images

Wenjun Tan ^{1,2}, Yue Yuan,¹ Anning Chen,¹ Lin Mao,¹ Yuqian Ke,¹ and Xinhui Lv¹

¹College of Computer Science and Engineering, Northeastern University, Shenyang 110819, China

²Cyberspace Institute of Advanced Technology, Guangzhou University, Guangzhou 510000, China

Correspondence should be addressed to Wenjun Tan; tanwenjun@cse.neu.edu.cn

Received 29 June 2018; Revised 25 November 2018; Accepted 12 December 2018; Published 17 January 2019

Academic Editor: Cesare Valenti

Copyright © 2019 Wenjun Tan et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Pulmonary vascular extraction from chest CT images plays an important role in the diagnosis of lung disease. To improve the accuracy rate of pulmonary vascular segmentation, a new pulmonary vascular extraction approach is proposed in this study. First, the lung tissue is extracted from chest CT images by region-growing and maximum between-class variance methods. Then the holes of the extracted region are filled by morphological operations to obtain complete lung region. Second, the points of the pulmonary vascular of the middle slice of the chest CT images are extracted as the original seed points. Finally, the seed points are spread throughout the lung region based on the fast marching method to extract the pulmonary vascular in the gradient image. Results of pulmonary vascular extraction from chest CT image datasets provided by the introduced approach are presented and discussed. Based on the ground truth pixels and the resulting quality measures, it can be concluded that the average accuracy of this approach is about 90%. Extensive experiments demonstrate that the proposed method has achieved the best performance in pulmonary vascular extraction compared with other two widely used methods.

1. Introduction

At present, computed tomography (CT) has become the most common imaging modality for the diagnosis of lung disease. In analysis of chest CT scans, pulmonary vascular extraction is often required before proceeding to diagnose disease and also is an important step in the treatment planning, and follow-up of lung diseases [1]. In computer-aided diagnosis of lung disease, pulmonary vascular extraction can reduce ambiguities and improve nodule detection performance [2] and is used to aid detection of other pulmonary anatomical structures such as airway tree, pulmonary tissue, lung vein, and artery [3–5]. Due to the complexity of the anatomical structure of the pulmonary vascular and the influence of other anatomical structures having similar intensities (e.g., tumor nodules and dense lesion) [6], there is a still more complicated task of extract pulmonary vascular from large scale chest CT images, especially for the small vascular.

For the large amount of chest CT images, manual extraction is an extremely time-consuming and tedious task for

doctors. The computerized semiautomatic and automatic methods may be helpful to reduce the doctor's effort. To address these issues, various methods have been extensively studied in recent years. Xu et al. [7] propose novel stagewise convolutional networks, followed by an orientation-based region-growing method, which learn discriminative features of pulmonary vessels automatically in a stage-by-stage manner by stagewise convolutional networks. Orkisz et al. [8] reported a vascular trees segmentation method by variational region growing. This process is performed within a lung mask, where the airways and bronchial walls were previously eliminated by adaptive multiscale morphological operations. Anna [9] introduced a 3D approach for segmentation of pulmonary vascular tree from CT thoracic scans based on author's experience in airway tree segmentation. Lai et al. [10] proposes an automatic integration segmentation approach for the vascular trees in low-dose CT scans. Zhu et al. [11] proposed a vascularity-oriented level set algorithm for pulmonary vessel segmentation in image-guided intervention therapy. Zhao et al. [12] proposed a

vessel segmentation method is proposed for lung images based on a random forest classifier and sparse autoencoder features. Chen et al. [13] proposed a method to simultaneously and separately segment the pulmonary nodules and blood vessels. First, a line structure enhancement (LSE) filter and a blob-like structure enhancement (BSE) filter were used to augment the initial selection of vessel regions and nodule candidates, respectively. Then a front surface propagation (FSP) procedure was employed for precise segmentation of blood vessels and nodules. Buelow et al. [14] presented an automated method for the extraction of the pulmonary vessel tree from multislice CT data by a seed point-based front-propagation algorithm. Rudyanto et al. [15] presented an annotated reference dataset containing 20 CT scans and proposed nine categories to perform a comprehensive evaluation of vessel segmentation algorithms from both academia and industry. This dataset is used in the VESSEL12 challenge held at the International Symposium on Biomedical Imaging (ISBI) 2012. The currently most general and extensive vascular segmentation method can be found in [16], with a highly detailed categorization of the existing work. However, because the pulmonary vessel trees are very complex and have a huge number of branches, the overall number of works on pulmonary vascular extraction is very limited for the computer-assisted diagnosis of pulmonary disease and the major methods for pulmonary extraction are still to be developed [7, 17, 18].

Coming forward to meet these needs, a new approach for the extraction of pulmonary vascular from chest CT images is proposed in this work. The proposed approach is based on author's previous work in airway tree extraction [19]. The main idea of the proposed method is to extract the whole lung region, enhance the extracted lung region, and then extract the pulmonary vascular.

This paper is organized as follows. First, in Section 1 a brief review of existing approaches to segmentation of the pulmonary vascular from chest CT images is introduced. Section 2 provides a detailed description of the proposed extraction approach. The image datasets, experiment software, and results of vascular extraction from chest CT image using the introduced method are presented and discussed in Section 3. Finally, Section 4 concludes the study.

2. Methods

The method of this study mainly includes the following steps: extract lung region, enhance the extracted lung region, and extract the pulmonary vascular. The flowchart of the method in this study is shown in Figure 1. The specific algorithms for each step are described in the following sections.

2.1. Lung Region Extraction Method. The anatomical structure of the lung in chest CT image is complex, as shown in Figure 2, if the pulmonary vascular is directly extracted in original CT images, it is difficult to remove the interference of other anatomical structures, such as bone, heart and muscle. Therefore, to segment accurately the pulmonary vascular, it is necessary to extract the lung region to remove other anatomical structures.

2.1.1. Extract Lung Tissue with Region-Growing and Maximum Between-Class Variance Methods. The chest CT image is divided into two categories of lung tissue and other anatomical structure, in which grayscale values are significantly different in the CT image. Region-growing method can segment effectively and correctly a specific set of grayscale values in an image based on the grayscale difference between pixel of that image. The region-growing method is an iterative image segmentation method with three elementary parts: seed points selection, definition of similarity, and criteria for convergence to terminate the iterative process [20]. The regions are grown from these seeds to adjacent points depending on a region membership criterion (e.g., grayscale intensity). Keep examining whether the adjacent points of seeds should be classified into the seed points until the criterion is not met any more [21].

In the previous work, we have used the region-growing method to extract airway from chest CT images [19]. The extracted airway result showed that this method is simple and effective. Therefore, we also use the region-growing method to extract the lung tissue in CT images in this study. This method requires seed points of lung tissue from a particular slice image, which is the middle slice of the patient's chest CT image in this work.

First, the seed points are extracted by the method of maximum between-class variance, which can divide the original image into two parts by using threshold: foreground and background. In the chest CT image, the background is lung tissue and the foreground is the other anatomical structure. T is set to be the segmentation threshold of the foreground and background, the number of foreground points accounts for w_0 of the image, the average grayscale value is u_0 , the number of background points accounts for w_1 of the image, the average grayscale value is u_1 , and the total grayscale value of the image is as follows:

$$u = w_0 u_0 + w_1 u_1. \quad (1)$$

The variance of foreground and background images:

$$\sigma^2 = w_0 (u_0 - u)^2 + w_1 (u_1 - u)^2. \quad (2)$$

Take u into formula (2) for calculating the variance of the two classes as follows:

$$\sigma^2 = w_0 w_1 (u_0 - u_1)^2. \quad (3)$$

The maximum threshold T_{\max} is obtained by using formula (3). The pixels of the image less than threshold T_{\max} are extracted as the seed points of lung tissue.

Second, the similarity definition is used to determine whether the unmarked pixels of image are added to the detection region. This definition refers to the difference of image intensity between the adjacent pixels. The similarity condition is formulated [22] as follows:

$$\begin{aligned} |I(x_k, y_k) - I(x, y)| < \theta, \\ (x_k, y_k) \in N_{26}. \end{aligned} \quad (4)$$

The unmarked voxel (x_k, y_k) in the 26-adjacent N_{26} can be added to the lung tissue region, if the difference between



FIGURE 1: The flowchart of this method.

the grayscale value of $I(x_k, y_k)$ and seed voxel $I(x, y)$ is less than the given threshold θ . And this voxel (x_k, y_k) is added to the seed queue voxel for next iteration. This process is shown in Figure 3.

Since this method relies on the difference between the grayscale value of the voxels in the 26-adjacent pixels, the given threshold θ plays an important role in this process. The extraction lung tissue regions are different with different given threshold. The threshold θ is calculated according to the following formula:

$$\theta = \frac{\max I_{\text{lung tissue}}(x, y) - \min I_{\text{lung tissue}}(x, y)}{2}, \quad (5)$$

where $\max I_{\text{lung tissue}}(x, y)$ and $\min I_{\text{lung tissue}}(x, y)$ are the maximum and minimum gray value and of the pixels of the extracted seed points by maximum between-class variance method, respectively.

The extracted result of the lung tissue is shown in Figure 4, which shows that there are many holes in the extracted lung region. The reasons are that the grayscale value of the vascular is different from the lung tissue.

2.1.2. Fill Holes of Extracted Lung Region with Morphological Operation. In this work, morphological operation method is used to fill the holes of lung tissue on the initial extraction region.

First, the contours of the lung are extracted in the all 2-D transverse images by 8-adjacent contour tracing method. There may be several contours in an image. The contours and the background are set to 1 and 0, respectively. And the set A is defined to represent the contours and the background pixels of the image. We formulate the filling "holes" function as

$$X_k = (X_{k-1} \oplus B) \cap A^c, \quad k = 1, 2, 3, \dots, \quad (6)$$

where $X_0 = p$, p is any point in the region, B is the 4-neighbourhood structural elements, \oplus is the dilation operator, and A^c is the complement set of A . At the end of the iteration, $X_k = X_{k-1}$. The union set of A and X_k is the region filling the holes.

The extracted region is expanded iteratively by the structural element. The number of iterations is manually based on experience. Then, the extracted region is corroded by the same structural element and the number of iterations of the expanding operation. The final result of the extracted lung region after filling holes is shown in Figure 5.

2.2. Vascular Enhancement Based on Fractional Differential. The actual chest CT image contains noise, which cause the edge of the pulmonary vascular of the image to be

unclear. If the pulmonary vascular are directly extracted from the original CT image, it is easy to cause inaccurate extraction of small pulmonary vascular. Therefore, the fractional order differential operator is proposed for enhancing vascular region in the extracted lung tissue(as shown in Figure 4) region before extracting pulmonary vascular.

The Grünwald–Letnikov (G-L) is used to define the numerical implementation of fractional differential of image [23–25]. The period of the unary signal $w(t)$ is $t \in [a, t]$, and the signal period $[a, t]$ is equally divided by a certain time interval (such as $h = 1, 2, \dots$), then $n = [(t - a)/h]^{h=1} = [t - a]$, and thus, the differential expressions with fractional differentials can be obtained:

$$\begin{aligned} \frac{d^\nu w(t)}{dt^\nu} &\approx w(t) + (-\nu)w(t-1) + \frac{-\nu(-\nu+1)}{2}w(t-2) \\ &+ \dots + \frac{\Gamma(-\nu+1)}{n!\Gamma(-\nu+n+1)}w(t-n). \end{aligned} \quad (7)$$

For the CT image, the differential expressions with fractional differentials can be obtained as follows:

$$\begin{aligned} \frac{\partial^\nu w(x, y)}{\partial x^\nu} &\approx w(x, y) + (-\nu)w(x-1, y) \\ &+ \frac{-\nu(-\nu+1)}{2}w(x-2, y) \\ &+ \dots + \frac{\Gamma(-\nu+1)}{(n-1)!\Gamma(-\nu+n)}w(x-n+1, y), \end{aligned} \quad (8)$$

$$\begin{aligned} \frac{\partial^\nu w(x, y)}{\partial y^\nu} &\approx w(x, y) + (-\nu)w(x, y-1) \\ &+ \frac{-\nu(-\nu+1)}{2}w(x, y-2) \\ &+ \dots + \frac{\Gamma(-\nu+1)}{(n-1)!\Gamma(-\nu+n)}w(x, y-n). \end{aligned} \quad (9)$$

The above differential expression constructs $3 * 3$ or $5 * 5$ differential operators to process the image to enhance the edge of the pulmonary vascular. Since the differential operator is isotropic, it does not enhance the small pulmonary vascular information with the characteristic direction. Thus, a fractional differential operator template can be constructed according to formulae (8) and (9). The operator has the following effect: it has little effect on the smooth region, where the surrounding gray value changes little; and the operator plays an enhancement role for the pixels whose surrounding gray value changes greatly.

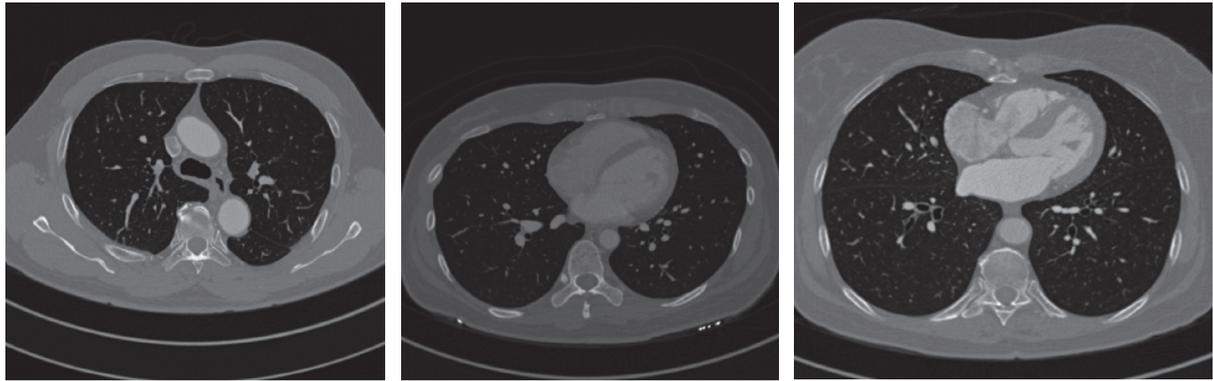


FIGURE 2: Chest CT images.

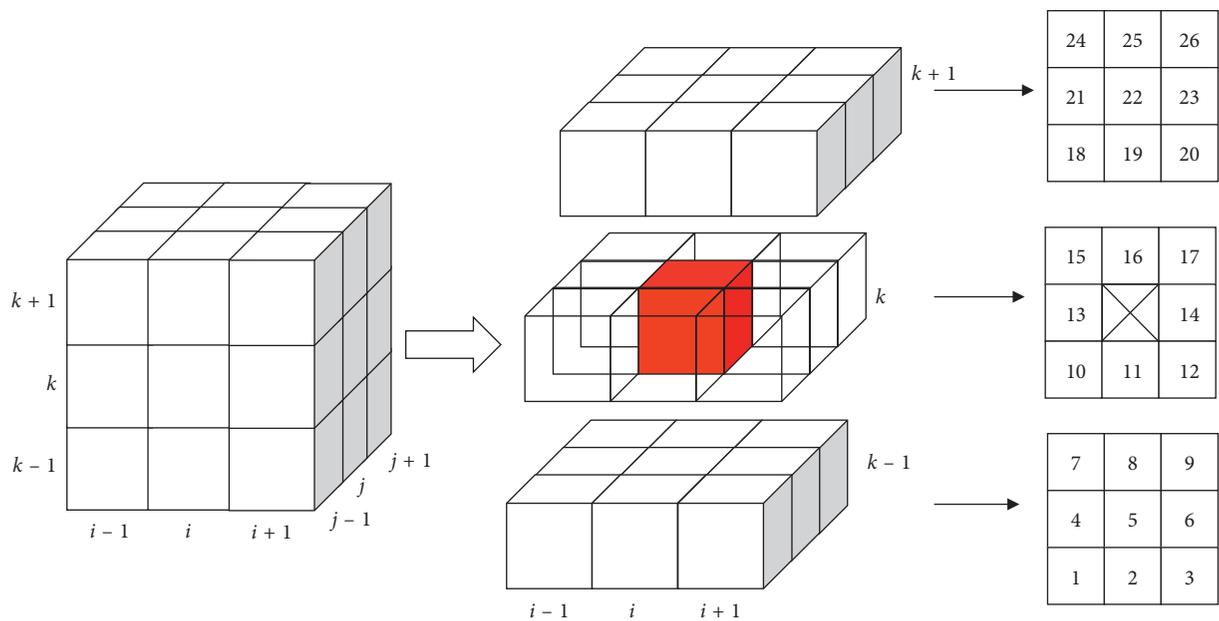


FIGURE 3: 26-adjacent pixels of region growing.

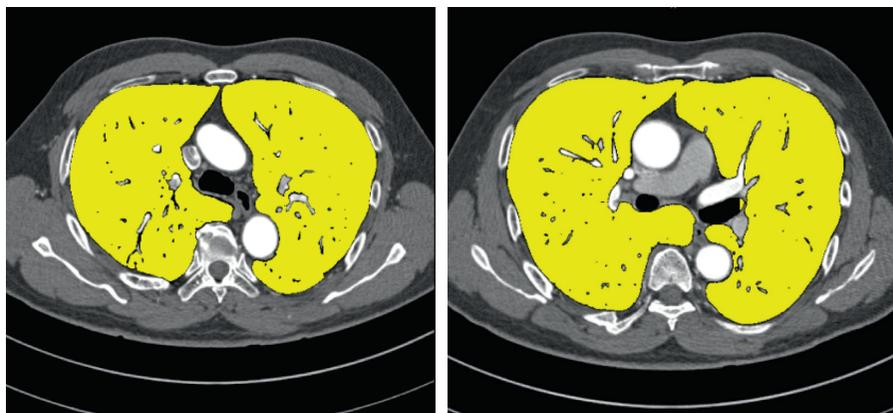


FIGURE 4: The extracted result of the lung tissue.

By analysing the characteristics and distribution of pulmonary vascular in the chest CT images, it is known that the pulmonary vascular are cylindrical and their directions

are mostly four diagonal directions. Thus, a 3×3 -differential operator template is constructed to enhance the region of vascular, as shown in Figure 6.

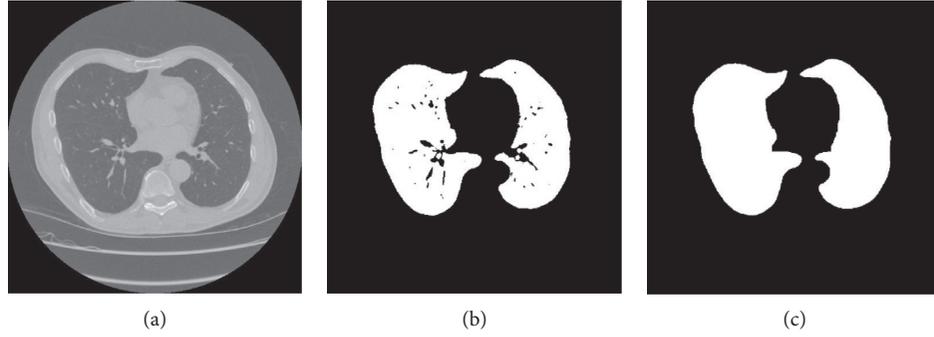


FIGURE 5: The final result of the extracted lung region: (a) original CT image, (b) lung tissue, and (c) the extracted lung region after filling holes.

0	0	$(v^2 - v)/2$
0	$-v$	0
1	0	0

(a)

$(v^2 - v)/2$	0	0
0	$-v$	0
0	0	1

(b)

0	0	1
0	$-v$	0
$(v^2 - v)/2$	0	0

(c)

1	0	0
0	$-v$	0
0	0	$(v^2 - v)/2$

(d)

FIGURE 6: 3×3 fractional differential operator template.

Since the operator is not symmetric, the following method should be used: the coordinates of the “1” of the operator template should coincide with the coordinates of the pixel to be subjected to the fractional differential operation. The templates in these four directions are convolved with the lung tissue region, and then the convolution results in each direction are weighted, and the results in the four directions are summed. The image after pulmonary vascular enhancement can be expressed by formula (10):

$$\begin{aligned}
 g(x, y) = & 4w(x, y) - v(w(x+1, y-1) + w(x-1, y+1) \\
 & + w(x-1, y-1) + w(x+1, y+1)) \\
 & + (v^2 - v)(w(x+2, y-2) + w(x-2, y+2) \\
 & + w(x-2, y-2) + w(x+2, y+2)).
 \end{aligned} \tag{10}$$

The lung region is processed using the above 3×3 fractional differential enhancement template. The different order can obtain different enhancement results. Therefore, we first study the experimental results of different fractional orders for chest CT image, which is shown in Figure 7. The results show that the pulmonary vascular in the CT image are enhanced as the differential order increases. At the same time, the result of small pulmonary vascular enhancement between 0.1 and 0.2 is the best. Although the

pulmonary vascular details are enhanced when the order is greater than 0.2, large pulmonary vascular are also suppressed, and the contrast between the large pulmonary vascular and other lung tissues is lowered. Thus, 0.2 fractional order is used to enhance the lung region of CT image in this work.

2.3. Extract Pulmonary Vascular by FMM. The complete lung region is extracted by the method of Section 2.1 and enhanced by fractional differential method. In this section, the pulmonary vascular is extracted with the fast marching method in the lung tissue region.

The fast marching method originates from the solution of the equation of the distance function. Suppose there is a curve φ moving in one direction. The FMM can be described as a family of schemes for computing the evolution of fronts. Things become interesting when the front evolves over time. In the context of the fast marching method, the speed does not have to be the same everywhere but the speed must always be non-negative. As a given point, the motion of the front is described by the equation known as the Eikonal equation, which can be expressed as [13]:

$$\begin{aligned}
 \|\nabla T(x)\|F(x) &= 1, \\
 T(\varphi) &= 0,
 \end{aligned} \tag{11}$$

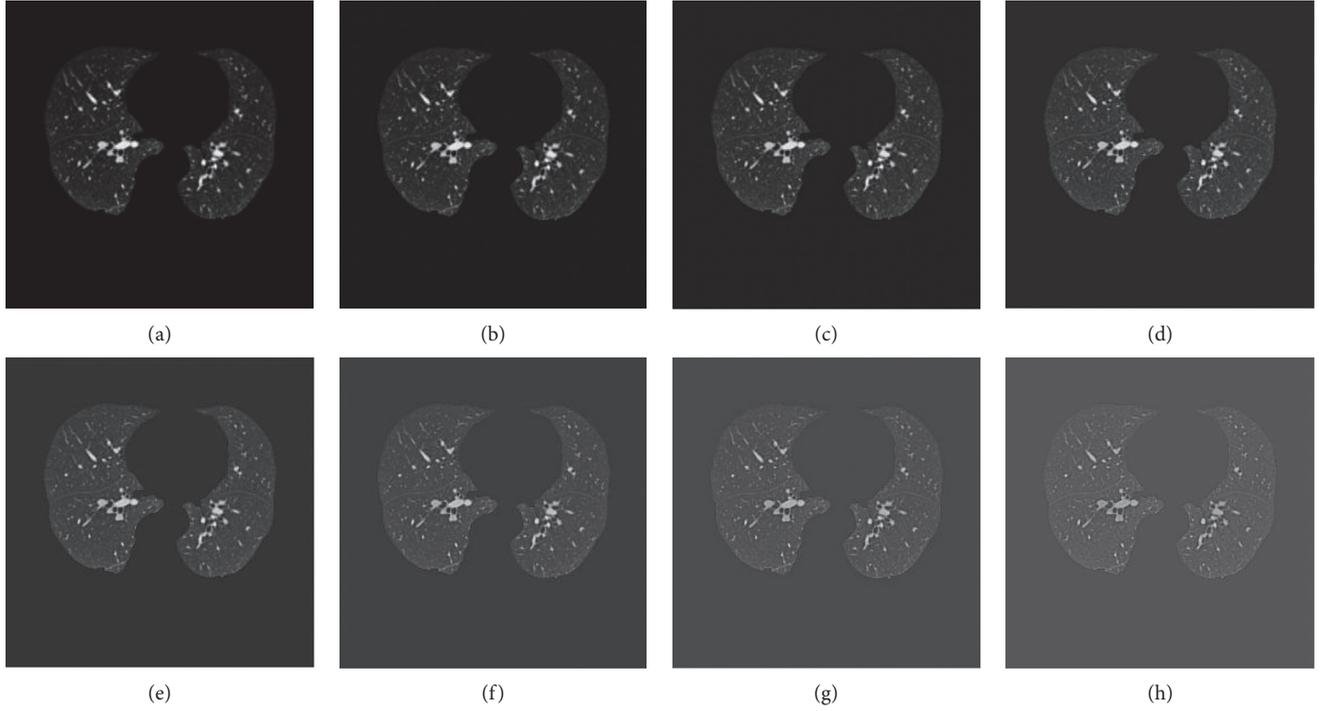


FIGURE 7: The result of different differential order: (a) original CT image, (b) 0.1 order, (c) 0.2 order, (d) 0.3 order, (e) 0.4 order, (f) 0.5 order, (g) 0.6 order, (h) 0.7 order.

where $T(x)$ is the arrival time of the front at point x and $F(x) \geq 0$ is the speed of the front at point x .

This is the general form of the equation of distance function. Godunov [26] gives a solution:

$$\max(D_{ij}^{-x}T, -D_{ij}^{+x}T, 0)^2 + \max(D_{ij}^{-y}T, -D_{ij}^{+y}T, 0)^2 = \frac{1}{V_{ij}^2}. \quad (12)$$

In this solution, D_{ij}^{-} and D_{ij}^{+} are backward difference and forward difference operators, i and j represent the two adjacent points of the curve φ , respectively.

If ∇T is approximated as a first-order finite difference operator, then formula (12) can be written as

$$\sum_{v=1}^2 \max\left(\frac{T - T_v}{\Delta v}, 0\right)^2 = \frac{1}{V^2}. \quad (13)$$

In that case, $\Delta_1 = \Delta_x$, $\Delta_2 = \Delta_y$, $T = T_{ij}$, $V = V_{ij}$, $T_1 = \min(T_{i+1,j}, T_{i-1,j})$, $T_2 = \min(T_{i,j+1}, T_{i,j-1})$.

A distance value for image must be found to estimate the length of the gradient $\|\nabla T\|$ is equal to $1/F$. The following formula (13) is proposed to solve $\|\nabla T\|$:

$$\|\nabla T\| = (D_{i,j}^2 + D_{i,k}^2 + D_{j,k}^2)^{1/2}. \quad (14)$$

The $D_{i,j}$, $D_{i,k}$, $D_{j,k}$ represent the gradient value in the three directions of Cartesian coordinate system, respectively. i , j , and k represent the three adjacent points.

Considering the case of adjacent points of three-dimensional CT images (as shown in Figure 8), formula (14) can be written as follows:

$$\begin{aligned} \|\nabla T\| = & \left(\max(V_A - V_B, V_A - V_C, 0)^2 \right. \\ & + \max(V_A - V_D, V_A - V_E, 0)^2 \\ & \left. + \max(V_A - V_F, V_A - V_G, 0)^2 \right)^{1/2}, \end{aligned} \quad (15)$$

where V_A is the unknown distance value and $V_B, V_C, V_D, V_E, V_F, V_G$ are the distance values at the neighbouring voxels. The distance value is the gray value of the pixels of CT images.

For speeding up in homogeneous regions of the image and slowly in regions with a high value of image gradient. The F of the chest CT volume data, be considered as gradient figure, is computed by the exponentiated gradient algorithm as follows:

$$F = e^{-k\|\nabla T\|}. \quad (16)$$

The solution process of FMM is described as follows:

Step 1. All the points are categorized into three categories: the processing points are the seed points (in the initial step, the number of the seed points is one; during the processing, the seed points may be many points); the boundary points are the points near the processing points; the pending points are all the remaining points.

Step 2. The arrival time $T(x)$ is calculated from the initial seed point to the boundary points and sort them according to the time from small to large.

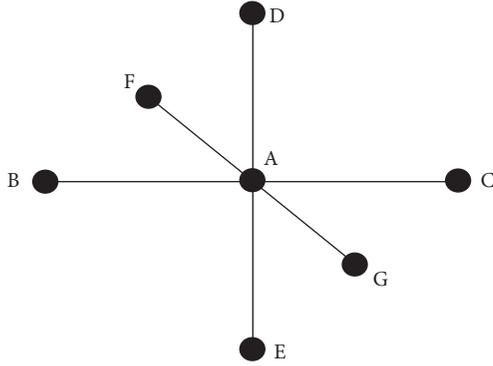


FIGURE 8: Adjacent points of three-dimensional CT images.

Step 3. The boundary points, less than a certain arrival time threshold, are marked as the next processing points of as the next processing points. Other boundary points are marked as the background points. The nearby points of the next processing points are marked as the new boundary points.

Step 4. The arrival time $T(x)$ of new boundary points are calculated and sorted. If the next processing points are empty, the processing is finished. Otherwise, it returns to (2).

The above processing can be considered that the curve spread along an equal timeline until it arrives at the front of a wall.

In this work, the full lung region of the chest CT images is extracted by the method of Section 2.1 and enhanced by fractional differential method. Then the points of pulmonary vascular are extracted with the maximum between-class variance method in the full lung region of the middle of slice of the chest CT images. In this case, the foreground is the pulmonary vascular. These extracted points are selected as initial seed points. And the minimum and maximum grayscale values of these pulmonary vascular points are calculated as the limited grayscale value threshold. According to the limited grayscale value threshold and the gradient field figure F, the seed points begin to spread through the lung region of images using the fast march method as the above solution process. The points of pulmonary vascular are extracted completely when the spread processing stops.

3. Results and Discussion

The 3 chest CT images dataset were used in this experiment with resolution of 512×512 , slice thickness less than 1.5 mm and slice number more than 350, which are from a hospital. In order to protect patient privacy information, the image dataset hides the hospital's and patient's name. The ground truth is manually drawn by the doctor for each database.

This method was implemented in Matlab 2015 on a PC with 4 Intel® Core i7-6700U CPUs 2.60 GHz, 8 GB DDR4 RAM and NVIDIA GeForce 940 Mx GPU with 2 GB video memory.

We will quantitatively analyse the results of the method for extraction of pulmonary vascular. By comparing the extraction results of the proposed method with the ground truth image, the accurate rate, the leakage rate and over rate of the proposed method are calculated to quantitatively evaluate the. The ground truth image, be manually drawn by the professional doctors on the original CT image, is provided in VESSEL12 challenge.

The evaluation criteria for pulmonary vascular segmentation are as follows:

$$\text{accuracy} = \frac{S \cap T}{T},$$

$$\text{leakage rate} = \frac{T - S \cap T}{T}, \quad (17)$$

$$\text{over rate} = \frac{S - S \cap T}{T},$$

where T and S is the number of pulmonary vascular pixels in the ground truth image and the extraction result by this method, respectively.

In order to observe the experimental results, the comparison results of the 3 datasets between the method in this paper and the ground truth is shown in Figure 9. Figures 9(a) and 9(c) © are the extraction results of this method, and Figures 9(b), 9(d), and 9(f) are the ground truth results. The blue circles indicate different places between the method and the ground truth. The results show that the method can completely extracted the pulmonary vascular, even the extraction results of small pulmonary vascular are also very good, and cause a little over segmentation in the edge of lung region.

The method is compared with the level set [11] and regional growth algorithm [8], which are widely used to extract pulmonary vascular from medical image. The accurate rate, the leakage rate, and the over rate of each segmentation method are calculated by comparing with the ground truth image.

The experimental results of an image dataset of different segmentation methods are shown in Figure 10. The different places of these results are shown by yellow and blue circles in the result images. The results show that the segmentation results of these four methods are equally good, but this method is a better accurate segmentation of the small pulmonary vascular and less over segmentation than the other methods. The blue circle of Figure 10 shows that this method can extract more the small pulmonary vascular. Meanwhile, the yellow circle of Figure 10 shows that this method is less over segmentation results than the other methods.

Table 1 and Figure 11 are the statistical analysis of the segmentation results of the three methods. The average of accurate rate of this method is 91.62%, which is better than the other two methods. And the average of leakage rate and over rate of this method is 8.38% and 3.43%, respectively. It is less than the other two methods.

In general, the running time of segmentation method is related to the slice thickness and the slice number. The average running time of this method is about 2 minutes,

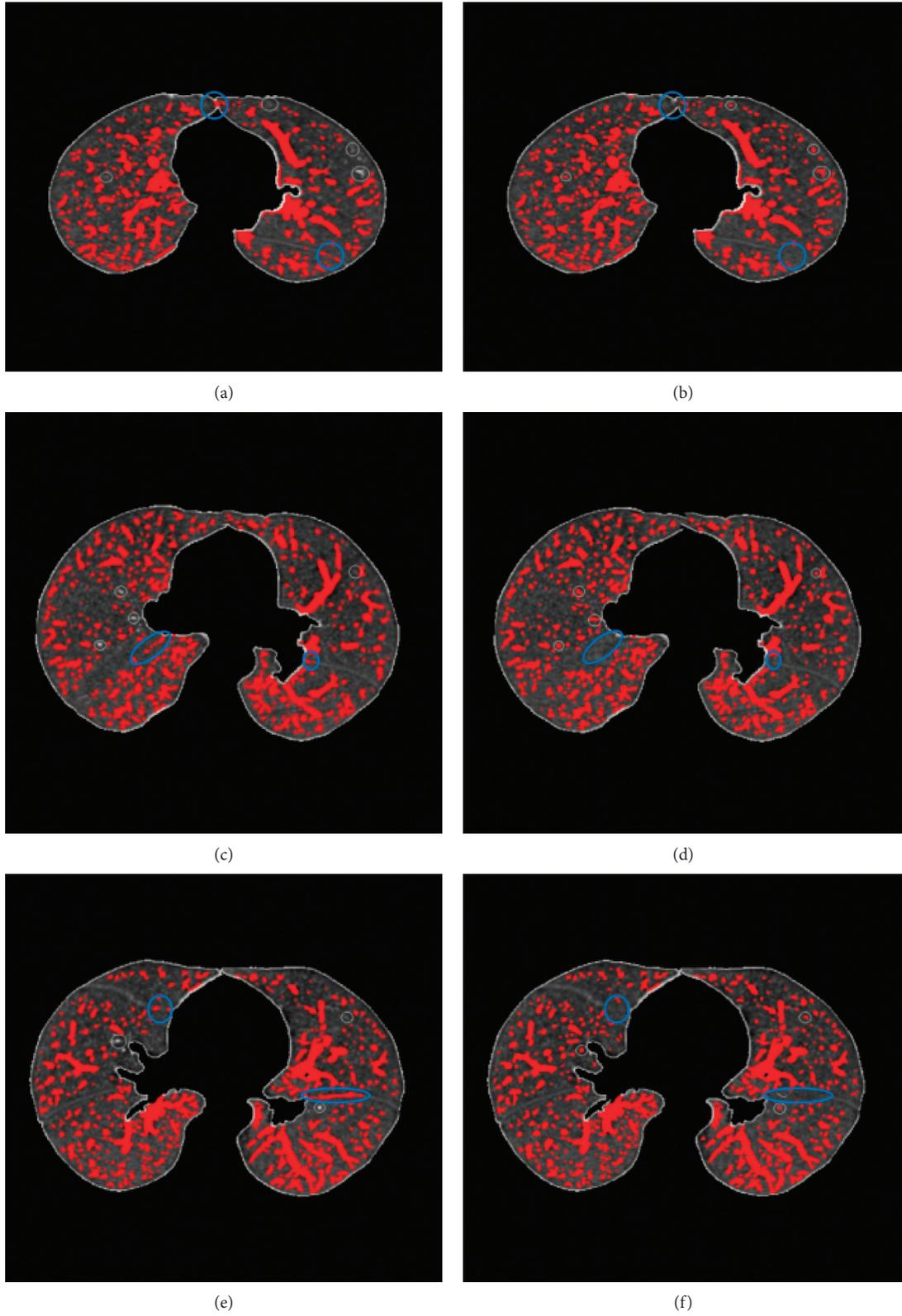


FIGURE 9: The comparison results between this method and the ground truth: (a–c) are the results of this method and (b–f) are corresponding ground truth.

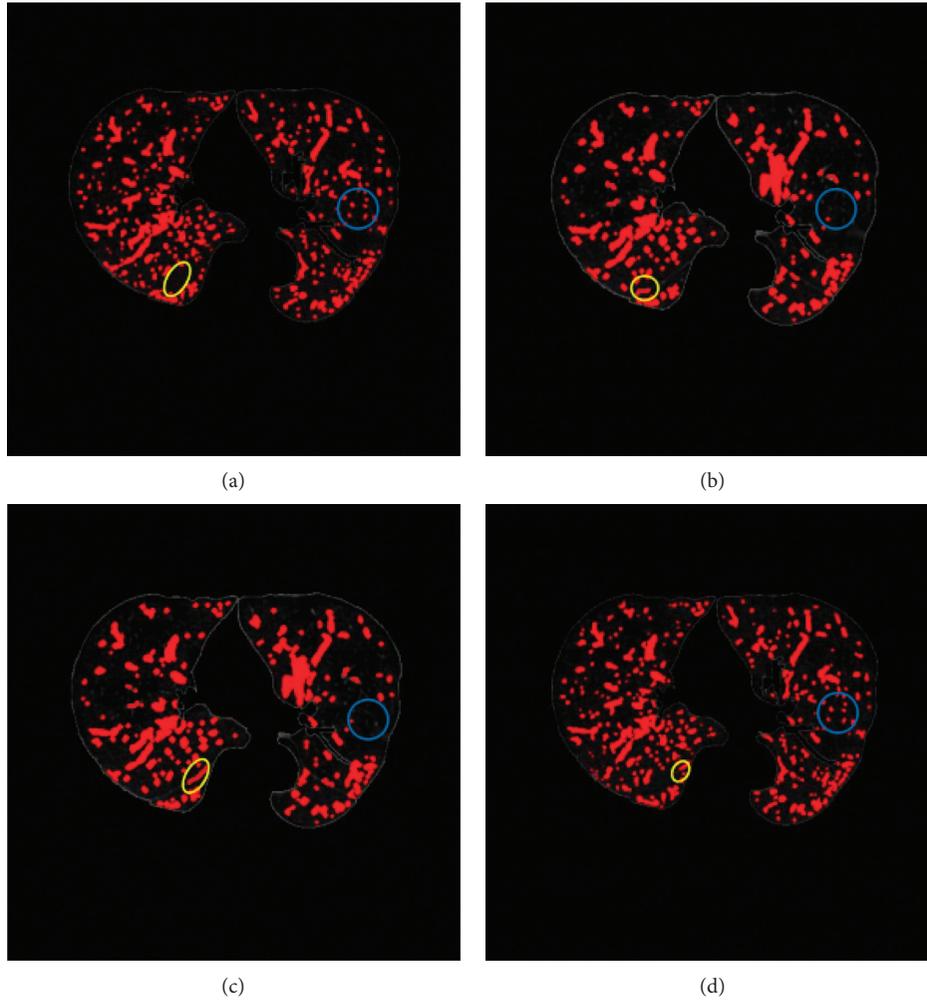


FIGURE 10: Compare the results of different segmentation methods and the ground truth: (a) ground truth, (b) result of level set method, (c) result of region-growing method, and (d) result of this method.

TABLE 1: The statistical analysis of the segmentation results of the three methods.

Dataset no.	Level method (%)			Region-growing method (%)			This method (%)		
	Accurate rate	Leakage rate	Over rate	Accurate rate	Leakage rate	Over rate	Accurate rate	Leakage rate	Over rate
1	90.42	9.58	4.43	89.53	10.47	4.11	91.87	8.13	3.15
2	88.51	11.49	3.78	88.17	11.83	4.06	90.53	9.47	3.42
3	90.36	9.64	3.61	89.24	10.76	3.56	92.45	7.55	3.71
Average	89.76	10.24	3.94	88.98	11.02	3.91	91.62	8.38	3.43

which can meet the requirements of clinical real-time operation. Furthermore, this method is completely automatic segmentation process without manual operation.

4. Conclusions

This study proposes a method for segmenting pulmonary vascular from chest CT images. The lung tissue region is extracted from chest CT images by region-growing method and maximum between-class variance. Then, the holes of the extracted region are filled by morphological operation. The points of the pulmonary vascular of

the middle slice of the CT images are extracted as the original seed points. Finally, the seed points are spread throughout the lung tissue based on the fast-marching method to extracted the pulmonary vascular with the gradient image. The experimental results show that this method can completely extracted the big and small pulmonary vascular and the accuracy is better than the level set method and region-growing method, and the leakage rate and over rate of this method is less than the two other methods. Furthermore, the running time of this method is about 2 minutes to meet the requirements of clinical real-time operation. This method may be used to promote the

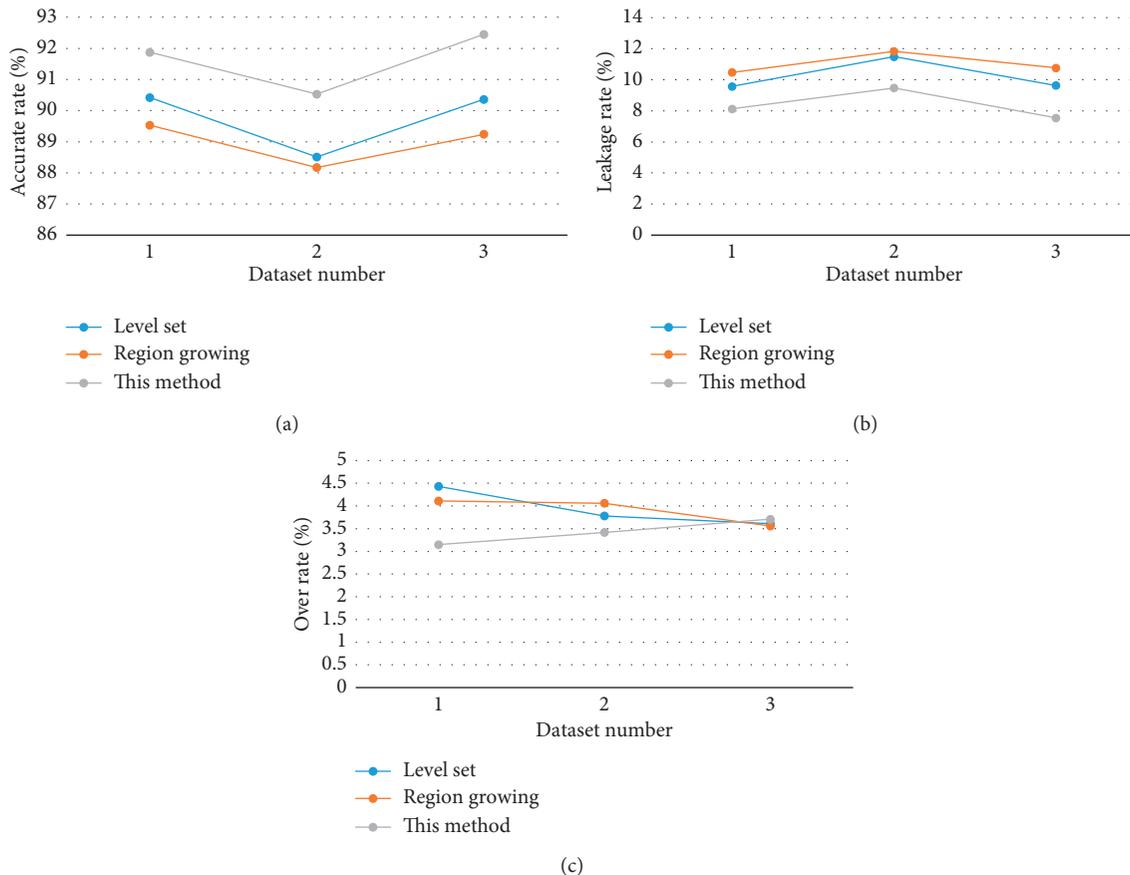


FIGURE 11: Comparison of segmentation results of three methods: (a) accurate rate, (b) leakage rate, and (c) over rate.

diagnosis technology of lung diseases for clinical application.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the National Key Research and Development Program of China (2018YFC1314501) and the Fundamental Research Funds for the Central Universities under grant nos. N150408001 and N161604006.

References

- [1] J. N. Kaftan and T. Aach, "Pulmonary vessel segmentation for multislice CT data: methods and applications," in *Lung Imaging and Computer Aided Diagnosis*, A. El-Baz and J. S. Suri, Eds., pp. 189–219, CRC Press, Boca Raton, FL, USA, 2011.
- [2] G. Agam, S. G. Armato, and C. Changhua Wu, "Vessel tree reconstruction in thoracic CT scans with application to nodule detection," *IEEE Transactions on Medical Imaging*, vol. 24, no. 4, pp. 486–499, 2005.
- [3] P. Lo, B. van Ginneken, J. M. Reinhardt et al., "Extraction of airways from CT (EXACT'09)," *IEEE Transactions on Medical Imaging*, vol. 31, no. 11, pp. 2093–2107, 2012.
- [4] P. Lo, J. Sparring, H. Ashraf, J. J. H. Pedersen, and M. de Bruijne, "Vessel-guided airway tree segmentation: a voxel classification approach," *Medical image analysis*, vol. 14, no. 4, pp. 527–538, 2010.
- [5] P. Nardelli, D. Jimenez-Carretero, D. Bermejo-Pelaez et al., "Pulmonary artery-vein classification in CT images using deep learning," *IEEE Transactions on Medical Imaging*, vol. 37, no. 11, pp. 2428–2440, 2018.
- [6] P. D. Korfiatis, C. Kalogeropoulou, A. N. Karahaliou, A. D. Kazantzi, and L. I. Costaridou, "Vessel tree segmentation in presence of interstitial lung disease in MDCT," *IEEE Transactions on Information Technology in Biomedicine*, vol. 15, no. 2, pp. 214–220, 2011.
- [7] Y. Xu, Z. Mao, C. Liu et al., "Pulmonary vessel segmentation via stage-wise convolutional networks with orientation-based region growing optimization," *IEEE Access*, vol. 6, pp. 71296–71305, 2018.
- [8] M. Orkisz, M. Hernández Hoyos, V. Pérez Romanello et al., "Segmentation of the pulmonary vascular trees in 3D CT images using variational region-growing," *IRBM*, vol. 35, no. 1, pp. 11–19, 2014.
- [9] F. Anna, "Segmentation of pulmonary vascular tree from 3D CT thorax scans," *Biocybernetics and Biomedical Engineering*, vol. 35, no. 2, pp. 106–119, 2015.

- [10] J. Lai, Y. Huang, Y. Wang et al., "Three-dimensions segmentation of pulmonary vascular trees for low dose CT scans," *Sensing and Imaging*, vol. 17, no. 1, 2016.
- [11] X. Zhu, Z. Xue, X. Gao et al., "Voles: vascularity-oriented level set algorithm for pulmonary vessel segmentation in image guided intervention therapy," in *Proceedings of IEEE International Symposium on Biomedical Imaging: from Nano to Macro, ISBI'09*, pp. 1247–1250, IEEE, Boston, MA, USA, June 2009.
- [12] B. Zhao, Z. Cao, and S. Wang, "Lung vessel segmentation based on random forests," *Electronics Letters*, vol. 53, no. 4, pp. 220–222, 2017.
- [13] B. Chen, T. Kitasaka, H. Honma et al., "Automatic segmentation of pulmonary blood vessels and nodules based on local intensity structure analysis and surface propagation in 3D chest CT images," *International Journal of Computer Assisted Radiology and Surgery*, vol. 7, no. 3, pp. 465–482, 2011.
- [14] T. Bülow, R. Wiemker, T. Blaffert et al., "Automatic extraction of the pulmonary artery tree from multi-slice CT data," in *Proceedings of conference on Medical Imaging 2005: Physiology, Function, and Structure from Medical Images*, vol. 5746, pp. 730–741, International Society for Optics and Photonics, San Diego, CA, USA, April 2005.
- [15] R. D. Rudyanto, S. Kerkstra, E. M. van Rikxoort et al., "Comparing algorithms for automated vessel segmentation in computed tomography scans of the lung: the VESSEL12 study," *Medical Image Analysis*, vol. 18, no. 7, pp. 1217–1232, 2014.
- [16] D. Lesage, E. D. Angelini, I. Bloch, and G. Funka-Lea, "A review of 3D vessel lumen segmentation techniques: models, features and extraction schemes," *Medical Image Analysis*, vol. 13, no. 6, pp. 819–845, 2009.
- [17] M. Paulinas, D. Miniotas, M. Meilūnas et al., "An algorithm for segmentation of blood vessels in images," *Elektronika ir Elektrotechnika*, vol. 83, no. 3, pp. 25–28, 2015.
- [18] X. Li, X. Wang, Y. Dai, and P. Zhang, "Supervised recursive segmentation of volumetric CT images for 3D reconstruction of lung and vessel tree," *Computer Methods and Programs in Biomedicine*, vol. 122, no. 3, pp. 316–329, 2015.
- [19] W. Tan, J. Yang, Z. Bian, Z. Gong, and D. Zhao, "Automatic extraction of 3D airway tree from multislice computed tomography images," *Journal of Medical Imaging and Health Informatics*, vol. 4, no. 5, pp. 768–775, 2014.
- [20] R. Adams and L. Bischof, "Seeded region growing," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 16, no. 6, pp. 641–647, 1994.
- [21] Q. Li, "Glioma segmentation with a unified algorithm in multimodal MRI images," *IEEE Access*, vol. 6, pp. 9543–9553, 2018.
- [22] K. C. Lin, "Fast image thresholding by finding the zero(s) of the first derivative of between-class variance," *Machine Vision and Applications*, vol. 13, no. 5-6, pp. 254–262, 2003.
- [23] D. Cafagna, "Fractional calculus: a mathematical tool from the past for present engineers (Past and present)," *IEEE Industrial Electronics Magazine*, vol. 1, no. 2, pp. 35–40, 2007.
- [24] K. S. Miller and B. Ross, *An Introduction to the Fractional Calculus and Fractional Differential Equations*, Wiley, New York, NY, USA, 1993.
- [25] C. Xiao, M. Staring, D. Shamonin, J. H. C. Reiber, J. Stolk, and B. C. Stoel, "A strain energy filter for 3D vessel enhancement with application to pulmonary CT images," *Medical Image Analysis*, vol. 15, no. 1, pp. 112–124, 2011.
- [26] S. K. Godunov, "A difference method for numerical calculation of discontinuous solutions of the equations of hydrodynamics," *Matematicheskii Sbornik*, vol. 89, no. 3, pp. 271–306, 1959.

Research Article

Fully Convolutional DenseNet with Multiscale Context for Automated Breast Tumor Segmentation

Jinjin Hai,¹ Kai Qiao,¹ Jian Chen ,¹ Hongna Tan,² Jingbo Xu,¹ Lei Zeng ,¹ Dapeng Shi,² and Bin Yan ¹

¹National Digital Switching System Engineering and Technological Research Center, Zhengzhou, Henan Province, China

²Department of Radiology, Henan Provincial People's Hospital, Zhengzhou, Henan Province, China

Correspondence should be addressed to Bin Yan; ybspace@hotmail.com

Received 29 June 2018; Revised 20 October 2018; Accepted 25 November 2018; Published 14 January 2019

Guest Editor: Cesare Valenti

Copyright © 2019 Jinjin Hai et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Breast tumor segmentation plays a crucial role in subsequent disease diagnosis, and most algorithms need interactive prior to firstly locate tumors and perform segmentation based on tumor-centric candidates. In this paper, we propose a fully convolutional network to achieve automatic segmentation of breast tumor in an end-to-end manner. Considering the diversity of shape and size for malignant tumors in the digital mammograms, we introduce multiscale image information into the fully convolutional dense network architecture to improve the segmentation precision. Multiple sampling rates of atrous convolution are concatenated to acquire different field-of-views of image features without adding additional number of parameters to avoid over fitting. Weighted loss function is also employed during training according to the proportion of the tumor pixels in the entire image, in order to weaken unbalanced classes problem. Qualitative and quantitative comparisons demonstrate that the proposed algorithm can achieve automatic tumor segmentation and has high segmentation precision for various size and shapes of tumor images without preprocessing and postprocessing.

1. Introduction

Breast cancer is the most common disease of women and has become the second disease which leads to death [1]. The number of breast cancer patients is increasing gradually since 1970s. Early detection of breast cancer is beneficial for improving the survival rate and survival quality. Mammography is the most conventional and noninvasive examination and is an effective screening method for early detection and diagnosis of breast cancer. Tumor segmentation provides morphological features and is an essential step for tumor analysis and classification.

Traditional studies on tumor segmentation mainly rely on gray-level and texture features which are distinct to partition mammogram into different regions. Region-based methods start from a set of manually located seed points or small patches containing suspicious region, such as region growing [2] and watershed methods [3]. Different location of seed points are set to segment tumor according to different image

preprocessing methods, such as Gaussian filtering [4, 5] and mathematical morphological operation [6]. Watershed methods mainly used some preprocessing algorithms to reduce the number of initial segmented basins [7, 8]. Active contour model, especially level set [9] is also used for breast masses segmentation. A radial gradient index- (RGI-) based segmentation method is applied to yield an initial contour closer to the lesion boundary location [10]. 3D radial-gradient index segmentation and 3D level set-based active contour algorithm [11] are also applied to 3D CT breast images. A feature embedded vector-valued contour-based level set method [12] is proposed to perform mammographic mass segmentation. It used level set method to obtain the initial boundaries on the smoothed mammogram as the shape constraint to design stopping function and integrated texture maps, gradient maps, and the original intensity map to obtain more accurate segmentation.

Region-based and contour-based methods are all unsupervised. There are also some supervised segmentation

algorithms based on deep networks. Dhungel et al. [13] combined multiple deep belief networks (DBNs), Gaussian mixture model (GMM) classifier, and a prior of location, size, and shape of the mass as potential functions and used structured SVM to learn a structured output and perform segmentation. They further used conditional random field (CRF) with tree reweighted belief propagation as structured prediction function to boost the segmentation performance [14]. The output of convolutional neural network (CNN) is also introduced as a complementary potential function in addition to the aforementioned potential functions, yielding state-of-the-art segmentation performance [15]. These methods all used two-stage training. An end-to-end network was proposed based on the mass region of interest (ROI) images [16]. It employed fully convolutional network to model potential function, followed by a CRF to perform structured learning, and integrated adversarial training to learn robustly from scarce mammographic images.

CNN has powerful ability of extracting abstracted features directly from the raw input data and achieves remarkable achievements in computer vision fields, such as image classification [17–19], object detection [20–22], and image segmentation [23–25]. Image segmentation methods based on CNN could get segmentation result through discriminating every pixel in the image. Long et al. [23] proposed the fully convolutional network (FCN) and performed fine tuning in an end-to-end manner based on pretrained VGG-Net [26] for image semantic segmentation. FCN replaced the fully connected layers with the convolutional layers in order to keep the location information, which adjusted the classification network for the segmentation task. Additionally, FCN employed the skip architecture to merge semantic features and detailed features and performed deconvolution to obtain more accurate segmentation results. Since that, the deconvolution operation had been widely used in many semantic segmentation networks. Hyeonwoo [24] proposed the symmetrical encoder-decoder architecture called DeconvNet. The DeconvNet employed successive unpooling layers which reconstruct the original size of activations through recording the locations of maximum activations selected during pooling operation and deconvolution layer with learned filters to generate dense pixelwise class probability map. SegNet [25] is another similar symmetrical network which used convolution after unpooling layer to refine the sparse feature maps. Due to the low localization precision of reconstructing the resolution directly from high-level features, U-Net [27] combined the upsampling output with the high-resolution features from the encoding path to improve the segmentation performance. Simon et al. [28] extended DenseNet [29] which achieved excellent results on image classification tasks to deal with the problem of semantic segmentation. The proposed Fully Convolutional DenseNet (FC-DenseNet) also employed skip connections that the feature maps from the downsampling path are concatenated with the corresponding feature maps in the upsampling path and achieved state-of-art results. In conclusion, the typical segmentation architecture built on CNN is mostly fully convolutional network with encoder-decoder architecture training in an end to

end manner. It mainly composes of a downsampling path responsible for extracting coarse semantic features and an upsampling path trained to recover the input image resolution.

Because of the larger size of digital mammograms, the segmentation time of active contour model, such as level set, is greatly increased when the whole image is automatically segmented. And most of traditional unsupervised methods rely on low-level features such as the image gray value, texture, gradient, and other information or the initial priori settings, such as the initial seed points of region growing and initial contour of level set method. But this cannot achieve fully automatic segmentation of breast tumor, and segmentation precision is influenced by the hand-crafted features and initial priori position. In addition, some image backgrounds are complex and similar to the characteristic of tumor region, such as the pectoral muscles, or the gray value of internal and external region of the tumor have small differences. So, many aforementioned segmentation methods are based on the tumor-centric candidate box [30, 31]. Although some supervised segmentation algorithms take into account the category information and can automatically extract tumor features to optimize the segmentation model, they also add priori information of the location, size, and shape of the tumor. And they are also based on the small tumor candidate box and cannot directly segment the entire large size mammograms.

The main goal of our proposed algorithm is to segment the breast tumor on the entire digital mammogram not on the tumor-centric rectangle region which is commonly used in most breast tumor segmentation algorithms. FC-DenseNet further exploited the feature reuse by upsampling the feature maps created by the preceding dense block and used skip connections to help the upsampling path recover spatially detailed information from the downsampling path. It outperforms current state-of-the-art results neither using pretrained parameters nor any further postprocessing. So, we extend FC-DenseNet to achieve automatic tumor segmentation. At the same time, considering that the multiscale information [32, 33] is benefit for improving the segmentation precision of different sizes of tumors, atrous spatial pyramid pooling (ASPP) [34] is added to FC-DenseNet semantic segmentation network. Without significantly increasing the number of learning parameters, ASPP extracts multiscale features by concatenating different sampling rates of atrous convolution [35] to enlarge the receptive field. Loss function is also improved to solve the extremely unbalanced class problem according to the proportion of tumor and background pixels in the entire image. We evaluate this algorithm in our collected digital mammogram dataset, and results demonstrate that our proposed algorithm has yielded better performance than other algorithms.

2. Methods

2.1. Review of FC-DenseNet. FC-DenseNet [28] is an extension of excellent DenseNet [29] classification network in semantic segmentation by adding an upsampling path to

recover the full input resolution. Downsampling path of FC-DenseNet is composed of dense block layer and transition down layer. Dense block layer is composed of batch normalization [36], followed by ReLU [37], a 3×3 same convolution (no resolution loss) and dropout with probability $p = 0.2$, which is different from DenseNet. A transition down layer is composed of a 1×1 convolution (which conserves the number of feature maps) followed by a 2×2 pooling operation. Upsampling path consists of dense block layer and transition up layer. Transition up includes a 3×3 transposed convolution with stride 2 to compensate for the pooling operation in the transition down. The upsampled feature maps are then concatenated to the ones with the same resolution from the downsampling path to form the input of a new dense block. But in order to prevent the linear growth of feature maps, the input of a dense block is not concatenated with its output. Thus, the transposed convolution is applied only to the feature maps obtained by the last dense block and not to all feature maps concatenated so far. The final layer in the network is a 1×1 convolution followed by a softmax nonlinearity function to predict the class label at each pixel.

2.2. ASPP-FC-DenseNet Segmentation Algorithm. As the shape and size of malignant tumor are various, as shown in Figure 1, the height and width of the tumor are mostly distributed in 200 and 800 pixel intervals. As the receptive field of single-image scale or small convolution kernel is relatively fixed, it is only effective to present the image features within the scope of the receptive field and cannot be well depicted on the edges of different sizes of tumor. This could influence the segmentation precision. Therefore, extracting multiscale information of the image is helpful to improve the segmentation precision of different sizes of tumors.

Multiscale images (image pyramid) [38] are commonly used as the input of network to extract features for each scale input, and the segmentation results of each scale are linearly interpolated and fused. But all layers of parallel CNN need to compute the features of multiple scale inputs, and the consumption of feature computation is large. Different sizes of convolution kernel have different receptive fields, and exploiting different convolution kernel size [39] to extract multiscale image features is an alternative approach. However, multiple parallel branches with different size of convolution kernels greatly increase the network width and the network learning parameters. Due to the small digital mammogram segmentation datasets, the network is easy to overfit. Therefore, exploiting multiple field-of-views to extract multiscale image features with relatively less increasing parameters is more requisite.

Spatial pyramid pooling (SPP) [40] is a common way to obtain multiscale image information, which was originally proposed to solve the problem of arbitrary input size of proposals in object detection. According to the requirement of the output dimension, SPP divides the input image with arbitrary size into the corresponding number of spatial bins, then the pooling operation is performed on each spatial bin.

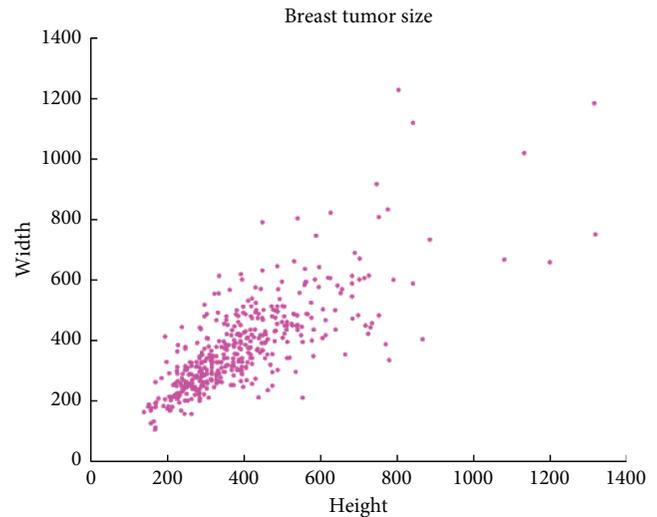


FIGURE 1: The distribution of breast tumor size.

The features of each spatial bin are combined to get fixed features output dimension. It can be seen that the output of SPP layer fuses the features of different level image bins, thus improves the recognition accuracy. But for image segmentation requiring pixel level semantic classification, multiple pooling operations lose the image details and spatial position information.

Atrous convolution [35] is an effective way to expand the network receptive field without increasing the size of convolution kernel and network parameters and is mainly operated by setting different atrous sampling rate. As shown in Figure 2, atrous convolution is performed standard convolution with filter 'with holes' according to the sampling rates, and the weights of the 'holes' are reset to 0. Atrous convolution with rate r introduces $r - 1$ zeros between consecutive filter values, effectively enlarging the kernel size of a $k \times k$ filter to $(k + (k - 1)(r - 1)) \times (k + (k - 1)(r - 1))$ filter without increasing the number of parameters. The atrous convolution operation could convolute every pixel of the input by setting a specific convolution stride. By setting the sampling rate, the small convolution kernel can also achieve the effect of large convolution kernel, thus expanding the receptive field of the network without requiring learning any extra parameters and increasing the amount of computation. Atrous convolution could arbitrarily enlarge the field-of-view of filters at any network layer. As shown in Figure 2, the 3×3 convolution kernel has the same receptive field of 5×5 and 7×7 convolution kernel by setting different sampling rates. According to the idea of spatial pyramid pooling, the extracted features of network are resampled using parallel atrous convolutional layers with different sampling rates, and then, the features extracted from each sampling rate are fused to generate the final result. As a consequence, the image features of different sizes of receptive fields are fused to be used to predict the object label, and this approach is called atrous spatial pyramid pooling (ASPP) [34].

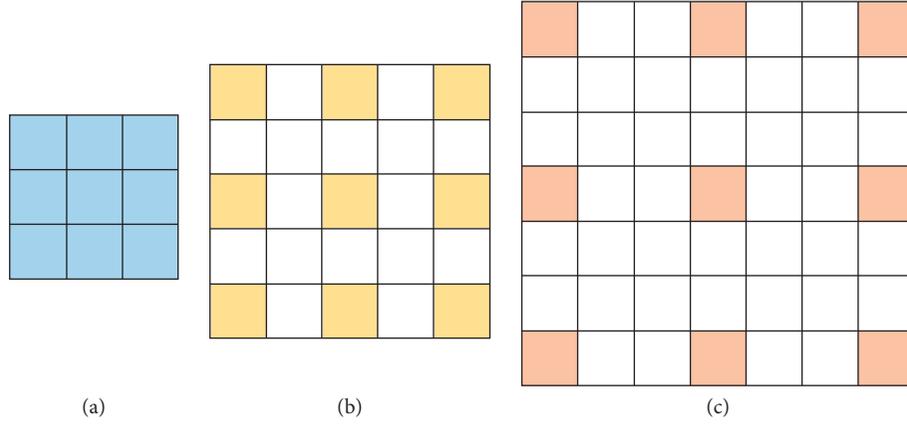


FIGURE 2: Illustration of atrous convolution with different atrous sampling rates in 1-D. (a) Rate = 1. (b) Rate = 2. (c) Rate = 3.

Because the size of digital mammograms is 4096×3328 , FC-DenseNet needs many downsampling operations to reduce the resolution of feature map to acquire the abstracted image features, and this increases the learning parameters and is easy to lead network overfit. In addition, it leads to the deep network and the large computation and memory consumption. So, we resized the input image to 512×512 . For the smaller tumors which have 200 pixels, the final size is almost 30 pixels. But FC-DenseNet has 5 pooling layers, and the down sample rate is 32, resulting in omitting the small size tumor. Although the upsampling path of FC-DenseNet is concatenated with the features extracted from the downsampling path through skip connection, but this still affects the final segmentation precision. Therefore, the downsampling operation of FC-DenseNet is reduced to 4. At the same time, the dense block between the first transition up and the last transition down of the original FC-DenseNet network is removed to ASPP module. The ASPP module consists of 1×1 convolution and atrous sampling rates of 6, 12, and 18, respectively, and the output feature maps of these 4 atrous convolutions are combined with the output of the downsampling operation before. Finally, the concatenated feature map pass through another 1×1 convolution. At the same time, the original pooling layer is changed to convolution layer with 3×3 kernel size and stride of 2, in order to reduce the information loss in the max pooling operation. We use FC-DenseNet with 56 layers which has 4 layers per dense block and a growth rate of 12 as based network. The proposed network (we name it ASPP-FC-DenseNet) finally has 4 transition down, 4 transition up and ASPP module, totally 47 layers, as shown in Figure 3.

In the high-resolution digital mammograms image, there is an extreme imbalance between foreground (tumor) and background classes, which causes the classifier to be more biased to the background class in training and leads to poor segmentation results. Therefore, we improve the simple softmax cross entropy loss function and take each class frequency of the image into consideration. Assume that the frequency of class l in the training data is f_l , and the sum of the frequency of all categories (background and tumor) is 1; that is, $\sum_l f_l = 1$. The inverse frequency of each class is added

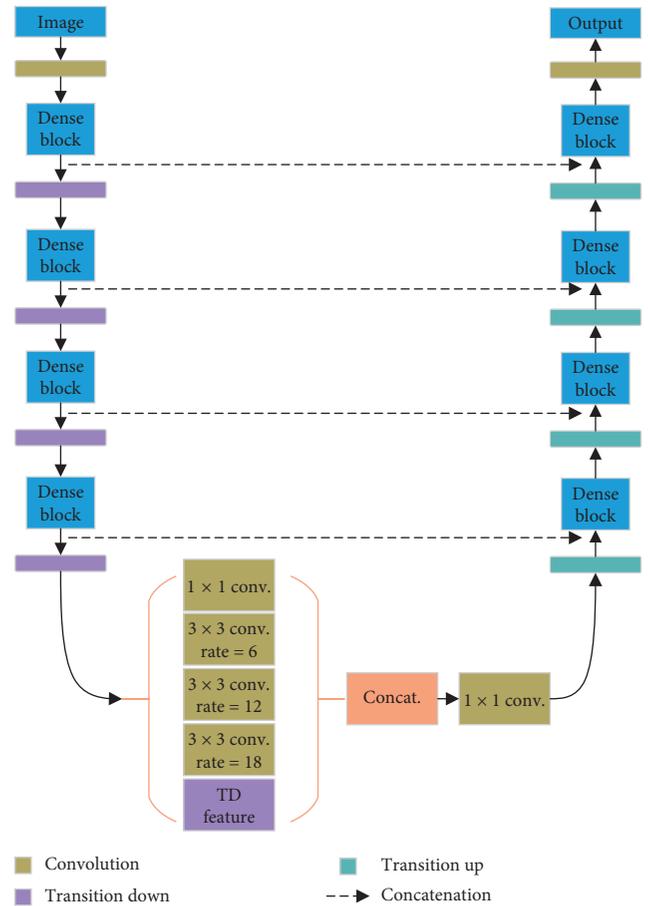


FIGURE 3: The proposed ASPP-FC-DenseNet.

to the cross entropy loss function, effectively strengthening each pixel of less frequency classes.

$$\text{loss}_{\text{ce}} = -\frac{1}{N} \sum_{i=1}^N \frac{1}{f_{y_i}} y_i \log p_i, \quad (1)$$

where N is the number of pixels of the image, y_i is the class label of pixel i , and p_i is the model's predicted probability for the pixel with the correct class label.

3. Results and Discussion

3.1. Data. The data used in our research are digital molybdenum target mammograms developed by the Department of Radiology in Henan Provincial People's Hospital. The mammograms dataset totally contains 190 patient cases, each of which contains the craniocaudal (CC) view and mediolateral oblique (MLO) view, as shown in Figure 4. That means there are 380 images in total. MLO and CC images are gray-level digitized mammograms with a resolution of 3328 (width) by 4096 (height) pixels saved as standard DICOM format. All the tumors in the mammograms were depicted by a professional radiologist in the hospital, as shown in Figure 4. We randomly divided the dataset into train set, validation set, and test set with no cross among them. The train set contained 230 images, and the validation and test set, respectively, contained 75 images.

3.2. Metrics. In this paper, we chose the Dice Index (DI) which is commonly used in most medical image segmentation task and pixel accuracy (PA) and Intersection Over Union (IOU) which is commonly preferred in natural image segmentation task to quantitatively evaluate the segmentation performance of the breast tumor segmentation algorithm. The calculation is shown as follows:

$$\begin{aligned} DI &= \frac{2 \times TP}{2 \times TP + FP + FN}, \\ PA &= \frac{TP}{TP + FN}, \\ IOU &= \frac{TP}{TP + FN + FP}. \end{aligned} \quad (2)$$

Among them, TP refers to the number of pixels that are correctly divided into tumors. FP is the number of background pixels that are wrongly judged as the tumor. TN is the number of pixels that are correctly identified as the background, and FN represents the number of tumor pixels that are identified as background.

3.3. Experimental Evaluation and Discussion. We evaluated the proposed model on our collected mammogram dataset. The initial learning rate of our proposed network is set to 0.001, and the Adam optimization algorithm [41] with default beta values is used to update the gradient and network parameters. Dropout with a rate of 0.2 and batch normalization are also used as a regularizer. The training batch size is set to 1, and we train our model for 100 epochs to compensate for the smaller batch size. Every pixel value in the mammograms is normalized into 0-1 and subtracts the pixel mean value as the network input.

In order to verify the performance of the proposed ASPP-FC-DenseNet, we compared the ASPP-FC-DenseNet with the original FC-DenseNet containing 5 downsampling operations on the test set. Figure 5 shows these two methods' segmentation results for different sizes of tumors. The ASPP-FC-DenseNet algorithm has a higher segmentation precision

compared with FC-DenseNet and has obvious advantage on the edge preservation of different sizes of tumor. Therefore, the fusion of multiscale image information can help to obtain the multilevel image features and improve the performance of image segmentation which needs pixel level semantic recognition. At the same time, FC-DenseNet still has high recognition accuracy for the first two mammogram images with small difference in the internal and external gray values of tumor. It can identify the tumor location accurately and verify the effectiveness of FC-DenseNet for breast cancer segmentation.

As shown in Table 1, the mean Dice Index of ASPP-FC-DenseNet algorithm on the tumor segmentation test set is 0.7697, the mean IOU is 0.6041, and the mean pixel accuracy is 0.7983. Both the Dice Index and the IOU of ASPP-FC-DenseNet algorithm have a small increase. The Dice Index is increased by 3.42%, the IOU is increased by 1%, and the pixel accuracy is almost not improved. Because the pixel accuracy is mainly concerned with the false negatives rate of tumor pixels, and the Dice Index and the IOU consider the false negatives rate and misdetection rate of the tumor pixels at the same time, which can more comprehensively illustrate the segmentation precision of the algorithm. Therefore, it also reflects that ASPP-FC-DenseNet has a competitive advantage in reducing the misdiagnosis rate of the tumor pixels in the same case.

We proposed the weighted cross entropy loss to mitigate the extreme imbalance between foreground (tumor) and background pixels counts. We also compared with normal cross entropy loss (no weighted) and dice loss which is recently proposed in medical image segmentation [42] to show the importance of the weighted cross entropy loss. The tumor segmentation results of ASPP-FC-DenseNet model with different loss functions are shown in Figure 6. From the segmentation results, weighted loss model has a lower false negative rate, and the segmented tumor contour is more accurate than other two loss models. For this reason, the dice coefficient of weighted loss is obviously higher which is same with the quantitative comparisons of Table 2. This shows that the weighted cross entropy loss has a better performance in class imbalance problem compared with common cross entropy loss. The computation of dice loss might lead the gradient and training unstable, and sometimes influence the performance.

We also performed experiments compared with the original PSPNet [43], deeplab v3+ [44], and U-Net to demonstrate the selected model superiority. The tumor segmentation results of different models are shown in Figure 7. Compared with other baseline models, our proposed ASPP-FC-DenseNet model has high segmentation precision for different size of tumors and complex backgrounds. All models have accurate tumor localization, but other three models could not obtain accurate tumor boundaries compared with our model. From these comparison results, we also found that U-Net had high false negative rate. Compared with U-Net, PSPNet and deeplab v3+ models all merge multiresolution image features and have a better segmentation performance. This also proves the importance of multiscale image features

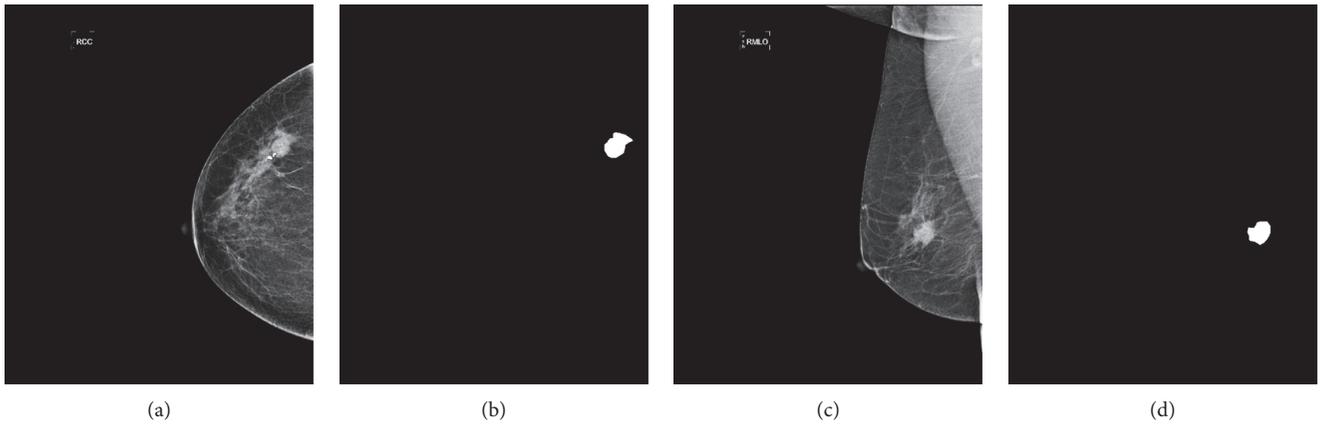


FIGURE 4: The mammogram data with different views and corresponding annotations. (a) CC view. (b) CC annotation. (c) MLO view. (d) MLO annotation.

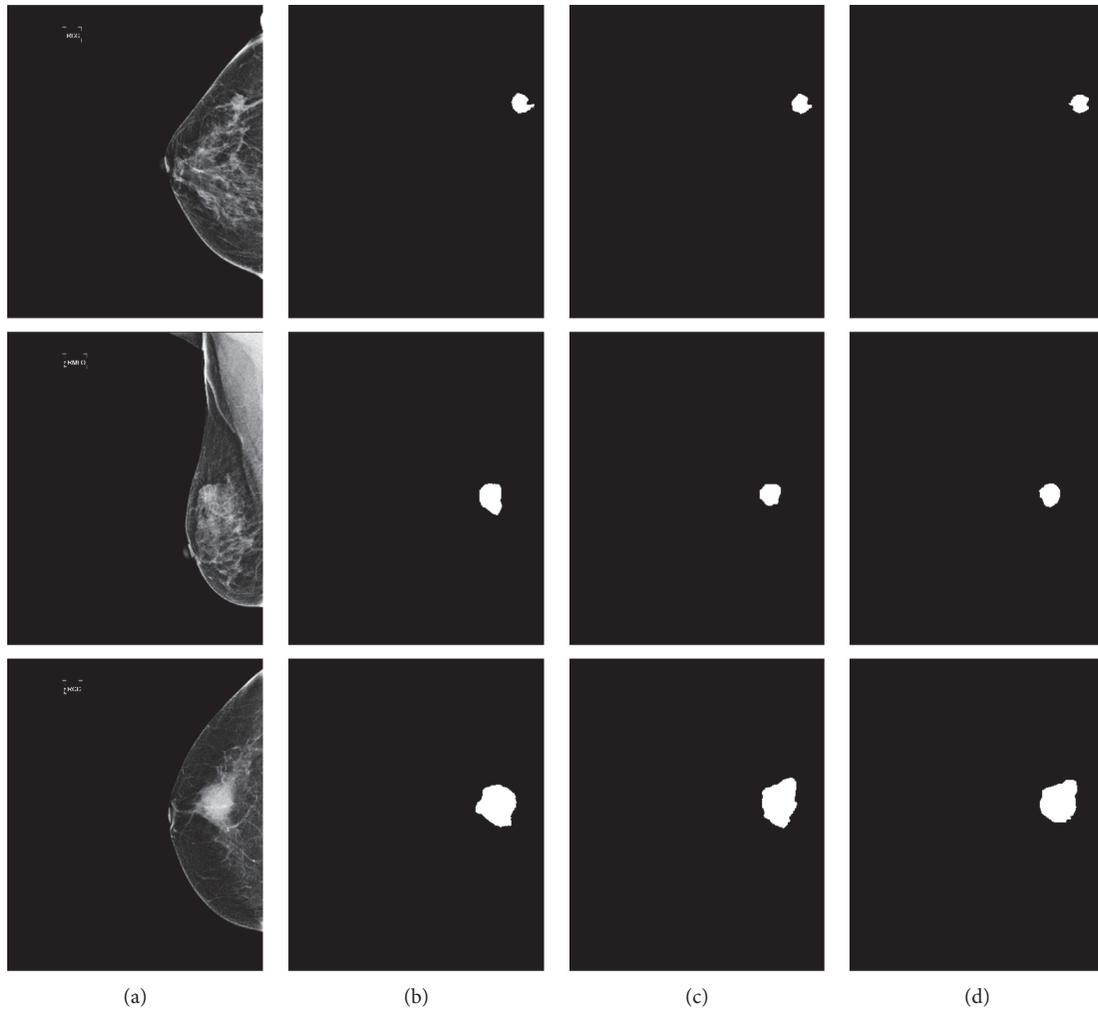


FIGURE 5: Segmentation results of different sizes of breast tumor. (a) Image. (b) FC-DenseNet. (c) ASPP-FC-DenseNet. (d) Ground Truth.

and verifies the advantage of added ASPP module at the same time.

The quantitative comparison results of these four models are shown in Table 3. Our model is obviously

superior to the other three models on the three evaluation metrics. The minimal improvement also reaches 4%. Deeplab v3+ model with encode-decode structure and ASPP module has higher segmentation precision compared

TABLE 1: The quantitative comparisons of the proposed and original FC-DenseNet algorithms.

Methods	DI	IOU	PA
FC-DenseNet	0.7355	0.5948	0.7968
ASPP-FC-DenseNet	0.7697	0.6041	0.7983

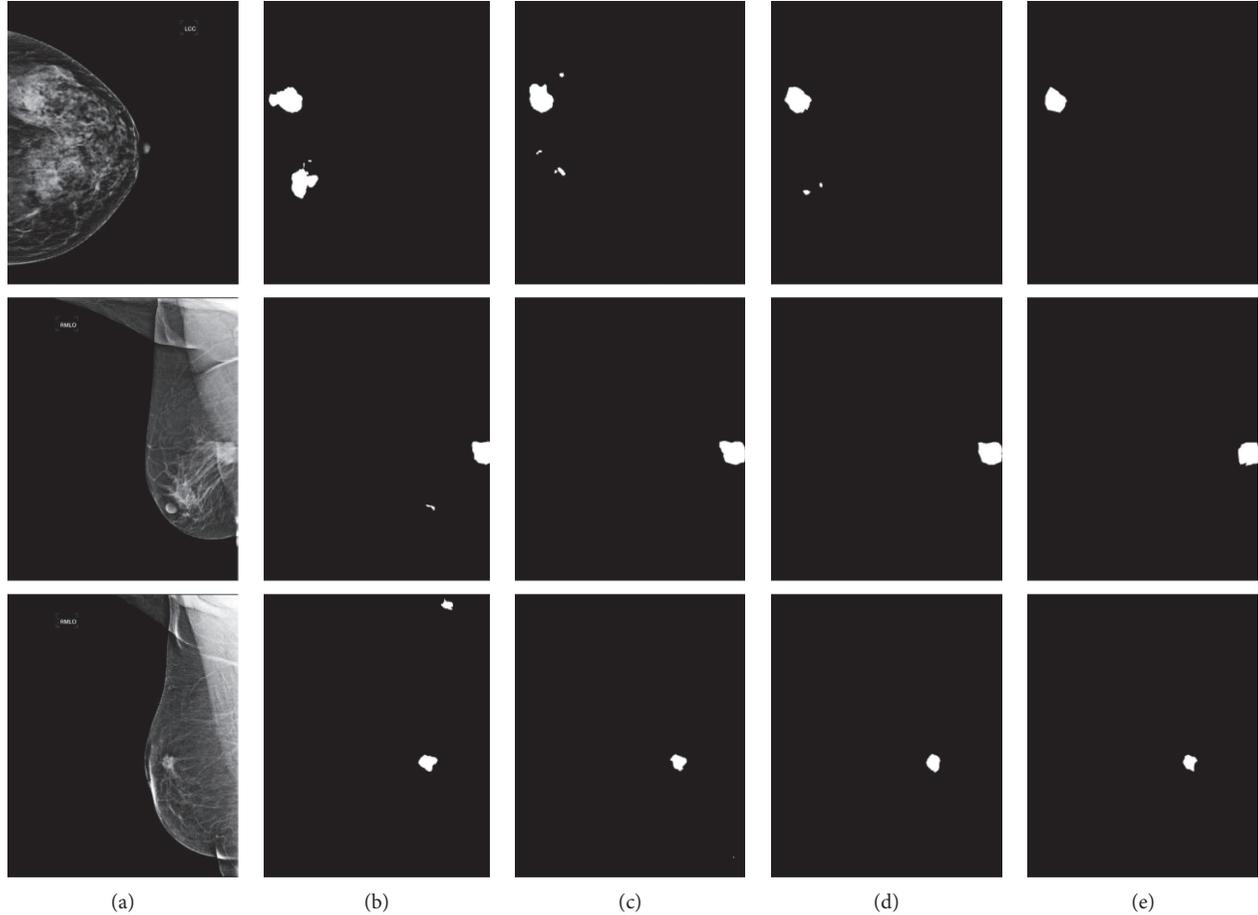


FIGURE 6: The tumor segmentation results of ASPP-FC-DenseNet model with different loss functions. (a) Image. (b) Dice loss. (c) No weighted loss. (d) Ours. (e) Ground Truth.

TABLE 2: The quantitative comparisons of ASPP-FC-DenseNet model with different loss functions.

Methods	DI	IOU	PA
No weighted loss	0.7151	0.5974	0.8015
Dice loss	0.7108	0.5920	0.7988
Ours	0.7697	0.6041	0.7983

with the other two models, which also demonstrated the advantage of these two structures. But the decode module of deeplab v3+ model used simple bilinear upsample operation which might lose detailed low-level features. Our model refers to U-Net decode module to recover image resolution step by step and concatenates with image features in the encode module. Therefore, it has a higher segmentation precision.

We also select the level set [45], graph cut [46], and threshold segmentation algorithm for qualitative and

quantitative comparison on the breast tumor segmentation test set.

The gray value of the pectoralis in the MLO view image is very close to the tumor, which affects the segmentation precision. Before using the level set, graph cut, and threshold segmentation algorithm, the pectoralis of the MLO image is removed first according to the location information and the gray threshold, as shown in Figure 8. Threshold segmentation algorithm is a simple image segmentation algorithm. We used a double threshold segmentation method. Firstly, the initial tumor region is obtained by the iterative threshold segmentation algorithm. Then the final threshold segmentation result is obtained by calculating the gray mean value of the first step of the tumor region as the threshold of the second step segmentation. For the segmentation results of these three contrast algorithms, the area of connected region is calculated and the isolated small connected region is deleted as the final segmentation result.

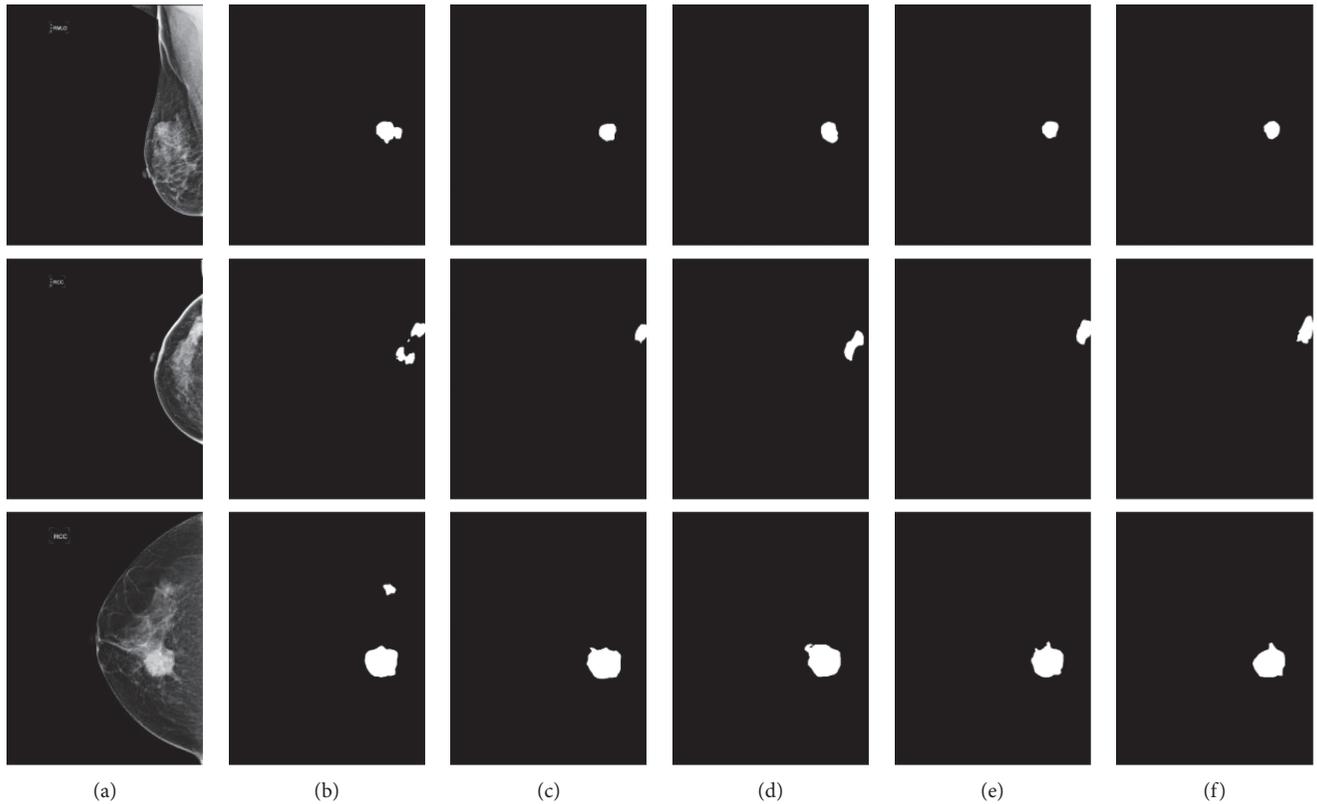


FIGURE 7: Tumor segmentation results of different CNNs. (a) Image. (b) U-Net. (c) PSPNet. (d) Deeplab v3+. (e) Ours. (f) Ground Truth.

TABLE 3: The quantitative comparisons of different CNNs.

Models	DI	IOU	PA
U-Net	0.6763	0.5608	0.7562
PSPNet	0.6785	0.5427	0.7202
Deeplab v3+	0.6827	0.5641	0.7072
ASPP-FC-DenseNet	0.7697	0.6041	0.7983

The segmentation results of our proposed algorithm are compared with other three segmentation algorithms, as shown in Figure 9. It is obvious that other three segmentation algorithms using preprocessing that removes the pectoralis and postprocessing still have a poor segmentation performance compared with proposed ASPP-FC-DenseNet segmentation algorithm. The other three segmentation algorithms can accurately locate the location of the tumor, but it has a high misdetection rate, especially for the tumor images with small difference and similar characteristics inside and outside the tumor. This is due to the differences of the tumor grayscale, texture, and other characteristics with the normal breast tissue, so it is easy to locate the tumor position. Although level set and graph cut are interactive segmentation algorithms, which can minimize the energy function by manually setting the initial location of the tumor as a priori, they are mainly based on the tumor grayscale and texture properties to update the evolution curves and lack the high-level semantic information of the image. For the non-tumor region whose characteristic is similar to tumor, these three algorithms have a poor segmentation result. The proposed

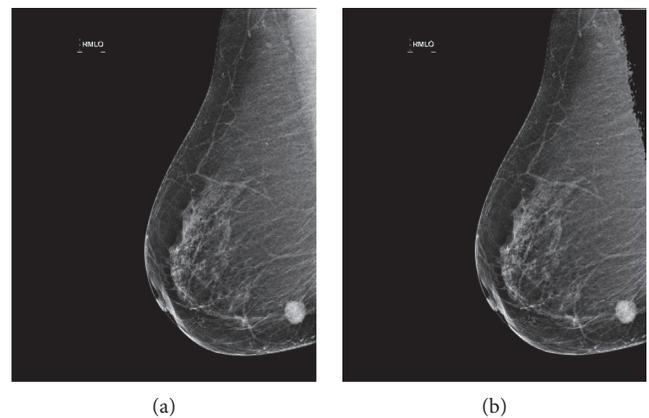


FIGURE 8: The original MLO mammogram and pectoralis deleted mammogram.

algorithm has powerful feature extraction and representation ability and can obtain more accurate segmentation results.

The quantitative comparison results of these four segmentation algorithms are shown in Table 4. The proposed ASPP-FC-DenseNet algorithm has significant improvement on the three evaluation metrics compared with the other three segmentation algorithms. Compared with the graph cut algorithm, ASPP-FC-DenseNet improved 30.34% on the Dice Index, increased by 25.50% on the IOU, and increased by 17.63% on the pixel accuracy. Even compared to the level set which has better segmentation performance in the three

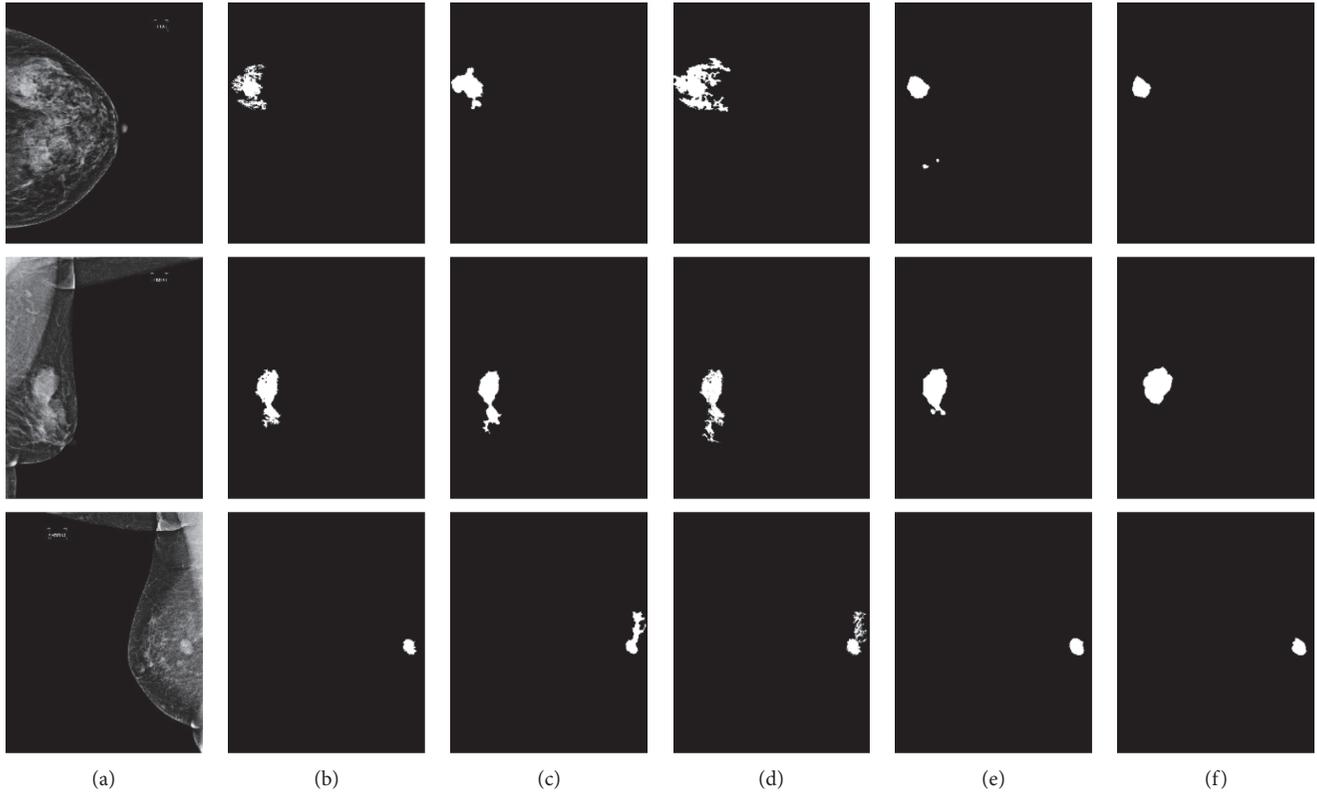


FIGURE 9: The tumor segmentation results compared with other segmentation algorithms. (a) Image. (b) Level set. (c) Grab cut. (d) Double threshold. (e) Ours. (f) Ground Truth.

TABLE 4: The quantitative comparisons of the proposed model and other algorithms.

Methods	DI	IOU	PA
Level set	0.5989	0.4893	0.6813
Grab cut	0.4663	0.3491	0.6220
Threshold	0.5464	0.4322	0.6440
ASPP-FC-DenseNet	0.7697	0.6041	0.7983

algorithms, the proposed algorithm also increased by 17.08% on the Dice Index. The IOU increased by 11.48%, and the pixel accuracy increased by 11.70%. The segmentation precision was improved significantly. At the same time, it can be seen that the proposed algorithm is more effective in improving the Dice Index and IOU, indicating that under the same misdetection rate, the proposed algorithm has a lower false negative rate compared with the other three algorithms and has a higher segmentation precision.

4. Conclusions

In this paper, a fully convolutional network ASPP-FC-DenseNet, which combines multiscale image information, is proposed to achieve automatic segmentation of breast tumor. The algorithm uses FC-DenseNet which further exploits the feature reuse by using skip connections to help the upsampling path recover spatially detailed information from the downsampling path. Considering that five pooling

layers in the network lead the small size tumor that cannot be identified easily, the number of pooling layers in the network is reduced to 4. Then the atrous spatial pyramid pooling module is added to the network after the last downsampling operation, which concatenates different field-of-views of image features through the combination of multiple sampling rates of atrous convolution. Finally, the loss function of the network is improved according to the proportion of the tumor pixels in the image, in order to weaken unbalanced class problem. Qualitative and quantitative experimental results prove that the algorithm proposed in this paper has high segmentation precision for various sizes and shapes of tumor mammograms without preprocessing and post-processing and achieves automatic tumor segmentation.

Data Availability

The digital mammogram data used to support the findings of this study have not been made available because of the patients' privacy.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the National Key R&D Program of China under grant 2017YFB1002502 and National

Natural Science Foundation of China (nos. 61701089 and 61601518).

References

- [1] A. McGuire, J. Brown, C. Malone, R. McLaughlin, and M. Kerin, "Effects of age on the detection and management of breast cancer," *Cancers*, vol. 7, no. 2, pp. 908–929, 2015.
- [2] S. W. Zucker, "Region growing: childhood and adolescence," *Computer Graphics and Image Processing*, vol. 5, pp. 382–399, 1976.
- [3] S. Beucher and C. Lantuejoul, "Use of watersheds in contour detection," in *Proceedings of International Workshop on Image Processing, Real-Time Edge and Motion Detection/Estimation*, Rennes, France, September 1979.
- [4] B. Zheng, Y.-H. Chang, and D. Gur, "Computerized detection of masses in digitized mammograms using single-image segmentation and a multilayer topographic feature analysis," *Academic Radiology*, vol. 2, no. 11, pp. 959–966, 1995.
- [5] N. Petrick, H.-P. Chan, B. Sahiner, and M. A. Helvie, "Combined adaptive enhancement and region-growing segmentation of breast masses on digitized mammograms," *Medical Physics*, vol. 26, no. 8, pp. 1642–1654, 1999.
- [6] M. R. Hejazi and Y. S. Ho, "Automated detection of tumors in mammograms using two segments for classification," in *Proceedings of Pacific-Rim Conference on Multimedia*, pp. 910–921, Springer, Berlin, Heidelberg, 2005.
- [7] V. T. Santos, H. Schiabel, C. u. E. Goes, and R. H. Benatti, "A segmentation technique to detect masses in dense breast digitized mammograms," *Journal of Digital Imaging*, vol. 15, pp. 210–213, 2002.
- [8] J. Herredsvela, T. O. Gulsrud, and K. Engan, "Detection of circumscribed masses in mammograms using morphological segmentation," in *Proceedings of International Society for Optics and Photonics Medical Imaging 2005: Image Processing*, vol. 5747, pp. 902–914, San Diego, CA, USA, April 2005.
- [9] C. Li, C. Xu, C. Gui et al., "Level set evolution without re-initialization: a new variational formulation," in *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 1, pp. 430–436, San Diego, CA, USA, June 2005.
- [10] Y. Yuan, M. L. Giger, H. Li, K. Suzuki, and C. Sennett, "A dual-stage method for lesion segmentation on digital mammograms," *Medical Physics*, vol. 34, no. 11, pp. 4180–4193, 2007.
- [11] H.-C. Kuo, M. L. Giger, I. Reiser et al., "Level set segmentation of breast masses in contrast-enhanced dedicated breast CT and evaluation of stopping criteria," *Journal of Digital Imaging*, vol. 27, no. 2, pp. 237–247, 2013.
- [12] Y. Wang, D. Tao, X. Gao, X. Li, and B. Wang, "Mammographic mass segmentation: embedding multiple features in vector-valued level set in ambiguous regions," *Pattern Recognition*, vol. 44, no. 9, pp. 1903–1915, 2011.
- [13] N. Dhungel, G. Carneiro, and A. P. Bradley, "Deep structured learning for mass segmentation from mammograms," in *Proceedings of 2015 IEEE International Conference on Image Processing (ICIP)*, pp. 2950–2954, Quebec, Canada, September 2015.
- [14] N. Dhungel, G. Carneiro, and A. P. Bradley, "Tree re-weighted belief propagation using deep learning potentials for mass segmentation from mammograms," in *Proceedings of 2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI)*, pp. 760–763, Brooklyn, NY, USA, April 2015.
- [15] N. Dhungel, G. Carneiro, and A. P. Bradley, "Deep learning and structured prediction for the segmentation of mass in mammograms," in *Proceedings of International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 605–612, Springer, Cham, Switzerland, October 2015.
- [16] W. Zhu and X. Xie, "Adversarial deep structural networks for mammographic mass segmentation," 2016, <http://arxiv.org/abs/1612.05970>.
- [17] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in Neural Information Processing Systems*, pp. 1097–1105, 2012.
- [18] C. Szegedy, W. Liu, Y. Jia et al., "Going deeper with convolutions," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–9, Boston, MA, USA, 2015.
- [19] K. He, X. Zhang, S. Ren et al., "Deep residual learning for image recognition," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778, Las Vegas, NV, USA, 2016.
- [20] J. Redmon, S. Divvala, R. Girshick et al., "You only look once: unified, real-time object detection," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 779–788, Las Vegas, NV, USA, 2016.
- [21] R. Girshick, "Fast R-CNN," in *Proceedings of IEEE International Conference on Computer Vision*, pp. 1440–1448, Santiago, Chile, December 2015.
- [22] S. Ren, K. He, R. Girshick et al., "Faster R-CNN: towards real-time object detection with region proposal networks," *Advances in Neural Information Processing Systems*, pp. 91–99, 2015.
- [23] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3431–3440, Boston, MA, USA, June 2015.
- [24] H. Noh, S. Hong, and B. Han, "Learning deconvolution network for semantic segmentation," in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 1520–1528, Santiago, Chile, 2015.
- [25] 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, 2017.
- [26] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, <http://arxiv.org/abs/1409.1556>.
- [27] O. Ronneberger, P. Fischer, and T. Brox, "U-net: convolutional networks for biomedical image segmentation," in *Proceedings of International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 234–241, Springer, Cham, Switzerland, 2015.
- [28] S. Jégou, M. Drozdal, D. Vazquez et al., "The one hundred layers tiramisu: fully convolutional densenets for semantic segmentation," in *Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 1175–1183, Honolulu, Hawaii, July 2017.
- [29] G. Huang, Z. Liu, K. Q. Weinberger et al., "Densely connected convolutional networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, vol. 1, p. 32, Honolulu, HI, USA, July 2017.
- [30] M. Hmida, K. Hamrouni, B. Solaiman et al., "Breast mass segmentation in mammograms combining fuzzy c-means and active contours," in *Proceedings of Tenth International*

- Conference on Machine Vision (ICMV 2017)*, International Society for Optics and Photonics, Vienna, Austria, 2018.
- [31] V. K. Singh, H. A. Rashwan, S. Romani et al., "Breast mass segmentation and shape classification in mammograms using deep neural networks," 2018, <http://arxiv.org/abs/1809.01687>.
 - [32] G. Sciortino, D. Tegolo, and C. Valenti, "Automatic detection and measurement of nuchal translucency," *Computers in Biology and Medicine*, vol. 82, pp. 12–20, 2017.
 - [33] D. Guastella and C. Valenti, "Cartoon filter via adaptive abstraction," *Journal of Visual Communication and Image Representation*, vol. 36, pp. 149–158, 2016.
 - [34] L. C. Chen, G. Papandreou, F. Schroff et al., "Rethinking atrous convolution for semantic image segmentation," 2017, <http://arxiv.org/abs/1706.05587>.
 - [35] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 4, pp. 834–848, 2018.
 - [36] S. Ioffe and C. Szegedy, "Batch normalization: accelerating deep network training by reducing internal covariate shift," 2015, <http://arxiv.org/abs/1502.03167>.
 - [37] V. Nair and G. E. Hinton, "Rectified linear units improve restricted Boltzmann machines," in *Proceedings of the 27th International Conference on Machine Learning (ICML-10)*, pp. 807–814, Haifa, Israel, June 2010.
 - [38] C. Farabet, C. Couprie, L. Najman, and Y. LeCun, "Learning hierarchical features for scene labeling," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 8, pp. 1915–1929, 2013.
 - [39] P. Moeskops, M. A. Viergever, A. M. Mendrik, L. S. de Vries, M. J. N. L. Benders, and I. Isgum, "Automatic segmentation of MR brain images with a convolutional neural network," *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1252–1261, 2016.
 - [40] K. He, X. Zhang, S. Ren et al., "Spatial pyramid pooling in deep convolutional networks for visual recognition," in *Proceedings of European Conference on Computer Vision*, pp. 346–361, Springer, Cham, Switzerland, September 2014.
 - [41] D. P. Kingma and J. Ba, "Adam: a method for stochastic optimization," 2014, <http://arxiv.org/abs/1412.6980>.
 - [42] F. Milletari, N. Navab, and S. A. Ahmadi, "V-net: fully convolutional neural networks for volumetric medical image segmentation," in *Proceedings of 2016 Fourth International Conference on 3D Vision (3DV)*, pp. 565–571, Stanford, CA, USA, October 2016.
 - [43] H. Zhao, J. Shi, X. Qi et al., "Pyramid scene parsing network," in *Proceedings of IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, pp. 2881–2890, Honolulu, Hawaii, July 2017.
 - [44] L. C. Chen, Y. Zhu, G. Papandreou et al., "Encoder-decoder with atrous separable convolution for semantic image segmentation," 2018, arXiv preprint [arXiv:1802.02611](https://arxiv.org/abs/1802.02611).
 - [45] C. Li, R. Huang, Z. Ding et al., "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, 2011.
 - [46] J. Yuan, E. Bae, and X. C. Tai, "A study on continuous max-flow and min-cut approaches," in *Proceedings of 2010 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2217–2224, San Francisco, CA, USA, June 2010.

Research Article

Analysis of the Cluster Prominence Feature for Detecting Calcifications in Mammograms

Alejandra Cruz-Bernal,^{1,2} Martha M. Flores-Barranco,¹ Dora L. Almanza-Ojeda ,^{1,3}
Sergio Ledesma ,³ and Mario A. Ibarra-Manzano ,^{1,3}

¹Laboratorio de Procesamiento Digital de Señales, Departamento de Ingeniería Electrónica, DICIS, Universidad de Guanajuato, Carr. Salamanca-Valle de Santiago KM. 3.5 + 1.8 Km., Salamanca 36885, Mexico

²Departamento de Ingeniería Robótica, Universidad Politécnica de Guanajuato, Av. Universidad Norte SN., Comunidad Juan Alonso, Cortazar 38496, Mexico

³Cuerpo Académico de Telemática, DICIS, Universidad de Guanajuato, Carr. Salamanca-Valle de Santiago KM. 3.5 + 1.8 Km., Salamanca 36885, Mexico

Correspondence should be addressed to Dora L. Almanza-Ojeda; dora.almanza@ugto.mx

Received 30 June 2018; Revised 22 October 2018; Accepted 6 November 2018; Published 30 December 2018

Guest Editor: Cesare Valenti

Copyright © 2018 Alejandra Cruz-Bernal et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In mammograms, a calcification is represented as small but brilliant white region of the digital image. Earlier detection of malignant calcifications in patients provides high expectation of surviving to this disease. Nevertheless, white regions are difficult to see by visual inspection because a mammogram is a gray-scale image of the breast. To help radiologists in detecting abnormal calcification, computer-inspection methods of mammograms have been proposed; however, it remains an open important issue. In this context, we propose a strategy for detecting calcifications in mammograms based on the analysis of the cluster prominence (*cp*) feature histogram. The highest frequencies of the *cp* histogram describe the calcifications on the mammography. Therefore, we obtain a function that models the behaviour of the *cp* histogram using the Vandermonde interpolation twice. The first interpolation yields a global representation, and the second models the highest frequencies of the histogram. A weak classifier is used for obtaining a final classification of the mammography, that is, with or without calcifications. Experimental results are compared with real DICOM images and their corresponding diagnosis provided by expert radiologists, showing that the *cp* feature is highly discriminative.

1. Introduction

Breast cancer is the top cancer that affects women both in developed and developing countries. Early detection of breast cancer increases treatment options and survival expectation [1]. Breast cancer statistics report that nearly 2 million of new cases were diagnosed in 2018; this represents about 12% of all new cancer cases and 25.3% of all cancers in women [2]. For an analysis of the efficiency in each stage of the cancer disease, the diagnosis, registration, and monitoring of diseases allow to validate the most appropriate treatments, including the optimization of costs [3]. The early detection for improving breast cancer outcome and survival remains an open issue.

The DICOM is the format used for registering a digitized mammographic image. Nowadays, mammography is a reliable method for breast cancer detection. In addition, several computer-aided detections (CAD systems) help mammogram processing to provide more accurate results [4]. A mammography is a low-energy radiography of the breast. The radiologist uses this method to localize morphological alterations and infers the presence or absence of anomalies, mainly small calcifications [5]. Breast calcifications are small spots of calcium salts in the breast tissue. The calcifications in the mammography appear as small white spots. There are two different types of calcifications, microcalcifications and macrocalcifications [6]. The macrocalcifications are large and coarse, mostly benign and

associated with the age. The microcalcifications can be early signs of breast cancer, with or without a visible mass. This mass can be a benign tumor, a cyst, or cancer.

The detection of microcalcifications in mammograms is suboptimal because it depends on the radiologist's experience, criterion, fatigue, and visual capability. As a result, radiologists fail to detect breast cancers due to misinterpretation of the lesion and can lead to a greater number of false-positive cases. Another aspect is the accuracy with which the radiologist considers the medical importance of the calcification regions because sometimes these region's sizes can be misinterpreted in the mammograms [7]. Microcalcifications are bright spots whose size oscillates between 0.1 mm and 1 mm [8], and usually, they are not easy to see. Furthermore, only certain suspicious microcalcifications (<0.5 mm) are alarming, as it is verified in a comparative study between cancer size measurements and the results of pathology [9]. As a result, radiologists fail to detect breast cancers.

Several methods have been developed to assist the radiologist in the detection of calcifications using mammography images [10, 11] or computed aid detection (CAD systems). CAD systems have had significant technology advances yield to detect segmentation and classify microcalcification clusters at digital mammograms. Thus, CAD systems have been used clinically for more than two decades as "second lector" in the diagnosis carried out by the radiologists [9, 12]. The use of CAD systems is popular due to the high sensitivity detection averaging up to 90% [13]. On the other hand, the images obtained during the acquisition of a mammography are low contrast, making the processing a challenging task. Some methods are proposed to segment several types of microcalcifications using texture features. Kim and Park [14] compared the surrounding region-dependence method (SRDM) to the other conventional texture-analysis methods with respect to detection of clustered microcalcifications in digitized mammograms. The performance results of the classification are usually evaluated using receiver operating curve (ROC) curve that describes the discrimination capacity of the approach [15]. Yadollahpour and Hamed in [16] presents a review of various methods considering texture analysis for mass and microcalcification detection in mammography used for early breast cancer detection. Jalalian et al. in [17] obtains statistical texture feature based on the co-occurrence matrix from a segmented volume of interest. The classification stage uses a multilayer perceptron neural network achieving high accuracy results.

Both applications Hough transform and threshold-based method were considered by Fanzinni et al. [18] as strong techniques that yield to group single microcalcifications, as "success events," into clusters using a set of expert codified rules. From this, it is possible to obtain the regions that contain the lesions of interest. High-frequency filters were used by Lauria et al. [19] as preprocessing filters to carry out the segmentation of possible suspicious areas at mammography. The microcalcification analysis yield to find and classify regions of interest (ROIs) by means of two neural networks: the first is a feedforward neural network, while the

second neural network uses the principal component method to end the process of classification. Samala et al. in [20] presents a study corresponding to the advantages and challenges for detection of microcalcifications in digital mammograms and digital breast tomosynthesis from a CAD systems perspective.

Basic concepts of mathematics define an inflection point as a point on the curve at which the sign of the curvature (i.e., the concavity) changes. In this approach, such changes on the curvature represent the fluctuations/variations of agglomerations in the image with respect to the specific level of the cp texture feature. In this paper, the analysis of the cluster prominence cp texture feature on mammography images is carried out to detect breast calcifications that may indicate cancer. The modelling of the last part of the cp histogram reveals the presence of microcalcifications with respect to another common mass on the breast tissue. The final classification of the mammography is performed using a weak classifier. The document is organized as follows: the next section describes the proposed strategy for detecting calcifications. Also, it explains the cp feature computation, the global and local interpolation, and the classification of the cp histogram. The experimental results and conclusions are, respectively, provided in the last part of the document.

2. Materials and Methods

The analysis of the cluster prominence (cp) texture feature is proposed to accurately detect calcifications in digital mammography. The proposed approach described here uses the mammography dataset collected at the General Hospital of Irapuato, Guanajuato. The dataset consists of 74 images: 22 diagnosed with calcifications and 52 diagnosed as normal tissue, in accordance with the classification system for breast mammogram, BI-RADS. An overview of our proposed approach is depicted in Figure 1, presenting mainly three stages: (1) feature extraction, (2) analysis of calcifications, and (3) classification of the mammography. First, the cp texture feature is computed from a mammogram using the sum and difference histograms (SDH) technique [21].

This feature is a measure of asymmetry for which a high value indicates large changes in the gray-scale levels of the image [22]. In other words, cp is a measure of how uniform is the gray-scale level distribution. In accordance with this, it is proposed that high cp values represent calcifications that can be more deeply analysed for validating the detection of calcifications.

The cp feature histogram is analysed using the Vandermonde technique. A global interpolation function $f(cp)$ that best describes the cp feature histogram is obtained. This function is used to analyse the range of the histogram with high possibilities of finding calcifications, by means of the n inflection points in $f(cp)$, here referred as Z_{cp} . Each inflection point represents different and noticeable agglomeration zones of the mammography with a high level of cp texture feature. However, a deeper analysis must be performed to search and validate the presence of calcifications on one specific zone of the high frequency. To accomplish this, the $g(cp)$ function is obtained from a local interpolation on the

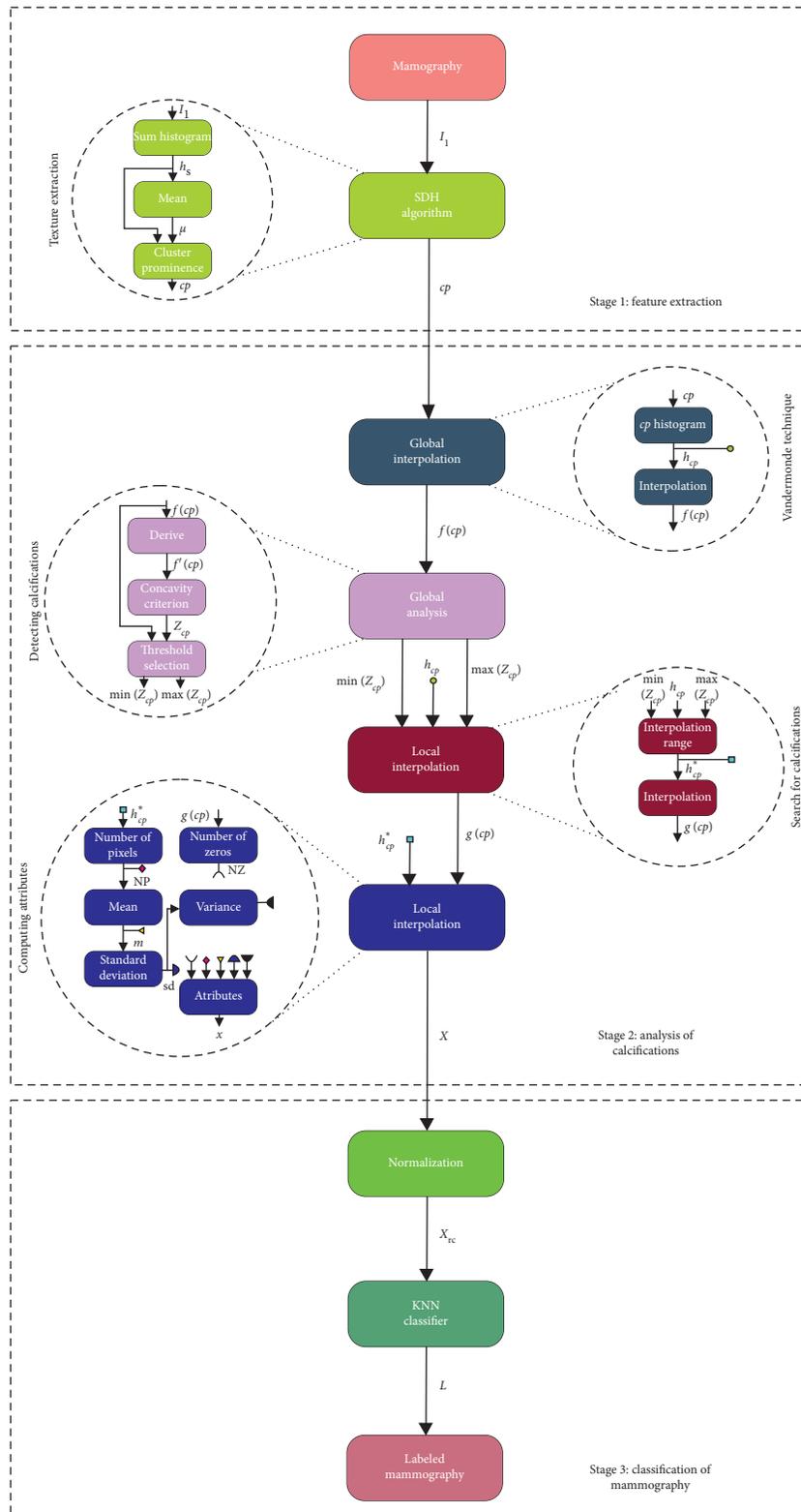


FIGURE 1: Block diagram of methodology to detect calcifications in mammograms.

high frequency range of the cp histogram. Using a specific range of the $g(cp)$ function, the attributes X vector is computed.

The stage of mammograms classification receives the normalized version of statistical attributes (including the number of zeros) of the X vector and then classified using the

K-nearest neighbour. The classification results refer the mammogram as “with calcifications” or “without calcifications.” In the following, these stages will be explained on detail.

2.1. Texture Feature Extraction Using SDH Algorithm. The sum and difference histogram (SDH) technique calculates histograms that collect the results of addition and subtraction of the gray-scale levels on a whole digital image [18]. The SDH requires basic arithmetic operations and less memory storage in comparison with other texture techniques. Furthermore, it stores important information about the image content.

Considering the mammogram image as a rectangular matrix of size $K \times L$, the sum image I_S is obtained by summing each pixel and their surrounded pixels are separated by a set of M relative displacements. The gray level at each pixel is quantified to N_g levels; therefore, the range of the I_S image is $[0, 2(N_g-1)]$. From the I_S image, the sum histogram (h_S) is calculated for a rectangular window of N elements ($N = \text{width} \times \text{height}$, 3×3) storing the cardinality at each coordinate (x, y) in the window with an intensity value i . Finally, the normalized sum histogram $\hat{P}_S(i)$ is given by the following equation:

$$\hat{P}_S(i) = \frac{h_S(i)}{N}. \quad (1)$$

The cp feature is given by the following equation:

$$cp = \sum_i (i - 2\mu)^4 \cdot \hat{P}_S(i). \quad (2)$$

As mentioned above, this feature is used to seek for calcifications in mammograms, as calcifications and agglomerations are related with high values in the cp texture feature.

The size of the region used for computing the cp feature is $D=3 \times 3$ because the calcifications are more visible using this size. The displacement was established only in the horizontal direction, although more directions were evaluated, i.e., 45° , 90° , and 135° without noticeable changes. The value was set at “1” because as mentioned above, the sizes of the microcalcifications are <0.5 mm [9]. This minimal resolution of “1” pixel allows to detect the microcalcifications.

2.2. Analysis of the cp Feature Histogram. In this section, it explains the theoretical context of the Vandermonde technique, and the global and local interpolations are performed for modelling the h_{cp} behaviour. The histogram of the cp feature (here denoted as h_{cp}) might be numerically modelled by a polynomial function $f(cp)$ that globally describes the histogram behaviour. Such $f(cp)$ function is obtained using the Vandermonde interpolation technique. After that, a local interpolation is performed (using the same interpolator) but only for high values of cp features.

2.2.1. Vandermonde Technique for Global Interpolation. The basic procedure to determine the coefficients a_0, a_1, \dots, a_n of the polynomial function

$$P_n(x) = a_0 + a_1x + a_2x^2 + \dots + a_nx^n \quad (3)$$

consists in interpolating the $m+1$ points $(x_0, y_0), (x_1, y_1), \dots, (x_m, y_m)$ yielding a linear system of equations

$$\begin{aligned} P_n(x_0) = y_0 &\longrightarrow a_0 + a_1x_0 + a_2x_0^2 + \dots + a_nx_0^n = y_0, \\ P_n(x_1) = y_1 &\longrightarrow a_0 + a_1x_1 + a_2x_1^2 + \dots + a_nx_1^n = y_1, \\ &\vdots \longrightarrow \vdots \\ P_n(x_m) = y_m &\longrightarrow a_0 + a_1x_m + a_2x_m^2 + \dots + a_nx_m^n = y_m, \end{aligned} \quad (4)$$

or in matrix form $V\bar{a} = \bar{y}$.

$$\begin{pmatrix} 1 & x_0 & x_0^2 & \dots & x_0^n \\ 1 & x_1 & x_1^2 & \dots & x_1^n \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & x_m & x_m^2 & \dots & x_m^n \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_m \end{pmatrix}. \quad (5)$$

The matrix V of this linear system is called *Vandermonde matrix*. As this matrix is nonsingular, the system $V\bar{a} = \bar{y}$ could be solved to obtain the coefficients $\bar{a} = (a_0, a_1, \dots, a_n)$. In this work, the $f(cp)$ function obtained from the Vandermonde interpolation is a polynomial function of order 13. This function is used to obtain the range of the histogram with high possibilities of finding calcifications. To accomplish this, the first derivative of $f(cp)$ is computed, while its n inflection points Z_{cp} :

$$Z_{cp} = \{cp_i \mid f'(cp) = 0, \quad i = 1, 2, \dots, n\}. \quad (6)$$

Figure 2 depicts in asterisks (*) the interpolated function $f(cp)$ and in circles (o) the corresponding inflection points of the function. Note that $f(cp)$ follows the shape of h_{cp} (continuous line), except at the highest values of the cp feature that shows minimal and maximal local frequency values. Therefore, a second interpolation must be performed to fit the h_{cp} behaviour in the last part of the function.

2.3. Global Analysis: Detecting Calcifications. Note in Figure 3 that to calculate the minimal and maximal threshold th_{\min} and th_{\max} , it is necessary to analyse two different behaviours on the last part of the $f(cp)$ function: (1) when the last inflection point in Z_{cp} is maximal and the function decreases; (2) when the last inflection point in Z_{cp} is minimal and the function increases. The behaviour 1 is illustrated by the first two graphs of Figure 3, and the last graph illustrates the second behaviour case.

In this work, we propose a concavity criterion to choose the threshold values. The concavity criterion is established as follows: the searching zone for calcifications is ranged from the penultimate inflection point to the last inflection point if the last inflection point is maximal, otherwise is ranged from

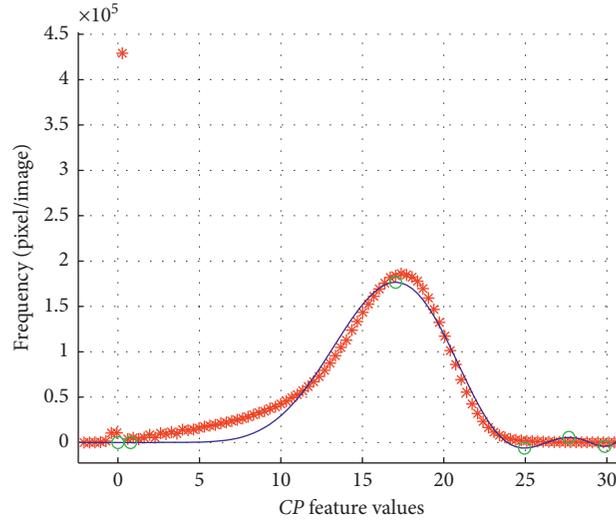


FIGURE 2: Global interpolation of h_{cp} . The asterisks * depict the discrete values of h_{cp} , the continuous blue line is the interpolated polynomial function $f(cp)$, and circles illustrate the inflection points.

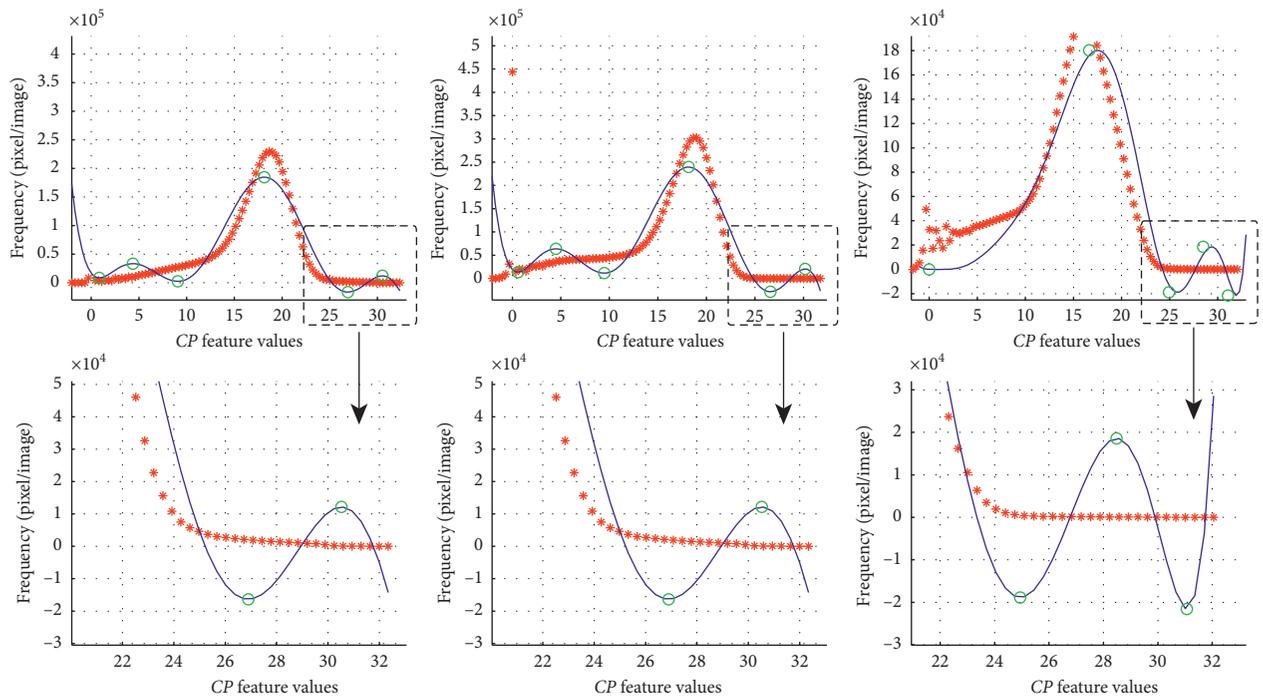


FIGURE 3: Behaviour of the high frequency part of different $f(cp)$ functions. First row: the histograms h_{cp} (red asterisks) and polynomial interpolation $f(cp)$ (blue continuous line) computed from different mammograms diagnosed with calcifications. Second row: a zoom inside of the last part of the graph. The circles represent Z_{cp} .

the antepenultimate inflection point if the last inflection point is minimal:

$$th_{\min} = \begin{cases} cp_{(n-1)}, & \text{if } f(cp_n) > f(cp_{n-1}), \\ cp_{(n-2)}, & \text{otherwise.} \end{cases} \quad (7)$$

In both cases illustrated in Figure 3, the maximal th_{\max} is the last inflection point. The range established by th_{\min} and th_{\max} is used for obtaining a second interpolation function of

h_{cp} . This function will be referred as $g(cp)$, and it is interpreted as the result of a local interpolation.

2.4. Local Interpolation: Search for Calcifications. The part of the histogram h_{cp} located in the ranges th_{\min} and th_{\max} is referred as h_{cp}^* . In this local interpolation, the function $g(cp)$ fits the h_{cp}^* curve values more accurately and is used to calculate the new inflection points of Z_{cp} .

In Figure 4, the global and local interpolation results are compared. Figures 4(a) and 4(b) show, respectively, the

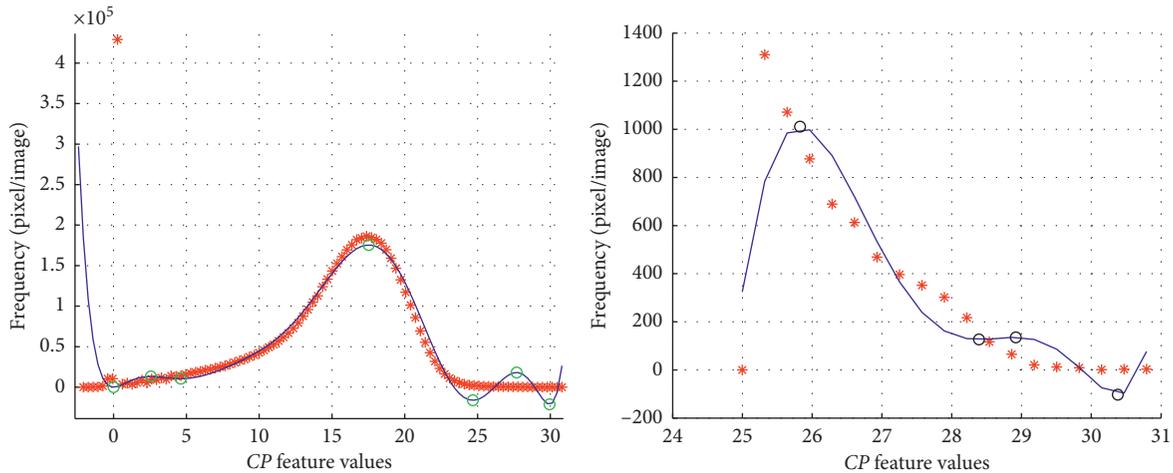


FIGURE 4: Comparing global and local interpolation: (a) global interpolation of h_{cp} ; (b) local interpolation of h_{cp}^* , where new inflection points were found and the variability of h_{cp} is noticeable.

global and local interpolation function $f(cp)$ and $g(cp)$. Note that after the local interpolation process, the function $g(cp)$ (the blue continuous line in Figure 4(b)) fits better the h_{cp}^* values (red asterisks) than in the last part of the global interpolation function of Figure 4(a). Furthermore, the new $g(cp)$ function includes additional inflection points in Z_{cp} (black circles in Figure 4(b)), which better describes the high frequency changes. This variability during second interpolation allows to suppose the presence of calcifications in the mammogram; however, the result will be delivered by the classifier.

2.4.1. Local Analysis for Computing Attributes. The $g(cp)$ function is used for computing some texture attributes (X vector) on the zone of the DICOM image in which the calcifications could be found. The attributes computed are (1) the number of zeros Z_{cp} and (2) the number of pixels contained in h_{cp}^* . From the second attribute, it is possible to compute three statistical values, like mean, standard deviation, and variance yielding 5 attributes in total. Such attributes are discriminant enough to be used in the classification process of normal and malign tissue.

2.5. Classification of Mammography. The five attributes are normalized by the centered-reduced data technique and used as the input to a KNN classifier based on Euclidean distance. The KNN classifier ranks third as the most used classifier in the last 20 years in the mammography analysis [23], and in our case, due to the model proposed here, this classifier is the most appropriate. The Euclidean distance is sufficiently discriminant because the attributes are not linear, allowing a greater separation among classes. To unbalance the single data-classes, K must be chosen as an odd number, in our case $K=3$. The classifier was trained and tested using leave-one-out cross validation (LOOCV). This validation technique ensures a very low error; additionally, it is typically used for small databases, providing a maximum adjustment of the training set and independent test set [23, 24].

3. Results and Discussion

As was described above, the mammography dataset consists of 74 images: 22 diagnosed with calcifications and 52 diagnosed as normal tissue, in accordance with the classification system for breast mammogram, BI-RADS. The mammograms are images stored as DICOM 3.0 format with a size of 4784×3517 pixels, as shown in Figure 5(a). Figure 5(b) shows the cp attribute of the mammogram, used in this work for modelling sharp changes in the intensity.

The results obtained from the global interpolation of the h_{cp} histogram are illustrated on Figure 6(a). Note that, this function highlights mass, conducts, calcifications, and healthy tissue at the same time. As we have mentioned above, a second interpolation technique is required for the high frequency zone of the cp histogram, yielding the $g(cp)$ function. Figure 6(b) illustrates the calcification detected by the $g(cp)$ function. However, this second analysis still needs to validate which of these spots found belong to abnormal tissue growth or mass, that is, if these spots are calcifications or not. To accomplish this, five attributes are computed and analysed from the $g(cp)$ function, and they are used as input to the KNN classifier.

Figure 7(a) illustrates the classification results obtained by an expert, and Figure 7(b) illustrates the results obtained using our method. Our method detects exactly the two calcifications indicated by the expert in Figure 7(a). The cp feature detects big changes of intensity on the mammography; therefore, it is possible to detect more calcification on Figure 7(b) than those detected by the human eye. Figures 7(c) and 7(d) show the clusters of the microcalcifications found.

A second result is illustrated in Figure 8. Here, the calcifications visually found by an expert are also entirely detected by the proposed method.

Figures 7 and 8 illustrate qualitatively the correct performance of the proposed approach with respect to mammograms previously diagnosed with calcifications. Following subsection will show the quantitative results and a comparative table to point out the relevance of our proposed strategy.

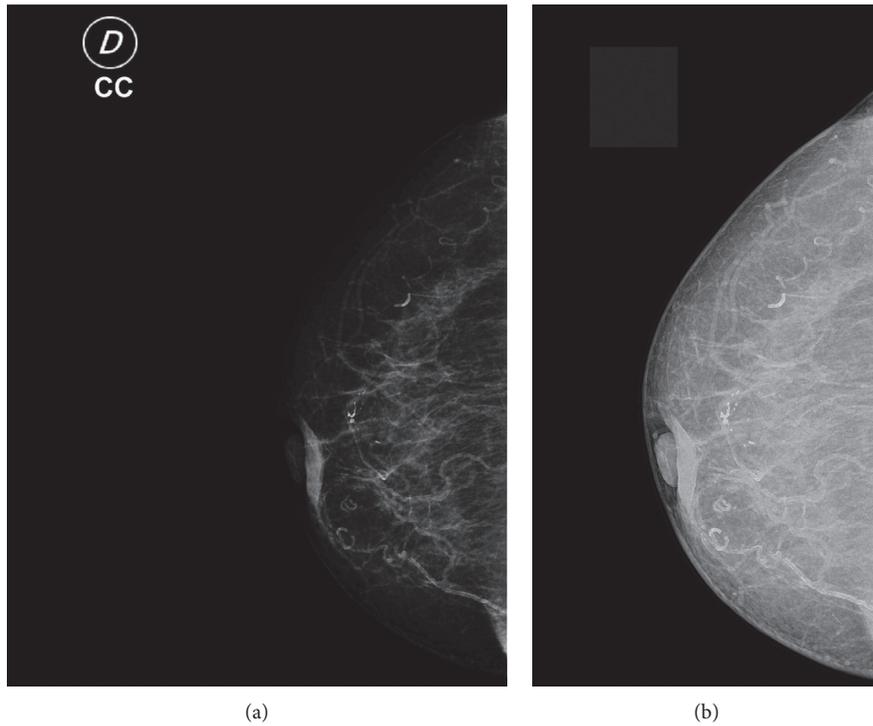


FIGURE 5: (a) Mammogram in the DICOM format, (b) *cp* feature of the mammogram (a).

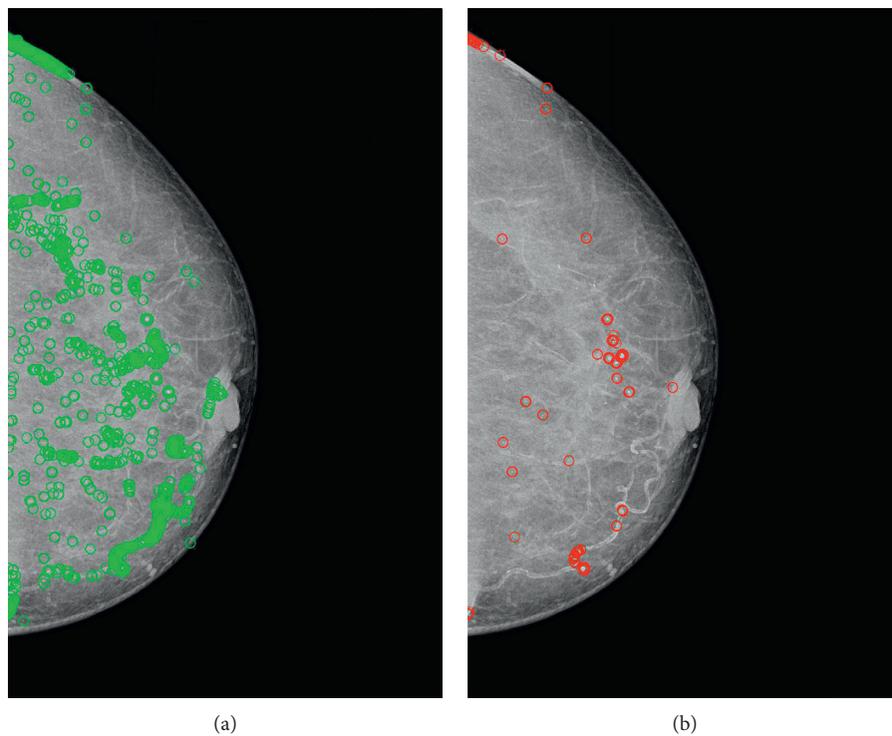


FIGURE 6: Calcification detection results: (a) calcifications and normal tissue detected with the global interpolation; (b) the local interpolation results only highlight the calcifications.

3.1. Metrics of Performance Analysis. The quantitative results obtained by the experimental tests are summarized in the confusion matrix shown in Table 1. The worst case of a

classifier is the false-negative (FN) score, that refers to the cases in which the classifier does not detect the calcification, and it exists. From Table 1, two false negatives (FNs) are

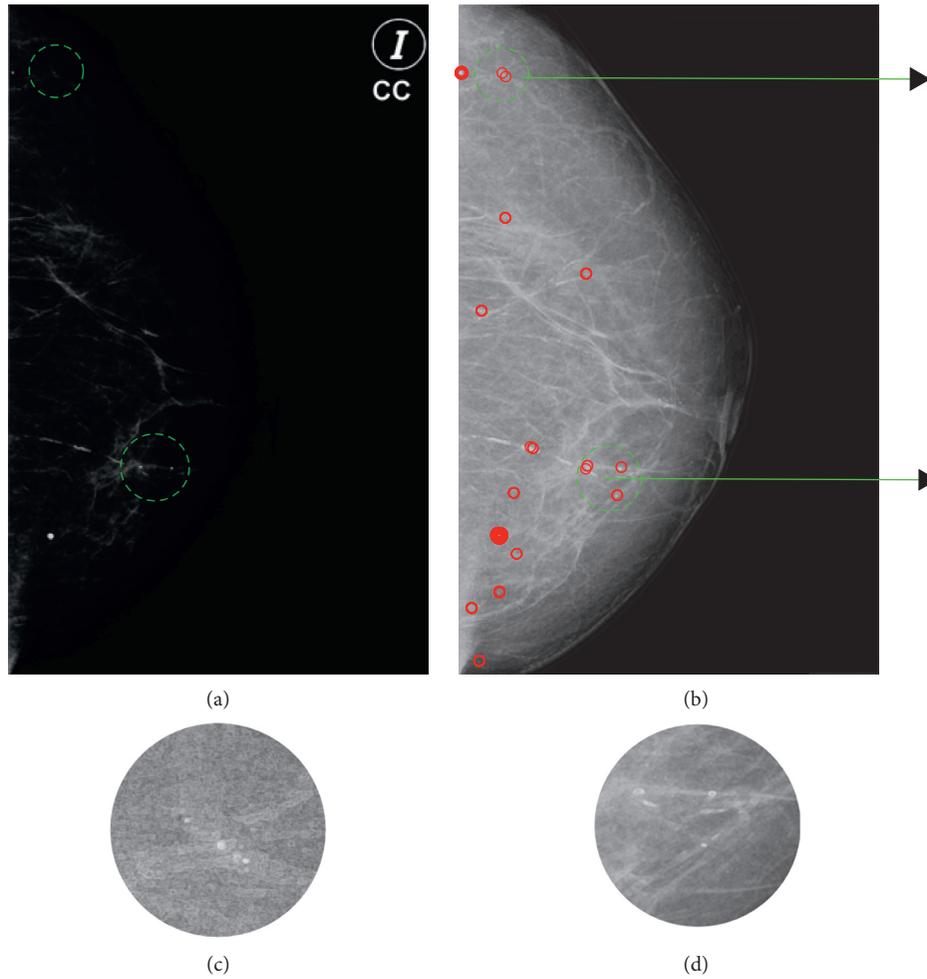


FIGURE 7: Classification results: (a) mammography in the DICOM format, the circles highlight the zones with calcifications found by an expert; (b) the cp attribute with the calcifications found by the proposed method; (c) and (d) zoom into the microcalcifications detected.

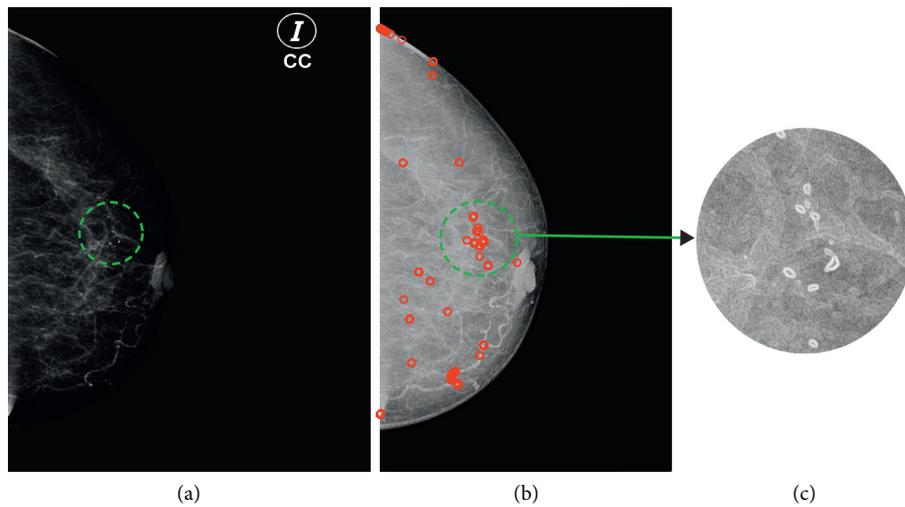


FIGURE 8: Classification results: (a) mammography in the DICOM format, circles highlight the zones with calcifications found by an expert; (b) cp attribute with the calcifications found by the proposed method; (c) zoom into the microcalcifications detected.

TABLE 1: Confusion matrix of the $g(cp)$ function for $K=3$ and using the Euclidean distance.

Desired/estimated	Normal tissue	Calcification	Specificity
Normal tissue	48 (TN)	4 (FP)	0.9230
Calcification	2 (FN)	20 (TP)	0.9090

TABLE 2: Metrics performance results of the proposed method for different orders of the polynomial function.

Order	Accuracy	Precision	Specificity	Sensitivity	% FA
9	0.8648	0.8333	0.9423	0.6818	0.0576
10	0.9054	0.8571	0.9423	0.8181	0.0576
11	0.8513	0.8333	0.9411	0.6818	0.0588
12	0.8378	0.9166	0.9807	0.5000	0.0192
13	0.9189	0.8333	0.9230	0.9090	0.0769
14	0.9189	0.8333	0.9230	0.9090	0.0769
15	0.9054	0.8226	0.9230	0.8636	0.0769

TABLE 3: Comparison between our methodology and state-of-the-art methods.

Compared methods	Dataset	Set-up	Accuracy	Sensitivity	FPI
Proposed strategy	Hosp. Irapuato	A	0.9189	0.9090	5.4054
PSOWNN [15]	Clinical	A	0.9367	0.9416	1.9006
DEOWNN [25]	MIAS	A	0.9353	0.9690	5.9754
Hough transform [18]	BCDR	A	0.9326	0.9178	3.9999
CALMA-ANN [19]	CALMA	A	0.9200	0.9200	4.9627
CAD-PPJ [20]	U of M.	A	0.8914	0.8500	1.7100
Texture feature + SLDA [26]	DDSM	S	0.8700	0.9333	10.000
Level set [27]	U. of M.	S	0.8500 ± 0.0200	—	—
Active contour [27]	U. of M.	S	0.8700 ± 0.0200	—	—

obtained from the experimental tests. On the other hand, four false positive (FP) have been obtained from the experimental tests, and this result indicates the presence of calcifications that does not exist, in accordance with the diagnosis of the medical expert.

As reported in the confusion matrix, the specificity obtained from the experimental results is 0.9230 for normal tissue detection and 0.9090 of sensitivity for calcifications detection. Such rate of classification is usually compared with other strategies proposed in the literature. The worst case during the classification stage is due to the presence of false-negative values. In this case, two of the FN cases were analysed, and in both cases, the mammography shows low brightness.

Now, the choice of the polynomial function order is essential to warranty the best accuracy results. Thus, Table 2 shows the performance evaluation of our approach measured with different metrics including accuracy, precision, specificity, sensitivity, and percentage of false alarms (FA), for different orders of the polynomial function. The best accuracy value is 0.9189, for the 13 and 14 order of the function; however, the 13 order was chosen for simplicity. For this order, the obtained precision value is 0.8333 which is not as high as we would expect due to the false positives. This means that the most common error of our strategy occurs when a calcification is detected, but it does not exist. Such results are also validated by the specificity and the sensitivity values: 0.9230 and 0.9090, respectively. The

percent of positive and negative incorrect detections (false alarms, FA) is 0.0769, and a low percentage of false negative is the most expected result.

A comparative analysis among different methods found in the state-of-the-art is shown in Table 3. The first column presents the techniques used by different authors including our proposed strategy, and the second column shows the dataset used for the experimental results. The third column indicates how the image processing is performed, which can be manual, semiautomated, and automated. The last column shows the achieved accuracy for each method.

Furthermore, in these methods, the algorithms processing is automated allowing the use of such strategies for real applications. However, we want to point out that in the case of methods that compute texture features, such as [20] and our approach, the best performance is shown by our strategy. Additionally, among these three methods, our approach is the only method that process images as automated.

Table 3 shows a comparative analysis between different related works in the state of the art. The first column lists the related methods, and all these methods use texture or appearance features for detecting calcifications and different classification methods. The dataset used for experimental tests is shown in the second column, and the third column indicates the execution mode of the method, that is automatic (A) and semiautomatic (S). Automatic refers that no intervention is necessary from the user; on the other hand,

the semiautomatic requires a minimum intervention from the user. Note that, only the methods proposed in [26, 27] are semiautomatic.

The fourth column shows the accuracy value that allows us to compare the performance between each of the methods. In particular, our proposed approach and the works proposed in [15, 18, 19, 25] show an accuracy higher than 0.9, indicating a high-performance evaluation. Although our proposal does not present the best accuracy, it is positioned in the middle of all the high-performed approaches, being a good and competitive solution in general. Note that, the approach with the best accuracy, presented in [15], only differs by 0.0178 with respect to ours, and this difference can be solved using a dataset with a greater number of samples. The fifth column shows the sensitivity, which evaluates the performance only when a calcification is detected. For most of the compared works, this measure is superior to 0.9, in particular the work proposed in [25] shows the highest sensitivity, and our proposed method is only 0.06 below it. Finally, the sixth column shows the false positives per image (FPI), which allows to measure the errors in calcifications detection, that is, when the normal tissue is detected as a calcification. For this measure, all listed references are less than 6, indicating a good FPI performance.

4. Conclusions

A method for detecting calcifications based on cluster prominence cp feature analysis on mammograms is proposed in this work. A deep analysis about the cluster prominence feature throws that is highly discriminative and allows the modelling of the calcifications in comparison with other attributes. The classification error obtained is low. Among this misclassification, the most common error of our strategy is the false detection, that is the false positives, which finally is less critical than the false-negatives results. A performance comparison demonstrates that our proposed strategy has better performance than similar works. This methodology is proposed as a tool for helping the radiologist during diagnosis. A web application is under construction for providing a more flexible support to the final user. Future works include further experimental tests of this approach using INbreast dataset and a new characterization of the cp feature for improving time performance results.

Data Availability

Mammograms database and their corresponding medical diagnosis were provided by the Hospital Regional of Irapuato based on a privacy agreement with the Digital Signal Processing Laboratory of the University of Guanajuato. The agreement avoids publicly sharing or distributing any kind of data or mammography provided by the hospital.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

The authors would like to thank the Hospital Regional de Irapuato for the mammograms dataset provided to perform the experimental test of this approach. This work was supported by project CIIC 003/2018 titled “Desarrollo de Modulo de Comunicacion basados en Linea Electrica (PLC) y Luz Visible (VLC) para la interaccion de Sistemas Autonomos” granted by the University of Guanajuato and the Research and Graduate Department (DAIP).

References

- [1] World Health Organization (WHO), “Breast cancer: prevention and control,” May 2018, <http://www.who.int/topics/cancer/breastcancer/es/index1.html>.
- [2] World Cancer Research Fund International, “Global cancer observatory,” October 2018, <http://gco.iarc.fr/today/>.
- [3] P. Katrakazas, V. Costarides, M. Tarousi et al., “Business process modelling for a Greek hospital’s medical equipment data center,” in *Proceedings of IEEE 31st International Symposium on Computer-Based Medical Systems (CBMS)*, pp. 328–332, Karlstad, Sweden, June 2018.
- [4] K. Suzuki, “Overview of deep learning in medical imaging,” *Radiological Physics and Technology*, vol. 10, no. 3, pp. 257–273, 2017.
- [5] Y. Yoshino, T. Miyajima, H. Lu et al., “Automatic classification of lung nodules on MDCT images with the temporal subtraction technique,” *International Journal of Computer Assisted Radiology and Surgery*, vol. 12, no. 10, pp. 1789–1798, 2017.
- [6] J. Wang, R. M. Nishikawa, and Y. Yang, “Quantitative comparison of clustered microcalcifications in for-presentation and for-processing mammograms in full-field digital mammography,” *Medical Physics*, vol. 44, no. 7, pp. 3726–3738, 2017.
- [7] F. J. Lopez-Aligue, A. Poveda-Pierola, I. Acevedo-Sotoca et al., “Detection of micro-calcifications in digital mammograms,” in *Proceedings of 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 3906–3909, Lyon, France, August 2007.
- [8] A. Jalalian, S. B. T. Mashohor, H. R. Mahmud, M. I. B. Saripan, A. R. B. Ramli, and B. Karasfi, “Computer-aided detection/diagnosis of breast cancer in mammography and ultrasound: a review,” *Clinical Imaging*, vol. 37, no. 3, pp. 420–426, 2013.
- [9] Y.-C. Cheung, H.-P. Tsai, Y.-F. Lo, S.-H. Ueng, P.-C. Huang, and S.-C. Chen, “Clinical utility of dual-energy contrast-enhanced spectral mammography for breast microcalcifications without associated mass: a preliminary analysis,” *European Radiology*, vol. 26, no. 4, pp. 1082–1089, 2016.
- [10] J. Arevalo, F. A. Gonzalez, R. Ramos-Pollán et al., “Representation learning for mammography mass lesion classification with convolutional neural networks,” *Computer Methods and Programs in Biomedicine*, vol. 127, pp. 248–257, 2016.
- [11] V. Hari, R. Jagathy, and R. Gopikakumari, “Enhancement of calcifications in mammograms using Volterra series based quadratic filter,” in *Proceedings of International Conference on Data Science & Engineering (ICDSE)*, pp. 85–89, Kochi, Kerala, July 2012.
- [12] L. Wei, Y. Yang, R. Nishikawa et al., “Improved cancer detection in automated breast ultrasound by radiologists using computer aided detection,” *European Journal of Radiology*, vol. 89, pp. 54–59, 2017.

- [13] H. D. Cheng, X. J. Shi, R. Min, L. M. Hu, X. P. Cai, and H. N. Du, "Approaches for automated detection and classification of masses in mammograms," *Pattern Recognition*, vol. 39, pp. 646–668, 2006.
- [14] J. K. Kim and H. Park, "Statistical textural features for detection of micro-calcifications in digitized mammograms," *IEEE Transactions on Medical Imaging*, vol. 18, no. 3, pp. 231–238, 1999.
- [15] J. Dheeba, N. Albert Singh, and S. Tamil Selvi, "Computer-aided detection of breast cancer on mammograms: a swarm intelligence optimized wavelet neural network approach," *Journal of Biomedical Informatics*, vol. 49, pp. 45–52, 2014.
- [16] A. Yadollahpour and S. Hamed, "Early breast cancer detection using mammogram images: a review of image processing techniques," *Biosciences Biotechnology Research Asia*, vol. 12, no. 1, pp. 225–234, 2015.
- [17] A. Jalalian, S. Mashohor, R. Mahmud, B. Karasfi, M. Iqbal Saripan, and A. R. Ramli, "Computer-assisted diagnosis system for breast cancer in computed tomography laser mammography (CTLM)," *Journal of Digital Imaging*, vol. 30, no. 6, pp. 796–811, 2017.
- [18] A. Fanizzi, T. M. A. Basile, L. Losurdo et al., "Hough transform for clustered microcalcifications detection in full-field digital mammograms," in *Proceedings of Applications of Digital Image Processing XL*, San Diego, CA, USA, November 2017.
- [19] A. Lauria, R. Palmiero, G. Forni et al., "The CALMA system: an artificial neural network method for detecting masses and microcalcifications in digitized mammograms," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 518, no. 1, pp. 391–393, 2004.
- [20] R. K. Samala, H.-P. Chan, L. M. Hadjiiski, and M. A. Helvie, "Analysis of computer-aided detection techniques and signal characteristics for clustered microcalcifications on digital mammography and digital breast tomosynthesis," *Physics in Medicine and Biology*, vol. 61, no. 19, pp. 7092–7112, 2016.
- [21] M. A. Ibarra-Manzano, M. Devy, and J. L. Boizard, "Real-time classification based on color and texture attributes on an FPGA-based architecture," in *Proceedings of Conference on Design and Architectures for Signal and Image Processing (DASIP)*, pp. 250–257, Edinburgh, Scotland, October 2010.
- [22] W. J. Curran and T. Liu, "Ultrasound glcm texture analysis of radiation-induced parotid-gland injury in head-and-neck cancer radiotherapy: an in vivo study of late toxicity," *Medical physics*, vol. 39, no. 9, pp. 5732–5739, 2012.
- [23] R. W. D. Pedro, A. Machado-Lima, and F. L. S. Nunes, "Is mass classification in mammograms a solved problem? - a critical review over the last 20 years," *Expert Systems with Applications*, vol. 119, pp. 90–103, 2019.
- [24] A. K. Jain, P. W. Duin, and J. J. Mao, "Statistical pattern recognition: a review," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 1, pp. 4–37, 2000.
- [25] J. Dheeba and S. Tamil Selvi, "An improved decision support system for detection of lesions in mammograms using differential evolution optimized wavelet neural network," *Journal of Medical Systems*, vol. 36, no. 5, pp. 3223–3232, 2011.
- [26] J. E. Ball and L. M. Bruce, "Digital mammographic computer aided diagnosis (cad) using adaptive level set segmentation," in *Proceedings of 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 4973–4978, Lyon, France, August 2007.
- [27] J. Shi, B. Sahiner, H.-P. Chan et al., "Characterization of mammographic masses based on level set segmentation with new image features and patient information," *Medical Physics*, vol. 35, no. 1, pp. 280–290, 2007.

Research Article

Fully Automated Lumen Segmentation Method for Intracoronary Optical Coherence Tomography

Elżbieta Pociask ¹, Krzysztof Piotr Malinowski ², Magdalena Ślęzak ³,
Joanna Jaworek-Korjakowska ⁴, Wojciech Wojakowski ⁵, and Tomasz Roleder ⁵

¹Department of Biocybernetics and Biomedical Engineering, AGH University of Science and Technology, Al. Mickiewicza 30, 30-059 Krakow, Poland

²Jagiellonian University Medical College, Faculty of Health Science, Institute of Public Health, Krakow, Poland

³KCRI, Krakow, Poland

⁴Department of Automatic Control and Robotics, AGH University of Science and Technology, Al. Mickiewicza 30, 30-059 Krakow, Poland

⁵Third Department of Cardiology, Medical University of Silesia, Katowice, Poland

Correspondence should be addressed to Elżbieta Pociask; elzbieta.pociask@gmail.com

Received 26 June 2018; Revised 5 October 2018; Accepted 10 December 2018; Published 26 December 2018

Guest Editor: Cesare Valenti

Copyright © 2018 Elżbieta Pociask et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Background. Optical coherence tomography (OCT) is an innovative imaging technique that generates high-resolution intracoronary images. In the last few years, the need for more precise analysis regarding coronary artery disease to achieve optimal treatment has made intravascular imaging an area of primary importance in interventional cardiology. One of the main challenges in OCT image analysis is the accurate detection of lumen which is significant for the further prognosis. **Method.** In this research, we present a new approach to the segmentation of lumen in OCT images. The proposed work is focused on designing an efficient automatic algorithm containing the following steps: preprocessing (artifacts removal: speckle noise, circular rings, and guide wire), conversion between polar and Cartesian coordinates, and segmentation algorithm. **Results.** The implemented method was tested on 667 OCT frames. The lumen border was extracted with a high correlation compared to the ground truth: 0.97 ICC (0.97–0.98). **Conclusions.** Proposed algorithm allows for fully automated lumen segmentation on optical coherence tomography images. This tool may be applied to automated quantitative lumen analysis.

1. Introduction

Today's medical practice diagnosis of coronary artery disease (CAD) is made using mostly invasive imaging modalities among which coronary angiography is the most popular one, being currently considered the standard during cardiac catheterization and hemodynamic assessment. However, coronary angiography produces "luminogram" delineating only the shape of the contrast-filled lumen without any information about plaque morphology or vessel wall [1]. This is why most recently angiography is accompanied by newer intravascular imaging techniques like IVUS and OCT which employ acoustic waves and near-infrared light, respectively, in

order to generate cross-sectional, volumetric images of coronary arteries [2]. OCT provides images of high contrast and very high spatial resolution (10–20 μm), 10 times higher than IVUS, thus allowing characterization of atherosclerotic plaques and assessment of coronary stenting including stent apposition and struts coverage [3, 4]. One of the main challenges in OCT image analysis is the accurate detection of lumen which is significant for the further prognosis.

This paper is organized in 4 sections as follows: Section 1.1 presents the motivation of this work and the review of the state of the art in the area of lumen segmentation. Section 2 specifies the overview of the implemented algorithm. The conducted statistical analysis, results, and discussion of the

achieved outcomes are presented in Section 3. At last, Section 4 closes the paper and highlights future directions.

1.1. Motivation. OCT images clearly depict the boundaries between lumen and vessel, which facilitate image interpretation. Currently image processing has been mainly conducted manually by Core Lab analysts, but due to large number of cross sections in OCT image sequence, this is usually a time-consuming process with high inter-observer variability [5]. However, above limitations can be addressed by introducing automatic image analysis including detection of lumen contours [6, 7]. Lumen segmentation is the first but crucial step in the image analysis process as it allows detection of stenosis and high-risk plaques. It has been addressed not only for OCT pullbacks but also for IVUS image sequences [8, 9].

1.2. Related Works. Automatic lumen contour detection can be a very challenging step as OCT images typically contain various artifacts like guide wire shadowing, motion artifacts, bifurcations, or nondiluted intraluminal blood. As the analysis of OCT images is a demanding process, many automatic methods have been developed for lumen detection in OCT [2, 10–16] in recent years.

These methods usually employ multistep image processing techniques including binarization approach [10, 11, 16], morphological operations [10, 11], intensity curve methods [16], Markov random field (MRF) model [2], or wavelet transform [12].

Different OCT technologies, various image textures, diffused and complex lesions and, furthermore, not well diluted blood from vessels have a huge impact on segmentation outcome as well as feature extraction in above described methods [2, 8–10]. Additionally, images with poor luminal or substantial luminal blood in contact with the arterial wall cannot be well delineated by active contour methods [14].

Methods developed and proposed by other research teams tend to be very accurate and have good computational cost. But, they applied them for high-quality images including only one kind of individual artifacts. Due to these obstacles, there is still room for a complex solution which could improve the segmentation process for most cases.

Diffused and complex lesions have motivated the development of the proposed method for OCT analyses with a new sequence of morphological operations, and interpolation methods which have been designed to reconstruct lumen object, resulting in a more accurate segmentation outcome, even in the presence of bifurcation structures and not well-diluted intraluminal blood. Most of the listed above methods can only be applied on the healthy or nonbifurcation images [11] or for good quality images without artifacts [10, 16].

Manual segmentation by independent observers is mainly used as the reference for particular method validation. To increase the value of our work, we have compared our outcomes with two widely used, commercially available systems (Medis medical imaging systems and St. Jude Medical OPTIS integrated system). Moreover, the test was conducted on the

same dataset for the results to be reliable. This algorithm achieved higher classification results compared to existing OCT segmentation programs, scoring 0.97 ICC in lumen area compared with a gold standard ground-truth method.

In this paper, we propose a fully automated method to segment the lumen area in run OCT pullbacks without excluding any frames. Our solution can be used to analyze poor quality images as well as images with diseased vessels and bifurcations.

2. Materials and Methods

The proposed automated lumen detection algorithm on intracoronary optical coherence tomography images consists of two main stages: preprocessing (image enhancement and artifacts removal) and lumen segmentation with contour correction. Image analysis has been implemented in Matlab software using the Image Processing Toolbox, where the flowchart of the proposed algorithm is presented in Figure 1.

2.1. Database Specification. The analyzed material is composed of 667 frames of different patients, from the Medical University of Silesia. Images used in this study were acquired by the FD-OCT system (C7-XR system OCT Intravascular Imaging System, Westford, MA) and two kinds of imaging catheters: the C7 Dragonfly and Dragonfly OPTIS catheter with automatic pullback, drive motor optical controller. The analyzed data were obtained with the pullback speed of 20 mm/s and 18 mm/s, respectively.

The chosen images contain a variety of vessel features like lumen irregularities caused by intraluminal masses (thrombus), branches, or different intensity profiles due to not well-diluted blood (Figure 2).

2.2. OCT Image Preprocessing. OCT images are inhomogeneous, complex (variation in degree of intensity and shape), and furthermore they contain extraneous artifacts, such as bright concentric circular rings and bright structure from guide wire with a characteristic shadow behind it. These types of artifacts appear in almost every frame which makes the advanced image analysis steps impossible. Therefore, the preprocessing stage is necessary to obtain the binary image of the intimal layer (the most inner layer of three layers building the vessel wall) without artifacts from the diagnostic catheter and improve the quality of the image for further analysis (Figure 3).

The proposed algorithm receives as an input each frame in turn from the whole OCT image dataset. Multiframe images have been saved in DICOM format (pullback run), and each frame is a 2D RGB image in a Cartesian coordinate system.

Firstly, all calibration markers and text remarks are removed from the image using a mask of the pixels that are colored. The analyzed RGB image is converted into grayscale with the NTSC 1953 standard, which converts RGB values to grayscale values by forming a weighted sum of the R, G, and B components.

After converting the RGB image to the grayscale image, the polar transform is applied, and further preprocessing

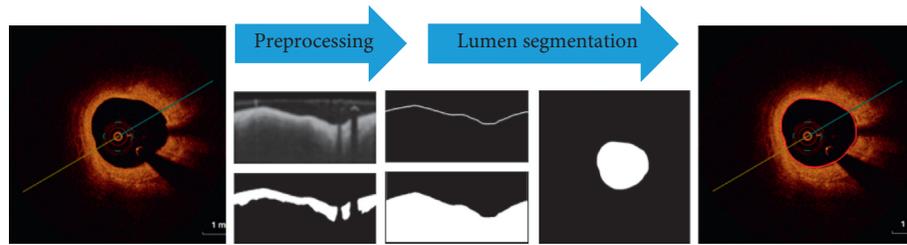


FIGURE 1: Proposed methodology for automatic lumen contour detection on OCT images. The flowchart shows the major steps of the detection process including preprocessing and segmentation with their outcomes.

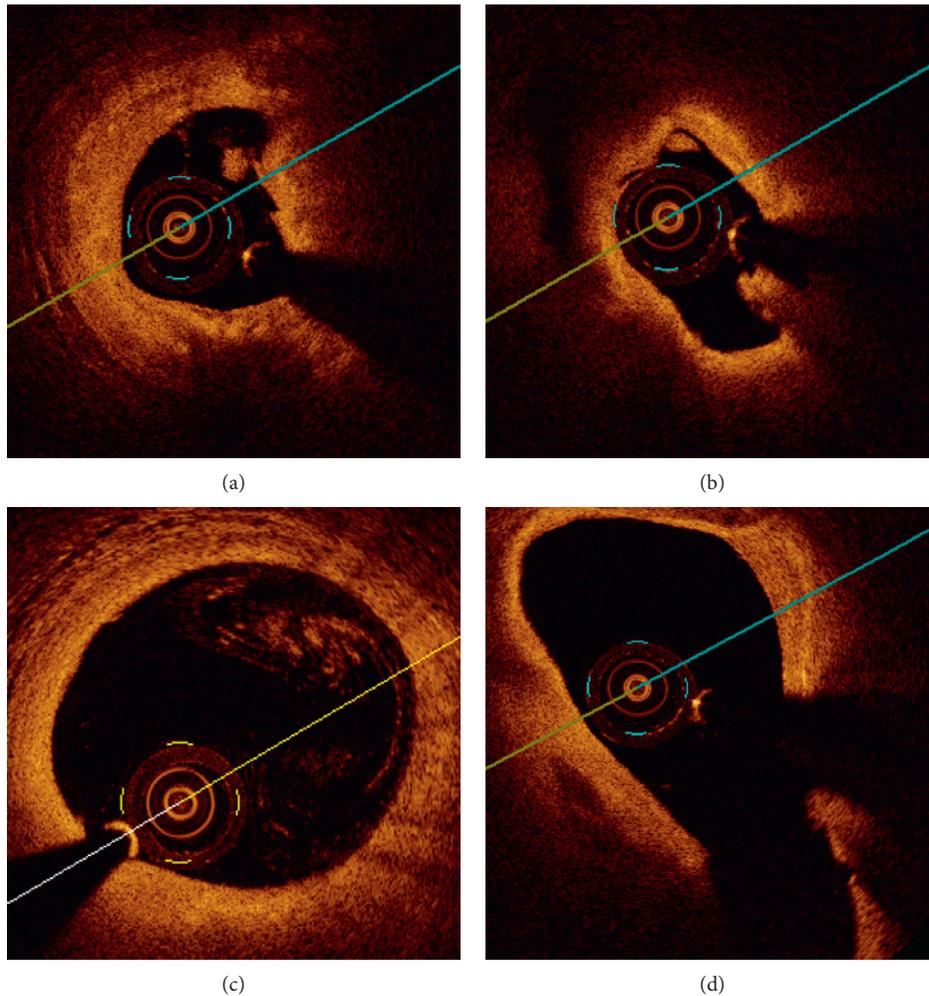


FIGURE 2: Examples of OCT cross-sectional view: (a), (b) lumen irregularities with visible thrombus, (c) residual blood, and (d) bifurcations.

stages are being performed in polar coordinates. This transformation allows us to convert the circular shape of the coronary artery visible in a cross-sectional view to a straightened structure. In mathematics, the polar coordinate system is a two-dimensional coordinate system in which each point on a plane is determined by a distance from a reference point and an angle from a reference direction.

The ring shape distortion from the imaging catheter in polar space is shown as the straightened structure on the left side of the image with the known size—Dragonfly catheter with a diameter of 2.7 French which gives 0.91 mm. Using the knowledge about spatial resolution of the image and

catheter diameter, we can calculate the region of structure and remove it from the image. Another significant artifact, which may limit the segmentation process, is a speckle noise from not well-diluted blood. Speckle noise may affect the lumen segmentation outcome by classifying it mistakenly as a tissue resulting in underestimated real lumen area.

In order to remove any destructive speckle effects without damaging borders, we use a median filter with a 5×5 window [17]. After median filtering, the Gaussian smoothing operator is used to “blur” the image, aiming at removing unnecessary details and reducing noise from background. The Gaussian smoothing operator is a 2-D convolution operator that uses a

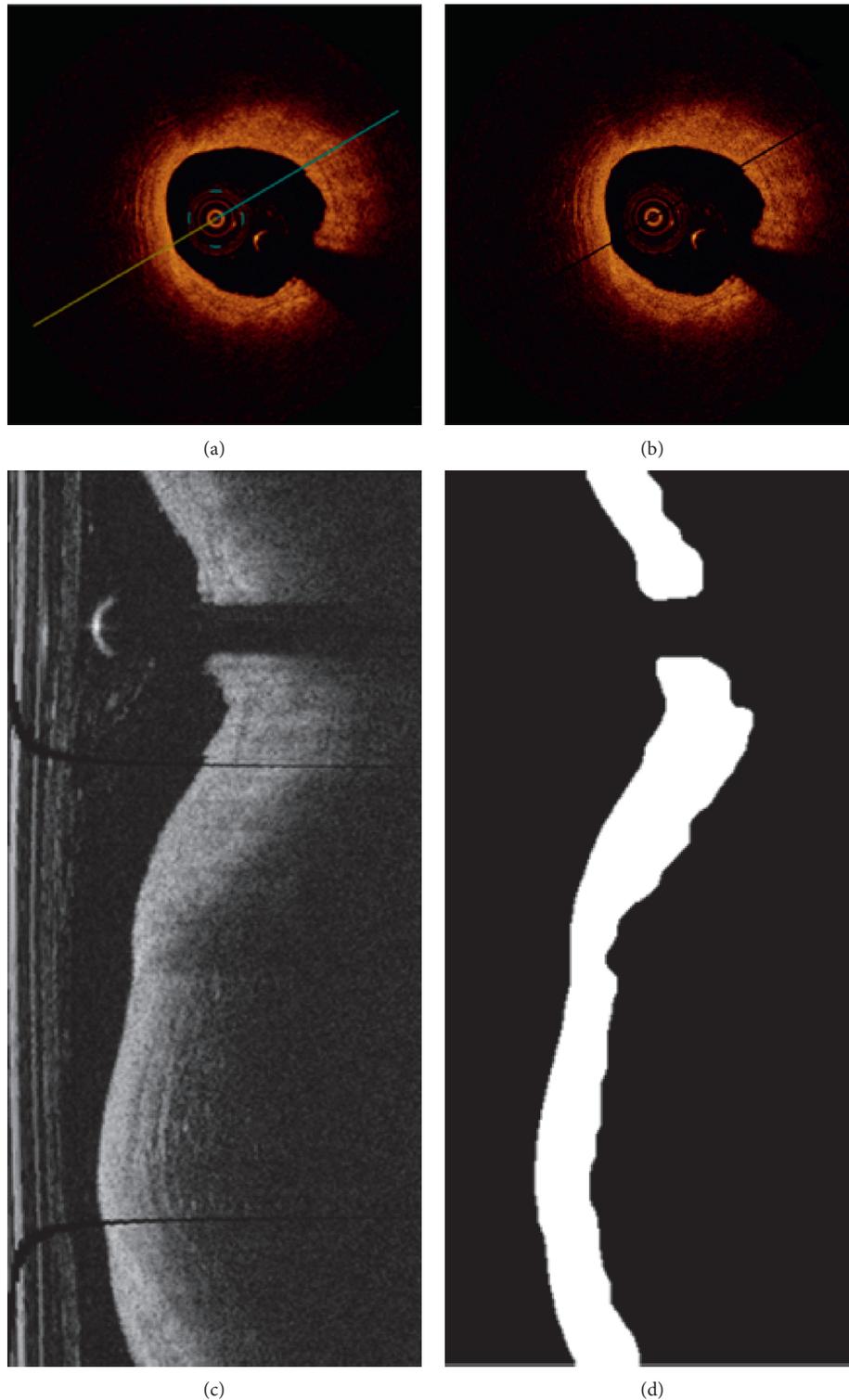


FIGURE 3: Output of the preprocessing step: (a) OCT input image, (b) artifact removal, (c) OCT image after polar transformation, and (d) primary segmentation of the lesion in polar coordinates.

kernel of Gaussian values representing the shape of a Gaussian (bell-shaped) hump.

The Gaussian filter is a low-pass filter, attenuating high-frequency signals. It calculates a weighted average of each pixel's neighborhood, with the average weighted more towards

the value of the central pixels, and a Gaussian distribution provides gentler smoothing and preserves the better edges [18].

An automatic thresholding on polar space is used to generate a new binary image with clearly separated region with high-gradient magnitude-intimal layer.

2.3. Lumen Segmentation and Contour Drawing. Methods and algorithms developed for segmentation of medical structures are specific to application, imaging modality, and type of body part to be studied. Because of image complexity, there is no perfect method to segment all of the medical structures with high efficiency. However, the success of the lumen segmentation step is crucial for the further analysis of OCT images and correct diagnosis.

The outcome of the preprocessing stage is a binary image with the primary segmented area that still contains small artifacts like insufficiently diluted blood close to the imaging catheter. To minimize the effect of artifacts on the final result, we subject the image to morphological opening and closing operations. While erosion and dilation have the major disadvantage of changing the size of our region of interest, opening and closing retain the interesting area. Opening and closing are basic methods of morphological noise removal. Opening removes small, unwanted objects from the foreground placing them in the background, while closing fills small holes and connects disjoint objects in the foreground, changing small areas of background into foreground [19]. Based on the lumen shape, we use a disk-shaped structuring element to preserve the circular nature of the object. The disk-shape element is a flat, structuring element, where R specifies the radius (Figure 4).

The radius was determined experimentally and set to 5. Figure 5 shows examples of the results of morphological opening and closing operations.

Artifacts from the imaging catheter and the guide wire were removed in the preprocessing stage. However, the shadow from guide wire makes the intimal layer discontinuities what can be observed in Figure 3(c). A similar effect is caused by bifurcations. The gap from guide wire shadow is usually of the same size, regular, and easily to be found and filled. More problematic are gaps caused by bifurcations which can vary in size, and additionally, the remained objects of the segmented lumen may have irregular shapes. Bifurcation results in lumen area distortion can be observed in Figure 6(d). The interpolation of remaining regions is necessary to draw the final lumen contour which should be as close as possible to the expected values. To solve this problem, we have applied a modified version of linear interpolation which is tailored to our needs.

In order to connect the parts of lumen, we analyze the boundary information (location and coordinates) of every disconnected part of intima layer (traces of the exterior boundaries of the object). We receive a cell array of boundary pixel coordinates of all the objects in the binary image [20]. To perform the linear interpolation, extreme points are calculated as presented in Figure 7.

Few of the extrema points are candidates to the contour points including bottom-left, top-left, and top-right. We analyze the objects from the top to the bottom. Following points are being interpolated: for the first, upper object, we select the left-bottom point, and for the second object which is located below, we choose between the top-left and top-right points. The final choice is determined by the value of Euclidean distance between extremes. The bigger the bifurcation is, the longer the

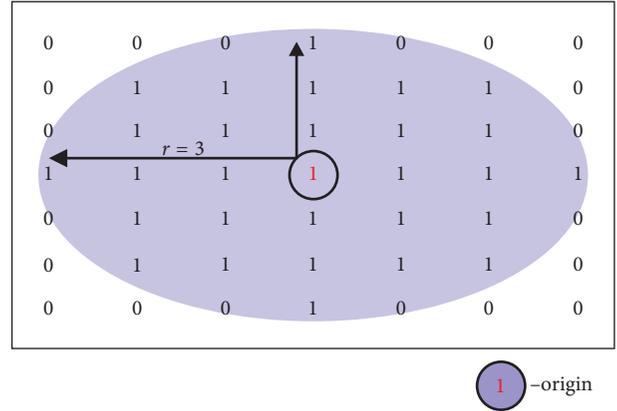


FIGURE 4: Illustration of a flat structuring element [20].

distance will be. The individual extreme coordinates are taken to calculate the distance and perform linear interpolation. We experimentally checked that the cutoff point for bifurcation is 2 mm size. Figure 6 presents some examples of chosen contour points. To avoid sharp contour reconstructions, additional points have been chosen by moving up and down from extreme points and finding the first white pixel in the current row.

After setting the contour points, a linear interpolation is used. In that way, all discontinuities (bifurcations, shadow from guide wire and from artifacts) are filled. Throughout this method, the lumen border line in the polar image was obtained (Figure 6(c)). Finally, the lumen border points are detected by the Sobel edge detection algorithm [21]. The Sobel operator performs a 2-D spatial gradient measurement and emphasizes regions of high spatial frequency that corresponds to edges. After all operations have been carried out, the resulting polar image is transformed into an image in Cartesian coordinates. Figure 8 shows each step of lumen segmentation.

As the segmentation outcome, the resulting contour does not have the smoothness that the vessel is expected to have. The Savitzky–Golay sliding polynomial filter with window width 35 and polynomial order 2 [22] is being applied.

Savitzky and Golay showed that a set of integers ($A_{-n}, A_{-(n-1)}, \dots, A_{(n-1)}, A_n$) could be derived and used as weighting coefficients to carry out the smoothing operation [23]. The use of these weighting coefficients, known as convolution integers, is exactly equivalent to fit the data to a polynomial. Therefore, the smoothed data point $(y_k)_s$ by the Savitzky–Golay algorithm is given by the following equation:

$$(y_k)_s = \frac{\sum_{i=-n}^n A_i y_{k+1}}{\sum_{i=-n}^n A_i}, \quad (1)$$

where A_i are weighting coefficients to perform the smoothing operation.

3. Results

The validation of the described fully automated lumen segmentation method has been performed on 667 intravascular optical coherence tomography frames from different patients.



FIGURE 5: Examples of the results after morphological opening and closing operations on OCT images in polar coordinates: (a) images after Gaussian filtering, (b) images after binarization, and (c) images after morphological opening and closing operations.

The data were provided by the Medical University of Silesia. Figure 9 presents the achieved results.

3.1. Statistical Analysis. Statistical analysis involves data obtained from four methods: our algorithm, two commercially available systems, and manual analysis (ground truth mask). Continuous parameters were reported as mean and median with the first and the third quartiles (Q1: 25%; Q3: 75%).

The Bland–Altman analysis was used to assess the agreement between two measurement methods. It is a method comparison technique proposed by Altman and Bland [24] based on the quantification of the agreement between two quantitative measurements by studying the mean difference and constructing limits of agreement.

The results for the particular measurements were presented as mean with 95% confidence interval and as median with the

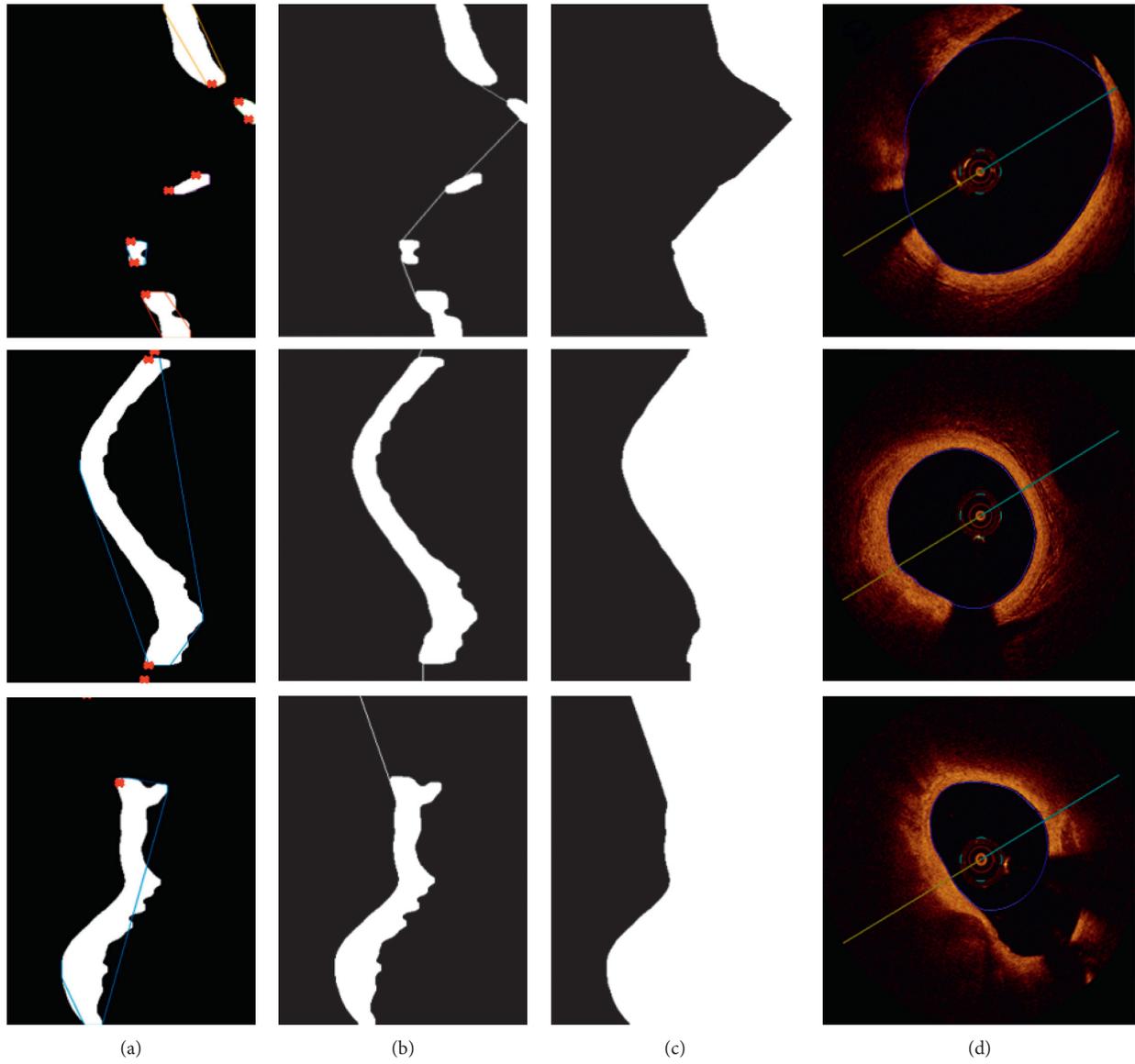


FIGURE 6: Examples of chosen extrema points needed to perform linear interpolation: (a) binary image after preprocessing and artifact removal with marked extrema, (b) extrema point connection (linear interpolation), (c) lumen segmentation outcome, and (d) input OCT image with lumen contour tracing. Images (a)–(c) are in polar coordinates, and image (d) is after transformation to Cartesian coordinates.

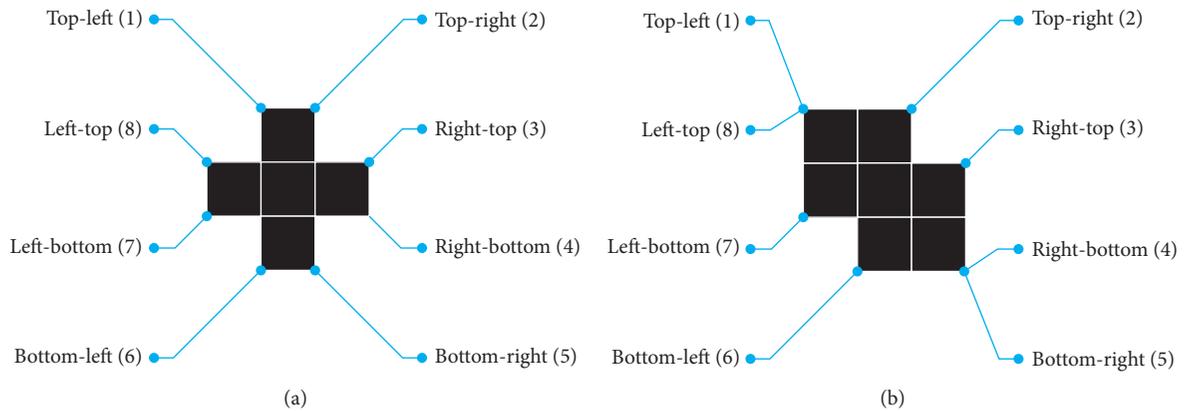


FIGURE 7: Illustration of marked extremes for two different regions [20].

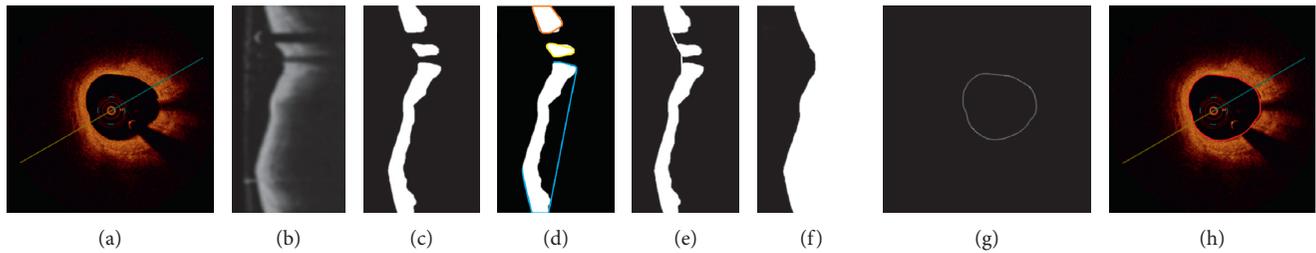


FIGURE 8: Lumen segmentation steps: (a) original image in Cartesian coordinates; (b) image after transformation to polar coordinates and after catheter removal applying Gaussian filtering; (c) image in polar coordinates after binarization and morphological operations, small artifacts are removed and small gaps filled; (d) Image in polar coordinates with marked extrema; (e) based on extrema, connection points are chosen and linear interpolation is applied to fill all gaps; (f) lumen segmentation outcomes; (g) segmented contour transformed back to Cartesian coordinates and after smoothing filter; (h) final image, cross-sectional view with marked contour.

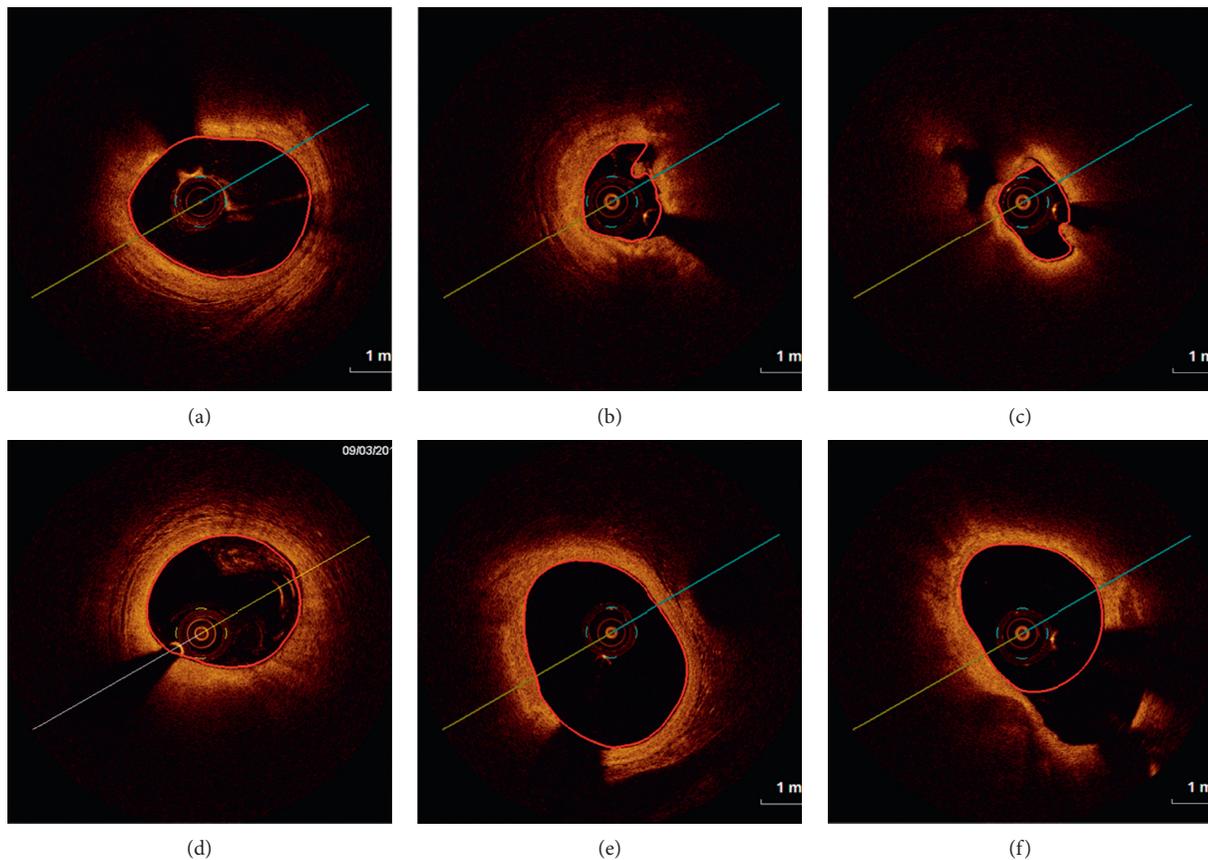


FIGURE 9: Results of the described lumen segmentation algorithm dedicated for OCT images. Presented images show six different cases including various artifacts and difficulties.

first and the third quartiles. Discrepancies between the first and the second analysis were calculated as absolute and relative differences and presented as means with 95% CIs. Intra-class correlations were calculated as the main measure of agreement along with the graphical representation as the Bland–Altman plots. Analyses for statistical computing were performed in R language (R Core Team 2017, Vienna, Austria).

3.2. Validation of Automated Lumen Segmentation. In order to validate the described algorithm, we compare four lumen detection methods: our solution, ground truth mask, and two commercially available systems including St. Jude

Medical and System (system 1) and Medis medical imaging systems (system 2). Manual segmentation has been performed by independent observers-interventional cardiologists with extensive clinical experience. Furthermore, our experts were involved in the development of methodology and results analysis.

The following parameters have been analyzed for each of the described methods: lumen area, mean lumen diameter, minimal lumen diameter, and maximal lumen diameter, and are collected in Table 1.

The results of the assessed parameters are collected in Tables 2–4 and presented by the Bland–Altman plots.

TABLE 1: Calculated parameters for each of the analyzed methods.

	Our method mean	System 1 mean	System 2 mean	Ground truth mean
Lumen area (mm ²)	5.99	6.04	5.76	5.89
Mean lumen diameter (mm)	2.72	2.72	2.66	2.68
Minimal lumen diameter (mm)	2.52	2.52	2.42	2.49
Maximal lumen diameter (mm)	2.91	2.93	2.87	2.88

TABLE 2: Statistical comparison of the parameters between our methodology and manual analyses by analyst (ground truth).

	Our method (first measure) mean (CI)	Ground truth (second measure) mean (CI)	Our method (first measure) median (IQR)	Ground truth (second measure) median (IQR)	Difference	Relative difference	ICC (95% CI)	ICC <i>p</i> value
Lumen area	5.99 (5.83–6.14)	5.89 (5.74–6.04)	6.45 (4.72–7.67)	6.29 (4.67–7.50)	0.10 (0.06–0.13)	-1.12% (-1.55% to -0.68%)	0.97 (0.97–0.98)	<0.0001
Mean lumen diameter	2.72 (2.68–2.76)	2.68 (2.64–2.72)	2.88 (2.41–3.13)	2.82 (2.43–3.09)	0.03 (0.02–0.04)	-1.15% (-1.48% to -0.83%)	0.96 (0.95–0.97)	<0.0001
Minimal lumen diameter	2.52 (2.48–2.56)	2.49 (2.45–2.53)	2.64 (2.18–2.96)	2.61 (2.10–2.93)	0.03 (0.02–0.03)	-1.11% (-1.44% to -0.78%)	0.98 (0.98–0.98)	<0.0001
Maximal lumen diameter	2.91 (2.87–2.96)	2.88 (2.84–2.92)	3.07 (2.54–3.30)	3.01 (2.58–3.25)	0.04 (0.02–0.05)	-0.81% (-1.23% to -0.39%)	0.91 (0.89–0.92)	<0.0001

TABLE 3: Statistical comparison of parameters between our methodology and commercially available system 1.

	Our method (first measure) mean (CI)	System 1 (second measure) mean (CI)	Our method (first measure) median (IQR)	System 1 (second measure) median (IQR)	Difference	Relative difference	ICC (95% CI)	ICC <i>p</i> value
Lumen area	5.99 (5.83–6.14)	6.04 (5.90–6.19)	6.45 (4.72–7.67)	6.50 (5.01–7.56)	-0.06 (-0.08–0.03)	1.67% (1.25%–2.10%)	0.99 (0.98–0.99)	<0.0001
Mean lumen diameter	2.72 (2.68–2.76)	2.72 (2.68–2.76)	2.88 (2.41–3.13)	2.87 (2.52–3.10)	-0.00 (-0.01–0.00)	0.34% (0.06%–0.63%)	0.98 (0.97–0.98)	<0.0001
Minimal lumen diameter	2.52 (2.48–2.56)	2.52 (2.48–2.56)	2.64 (2.18–2.96)	2.52 (2.48–2.56)	0.00 (-0.01–0.01)	0.16% (-0.16%–0.49%)	0.98 (0.97–0.98)	<0.0001
Maximal lumen diameter	2.91 (2.87–2.96)	2.93 (2.89–2.97)	3.07 (2.54–3.30)	3.08 (2.66–3.28)	-0.02 (-0.03–0.00)	0.82% (0.43%–1.21%)	0.94 (0.93–0.95)	<0.0001

TABLE 4: Statistical comparison of parameters between our methodology and another commercially available system 2.

	Our method (first measure) mean (CI)	System 2 (second measure) mean (CI)	Our method (first measure) median (IQR)	System 2 (second measure) median (IQR)	Difference	Relative difference	ICC (95% CI)	ICC <i>p</i> value
Lumen area	5.99 (5.83–6.14)	5.76 (5.61–5.92)	6.45 (4.72–7.67)	6.13 (3.98–7.48)	0.22 (0.18–0.27)	-4.23% (-4.88% to -3.57%)	0.95 (0.93–0.97)	<0.0001
Mean lumen diameter	2.72 (2.68–2.76)	2.66 (2.61–2.70)	2.88 (2.41–3.13)	2.79 (2.25–3.09)	0.06 (0.05–0.07)	-2.42% (-2.85% to -1.99%)	0.94 (0.92–0.96)	<0.0001
Minimal lumen diameter	2.52 (2.48–2.56)	2.42 (2.37–2.46)	2.64 (2.18–2.96)	2.56 (2.01–2.91)	0.10 (0.09–0.12)	-4.96% (-5.53% to -4.39%)	0.95 (0.85–0.97)	<0.0001
Maximal lumen diameter	2.91 (2.87–2.96)	2.87 (2.83–2.91)	3.07 (2.54–3.30)	3.00 (2.50–3.26)	0.04 (0.02–0.06)	-1.11% (-1.58% to -0.63%)	0.89 (0.88–0.91)	<0.0001

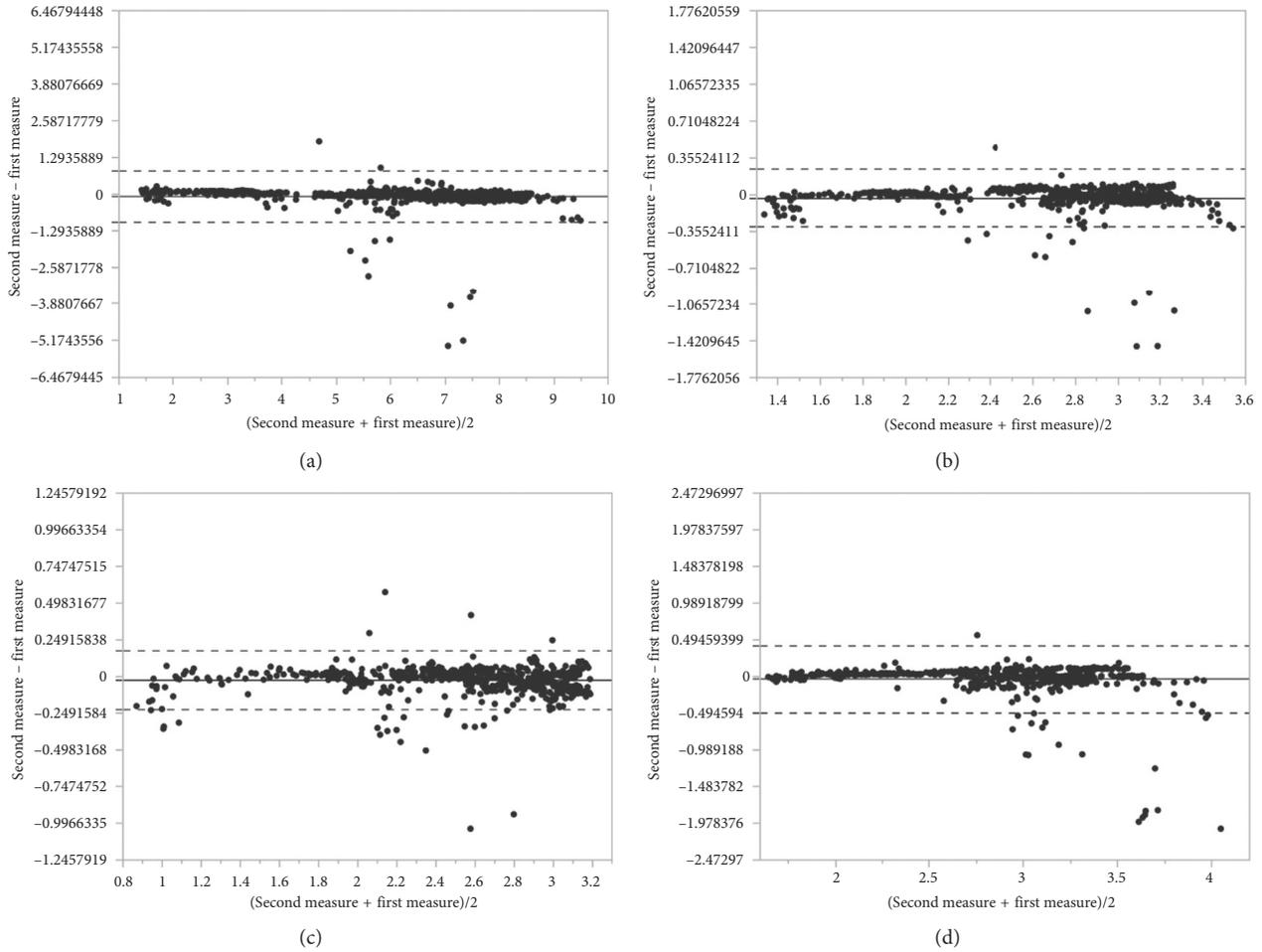


FIGURE 10: Bland–Altman plot for (a) lumen area, (b) mean lumen diameter, (c) minimal lumen diameter, and (d) maximal lumen diameter between our method and ground truth method.

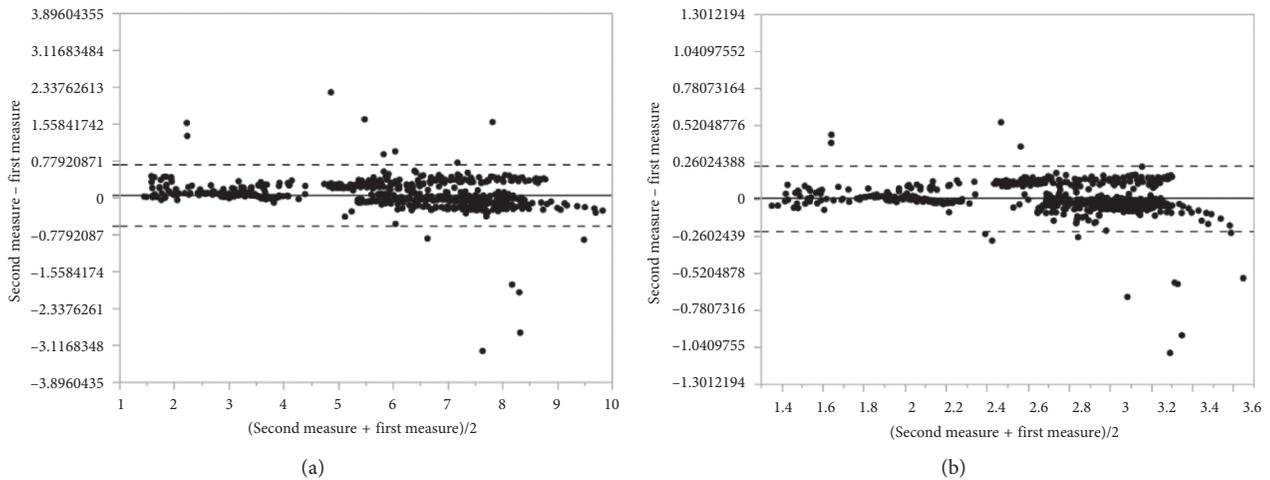


FIGURE 11: Continued.

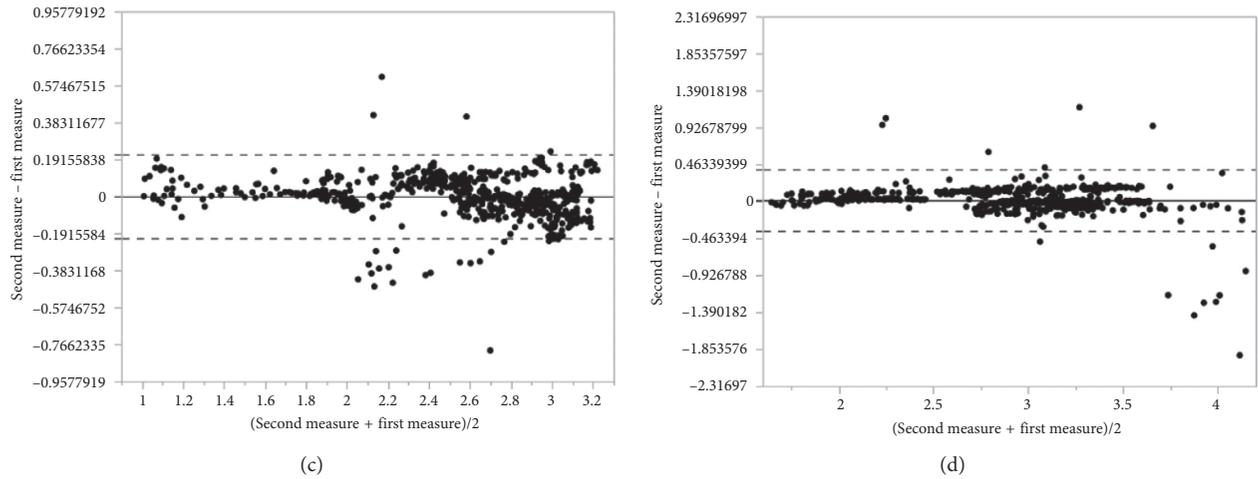


FIGURE 11: Bland–Altman plot for (a) lumen area, (b) mean lumen diameter, (c) minimal lumen diameter, and (d) maximal lumen diameter between our method and automated lumen detection proposed by commercially available system 1.

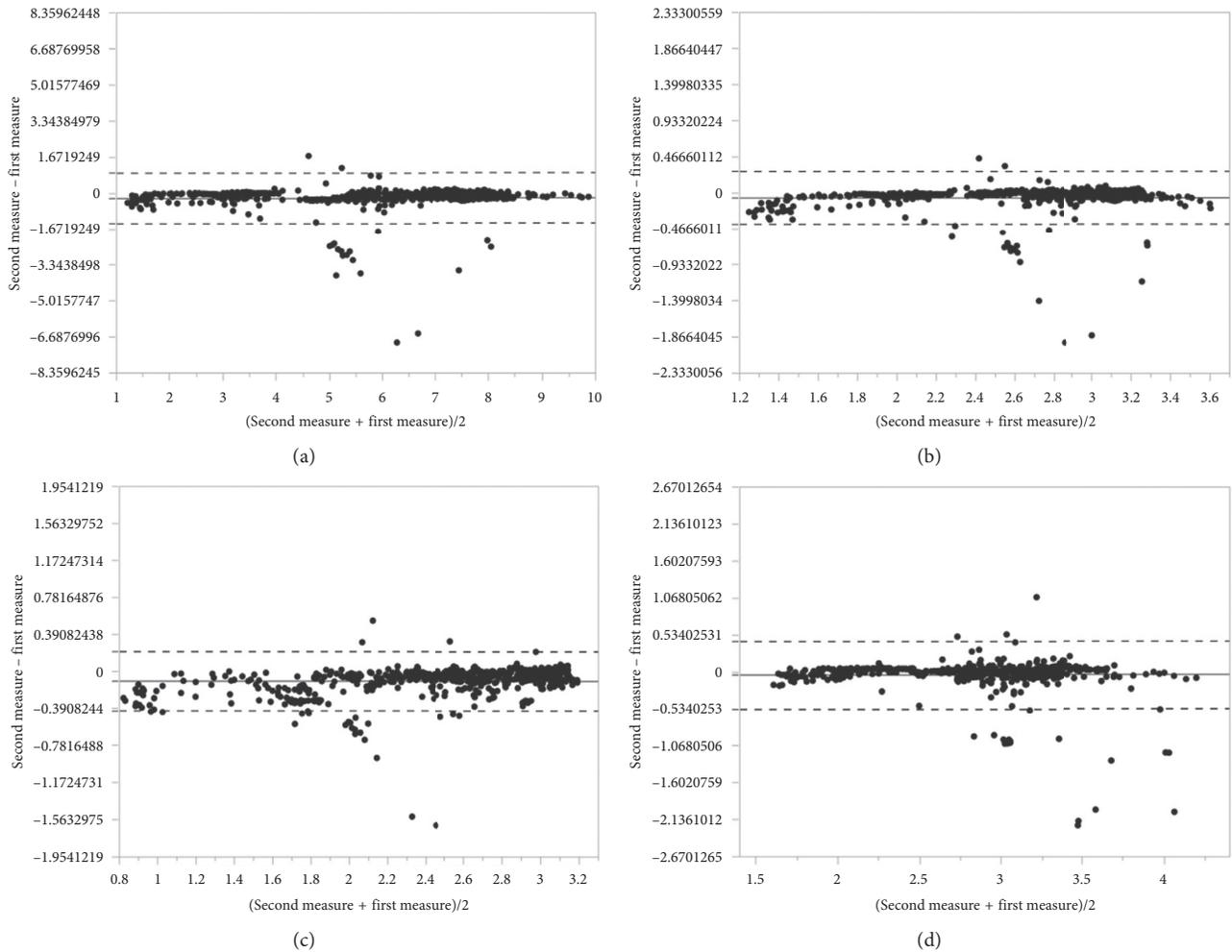


FIGURE 12: Bland–Altman plot for (a) lumen area, (b) mean lumen diameter, (c) minimal lumen diameter, and (d) maximal lumen diameter between our method and automated lumen detection proposed by commercially available system 2.

To enable the analysis of statistical data, the following parameters have been collected additionally.

The relative difference is calculated using the following equation:

$$RD = \frac{\sum_{i=1}^N (I_i - O_i / \max(O_i, I_i))}{N} * 100\%. \quad (2)$$

The absolute relative difference is calculated using

$$ARD = \frac{\sum_{i=1}^N (|I_i - O_i| / \max(O_i, I_i))}{N} * 100\%, \quad (3)$$

where N is the total number of frames, i is the number of current frame, O_i is the value for 1st measurement, and I_i is the value for 2nd measurement.

3.3. Discussion of the Results. The lumen detection was performed on a desktop computer with an Intel Core i5-4200, 1.60 GHz processor, 8 GB RAM, Windows 10 64 bits, and Matlab (R2016b). The average time of the lumen contour detection was 1.099 s. The average time of manual segmentation of a slice was approximately 60 seconds. As it is shown, the computer-aided segmentation systems is much faster than the manual segmentation; furthermore, it is objective to the same cases and also very accurate. To validate our proposed method, we tested the same dataset with results from manual analyses and two commercially available tools for automatic lumen detection. We achieved high correlation in lumen area compared with a gold standard ground-truth method (manual analyses performed by an expert): 0.97 ICC. The results in the literature [2, 11, 12] reported an absolute difference of the mean lumen area of 0.1 mm^2 . De Macedo et al. [11] obtained absolute difference of mean lumen area of 0.17 mm^2 . Our proposed method showed similar results (absolute difference of mean lumen area of 0.1 mm^2) to those presented previously published methods, but what is worth to highlight in our validation process is that all frames were included to analyse even frames containing complex plaque, artifacts from residual blood, or bifurcations with diameters $> 2 \text{ mm}$. Furthermore, as we can see the parameters calculated by our methods are similar to obtained results from both commercially available systems (system 1 and system 2), the systems have not been described in any paper. Our algorithm can be easily implemented again and tested on a new dataset.

The Bland–Altman plots (Figures 10–12) indicate a good agreement between used methods, where the solid line denotes the mean difference between the first and the second measurement, while the dashed lines indicate ± 1.96 standard deviation. Most points plotted are between the solid line (mean diff) and the dashed line (mean $\pm 2 * \text{standard deviation}$).

An absolute difference of mean lumen area calculated between our method and automated lumen detection proposed by system 1 is of 0.06 mm^2 compared with system 2 results of absolute difference of mean lumen area of 0.22. Although the lumen areas are similar in all methods and there is high ICC between our method and the others (0.95–0.99), the lumen diameters are shown with lower ICC, especially between our method and system 2.

In terms of limitations, our method was not tested on images with the presence of stents which could have a negative impact on our algorithm. This limitation may be solved in the future by developing methods to extract the struts and fill the artifacts from strut shadows.

4. Conclusion

We presented a fully automated methodology which is able to detect and draw correctly lumen contours in OCT images including frames with bifurcations and artifacts from blood. The automated method was validated using the manual analyzes performed by an Expert as a gold standard as well as commercially available tools. The results suggest that our method can be a useful tool for vessel segmentation and further analysis. The achieved results indicate that the proposed algorithm fulfills the requirement.

Data Availability

No data were used to support this study.

Conflicts of Interest

The authors declare that there are no conflicts of interests regarding the publication of this paper.

Acknowledgments

Scientific work partly supported by the AGH research grants, 15.11.120.884 and 15.11.120.641, and the Medical University of Silesia. The data were provided by the Medical University of Silesia.

References

- [1] K. H. Chan and M. K. C. Ng, “Is there a role for coronary angiography in the early detection of the vulnerable plaque?,” *International Journal of Cardiology*, vol. 164, no. 3, pp. 262–266, 2013.
- [2] S. Tsantis, G. C. Kagadis, K. Katsanos, D. Karnabatidis, G. Bourantas, and G. C. Nikiforidis, “Automatic vessel lumen segmentation and stent strut detection in intravascular optical coherence tomography,” *Medical Physics*, vol. 39, no. 1, pp. 503–513, 2012.
- [3] F. Prati, E. Regar, G. S. Mintz et al., “Expert review document on methodology, terminology, and clinical applications of optical coherence tomography: physical principles, methodology of image acquisition, and clinical application for assessment of coronary arteries and atherosclerosis,” *European Heart Journal*, vol. 31, no. 4, pp. 401–415, 2009.
- [4] H. G. Bezerra, M. A. Costa, G. Guagliumi, A. M. Rollins, and D. I. Simon, “Intracoronary optical coherence tomography: a comprehensive review,” *JACC: Cardiovascular Interventions*, vol. 2, no. 11, pp. 1035–1046, 2009.
- [5] A. Wang, J. Eggermont, N. Dekker et al., “Automatic stent strut detection in intravascular optical coherence tomographic pullback runs,” *International Journal of Cardiovascular Imaging*, vol. 29, no. 1, pp. 29–38, 2012.
- [6] E. Kurgan and P. Gas, “Estimation of temperature distribution inside tissues in external RF hyperthermia,” *Przegląd Elektrotechniczny*, vol. 86, no. 1, pp. 100–102, 2010.

- [7] R. Ogiela, R. Tadeusiewicz, and R. Ogiela, "Cognitive computing in intelligent medical pattern recognition systems, intelligent control and automation," in *Intelligent Control and Automation*, pp. 851–856, vol. 344 Springer, Berlin, Germany, 2006.
- [8] M. C. Moraes and S. S. Furuie, "Automatic coronary wall segmentation in intravascular ultrasound images using binary morphological reconstruction," *Ultrasound in Medicine & Biology*, vol. 37, no. 9, pp. 1486–1499, 2011.
- [9] L.-D. Gao, S.-X. Hu, H. Zhang et al., "Correlation analysis of EV71 detection and case severity in hand, foot, and mouth disease in the human province of China," *PloS One*, vol. 9, no. 11, Article ID e100003, 2014.
- [10] G.-A. Cheimariotis, Y. S. Chatzizisis, V. G. Koutkias et al., "ARC OCT: automatic detection of lumen border in intravascular OCT images," *Computer Methods and Programs in Biomedicine*, vol. 151, pp. 21–32, 2017.
- [11] M. M. T. Macedo, C. K. Takimura, P. A. Lemos, and M. A. Gutierrez, "A robust fully automatic lumen segmentation method for in vivo intracoronary optical coherence tomography," *Research on Biomedical Engineering*, vol. 32, no. 1, pp. 35–43, 2016.
- [12] M. Moraes, D. A. Cardenas, and S. Furuie, "Automatic lumen segmentation in IVOCT images using binary morphological reconstruction," *Biomedical Engineering Online*, vol. 12, no. 1, p. 78, 2013.
- [13] K. Sihan, C. Botha, F. Post et al., "A novel approach to quantitative analysis of intravascular optical coherence tomography imaging," in *Proceedings of 35th Annual Computers in Cardiology Conference*, Bologna, Italy, September 2008.
- [14] K. Sihan, C. Botha, F. Post et al., "Fully automatic three-dimensional quantitative analysis of intracoronary optical coherence tomography," *Catheterization and Cardiovascular Interventions*, vol. 74, no. 7, pp. 1058–1065, 2009.
- [15] G. J. Ughi, T. Adriaenssens, W. Desmet, and J. D'hooge, "Fully automatic three-dimensional visualization of intravascular optical coherence tomography images: methods and feasibility in vivo," *Biomedical optics express*, vol. 3, no. 12, pp. 3291–3303, 2012.
- [16] H. M. Kim, S. H. Lee, C. Lee, J. W. Ha, and Y. R. Yoon, "Automatic lumen contour detection in intravascular oct images using otsu binarization and intensity curve," in *36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Chicago, IL, USA, August 2014.
- [17] E. D. L. Arias-Castro and D. L. Donoho, "Does median filtering truly preserve edges better than linear filtering?," *Annals of Statistics*, vol. 37, no. 3, pp. 1172–1206, 2009.
- [18] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, Springer, Berlin, Germany, 3rd edition, 2015.
- [19] S. E. Umbaugh, *Computer Vision and Image Processing*, Prentice Hall, Upper Saddle River, NJ, USA, 1998.
- [20] R. C. Gonzalez and R. E. Woods, *Digital Image Processing Using Matlab*, Prentice Hall, Upper Saddle River, NJ, USA, 2004.
- [21] I. Sobel, "An isotropic 3×3 image gradient operator," in *Machine Vision for Three-Dimensional Scenes*, Academic Press, Cambridge, MA, USA, 1990.
- [22] W. H. Press and S. A. Teukolsky, "Savitzky-Golay smoothing filters," *Computer in Physics*, vol. 4, no. 6, p. 669, 1990.
- [23] S. J. Orfanidis, *Introduction to Signal Processing*, Academic Press, Cambridge, MA, USA, 1996.
- [24] D. Giavarina, "Understanding Bland Altman analysis," *Biochemia Medica*, vol. 25, no. 2, pp. 141–151, 2015.

Research Article

An Application for Skin Macules Characterization Based on a 3-Stage Image-Processing Algorithm for Patients with Diabetes

Cynthia Lourdes Toledo Peral ¹, Francisco José Ramos Becerril,²
Gabriel Vega Martínez,³ Arturo Vera Hernández ⁴, Lorenzo Leija Salas,⁴
and Josefina Gutiérrez Martínez ¹

¹División de Investigación en Ingeniería Médica, Instituto Nacional de Rehabilitación “Luis Guillermo Ibarra Ibarra”, Calz. México–Xochimilco No. 289, Col. Arenal de Guadalupe, Tlalpan, C.P. 14389 Ciudad de México, Mexico

²Servicio de Rehabilitación Cardíaca, Instituto Nacional de Rehabilitación “Luis Guillermo Ibarra Ibarra”, Calz. México–Xochimilco No. 289, Col. Arenal de Guadalupe, Tlalpan, C.P. 14389 Ciudad de México, Mexico

³Subdirección de Medicina del Deporte, Instituto Nacional de Rehabilitación “Luis Guillermo Ibarra Ibarra”, Calz. México–Xochimilco No. 289, Col. Arenal de Guadalupe, Tlalpan, C.P. 14389 Ciudad de México, Mexico

⁴LAREMUS, Sección Bioelectrónica, Departamento de Ingeniería Eléctrica, Centro de Investigación y Estudios Avanzados del Instituto Politécnico Nacional, Av. Instituto Politécnico Nacional 2508, Col. San Pedro Zacatenco, Gustavo A. Madero, C.P. 07360 Ciudad de México, Mexico

Correspondence should be addressed to Josefina Gutiérrez Martínez; josefina_gutierrez@hotmail.com

Received 26 June 2018; Accepted 28 November 2018; Published 16 December 2018

Academic Editor: Vincenzo Conti

Copyright © 2018 Cynthia Lourdes Toledo Peral et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Diabetic skin manifestations, previous to ulcers and wounds, are not highly accounted as part of diagnosis even when they represent the first symptom of vascular damage and are present in up to 70% of patients with diabetes mellitus type II. Here, an application for skin macules characterization based on a three-stage segmentation and characterization algorithm used to classify vascular, petechiae, trophic changes, and trauma macules from digital photographs of the lower limbs is presented. First, in order to find the *skin region*, a logical multiplication is performed on two skin masks obtained from color space transformations; dynamic thresholds are stabilised to self-adjust to a variety of skin tones. Then, in order to locate the *lesion region*, illumination enhancement is performed using a chromatic model color space, followed by a principal component analysis gray-scale transformation. Finally, characteristics of each type of macule are considered and classified; *morphologic properties* (area, axes, perimeter, and solidity), *intensity properties*, and a set of *shade indices* (red, green, blue, and brown) are proposed as a measure to obviate skin color differences among subjects. The values calculated show differences between macules with a statistical significance, which agree with the physician’s diagnosis. Later, macule properties are fed to an artificial neural network classifier, which proved a 97.5% accuracy, to differentiate between them. Characterization is useful in order to track macule changes and development along time, provides meaningful information to provide early treatments, and offers support in the prevention of amputations due to diabetic feet. A graphical user interface was designed to show the properties of the macules; this application could be the background of a future *Diagnosis Assistance Tool* for educational (i.e., untrained physicians) and preventive assistance technology purposes.

1. Introduction

Diabetes is a rapidly growing chronic disease with a 20% prevalence and which is catalogued as a noncommunicable disease [1]. Diabetes mellitus type II is characterized by insulin resistance. Insulin is a hormone that helps deliver glucose to cells, e.g., to muscular cells where it is metabolized as energy

[2]. Insulin resistance is a sign of diabetes development. In this process, which is called hyperglycemia, glucose is not delivered to the cells and builds up in the body.

According to the 2014 diabetes report from the World Health Organization (WHO) [3], in the world there are 422 million people living with diabetes. In developing countries, the prevalence is increasing. In Mexico, the Health Department via

the 2016 National Health and Nutrition Survey (ENSANUT 2016) [4] reported that 9.4% of Mexican adults (i.e., 6.5 millions) have been diagnosed with diabetes. However, in 2017, the International Diabetes Federation (IDF) [5] reported that there are 12 millions of Mexican adults living with diabetes, but 37.5% are not aware that they have this disease.

Comorbidities such as obesity, hypertension, and dyslipidemia, among others, are precipitating factors to develop diabetes [6]. Even more, when these comorbidities are present along with diabetes, a rapid deterioration of body functions could arise and persist; diabetic retinopathy and diabetic foot [4] can cause blindness or lead to amputations which lead to disabilities.

Diabetes is associated, in the long-term, with degenerative processes that affect the cardiovascular and nervous system, as well as the eyes and skin [7]. From 30 to 70% of patients with diabetes develop skin problems [7, 8]. Neuropathy, microangiopathy, and macroangiopathy are the main predisposing factors for diabetic foot. Their evolution leads to blood flow reduction and ischemia, structural and functional damage, and an overloaded extremity due to the lack of sensitivity; all these put the foot at risk. Moreover, just adding up anything like a simple trauma or an infection could lead to ulcers, lesions, and even necrosis [9].

Although microangiopathy and macroangiopathy are major contributors to complications like skin lesions or diabetic foot, metabolic disruptions also have a significant direct effect, especially in alterations of the skin [7]. Some of these manifestations are called macules [10], which are defined as a flat, distinct, and discolored area of skin. Other manifestations may include lack of body hair, yellowish coloration, callous formation, onychomycosis, foot and toe deformation, and others [7, 8]. Even though macules occur commonly, they are not taken into account as a diagnose element [11, 12], nor are they registered as information that could lead to an early diabetic foot diagnosis [12, 13].

Relevantly, microangiopathy and macroangiopathy are also the cause of most skin manifestations found in patients with diabetes mellitus who have not been diagnosed with diabetic foot [14].

In the case of diabetes mellitus, skin manifestations have not been accounted as an important aspect of the disease [15]. There is a high prevalence of skin disorders among these patients as a matter of fact, and various authors report that these disorders are present in ~70% of their patients [14].

Kiziltan et al. [14] state that diabetic dermopathy is more common on patients with neuropathy or large vessel disease; also, they report it as frequently present in patients with signs and symptoms of polyneuropathy. Pavicic and Korting [16] report that peripheral arterial obstructive disease (PAOD) is up to 6 times more frequent in patients with diabetes and PAOD, neuropathy, and macroangiopathy are key pathophysiological factors in its development.

Several related studies report that 73% [15] to 80% [17] of the sampled patients present skin lesions or changes. Diabetic dermopathy always comes first as the most common skin manifestation in patients with diabetes. Pavicic and Korting [16] also state that the increasing duration of the

disease rise the possibility of skin involvement; 45% of patients suffering diabetes for more than 20 years developed a peripheral vascular disease, and 75–82.1% presented xerosis, which could cause skin tears [16].

Any change in skin pigmentation is called a *macule*. Macules can be erythematous (originated by blood vessels dilation or formation of new vessels), pigmentosae (which can be hyperpigmented, hypopigmented, or achromatic), or artificial, among others. Vascular macules occur as a secondary reaction, e.g., to medication, due to peripheral venous insufficiency or trauma [18].

A *vascular macule* is the one originated from a micro- or macrovascular problem, where the vessels underneath the skin are affected. These macules are rounded and are of reddish to brown color; normally, they present a diameter of 1 cm, but they can be smaller. *Petechiae* are very small (the size of a pinhead), reddish, rounded spots that appear on the shins and usually are a secondary effect of treatment with acetylsalicylic acid. *Macules due to trophic changes* are present when the patient has chronic venous insufficiency. They are darker patches of skin, have a larger area than other macules, and appear mainly in the ankles and shins. *Macules due to trauma* are the evidence (different than a scar or scab) of a traumatic event such as a hit in the shin. They are of brownish color, and the shape varies according to the trauma presented. This macule lingers in the skin of the patient with diabetes for a longer period of time than it would on a healthy patient.

All these skin manifestations are present previous to a diabetic foot diagnosis; patients can present them all at the same time, and they are generally overlaid. These macules appear in different parts of the leg and have large areas with undefined borders. Their localization and ulterior segmentation represent a challenging task, but the results can eventually be used as a tool for macule characterization, foot health prognosis, and even for amputation risk assessment.

Regarding the algorithms for image processing, these types of macules are not evaluated or processed until they become lesions or ulcers [19]. Computer aid diagnosis has been used in skin lesions for dermatology and dermoscopy (e.g., carcinomas and melanomas) [20–22] by means of support vector machines [21], support vector classification [20], or seeded region growing [22], but not in the prevention of diabetic foot development. Generally, tools for assessing skin problems due to diabetes mellitus type 2 are focused on advanced lesions and use questionnaires [18] that evaluate lesions like ingrown toenails, ulcers, calluses, or fissures, which take place after the diabetic foot diagnosis.

In this paper, we present the design of a graphical user interface (GUI) developed in Matlab® as an application for the characterization of skin macules. The GUI is based on a segmentation algorithm that applies image-processing techniques in order to find the region of interest (ROI) and characterize the macules present in images of the leg and foot of patients with diabetes mellitus type 2. We also present a statistical study of the calculated properties and a classifier of the 4 types of macules.

2. Materials and Methods

The first step was to acquire color digital photographs of skin macules “*skin images*” from the lower limb. For this purpose, a device called Wireless Image Acquisition System (WIAS) [23] was used. The device included a digital wireless camera (Sony DCS-QX100, 18MP), which provided an RGB image (Figure 1). Zoom and flash were never used in order to avoid changes in resolution or capturing bright areas, respectively. Changes in area, shape, and coloration of macules were document by the *skin images*.

The macules studied in this work were vascular macules, petechiae, macules due to trophic changes, and macules due to trauma. The study was performed at the Cardiac Rehabilitation Service of a National Institute in Mexico City. *Skin images* were processed using the Image Processing Toolbox of Matlab®. They were taken from 19 Mexican patients diagnosed with diabetes mellitus type II, but not yet with diabetic foot, who gave their signed informed consent.

Segmentation and characterization were performed through a proposed 3-stage image-processing algorithm, as described below:

Stage 1. (skin region). Skin identification in the color photography.

Stage 2. (Skin region-lesion region). Identification of possible lesions.

Stage 3. (lesion region). Characterization of macules.

2.1. Stage 1 (Skin Region). The aim of using the WIAS device was to be able to acquire repeatable digital photographs from areas of interest form the lower limb, and these images were called *skin images*. The color *skin image* contained elements that were not of interest, e.g., the robe, the bed clothing, and other background components. So, the first objective was to segment the legs of the patient from it.

A color image can be transformed to different color spaces [24] (i.e., domain transformation) in order to enhance the characteristics of interest, i.e., the differences between skin and nonskin and the similarity among different skin tones. If we see the *skin image* as a matrix, size is determined by the resolution of the camera. The image has 3 levels of depth; each level corresponds to one RGB color matrix, and each cell in these matrices corresponds to a pixel, whose value is the level of intensity in 8 bits.

In Stage 1, the first step was to transform the image from RGB to HSV color space. RGB describes an image for the amount of red, green, and blue in it. HSV color space does the same but in terms of Hue, saturation, and value. The algorithm [25] is described by equation (1). The RGB values should be normalized to the range [0 1]:

$$\text{HSV} = \left\{ \begin{array}{l} V = \max(R, G, B), \\ S = \frac{V - X}{V}, \quad \text{where } X = \min(R, G, B), \\ H = \left\{ \begin{array}{ll} \frac{5 + ((V - B)/(V - X))}{6} & \text{for } R = V \text{ and } G = X, \\ \frac{1 - ((V - G)/(V - X))}{6} & \text{for } R = V \text{ and } G \neq X, \\ \frac{1 + ((V - R)/(V - X))}{6} & \text{for } G = V \text{ and } B = X, \\ \frac{3 - ((V - B)/(V - X))}{6} & \text{for } G = V \text{ and } B \neq X, \\ \frac{3 + ((V - G)/(V - X))}{6} & \text{for } R \neq V, G \neq V, \text{ and } R = X, \\ \frac{5 - ((V - R)/(V - X))}{6} & \text{for } R \neq V, G \neq V, \text{ and } R \neq X, \end{array} \right. \end{array} \right. \quad (1)$$

where V represented the brightness value, S is the saturation, and H is the Hue matrix. The R , G , and B values had to be divided by 255 (e.g., $R = R/255$) in order to satisfy the normalization condition. The Hue matrix was selected (this

property allowed the differentiation between ROIs and background), a fixed threshold was set, and an intensity range was determined to find a tone—set of values (equation (2)). This became the first skin mask:

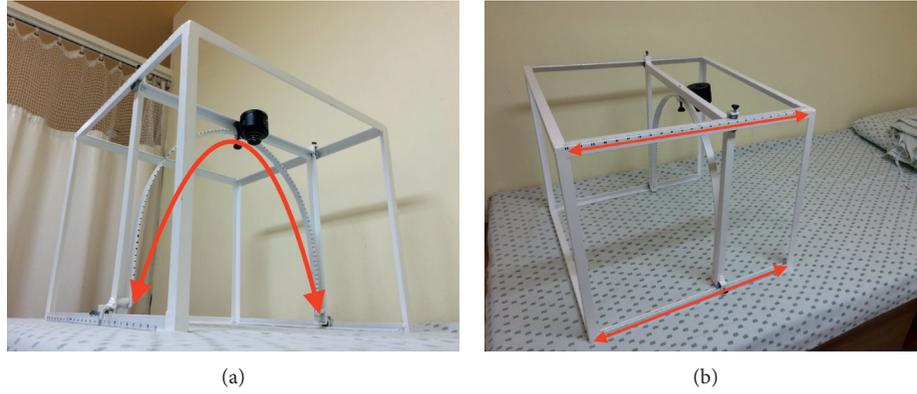


FIGURE 1: Description of WIAS. The camera is located on top of the device, and it can be moved along the arc; also, the frame can be slid horizontally. (a) Bottom view of the device. The arrow shows how the camera slides on the arc. (b) Top view of the device. The arrows show how the frame slides.

$$\text{Skin Mask 1} = \begin{cases} \text{hue } (m, n), & 0.01 \leq \text{intensity} \leq 0.1, \\ 0, & \text{else.} \end{cases} \quad (2)$$

With the Skin Mask 1, it was not possible to identify a wide range of skin tones, so it was necessary to make the algorithm more robust. Therefore, a second color space transformation was applied using the YCbCr color space transformation matrix determined as follows [26]:

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 0 \\ 128 \\ 128 \end{bmatrix} + \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.169 & -0.331 & 0.500 \\ 0.500 & -0.419 & -0.081 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}. \quad (3)$$

Then, a dynamic range was used. The histograms for the Cb and Cr matrix were calculated; then, these values were used to set dynamic limits in order to process different skin colors and tones in a wide range. This means that depending on the histogram values, the algorithm would adjust the threshold, so it would tune itself to the skin tone of the patient. The values found within the dynamic range outlined the second skin mask:

$$\text{Skin Mask 2} = \begin{cases} 1, & \min(Cb) \leq Cb \leq \text{mean}(Cb) \text{ and} \\ & \text{mean}(Cr) \leq Cr \leq \max Cr, \\ 0, & \text{else,} \end{cases} \quad (4)$$

where the limit values of Cb and Cr used in the equation changed for every skin tone found.

Then, in order to link the data from both color spaces, the HSV and YCbCr masks were added in an AND operation; this allowed for the resulting mask to work in a wide range of skin tones. This yielded a more robust algorithm for this stage and a more precise *skin region*.

2.2. Stage 2 (Lesion Region). Once the *skin region* was segmented, *skin lesions* had to be identified. From the raw image in RGB color space, pixel values had to be amplified, so they became darker or lighter as they corresponded to healthy or

damaged regions. For this purpose, the process described below was followed.

The color space *CIE 1976 L*a*b** was used to handle luminosity [27], in order to saturate the intensity values. This transformation was derived from the following equations [28]:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.4125 & 0.3576 & 0.1804 \\ 0.2127 & 0.7152 & 0.0722 \\ 0.0193 & 0.1192 & 0.9502 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix},$$

$$L = \begin{cases} 116 \left(\frac{Y}{Y_n} \right)^{1/3} - 16, & \text{if } \frac{Y}{Y_n} > 0.008856, \\ 903.3, & \text{else,} \end{cases} \quad (5)$$

$$a = 500 \left[\frac{X^{1/3}}{X_n} - \frac{Y^{1/3}}{Y_n} \right],$$

$$b = 200 \left[\frac{Y^{1/3}}{Y_n} - \frac{Z^{1/3}}{Z_n} \right].$$

For this stage, the *L* matrix corresponded to luminosity from black to white and was the one selected [29]. The resulting saturated image was then reenhanced by converting it to grayscale using principal component analysis (PCA) [30]. *Lesion region* was calculated using the histogram of the PCA grayscale image, where a threshold was set to find the damaged areas. This threshold also shifted depending on the tones detected from the healthy and the lesioned skin, but it took approximately 10% of the values found in the image (Figure 2).

2.3. Stage 3 (Characterization). Characterization of the damage in the lower limbs of the patients was performed in 2 stages:

- (1) Data values of extracted features at the segmented *lesion region* were classified into 2 types: morphologic properties—area, major axis, minor axis,

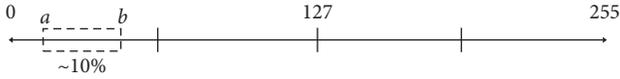


FIGURE 2: Selected PCA grayscale values. The values from a to b represented approximately 10% of the total and were the ones used to determine *lesion region*.

perimeter, solidity—and intensity properties—maximum intensity and minimum intensity.

- (2) The *Shade Index* (ShI) was a parameter used to measure color variations from the RGB raw image. An equation was designed for each color, where equation (6) was used for the Shade Index Red (ShI_R), equation (7) for the Shade Index Green (ShI_G), and equation (8) for the Shade Index Blue (ShI_B):

$$\text{ShI}_R = \frac{\text{mean}(M_{\text{red}})}{\text{mean}(\text{HS}_{\text{red}})}, \quad (6)$$

$$\text{ShI}_G = \frac{\text{mean}(M_{\text{green}})}{\text{mean}(\text{HS}_{\text{green}})}, \quad (7)$$

$$\text{ShI}_B = \frac{\text{mean}(M_{\text{blue}})}{\text{mean}(\text{HS}_{\text{blue}})}, \quad (8)$$

where M_{red} is the red component of the area inside the segmented macule, M_{green} is the green component of the area inside the segmented macule, and M_{blue} is the blue component of the area inside the segmented macule in RGB; HS_{red} is the red component of an area of healthy skin around the macule, HS_{green} is the green component of an area of healthy skin around the macule, and HS_{blue} is the blue component of an area of healthy skin around the macule, in RGB. Finally, a *Shade Index Brown* (ShI_{BR}) (equation (9)) was used to identify brownish changes in the skin:

$$\text{ShI}_{\text{BR}} = \frac{\text{mean}(M_{\text{red}}) + \text{mean}(M_{\text{blue}})}{\text{mean}(\text{HS}_{\text{red}}) + \text{mean}(\text{HS}_{\text{blue}})}. \quad (9)$$

Figure 3 shows the flow diagram for the 3-stage algorithm for *skin* and *lesion region* segmentation in addition to the characterization feature.

It was necessary to find out if the differences were statistically significant among the values calculated for the extracted features in the algorithm for each type of macule. In order to validate this, Student's *t*-test was performed using SPSS v17 with a confidence interval of 95% ($p < 0.05$).

Also, a classifier was designed in order to identify each macule by means of building an artificial neural network and the feature vectors that characterize each of them. 60% of the data was used to train the network and 40% to test it.

3. Results and Discussion

The 3-stage image-processing algorithm reported in this paper is composed by segmentation of skin and its lesions, as

well as the values of the features obtained from the *shade indices*.

Using the *skin images* acquired with the WAIS, the specialist classified the macules found in the patients as vascular and petechiae, due to trophic changes, or due to trauma macules. The results of image processing for the segmentation of *skin region* are shown in Results for Stage 1. Segmentation of *lesion region* is later displayed in Results for Stage 2, and the features of macule characterization are obtained and analyzed in Results for Stage 3.

3.1. Results for Stage 1 (Skin Region). An example of the histogram obtained after the YCbCr color space transformation used to find the dynamic range that self-adjusted to a wide variety of skin tones is shown in Figure 4. Mean value for each matrix fell in the valley of the histograms; the first section of Cb and the second section of Cr were selected in order to find the values that outlined the second skin mask.

Skin region was obtained from *skin image*, as shown in Figure 5. The background was eliminated with the intention of avoiding segmentation errors due to, e.g., the logo of the bed sheets or any object in the back.

3.2. Results for Stage 2 (Lesion Region). Figure 6 shows examples of different *lesion regions* (which include vascular, petechiae, trophic changes, and trauma macules) found in 4 patients. These images were the result of applying the novel proposed algorithm. From these examples, it was noticeable that some areas could be overseen in the RGB images, but after the processing enhancement with the *CIE 1976 L*a*b** color space transformation, the selection of the luminosity matrix, and the PCA gray-scale transformation, these hidden macules are now within the spectrum of the dynamic range selected from the histogram. From this stage, a general state of health of the extremity was calculated and displayed as percentage of damage (29% for patient no. 1, 24% for patient no. 2, 31% for patient no. 3, and 21% for patient no. 4).

3.3. Results for Stage 3 (Characterization of Features). In order to characterize the macules (vascular, petechiae, due to trophic changes, or due to trauma), feature extraction for morphologic properties, intensity properties, and *Shade Indices* was performed in 82 macules obtained from the *lesion regions* found. Table 1 shows the values obtained.

By means of statistical analysis, significant differences ($p < 0.05$) were found among the macules studied; these p values are shown in Table 2. According to it, petechiae and vascular macules can be differentiated through *morphologic properties* and *Shade Indices* (except ShI_B). Differences between petechiae and macules due to trophic changes can be found comparing their *morphologic properties*; *morphologic properties* and *Shade Indices*, ShI_R, were significantly different for petechiae and trauma macules. Vascular macules and those resulting from trophic changes can only be differentiated through their *morphologic properties*, while trauma macules can be differentiated comparing

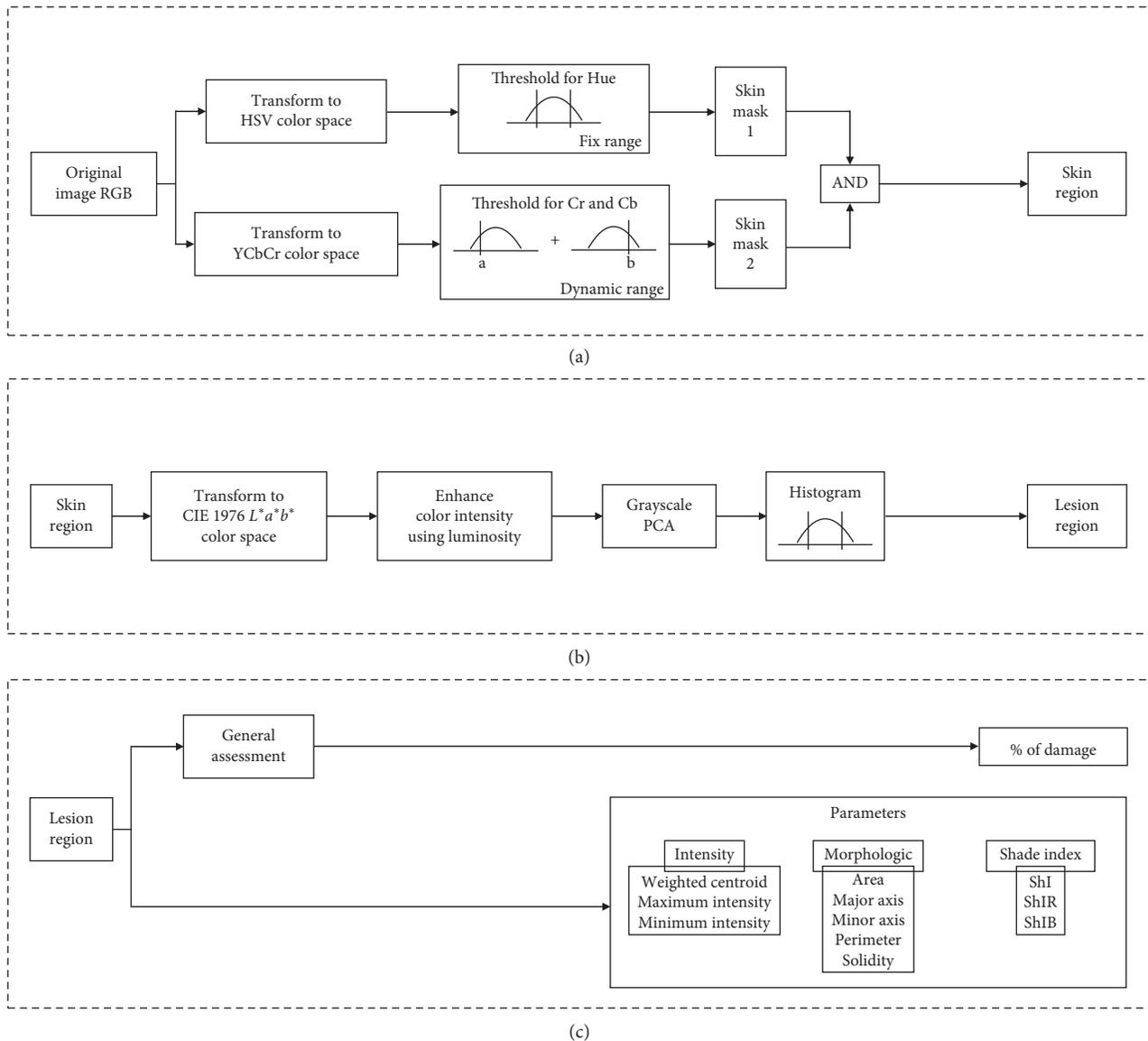


FIGURE 3: Three-stage algorithm: (a) Stage 1—processing from *skin image* to segment a *skin region* using the AND of two skin masks (via HSV and YCbCr color space transformation). (b) Stage 2—*lesion region* segmentation from the *skin region* by means of CIE 1976 $L^*a^*b^*$ transformation followed by luminosity enhancement and PCA grayscale transformation. (c) Stage 3—Characterization of the *lesion region*. Calculation of the parameters for *intensity*, *morphologic* and the *shade indices* for macules.

all properties expected for *solidity* and *minimum intensity*. Macules due to trophic changes and trauma can be differentiated using the *Shade Indices*: ShI_G , ShI_B , and ShI_{BR} , and 4 other properties.

The concatenation of properties, calculated for each macule evaluated, form the feature vector for each example. Figures 7(a) and 7(b) show the average value for each property and macule, or the average feature vector.

So, in order to identify each macule, the proposed architecture is a feedforward backpropagation network with 2 hidden layers and 4 neurons per layer; the transfer functions are hyperbolic tangent sigmoid and logsigmoid. The training function updates weight and bias values according to the Levenberg–Marquardt optimization [31]. In order to train the classifier, an 11×40 matrix was built,

where each type of macule yielded 10 examples; 60% of the data was used for the training, and the remaining 40% was used to test the network. The results were displayed through a confusion matrix (Figure 8(a)), where the coincidence between one of the 10 feature vectors and the target class was demonstrated. The correct identification of the data corresponded to 97.5%. A linear regression of the data (Figure 8(b)) shows the relation between the target data and the results obtained from the network, where $R = 0.95054$ indicates that the model was capable of identifying ~95% of the segmented lesions.

3.4. *Skin Macules Characterization (SMaC) Software*. A Matlab® GUI, called skin macules characterization (SMaC)

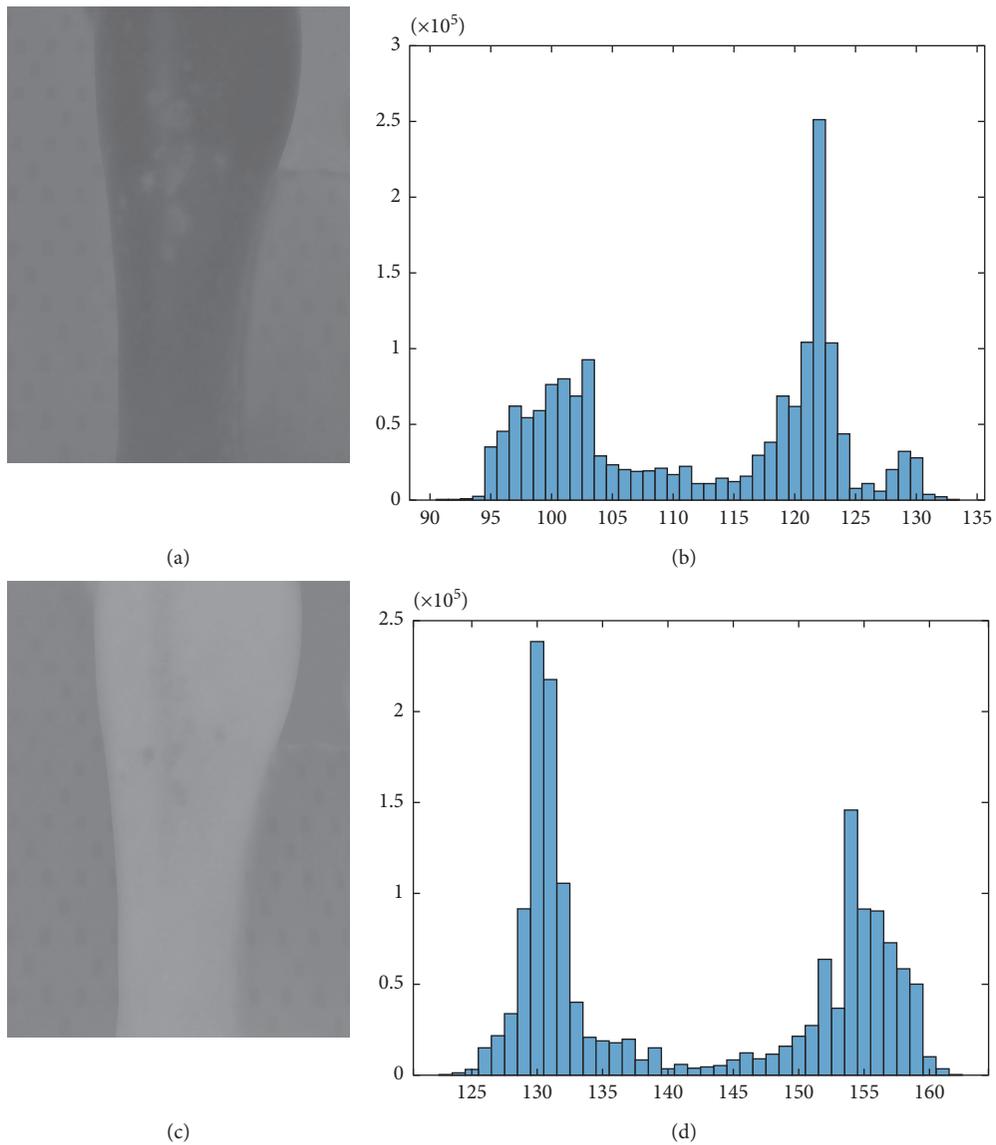


FIGURE 4: Histograms used to find the dynamic range: (a) image obtained for the Cb component from the YCbCr color space transformation; (b) Cb matrix histogram; (c) image obtained for the Cr component from the YCbCr color space transformation; (d) Cr matrix histogram.



FIGURE 5: Stage 1 (*skin region*) segmentation example: (a) raw *skin image*; (b) segmented *skin region*.

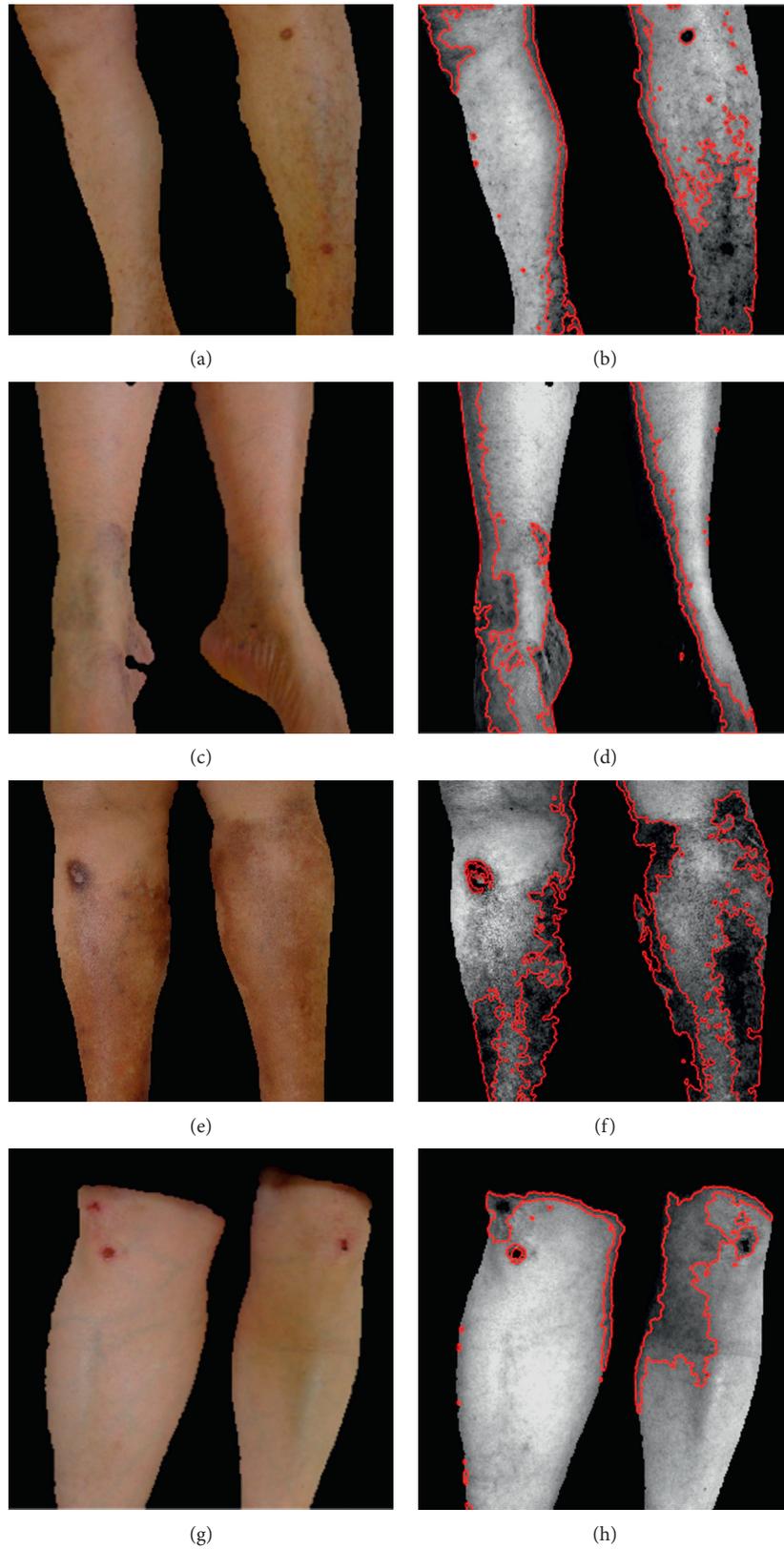


FIGURE 6: Example of segmentation of *skin regions* and *lesion regions* found in Stage 2 for four patients: (a) patient no. 1, *skin region*; (b) patient no. 1 shows 29% damage due to vascular damage and a trauma; (c) patient no. 2, *skin region*; (d) patient no. 2 shows 24% damage due to petechiae and vascular damage; (e) patient no. 3, *skin region*; (f) patient no. 3 shows 31% damage due to trophic changes, a trauma, and vascular damage; (g) patient no. 4, *skin region*; (h) patient no. 4 shows 21% damage due to traumatic lesions and vascular damage.

TABLE 1: Characteristic feature values found for vascular, petechiae, trophic changes, and trauma macules.

Macule	Vascular ($n = 47$)	Petechiae ($n = 10$)	Due to trophic changes ($n = 10$)	Due to trauma ($n = 15$)
<i>Morphologic properties</i>				
Area (pixels)	1010.29 ± 1893.99	18.50 ± 5.99	2152.90 ± 1541.85	2133.60 ± 3684.09
Major axis (pixels)	46.97 ± 48.96	6.48 ± 1.60	109.97 ± 54.13	59.08 ± 63.30
Minor axis (pixels)	23.76 ± 17.75	3.88 ± 1.08	30.67 ± 13.49	29.85 ± 24.87
Perimeter (pixels)	143.90 ± 154.46	15.83 ± 4.76	289.46 ± 116.72	166.16 ± 173.27
Solidity (%)	72.56 ± 12.76	88.00 ± 12.97	70.40 ± 11.96	79.87 ± 9.31
<i>Intensity properties</i>				
Maximum intensity	0.55 ± 0.18	0.61 ± 0.14	0.47 ± 0.16	0.62 ± 0.13
Minimum intensity	0.35 ± 0.18	0.55 ± 0.14	0.28 ± 0.19	0.31 ± 0.21
<i>Shade index</i>				
ShI _R	0.90 ± 0.06	0.95 ± 0.03	0.88 ± 0.07	0.86 ± 0.08
ShI _G	0.84 ± 0.08	0.90 ± 0.07	0.84 ± 0.09	0.74 ± 0.14
ShI _B	0.82 ± 0.11	0.87 ± 0.09	0.84 ± 0.08	0.68 ± 0.15
ShI _{BR}	0.87 ± 0.11	0.92 ± 0.04	0.87 ± 0.07	0.79 ± 0.10

TABLE 2: p values for the differences among features data distribution for each macule relation.

Macule	Petechiae vs vascular	Petechiae vs trophic changes	Petechiae vs trauma	Vascular vs trophic changes	Vascular vs trauma	Trophic changes vs trauma
<i>Morphologic properties</i>						
Area	0.0001*	0.0001*	0.043*	0.0001*	0.002*	0.986
Major axis	0.0001*	0.0001*	0.006*	0.0001*	0.012*	0.043*
Minor axis	0.0001*	0.0001*	0.001*	0.0001*	0.003*	0.916
Perimeter	0.0001*	0.0001*	0.005*	0.0001*	0.011*	0.044*
Solidity	0.003*	0.005*	0.080	0.457	0.054	0.036*
<i>Intensity properties</i>						
Maximum intensity	0.129	0.059	0.776	0.406	0.026*	0.015*
Minimum intensity	0.0001*	0.002*	0.002*	0.343	0.632	0.682
<i>Shade index</i>						
ShI _R	0.0001*	0.023*	0.0001*	0.598	0.038*	0.373
ShI _G	0.024*	0.157	0.0001*	0.668	0.004*	0.035*
ShI _B	0.104	0.435	0.0001*	0.448	0.001*	0.003*
ShI _{BR}	0.004*	0.072	0.0001*	0.899	0.001*	0.035*

*Statistically significant p values ($p < 0.05$).

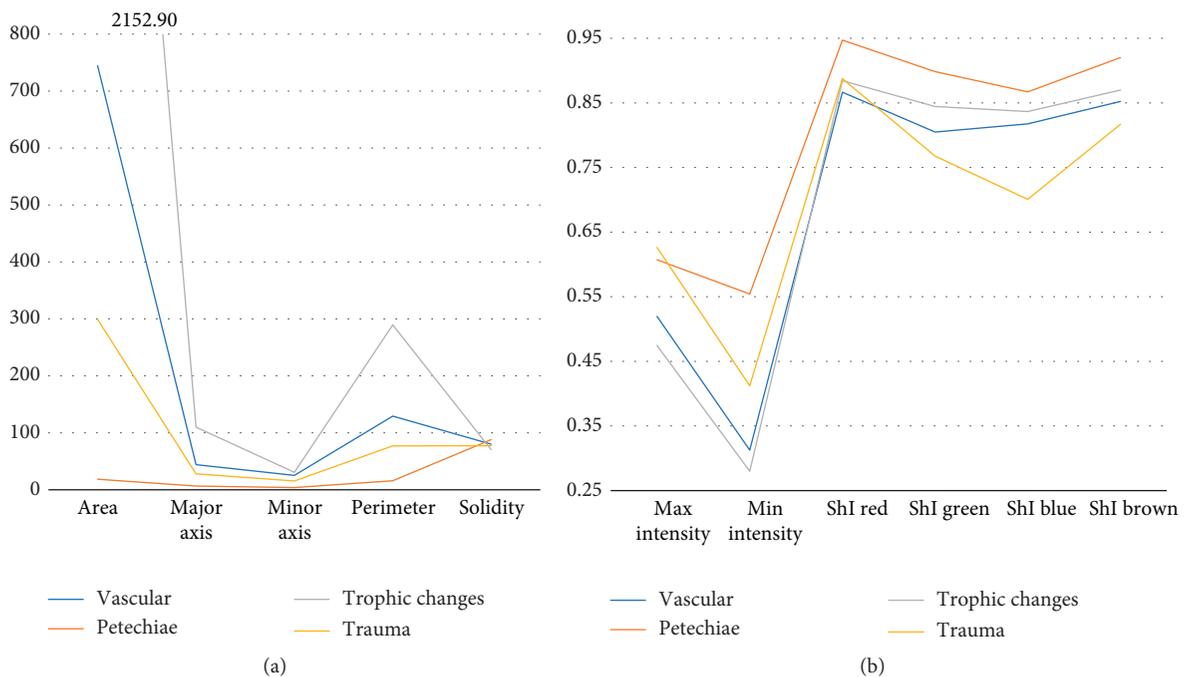


FIGURE 7: Average values of (a) morphologic properties and (b) intensity properties and shade indices.

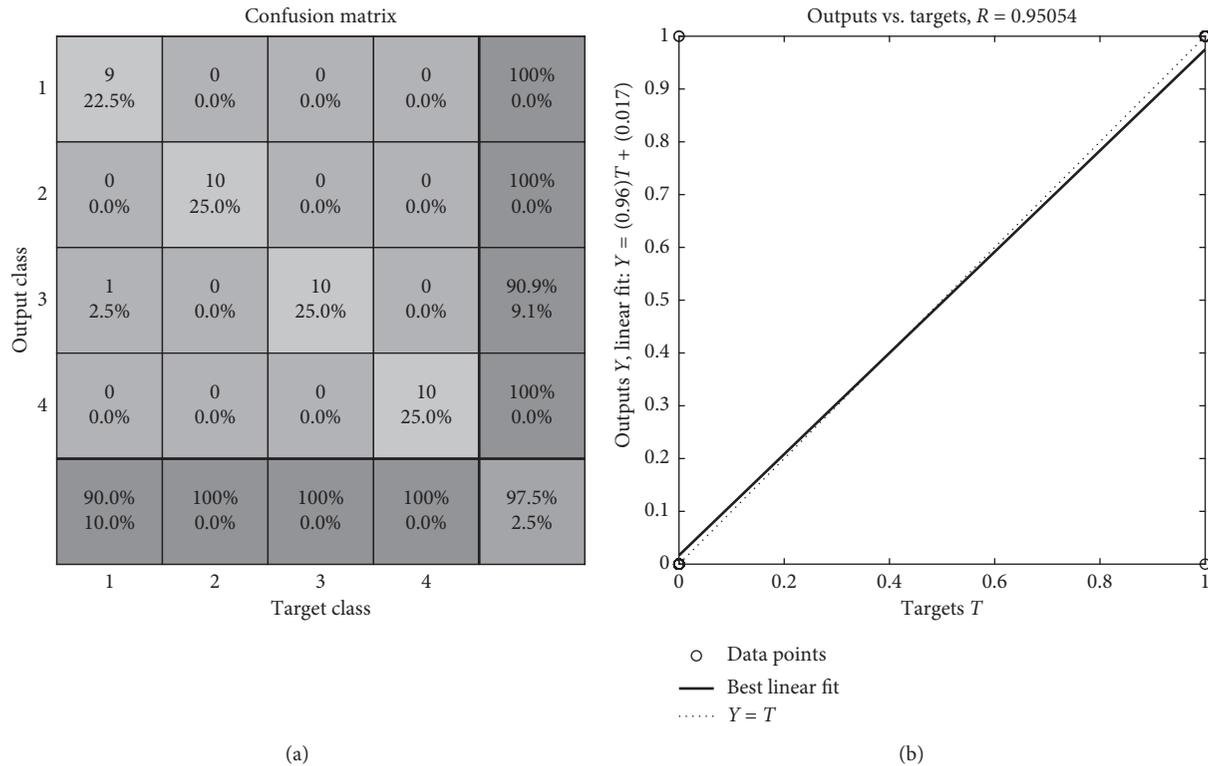


FIGURE 8: Classifier results. (a) Confusion matrix for the macule classifier. Class 1: vascular macule. Class 2: petechiae. Class 3: trophic changes. Class 4: trauma. (b) Linear regression that shows the relation between the elements of the network response and the corresponding targets.

Software©, showed the results of the 3-stage image-processing algorithm for segmentation of the skin macules, feature extraction, and characterization of 4 types of macules: vascular, petechiae, trophic changes, and trauma, as seen in Figure 9. This software was registered at the National Copyright Institution of Mexico (INDAutor, no. 03-2017-071912253000-01).

Macule images need special processing algorithms as they are very peculiar and present different features depending on the patient. There are no algorithms reported to address this particular problem. Moreover, because of the wide range of human skin color tones, the major challenge to overcome was to find the macules in spite of the changes in illumination among the *skin images*.

Color space transformations became a useful tool to find different views of the image that allowed enhancing characteristics that were convenient to solve the segmentation problem. In the HSV color space transformation, the Hue values selected showed a good performance with medium skin colors, but it depended of the light in the room. In order to address this situation, a second color space transformation was applied (YCbCr). This color space extracted red and blue components from the image; since macule color varies from red to brownish, these components became very helpful for macule location and segmentation.

In this case, a fix range for the Cb and Cr values was not useful, even when it is the method of choice in the literature [26, 28] because it limits the variety of human skin tones detected to a small selection. However, the dynamic range proposed in this paper allowed the algorithm to adjust to a

wide range of skin tones, which increases its usefulness meaningfully. The minimum and maximum values taken from the Cb and Cr matrices represented the illumination range of the image. Histogram values allowed the algorithm to self-tune to the specific image and hence to the specific skin tone and illumination, maintaining the simplicity and efficiency of the algorithm without adding the computational cost of neural networks. The position of the camera can be adjusted using the *WIAS* device in order to avoid areas with too much brightness or intense illumination.

A normal grayscale transformation was not useful for skin segmentation because it equalizes the distribution of gray levels, which is counterproductive for this scenario. On the contrary, PCA generates an image in gray levels within the limit values of the histogram of one specific image every time, it gets rid of the healthy tissue in the image and keeps the sections with clear manifestations of saturation, and these sections are classified and selected as lesioned skin. This technique helps take advantage of the illumination enhancements achieved by the *CIE 1976 L*a*b** transformation used in Stage 2. And, again, to set thresholds and ranges through histogram, values allow the algorithm to adjust to the particularities of the *lesion region* found without the need to use more complex processing in order to classify between healthy and lesioned skin. So, this algorithm is a simple and efficient solution for processing macules, which can have multiple applications.

Ideally, in an image identification process posterior to a segmentation, it is important to have a feature or property that allows distinguishing between classes of a group of data.

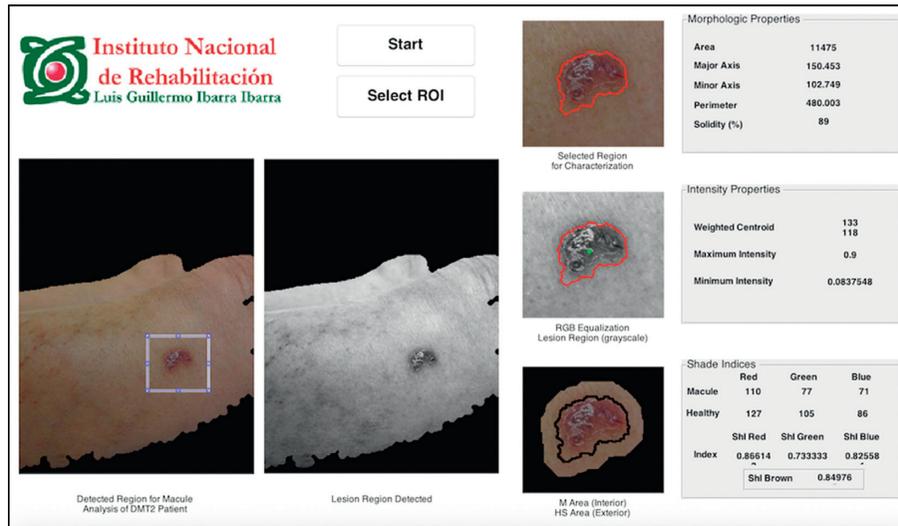


FIGURE 9: SMaC Software© for segmentation and characterization. This example shows the segmentation and characterization of a trauma macule.

This can be complicated depending on the characteristics of what is being identified in the image, so quantitative parameters are preferred to guarantee a more robust result.

Therefore, the macules measured were characterized using *morphologic properties*, to define shape and geometry; *intensity properties*, to establish maximum and minimum values of the pixels within the lesion and to separate them from the healthy skin; and the proposed *Shade Indices*, to identify lesions by color tone. During the development of the characterization (Stage 3), it became evident that, e.g., color red (in the RGB image) did not look the same in every skin tone or even with different illumination, so the reference value could not be fixed. The solution to overcome these problems was to come up with a novel set of *Shade Indices*, where the healthy skin around the macule was used as a reference for color tone shifts. On the contrary, ShI_R and ShI_{BR} turned out to be the indices that helped differentiating the most between the macules studied.

Morphologic properties are features of the macule in which it was possible to point out geometric and shape variations among most of them. *Intensity properties* did not seem to have a considerable input for classification of data since their p values were significant in less than 40% of the relations studied.

In general, from this analysis, it was comprehensible how complex the problem was since different kinds of macules were present at the same time and, even more, they were overlaid. From the image segmentation and processing point of view, there was a high difficulty to isolate the *lesion region* to provide an accurate assessment. Nevertheless, with the macule properties chosen and calculated, it was possible to classify each type of macule with 97.5% of accuracy.

With the use of the SMaC Software© characterization and latter classification, the macules of patients with diabetes can be measured and tracked along the development of the disease in order to prevent further disabilities and comorbidities. The use of this software can be especially beneficial for those physicians who do not have specialized training or

enough expertise to identify specific macules; it can also be used as an educational tool. The perception of the importance on skin manifestations that appear previous to ulcers or amputations must be changed since they seem to be the first symptom of endothelial decay and vascular damage which lead to worse symptoms of diabetic foot and, eventually, to amputation. From the clinical perspective, the origin of skin and limb damage is multifactorial, but it relates mainly to endothelial decay.

In the future, we aim to turn this GUI Software into a *Diagnosis Assistance Tool*, which would include clinical variables and other diabetic foot manifestations in order to gather enough data to eventually form a database of patients with diabetes, for preventive purposes.

4. Conclusion

Nowadays, lower limb skin manifestations are not taken into account in the general evaluation of the state and development of diabetes mellitus type II, even when they have an underlying vascular origin. This paper presents the application of an algorithm for the segmentation, characterization, and classification of skin manifestations from photographic images and the identification of them in the lower limbs of diabetic patients. An efficient algorithm for image processing of skin macules characterization performed by means of extracting *morphologic* and *intensity properties* is proposed, along to a new set of *Shade Indices* used to assess color shifts in different skin tones. From the three sets of features, *morphologic properties* and *Shade Indices* resulted statistically significant in order to differentiate among macules of various origins. The indices described here are a new way to assess changes in color for different skin tones, which increase the usefulness of the application. The properties extracted are used as feature vectors for the input of a classification network which resulted in a 97.5% accuracy for the 4 types of macules studied in this paper: vascular, petechiae, trophic changes,

and trauma. The SMaC Software© was designed to bring the proposed algorithm as a tool for the physician in order to aid in the identification and assessment of skin lesions in the lower limbs of patients with diabetes.

Data Availability

Photographic images of macules at the lower limbs of patients with diabetes used to support the findings of this study are restricted by our Institutional Ethics and Research Review Board in order to protect patients' privacy. The data may be released upon petition to Research Review Board who establishes the criteria to access confidential data.

Conflicts of Interest

There are no conflicts of interest to declare.

Acknowledgments

This work was supported by the National Council for Science and Technology (Consejo Nacional de Ciencia y Tecnología—CONACYT) (CONACYT SALUD-2013-1-201590). The authors would also like to thank Dr. Verónica Álvarez for her contributions regarding dermatopathology.

References

- [1] Population Reference Bureau, "Noncommunicable diseases in Latin America and the Caribbean: youth are key to prevention," July 2016, <http://www.prb.org/Publications/Datasheets/2013/noncommunicable-diseases-latinamerica-youth-datasheet.aspx>.
- [2] W. F. Ganong, *Balance Energético, Metabolismo y Nutrición. Fisiología Médica*, Editorial El Mundo Moderno, Madrid, Spain, 2006.
- [3] World Health Organization, April 2018, <http://www.who.int/diabetes/global-report/es/>.
- [4] Encuesta Nacional de Salud y Nutrición 2016, May 2018, <http://ensanut.insp.mx>.
- [5] International Diabetes Federation, "IDF diabetes atlas," May 2018, <http://www.idf.org/diabetesatlas>.
- [6] Organización Panamericana de la Salud, "Intervenciones para la prevención y manejo de la Diabetes y la Obesidad," May 2016, http://www.paho.org/hq/index.php?option=com_content&view=article&id=6721&Itemid=39451.
- [7] J. Viadé and J. Royo, "Piel y diabetes," in *Pie Diabético: Guía Para la Práctica Clínica*, pp. 73–82, Médica Panamericana, Madrid, Spain, 2nd edition, 2013.
- [8] E. C. Rodríguez, *Manifestaciones Cutáneas de la Diabetes Mellitus*, Centro Dermatológico Federico Lleras Acosta, Univ Med 44, Bogotá, Colombia, 2003.
- [9] J. I. Blanes, *Etiopatología del Pie Diabético, Capítulo III. Tratado del Pie Diabético*, Centro de Documentación del Grupo Esteve, Spain, 2002.
- [10] T. P. Habif, "Principles of diagnosis and anatomy," in *Clinical Dermatology*, Mosby Elsevier, St. Louis, MO, USA, 5th edition, 2009.
- [11] F. R. A. Hogg, G. Peach, P. Price, M. M. Thompson, and R. J. Hinchliffe, "Measures of health-related quality of life in diabetes-related foot disease: a systematic review," *Diabetologia*, vol. 55, no. 3, pp. 552–565, 2012.
- [12] D. G. Armstrong, L. A. Lavery, and L. B. Harkless, "Validation of a diabetic wound classification system: the contribution of depth, infection, and ischemia to risk of amputation," *Diabetes Care*, vol. 21, no. 5, pp. 855–859, 1998.
- [13] C. Liu, F. van der Heijden, M. E. Klein, J. G. van Baal, S. A. Bus, and J. J. van Netten, "Infrared dermal thermography on diabetic feet soles to predict ulcerations: a case study," in *Proceedings of Advanced Biomedical and Clinical Diagnostic Systems, XI*, vol. 8572, San Francisco, CA, USA, March 2013.
- [14] M. E. Kiziltan, G. Benbir, and M. A. Akalin, "Is diabetic Dermopathy a sign for severe neuropathy in patients with diabetes mellitus? Nerve conduction studies and symptom analysis," *Clinical Neurophysiology*, vol. 117, no. 8, pp. 1862–1869, 2006.
- [15] N. Chatterjee, C. Chattopadhyay, N. Sengupta, C. Das, N. Sarma, and S. Pal, "An observational study of cutaneous manifestations in diabetes mellitus in a tertiary care Hospital of Eastern India," *Indian Journal of Endocrinology and Metabolism*, vol. 18, no. 2, pp. 217–220, 2014.
- [16] T. Pavicic and H. C. Korting, "Xerosis and callus formation as a key to the diabetic foot syndrome: dermatologic view of the problem and its management," *JDDG*, vol. 4, no. 11, pp. 935–941, 2006.
- [17] S. Furqan, L. Kamani, and A. Jabbar, "Skin manifestations in diabetes mellitus," *Journal of Ayub Medical College Abbotabad*, vol. 26, no. 1, 2014.
- [18] M. G. Woodbury, R. Gary Sibbald, B. Ostrow, R. Persaud, and J. M. Lowe, "Tool for rapid & easy identification of high risk diabetic foot: validation & clinical pilot of the simplified 60 second diabetic foot screening tool," *PLoS One*, vol. 10, no. 6, Article ID e0125578, 2015.
- [19] C. Toledo, J. F. Ramos, A. Vera, L. Leija, and J. Gutierrez, "Non-invasive imaging techniques to assess diabetic foot ulcers: a state of the art review," in *Proceedings of the Pan American Health Care Exchanges Conference*, Brasilia, Brazil, April 2014.
- [20] M. J. M. Vasconcelos, L. Rosado, and M. Ferreira, "A new risk assessment methodology for dermoscopic skin lesion images," in *Proceedings of IEEE International Symposium on Medical Measurements and Applications (MeMeA)*, pp. 570–575, Torino, Italy, May 2015.
- [21] W. Y. Chang, A. Huang, C. Y. Yang et al., "Computer-aided diagnosis of skin lesions using conventional digital photography: a reliability and feasibility study," *PLoS One*, vol. 8, no. 11, Article ID e76212, 2013.
- [22] J. Jaworek-Korjakowska, "Novel method for border irregularity assessment in dermoscopic color images," *Computational and Mathematical Methods in Medicine*, vol. 2015, Article ID 496202, 11 pages, 2015.
- [23] C. Toledo, J. F. Ramos, and A. Vera, L. Leija and J. Gutierrez, "Técnicas de segmentación aplicadas a las máculas de la piel como método de detección temprana de la progresión del pie diabético," *Investigación en Discapacidad*, vol. 2014, no. S1, p. S60, 2014.
- [24] M. C. Shin, K. I. Chang, and L. V. Tsap, "Does color space transformation make any difference on skin detection?," in *Proceedings of 6th IEEE Workshop on Applications of Computer Vision*, pp. 1–5, Orlando, FL, USA, December 2002.
- [25] A. R. Smith, "Color gamut transform pairs," in *Tutorial: Computer Graphics*, pp. 376–383, IEEE Computer Society Press, Los Alamitos, CA, USA, 2nd edition, 1982.
- [26] K. B. Shaik, P. Ganesan, V. Kalist, B. S. Sathish, and J. M. M. Jenitha, "Comparative study of skin color detection and segmentation in HSV and YCbCr color space," *Procedia Computer Science*, vol. 57, pp. 41–48, 2015.

- [27] D. J. Bora, A. K. Gupta, and F. A. Khan, "Comparing the performance of L*A*B* and HSV color spaces with respect to color image segmentation," *International Journal of Emerging Technology and Advanced Engineering*, vol. 5, no. 2, 2015.
- [28] A. Kaur and B. V. Kranthi, "Comparison between YCbCr color space and CIELab color space for skin color segmentation," *International Journal of Applied Information Systems*, vol. 3, no. 4, pp. 30–33, 2012.
- [29] C. Kanan and G. W. Cottrell, "Color-to-Grayscale: does the method matter in image recognition?," *PLoS One*, vol. 7, no. 1, Article ID e29740, 2012.
- [30] A. A. Gooch, S. C. Olsen, J. Tumblin, and B. Gooch, "Color2Gray: saliency-preserving color removal," February 2016, <http://www.cs.northwestern.edu/~ago820/color2gray/>.
- [31] S. Roweis, "Levenberg-marquardt optimization," June 2018, <http://www.cs.toronto.edu/~roweis/notes/lm.pdf>.

Research Article

Automated Region Extraction from Thermal Images for Peripheral Vascular Disease Monitoring

Jean Gauci ¹, Owen Falzon ¹, Cynthia Formosa ², Alfred Gatt ², Christian Ellul,²
Stephen Mizzi ², Anabelle Mizzi ², Cassandra Sturgeon Delia,³ Kevin Cassar,³
Nachiappan Chockalingam ⁴, and Kenneth P. Camilleri ¹

¹Centre for Biomedical Cybernetics, University of Malta, Msida MSD2080, Malta

²Department of Podiatry, University of Malta, Msida MSD2080, Malta

³Department of Surgery, University of Malta, Msida MSD2080, Malta

⁴Faculty of Health Sciences, Staffordshire University, Staffordshire ST4 2DE, UK

Correspondence should be addressed to Jean Gauci; jean.gauci@um.edu.mt

Received 18 June 2018; Revised 26 October 2018; Accepted 15 November 2018; Published 13 December 2018

Guest Editor: Orazio Gambino

Copyright © 2018 Jean Gauci et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This work develops a method for automatically extracting temperature data from prespecified anatomical regions of interest from thermal images of human hands, feet, and shins for the monitoring of peripheral arterial disease in diabetic patients. Binarisation, morphological operations, and geometric transformations are applied in cascade to automatically extract the required data from 44 predefined regions of interest. The implemented algorithms for region extraction were tested on data from 395 participants. A correct extraction in around 90% of the images was achieved. The process of automatically extracting 44 regions of interest was performed in a total computation time of approximately 1 minute, a substantial improvement over 10 minutes it took for a corresponding manual extraction of the regions by a trained individual. Interrater reliability tests showed that the automatically extracted ROIs are similar to those extracted by humans with minimal temperature difference. This set of algorithms provides a sufficiently accurate and reliable method for temperature extraction from thermal images at par with human raters with a tenfold reduction in time requirement. The automated process may replace the manual human extraction, leading to a faster process, making it feasible to carry out large-scale studies and to increase the regions of interest with minimal cost. The code for the developed algorithms, to extract the 44 ROIs from thermal images of hands, feet, and shins, has been made available online in the form of MATLAB functions and can be accessed from <http://www.um.edu.mt/cbc/tipmid>.

1. Introduction

Body temperature has long been used as a natural indicator of diseases [1]. Normally, the body maintains a constant core temperature of 37°C. The body naturally responds to changes in the peripheral regions, by sweating and other physiological processes, to control the core temperature. Diseases, such as peripheral arterial disease (PAD) and local infection, may affect this thermoregulatory process resulting in cooler extremities or increased peripheral temperature [2].

Peripheral arterial disease is a condition in which plaque builds up in the arteries carrying blood from the heart to the peripheries. Over time, the extent of plaque increases leading to narrowing of the lumen of the arteries,

restricting blood flow to the peripheries and thus reducing the ability of the body to regulate the peripheral temperature [3]. Diabetic patients are more likely to suffer from PAD and its complications, with the lower limbs being the most affected area in what is commonly known as the diabetic foot [4]. Assessing the temperature in the lower limbs in patients suffering from PAD is critical to detect any arising complications, such as local infection, which may eventually lead to amputation of part of or the whole limb. Presently, the clinical practice of temperature assessment involves manual palpation of the lower limbs by the clinician. However, this method is unable to detect small temperature changes in the foot and may be affected by various exterior factors [2].

Thermal imaging detects the long-wave infrared radiation emitted from a surface of interest. The amount of radiation can be accurately related to the temperature of the same surface [5]. Human skin, with an emissivity value of 0.98, is a particularly suitable material for temperature measurement using thermal imaging [6]. This is reflected in the increase in the study and use of thermal imaging in medical and clinical applications [1, 7, 8]. One of the potential medical application areas is the monitoring of peripheral arterial disease, and its complications, in the diabetic foot [1, 9]. Compared to the present method of detecting differences in temperature in the diabetic foot using palpations of the plantar aspect of the foot, thermal imaging can detect very small changes in the surface temperature. Moreover, it can easily detect these changes in multiple places as opposed to the manual method which can monitor one region at a time. Furthermore, thermal imaging gives an objective value for the temperature in contrast with the subjective temperature estimate obtained from manual palpation.

Areas of increased temperature in the foot are predictive of the development of ulcerations. A difference in temperature in contralateral regions of the plantar aspect greater than 22°C is generally considered to be abnormal [10]. Benbow et al. have concluded that thermography is a simple, inexpensive, and noninvasive method for the identification of the neuropathic foot and the increased risk of ulceration [11].

Oe et al. found that when examining the thermal images in a case study of a subject with diabetes and a foot ulcer, a high-temperature area could be identified extending from the fourth toe (where the ulcer was present) to the ankle [12]. The patient was later diagnosed with osteomyelitis in the ulcer on the left foot. Based on these observations, thermography could prove useful for screening of foot ulcers for further complications. Tamaki et al. confirmed the predictive power of thermal imaging in detecting osteomyelitis in diabetic feet by studying 18 subjects with osteomyelitis and obtaining a 100% positive predictive value [13].

Nagase et al. [14] have monitored thermographic patterns in the plantar surface of the foot and compared the patterns between nondiabetic subjects and diabetic subjects. They have determined 20 different possible patterns which the plantar thermographic pattern can follow using the four different angiosomes in the plantar region of the foot. Both healthy control subjects and diabetic subjects were screened, and each of these was categorised to one of the 20 patterns. While the control group subjects were generally categorised in two of the patterns, the diabetic group had a larger distribution across the 20 patterns. This suggests that, by monitoring the thermal patterns in the plantar aspect of the foot, it is possible to identify any abnormalities and possibly detect complications in the diabetic foot.

In our previous work, Gatt et al. [2], we have determined the normative heat pattern distribution in hands, feet, and shins in a population of healthy adults. We identified 44 regions of interest (ROIs) across the three body regions from which the thermal patterns were established. Our aim was to repeat this study on a population of subjects which suffer

from diabetes and PAD to identify any difference in the thermal patterns between the two populations. Such a study would help us to better understand the relationship between PAD and how it affects the peripheral temperature. However, such an analysis would require a large number of subjects. In the study of Gatt et al. [2], the temperature values were extracted manually by trained clinicians using an area demarcation software tool, namely, FLIR Altair software. This process was time-consuming, especially considering the large number of ROIs that had to be extracted. Additionally, regions extracted had a degree of subjectivity and non-repeatability owing to the human element in the extraction process. An automated approach that allows for the rapid and consistent segmentation of regions of interest from thermal images is desirable as a first step in the monitoring of PAD using thermography. In this work, we will focus on the technical aspect of this segmentation process. Specifically, a method for automatic extraction of temperature data from ROIs on thermal images of human hands, feet, and shins is presented.

2. Related Work

The literature in the field of thermal imaging for the monitoring of diabetic feet, and medical thermography in general, focuses on the detection of abnormal regions of temperatures which may be indicative of emerging ulcerations or other pathologies. This is often done through asymmetry analysis in which regions from one foot are compared to the corresponding regions on the contralateral foot, with a deviation of more than 22°C being considered substantial to warrant further investigation [15–17]. In such studies, the body regions are segmented from the background such that corresponding contralateral regions can be compared. This step is generally carried out using either thresholding or clustering method [17, 18]. Once the body region is segmented from the background, the two contralateral regions are transformed such that corresponding pixels are aligned. This step is usually carried out using either rigid or nonrigid transformation based on automatically detected corresponding points of the body.

A further step beyond segmenting the body regions from the background is to subsegment the identified body regions. Yoon et al. [19] analysed thermal images of human arms, and automatically extracted the forearm region from the rest of the arm. Their method was shown to correctly extract the forearm regions; however, no further analysis was carried out on the segmented areas.

Herry [20] also segmented regions from thermal images of hands, with the fingers and the palm being identified. The subsegmented regions were visually inspected to provide qualitative results, but no quantitative performance measures were provided.

Blank and Kargel [21] proposed a method to identify finger and palm regions from thermal images of human hands using morphological opening with varying radii circular structuring elements. In an iterative process, the radius of the structuring element was increased and the area and circumference of the structure that remained after

morphological opening were recorded. By detecting the two largest peaks in the first derivative of the change in area, dA/dr , the finger and palm regions could be identified. However, the derivative of the signal is susceptible to minor local variations and may result in false detections. Blank and Kargel tested their method on thermal images from seven subjects. When we tested this method on a larger test database, this method failed to reliably extract the correct points. In the work of Blank and Kargel, the average finger and palm temperatures were next computed from all pixels in these regions. The method is pose independent but critically depends on good background-foreground segmentation, as any areas which are segmented incorrectly will change the value and position of the peaks. Since in our work the hand thermal images have a good foreground to background contrast, this method was adopted with some variations to segment the fingers from the rest of the hand for subsequent ROI extraction.

The same method was also used by Gauci et al. [22] to subsegment the fingers and the palm from the rest of the hand in thermal images.

Gauci et al. [22] automatically extracted temperature values from anatomical regions of interest from thermal images of human hands, feet, and shins. The methods disclosed in this paper are an improvement of the algorithms proposed in this work.

Most of the reviewed studies stop at the background-foreground segmentation or at the subsegmentation stage, from which the mean temperature of the whole body region is computed [21]. However, Nagase et al. have shown that certain pathologies are characterised by localized variations in temperature [14]. A localized temperature variation might not substantially affect the average temperature of the whole body region, and thus the detectability of this local temperature variation is diminished if only the mean temperature is considered. Furthermore, the mean temperature of the whole body region does not provide any spatial information on the local temperature variations. Therefore in order to detect the pathology, temperature measurements need to be extracted from localized regions of interest rather than from the whole body region.

Scientific studies on the thermal patterns associated with normal and pathological subjects would benefit from automated image analysis techniques in order to allow for the processing of large volumes of data. Similarly, deployment of such tools in a clinical setting dealing with sizeable number of patients would benefit from automated techniques that would provide the physician with readily available results. This motivates the efforts towards automated techniques for the extraction of reliable temperature data from regions of interest. In this work, we extract consistent local regions of interest from the subsegmented body regions in order to measure specific local temperatures.

3. Methods

3.1. Data Acquisition. For the purpose of this work, thermal images of the volar aspect of hands, the anterior aspect of shins, and the dorsal aspect of the feet were considered.

Thermal and visual images as shown in Figure 1 were acquired from each participant. A clutter-free background is assumed for the acquired images. In total, 44 anatomical regions of interest (ROIs) were considered as shown in Figure 2. These ROIs are linked to regions of the foot which are at higher risk of developing ulcerations, while the regions on the hands and the shins serve as baseline data as these regions are relatively unaffected by PAD. Since in our work, we are considering only healthy limbs without any amputated peripherals, we have a set of 44 ROIs for each subject that need to be found in the images and labelled uniquely with 44 distinct ROI labels. 395 subjects gave their informed consent to participate in this study which was approved by the University Research Ethics Committee at the University of Malta.

A FLIR SC7200M infrared thermal camera, which has a spatial resolution of 320×256 pixels and a thermal sensitivity of 20 mK with an accuracy of $\pm 1^\circ\text{C}$, was used to acquire thermal images. Standard medical thermal image acquisition protocol was followed to acquire the thermal images [23, 24]. The camera was kept perpendicular to the surface being imaged and at a distance between 1.5 m and 2 m from the surface of interest. A corresponding visual image was acquired for every thermal image using a Canon EOS 1100D digital visual camera with a resolution of 12 MP. It is worth mentioning that the same anatomical portion of the corresponding body region was captured in all images. Specifically, in the hands, it is important to capture all the fingers and the palm and that these are centred in the image such that some of the forearm is also included. In the case of the shin images, the section of the shins from just below the knee till the toes has to be in full view and, in the case of the feet, the whole plantar aspect of the feet, including the toes, must be acquired. This is the only operator-dependent procedure in this algorithm.

Since the thermal and visual images were captured using different acquisition devices, images of the same region are not aligned. When capturing the images, the subject was asked to remain stationary during the acquisition of the thermal and visual images. Both images were captured from approximately similar poses and distances and perpendicularly to the surface of interest; thus, the two images are sufficiently close to orthographic and with little misalignment between the two images.

The segmentation methods developed herein exploit the shape and geometry of the body regions of interest. Since the body regions considered in this work vary significantly in shape, three different shape-based methods were developed, the hand images, the feet images, and the shin images, respectively.

3.2. Hands Segmentation and Extraction Algorithm. The first step for every image, whether hand, foot, or shin, is that of identifying the body regions from the background in the thermal image. The hands usually have a significantly higher temperature than the background and consequently a higher intensity in the thermal image. However, regions at the edges of the hand, such as the finger tips, might have a significantly

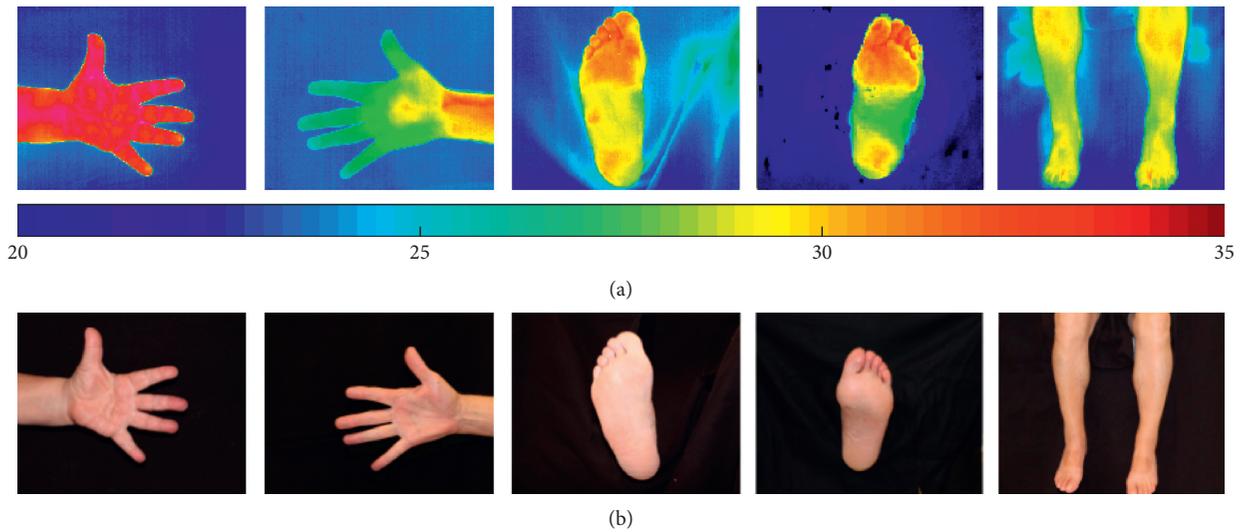


FIGURE 1: Sample thermal (a) and visual (b) images acquired in this study.

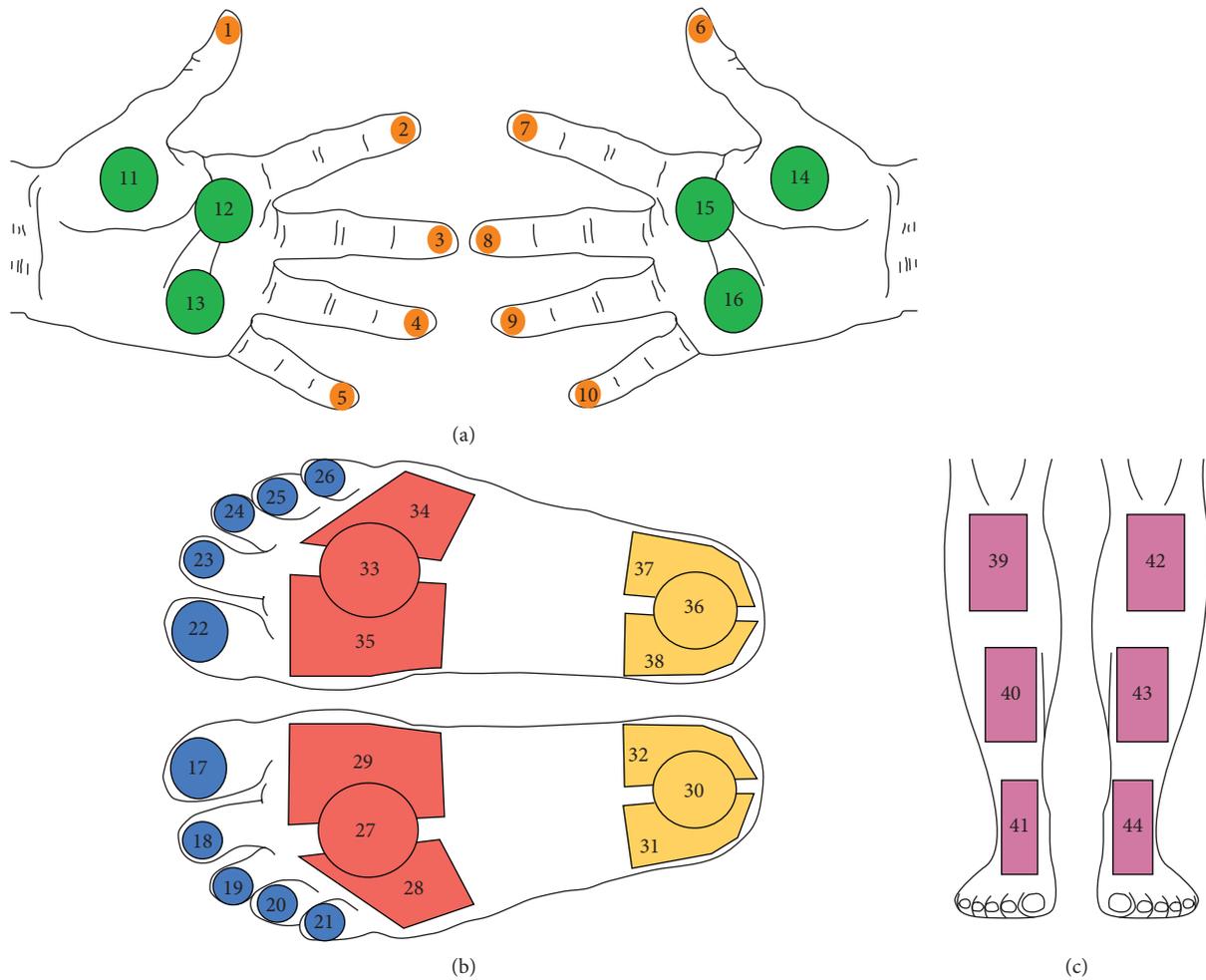


FIGURE 2: 44 anatomical regions of interest identified on the hands (a), feet (b), and shins (c).

lower temperature than central regions such as the palm of the hand. For this reason, local adaptive thresholding for image binarisation was used so that the threshold is adapted

locally for each pixel according to the intensity of the pixel neighbourhood. For the thresholding process, a Gaussian weighted mean statistic in a neighbourhood which is 1/16th

the size of the image was used. Figure 3 shows the segmentation of a sample hand thermal image.

Besides segmenting the hand from the background, it is necessary to identify the fingers. This was done using an adaptation of the iterative morphological opening approach proposed by Blank and Kargel [21]. Specifically a morphological opening operation using a disc-shaped structuring element (SE) was used to identify the finger pixels from the rest of the hand. The morphological opening operator was applied in an iterative manner, increasing the radius r of the disc-shaped SE at each iteration. For every iteration, the area $A(r)$ of remaining foreground pixels was calculated. Figure 4 shows the plot of $A(r)$ versus the increasing radii of the SE, r , together with the actual remaining pixels on the image. Three salient features can be identified in the plot, namely, the sudden drop in area when the SE radius is large enough to remove the fingers from the image; the long plateau segment following this drop, where there is only minor smoothing of the palm region; and another sharp, but shallower, drop when the SE radius is large enough to remove the forearm from the image. The radius of the SE at the end of the first drop in area is used to identify the fingers from the rest of the hand. To avoid false detections arising due to the susceptibility of the first derivative, dA/dr , to local variations, the required radius was determined based on the variance of $A(r)$ across the increasing r since this variance is less susceptible to local variations due to the averaging effect. Specifically, a variance measure $V(r)$ at every point along $A(r)$ was calculated. $V(r)$ is defined as

$$V(r) = \frac{1}{N-1} \sum_{r'=r-N/2}^{r+N/2} (A(r') - \mu(A(r')))^2, \quad (1)$$

where N is the size of a moving window around r . The midpoint of the signal $A(r)$ was observed to generally lie on the second salient feature of the signal (i.e., the long plateau segment following the first drop; refer to Figure 4). Thus the midpoint of $A(r)$ was used as a simple heuristic to provide a starting point for the process of identifying the required radius. Starting from this midpoint and moving towards $r = 0$, the average variance of the first three neighbourhoods was used to establish a reference value. The first neighbourhood whose variance was greater than the reference by a predefined factor was taken to be the radius of SE required to identify the fingers. Using this radius, a binary image which contains only the fingers in the image can be obtained, as shown in Figure 5.

From the finger binary image, shown in Figure 5(b), each finger was identified by analysing the angle between the lines from the centroid of each finger blob to the centroid of the palm. The largest angle corresponds to the angle between the thumb and the little finger since these two fingers are the most further apart in the hand. Since each hand image is labelled as a left hand or a right hand, the direction of the angle between the two lines, clockwise or anticlockwise, can be used to distinguish the thumb from the little finger. The smallest angle from the identified thumb then corresponds to the angle between the thumb and the index finger. Using similar computations, all five fingers can be identified in a manner which is rotation independent.

A template which contains the three ROIs on the palm of the hand was defined, as shown in Figure 6(a). The template has three anchor points, circled in red in Figure 6(a). To determine the 3 ROIs on the palm, the 3 corresponding points on the thermal image are determined by detecting the intersection points between the major axis of the thumb, index, and little finger blobs and the edges of the respective blobs, selecting the intersection point closest to the palm centroid. This process is shown in Figure 6(b), with the selected intersection point shown in blue. An affine transformation, mapping these three points to the corresponding anchor points on the template is determined to fit the template on the palm of the image being processed and determine the three ROIs on the palm.

To extract the five ROIs on the finger tips, the Hough transform was applied to the edges of each individual finger. If more than one circle is returned, the circle whose centre is closest to the fingertip of the finger, shown in green in Figure 6(b), is used and all other circles are discarded. Figure 7 shows the hand ROIs extraction process.

3.3. Shins Segmentation and Extraction Algorithm. Similar to the hands, the shins in the thermal images have a higher intensity than the background, and thus thresholding can be used to identify the shins from background. Local adaptive thresholding was once again used. Figure 8 shows this procedure being applied on a sample shins thermal image.

Following binarisation, the medial line of each shin is found and the angle which this line makes with the horizontal axis of the image is computed. This angle is used as the orientation of the limb being processed. Three points are identified on the medial line which will serve as reference points for the placement of the ROIs on the shins. These points lie at the top of the shin, at the centre of the shin, and at 3/4 of the length of the shin according to the predefined position of the shin ROIs as shown in Figure 2. Each point is then used to place a rectangular ROI whose length and height are automatically scaled according to the local limb dimensions, which is rotated to follow the orientation of the limb being processed. Figure 9 illustrates the process of extracting six rectangular ROIs on a sample thermal image.

3.4. Feet Segmentation and Extraction Algorithm. The temperatures in the feet are generally lower than the temperatures in the hands and the shins, and are usually close to that of the background temperature. Due to this lower temperature in the feet, there is a similarity in intensity between foreground and background regions in the thermal image making the background-foreground segmentation very challenging. On the contrary, the corresponding visual images offer a significant and consistent difference between the feet and the background and can be accurately segmented using Otsu's thresholding method. Therefore, the binarised visual image was used to segment the thermal image.

The two images first had to be aligned. Since the visual and thermal images were captured almost simultaneously and close to perpendicular to the surface of interest, the difference between the two images mainly consists of

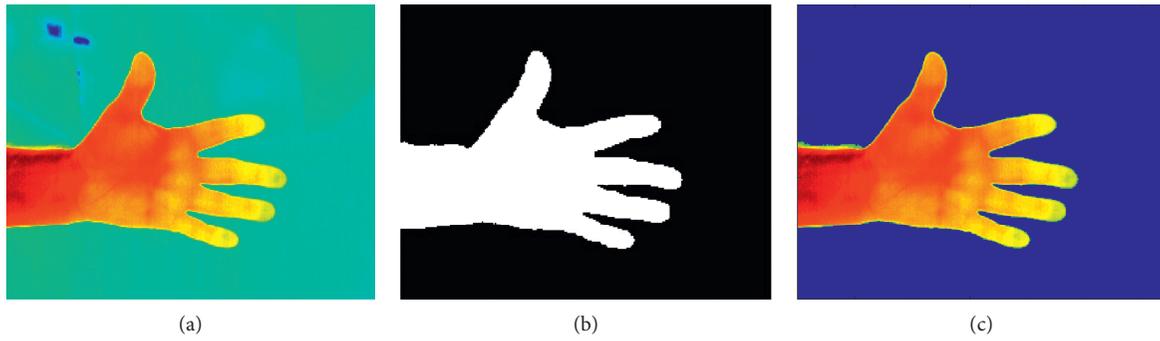


FIGURE 3: Segmentation of a sample hand thermal image using local adaptive thresholding. (a) Original thermal image. (b) Binary mask. (c) Segmented image.

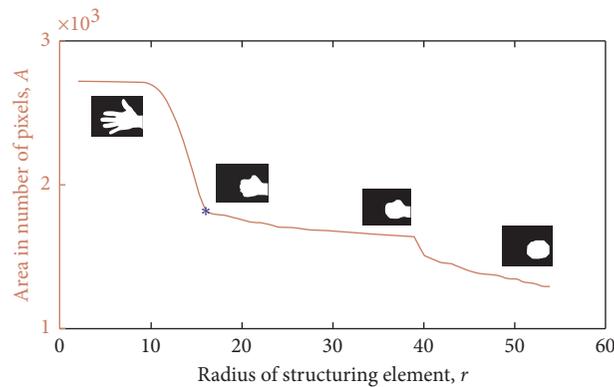


FIGURE 4: The number of remaining foreground pixels A , at each iteration with radius r , together with the resulting image, and the detected radius at which the finger removal process ends (blue).



FIGURE 5: Identification of the finger pixels. (a) Detected finger pixels (green). (b) Fingers binary image.

rotation and scaling, with possibly some skewness. Thus, an affine transform is sufficient to align the two images. The visual image was aligned to the thermal image using an affine transformation based on, at least, three corresponding points on the visual and thermal images which were manually selected by the user. Figure 10 shows the process of aligning a visual image to the corresponding thermal image. Figure 11 illustrates the segmentation of a sample foot thermal image.

The segmented foot thermal image was rotated such that the major axis of the foot, determined by eigenvalue decomposition, was vertically aligned. The aligned and segmented visual image was once again used in the extraction of the five ROIs on the toes. The edges around and between the toes were stronger and more consistent in the visual images than in the thermal images allowing for more reliable region extraction. A Laplacian of the Gaussian filter was used to extract the edges of the toes and nearby regions, in the upper



FIGURE 6: Palm ROIs extraction process. (a) Template used to extract the palm ROIs with three anchor points circled in red. (b) Detection of point corresponding to the anchor point on the thumb image. The blue intersection point is selected.

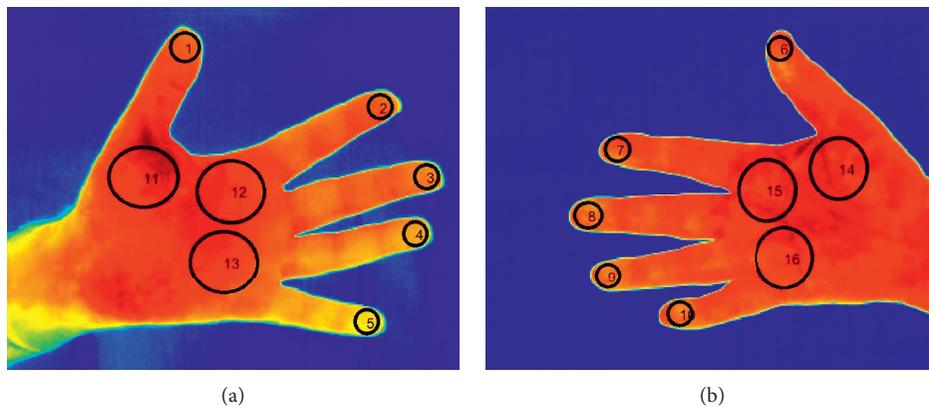


FIGURE 7: Identification of the ROIs on the hand.

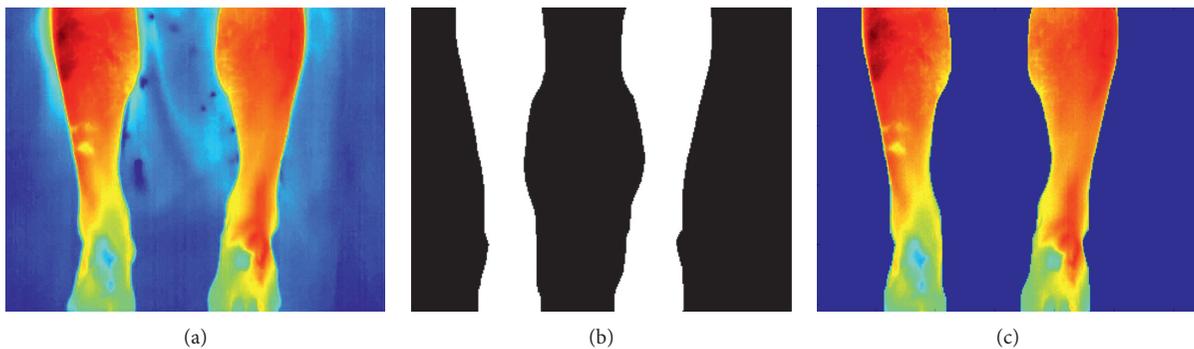


FIGURE 8: Segmentation of a shin thermal image using local adaptive thresholding. (a) Original thermal image. (b) Binarised thermal image. (c) Segmented thermal image.

half of the vertically aligned foot. The detected edges were binarised, skeletonised, and the outer contours of the foot were removed as shown in Figure 12(a). The resulting edge map contains both the edges around and between the toes, and spurious, unwanted edges arise due to illumination changes and shadows on the foot in the visual image. In contrast to the edges around the toes, the unwanted edges are characterised by a lower intensity and a relatively smaller size. Consequently, these edges were removed by

thresholding on the edge size and corresponding mean intensity, thereby retaining only the edges around and between the toes, as shown in Figure 12(b).

The outer contours are next restored, and the resulting edge map is used to form closed contours around the toes. Edgels whose connectivity with their edge contours is through only one side are considered as endpoint edgels. These endpoint edgels were connected to the closest edgel in the edge map which is not on the same contour as the

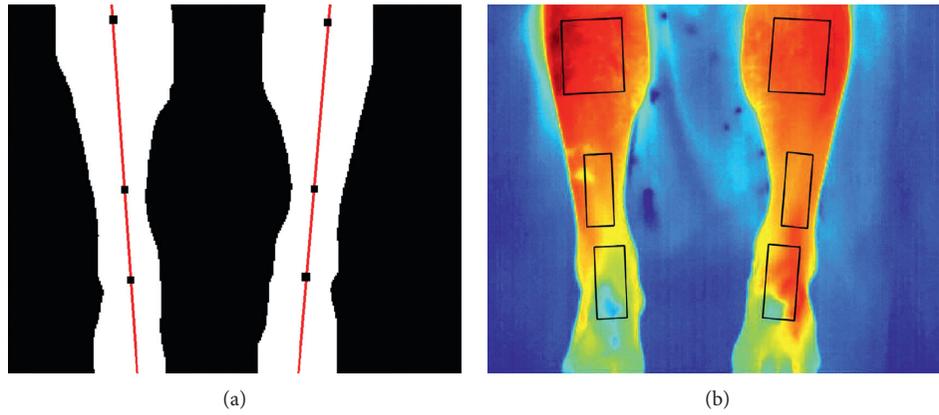


FIGURE 9: Shin ROIs extraction. (a) Extracted medial line (red) and identified centre points. (b) Identified and rotated rectangular ROIs.

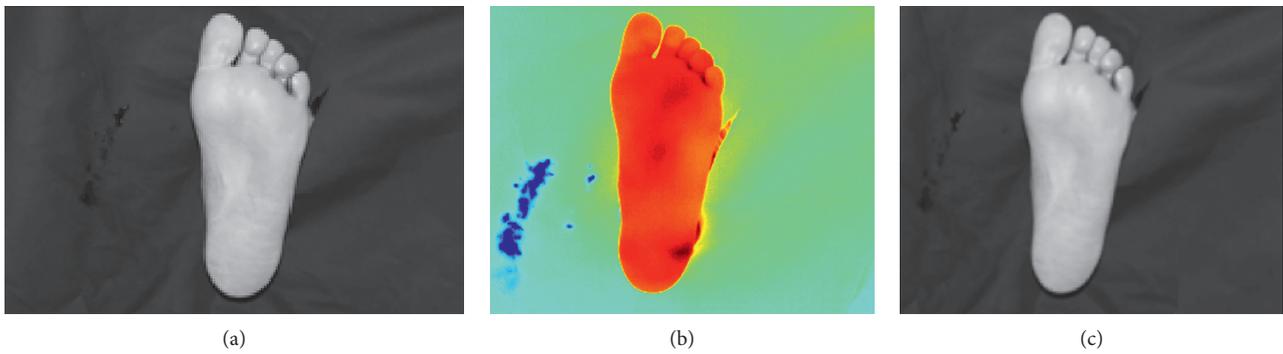


FIGURE 10: Alignment of the visual image to the thermal image. (a) Original visual image. (b) Original thermal image. (c) Aligned visual image.



FIGURE 11: Segmentation of a thermal image of the foot. (a) Aligned visual image. (b) Segmented visual image. (c) Segmented thermal image.

endpoint to form new contours as shown in Figure 13. The five largest closed regions detected in this manner are assumed to correspond to the toes, as seen in Figure 14(b). A morphological opening operator, using a disc-shaped structuring element, is applied to the resulting binary image to extract the remaining regions. Figure 14 shows the results of this algorithm.

To extract the circular regions on the ball of the foot (BOF), corresponding to regions 27 and 33 in Figure 2, an iterative process is applied which starts at the bottom row of

the vertically aligned foot and moves up, towards the toes, one row at a time. In each iteration, a circle, C , is defined as

$$C = (c(W_r), r, \beta W_r), \quad (2)$$

where r is the y coordinate of the circle centre, set to the current iteration row; $c(W_r)$ is the x coordinate of the circle centre, set to the midpoint of the horizontal width of the foot at row r , W_r ; and βW_r is the circle diameter where β is defined according to the desired size of the circular ROI according to

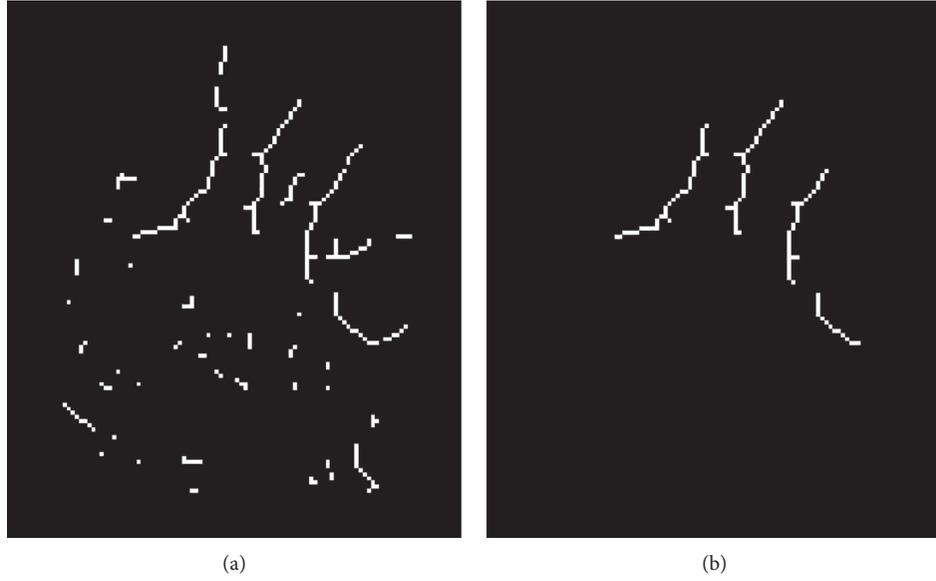


FIGURE 12: Extraction of edges close to the toes regions (a) and removal of unwanted edges (b). Remaining edges correspond to edge pixels in the regions between the toes.

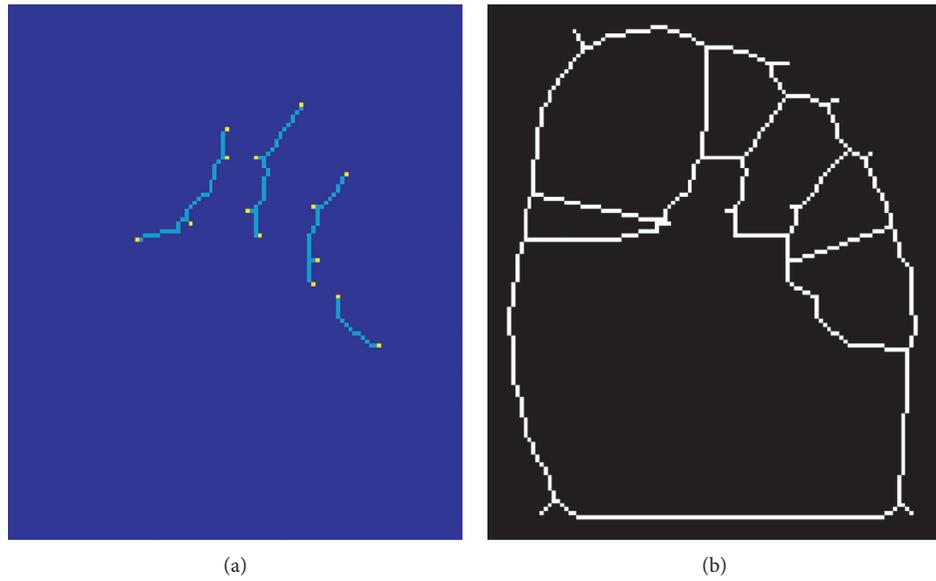


FIGURE 13: Detection of edge endpoint and connection of edges to form closed loops around the toes. (a) Detected edge endpoints. (b) Extended edges.

Figure 2. The iterative process terminates at row r^* , when the Euclidean distance $d(\cdot)$ between the centre of circle C , denoted by C_{xy} , and the middle row of the toe ROIs, R , satisfies

$$d(C_{xy}, R) = \tau, \quad (3)$$

where τ is set according to the position of the desired ROIs according to Figure 2. Once the candidate row, r^* , is chosen, two quadrilateral regions, regions 28–29 and 34–35 in Figure 2, are defined. Specifically, their centre points are set to

$$\begin{aligned} Q1_c &= \left(C_x - \frac{\beta W_{r^*}}{2} - \frac{W_{\text{rect}_1}}{4}, r^* \right), \\ Q2_c &= \left(C_x + \frac{\beta W_{r^*}}{2} + \frac{W_{\text{rect}_2}}{4}, r^* \right), \end{aligned} \quad (4)$$

where $C_x = c(W_{r^*})$ is the x coordinate of the circle centre and $W_{\text{rect}_1} = W_{\text{rect}_2} = \alpha W_{r^*}$ is the width of the quadrilaterals. α is set according to the desired size of the ROIs, as shown in

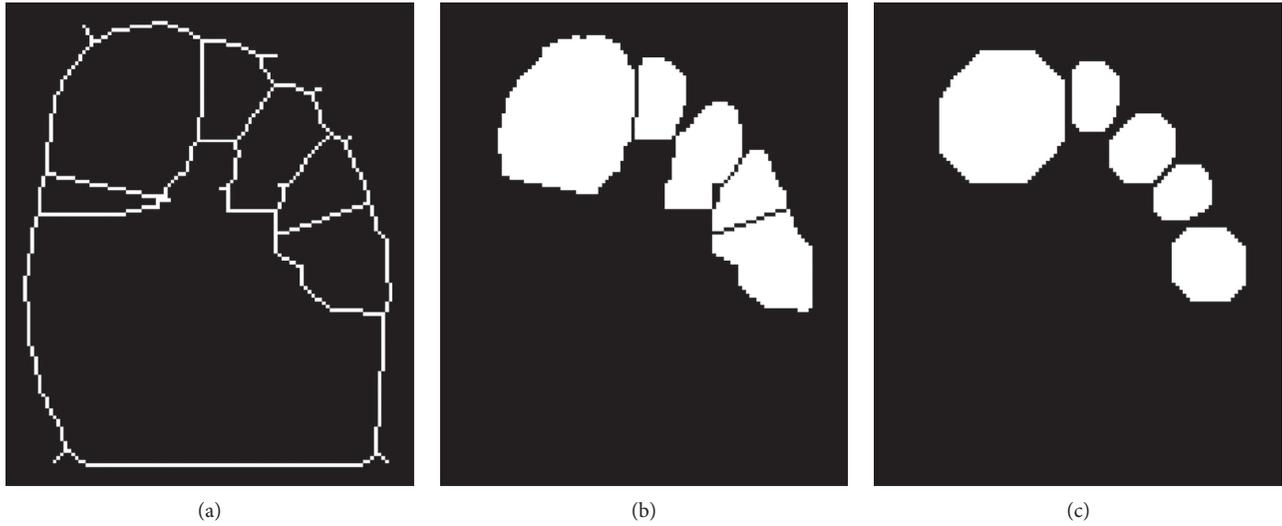


FIGURE 14: Regions completely contained within the extended edges are detected and shaped using morphological opening. (a) Extended edges. (b) Detected toes regions. (c) Disc-shaped regions.

Figure 2. Figure 15 summarizes the placement procedure of the quadrilateral ROIs on the ball of the foot.

If there is any overlap between the extracted ROIs and the toe ROIs, vertices A and B in Figure 15 are moved horizontally downwards. If there is an overlap between the extracted ROIs and the background, vertices A and D are moved diagonally towards the quadrilateral centre. Since these ROIs are on the ball of the foot which typically has a convex shape, a diagonal movement is bound to remove the overlap whilst retaining the quadrilateral shape.

To extract the three ROIs on the heel of the foot, corresponding to regions 30–32 and 36–38 in Figure 2, a similar procedure is carried out. The only difference to the extraction of the ROIs on the ball of the foot is the placement of the circle centre. To find the centre of the circle on the heel of the foot, the circle Hough transform is applied to the edges of the thermal image of the foot. Circles which have a diameter comparable to the width of the foot in the image are considered. The circle with the centre at the lowest position corresponds to the heel. This row position is also used to place the heel ROIs, and the same process used for the regions on the ball of the foot is repeated. Figure 16 shows the results of extracting the 11 ROIs on feet thermal images.

The methods presented in this paper are publicly available, in the form of MATLAB functions, together with some sample thermal images and may be accessed from <http://www.um.edu.mt/cbc/tipmid>.

4. Results

4.1. Segmentation and ROIs Extraction. The algorithms were applied to the acquired thermal images, in a pipeline, one for each body region, and at every stage of the pipeline, a visual inspection was carried out to verify the integrity of the segmentation and the correct placement and sizing of the ROI labels. Based on this visual inspection, it was possible to take a decision to determine whether an ROI was extracted

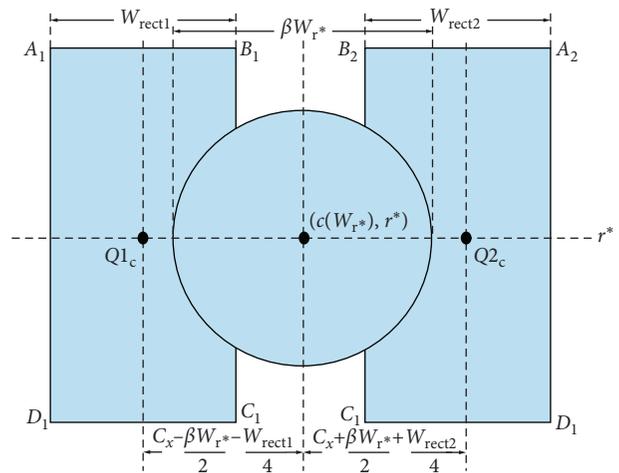


FIGURE 15: Positioning of the circular and quadrilateral ROIs on the ball of the foot and the heel.

correctly or otherwise. The decision logic used for this purpose is as follows:

- (i) A background-foreground segmentation was considered correct if it retained the integrity of the shape of the body region without any missing or additional features in the segmented image
- (ii) An extracted ROI was considered correct if it followed the correct placement and size with respect to the dimensions of the body region as defined in the template shown in Figure 2 and which does not overlap with the background regions or any other ROIs on the same image

One feature to be noted in the pipeline is that an error in one of the stages of the pipeline does not necessarily imply that the following algorithm will also fail.

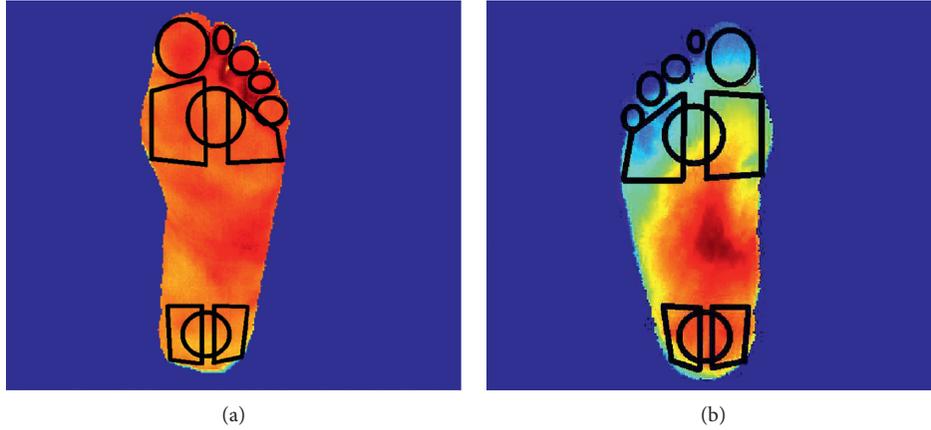


FIGURE 16: Identification of the ROIs on the feet.

The success rate was determined by obtaining the number of correct and incorrect ROI labels. In this study, a true positive (TP) would be a correctly localized ROI which is assigned its respective, distinct label; a false positive (FP) is an incorrectly localized ROI. Since, every image contains a specific number of ROIs and all ROI labels for an image are being assigned, negatives are undefined. Specifically, hand images contain eight ROIs, shin images contain six ROIs, and feet images have 11 ROIs, with each ROI having a specific ROI label.

As shown in Table 1, out of 155 hand images, 97.9% of the finger ROIs and 94% of the palm ROIs were correctly localized and labelled. The main source of failure was the background-foreground segmentation, which occurred in two hand images, and the misidentification of the fingers from the rest of the hand, which occurred in 11 other images.

As shown in Table 2, 99% of the shin ROIs were correctly localized and labelled. The main source of failure was the background-foreground segmentation which failed in three images due to a close similarity in temperatures between the foreground and background temperatures in the lower parts of the shins.

As shown in Table 3, 77.5% of the toe ROIs, 73.7% of the ball-of-the-foot ROIs, and 79.5% of the heel ROIs were correctly localized and labelled. The main sources of errors were the segmentation of the visual image, which failed in four images, and the misalignment of the thermal and visual images, which failed in eight images.

4.2. Method Validation via Interrater Reliability. To validate the segmentation results, an interrater reliability test was carried out. A set of 60 correctly segmented thermal images consisting of 20 hand images, 20 shin images, and 20 foot images were selected. Four human raters were asked to manually demarcate the ROIs using FLIR Altair area demarcation tool. The automated ROIs extraction method developed herein was used as the fifth rater in this study. The objective of this interrater analysis was to establish whether the results from the ROI extraction methods are comparable to the results obtained from

TABLE 1: Success rate of hand ROIs extraction.

Test database			
155 images	775 finger ROIs	465 palm ROIs	1240 total ROIs
Correct	759 (97.9%)	437 (94%)	1196 (96.5%)
Incorrect	16 (2.1%)	28 (6%)	44 (3.5%)

TABLE 2: Success rate of shin ROIs extraction.

Test database	
134 images	804 shin ROIs
Correct	796 (99%)
Incorrect	8 (1%)

TABLE 3: Success rate of foot ROIs extraction.

Test database				
171 images	855 toe ROIs	513 BOF ROIs	513 heel ROIs	1881 total ROIs
Correct	663 (77.5%)	378 (73.7%)	408 (79.5%)	1449 (77%)
Incorrect	192 (22.5%)	135 (26.3%)	105 (20.5%)	432 (23%)

human raters. The latter currently constitute the gold standard for region extraction. A spatial overlap coefficient, the Dice similarity coefficient (DSC) [25], was used to determine the similarity between the extracted regions. The DSC is defined as

$$DSC(A, B) = \frac{2|A \cap B|}{|A| + |B|}, \quad (5)$$

where A and B are the two target regions. The 44 ROIs were categorised into six groups as follows: fingers, palms, shins, toes, ball of the foot (BOF), and heel ROIs. The Dice coefficient was computed between corresponding regions between pairs of human raters and between the algorithm and the human raters. For each ROI group, there are six possible pairwise combinations between human raters, and four combinations between human raters and algorithm. The mean and standard deviation (SD) of the similarity

coefficients for each group of ROIs was extracted for human rater pairs and for algorithm-human rater pairs as shown in Table 4.

Referring to Table 4, it is noted that, for all ROI groups, except for the toes, the mean algorithm-human rater coefficient is similar to the mean human rater coefficient. This implies that the algorithm may equivalently replace a human rater.

For the toes group, a mean human rater coefficient of 0.621 (SD 0.182) and an algorithm-human rater coefficient of 0.427 (SD 0.170) were achieved, suggesting that the algorithm extracts ROIs that are substantially different than those extracted by humans. Through visual inspection of the ROIs was extracted by the algorithm, it was determined that the ROIs correctly cover the whole toe area, whereas the human rater ROIs are limited to circular ROIs, typically centred on the toes. Despite the discrepancy between the humans and the algorithm, it is noted that the algorithm ROIs are still valid and correct as shown in Figure 17.

Since the purpose of the ROI extraction is to measure the ROI temperature from the radiometric data, the mean temperature of the ROIs extracted by the algorithm was compared to that extracted by the human raters. Table 5 shows the mean difference between the temperatures extracted by humans and by the algorithm. Since in the diabetic foot, abnormalities are associated with temperature differences of 2.2°C or greater, we considered a difference in temperature between the two sets of extracted ROIs of quarter of the 2.2°C reference as being negligible. A sign test confirmed that there was no statistically significant difference greater than 0.55°C at an alpha value of 0.05 between the mean temperature extracted by the algorithm and that obtained by the human raters, as shown in Table 5. This result demonstrates that the mean temperature extracted by the algorithm is, in substance, the same as that extracted by the human raters.

5. Discussion

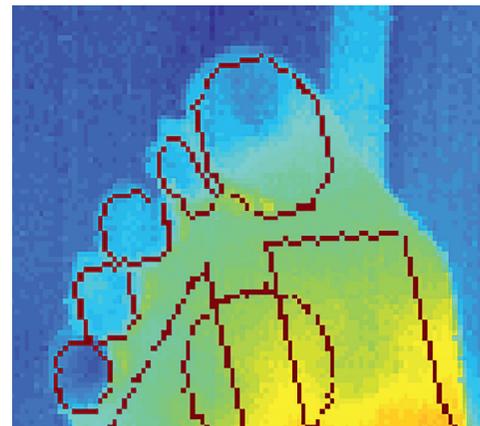
The implemented algorithms were found to be effective for a large proportion of the thermal images on which they were tested. From Tables 1–3, it is noted that 484 ROIs required a manual intervention from a total of 3925 ROIs, which means an intervention in 12.3% of the ROIs. Therefore, 87.7% of the ROIs were extracted without any need for a manual intervention. The extraction of the foot ROIs was the most challenging out of the three algorithms due to the highly overlapping temperature ranges of the feet and the background in the captured thermal images.

Failed attempts at extracting the ROIs can be split in two categories: missed ROIs, in which the desired ROI has not been detected and there is no information about the temperature in the region, and inaccurate ROIs, in which the ROI is offset from the desired location.

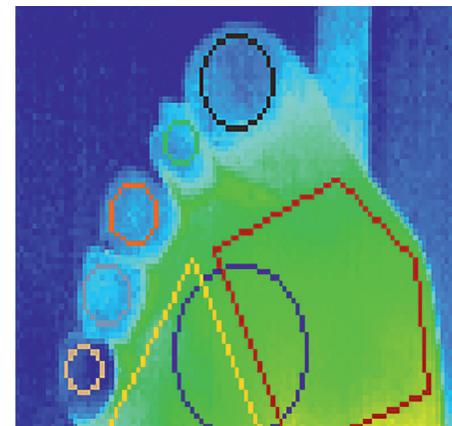
When executed using MATLAB on a system with a 2.7GHz dual core processor and 8 GB of RAM, the algorithms required 1.4 seconds to process thermal images of the hands, less than 0.1 seconds for thermal images of the

TABLE 4: Dice similarity coefficients mean and standard deviation.

ROI group	Human rater pairs		Algorithm-human pairs	
	Size	Mean (SD)	Size	Mean (SD)
Fingers	600	0.756 (0.124)	400	0.688 (0.142)
Palms	360	0.698 (0.175)	240	0.691 (0.171)
Shins	720	0.552 (0.237)	480	0.512 (0.242)
Toes	600	0.621 (0.182)	400	0.427 (0.170)
BOFs	360	0.676 (0.192)	240	0.700 (0.148)
Heels	360	0.721 (0.134)	240	0.702 (0.128)



(a)



(b)

FIGURE 17: Comparison of automatic ROIs (a) and manual ROIs (b) extracted on the toes for the same image. Although the agreement between these regions is poor, the automatically extracted ROIs are more desirable.

TABLE 5: Mean temperature differences and statistical test results.

ROI group	Algorithm-human differences		ρ value
	Mean	(SD)	
Fingers	0.077	(0.079)	<0.02
Palms	0.159	(0.161)	0.023
Shins	0.364	(0.442)	0.08
Toes	0.145	(0.139)	0.09
BOF	0.132	(0.148)	0.007
Heels	0.117	(0.111)	0.007

shins and 2.2 seconds for thermal images of the feet. Since for each subject, two hand thermal images, one thermal image of the shins, and two feet thermal images were acquired, and the algorithms required 7.3 seconds to process the five thermal images. The algorithm which processes thermal images of the feet also required a manual user intervention to register the thermal and visual images, which typically could be carried out in less than a minute. Therefore, the whole process of extracting the 44 desired ROIs from the five thermal images took around one minute to be completed. A corresponding process of extracting these regions manually, using an area demarcation tool, took an experienced user around 10 minutes. Therefore besides automating the area extraction process, significant reduction in the time required for the operation is achieved.

Interrater reliability tests showed that the ROIs extracted by the algorithm are within tolerance of the ROIs extracted by humans. This shows that the algorithms presented in this work can be used to extract temperature data with similar reliability and accuracy as humans.

6. Conclusion

Several medical applications of thermography require the extraction of temperature data from specific anatomical areas. Most literature in the field requires the clinicians to manually extract these data using area demarcation tools. This operation is subjective, may lead to nonrepeatable results, and is also very lengthy, especially for a large number of images and ROIs. Furthermore, the temperatures considered for such studies are averaged over large regions. This work proposes a set of algorithms to automate temperature extraction from local regions of interest. A success rate of around 90% was obtained for each algorithm, with minimal user intervention. In addition to the automation of the data extraction process, the algorithms also provide a significant reduction in the time required for the operation. The context for the algorithms is the detection of early signs of complications in the diabetic foot but can be extended to any application which requires temperature data from the same regions in the hands, shins, or feet. Future work should aim to enhance the algorithm to reduce the number of missed ROIs and to reduce the manual intervention required in the ROI extraction process.

Data Availability

The methods presented in this paper have been made available online, in the form of MATLAB functions, together with some sample thermal images and may be accessed from <http://www.um.edu.mt/cbc/tipmid>.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this article.

Acknowledgments

Research was supported by the Malta Council for Science and Technology through the National Research and Innovation Programme 2013 (Grant number R&I-2013-028).

References

- [1] B. B. Lahiri, S. Bagavathiappan, T. Jayakumar, and J. Philip, "Medical applications of infrared thermography: a review," *Infrared Physics and Technology*, vol. 55, no. 4, pp. 221–235, 2012.
- [2] A. Gatt, C. Formosa, K. Cassar et al., "Thermographic patterns of the upper and lower limbs: baseline data," *International Journal of Vascular Medicine*, vol. 2015, Article ID 831369, 9 pages, 2015.
- [3] K. Ouriel, "Peripheral arterial disease," *Lancet*, vol. 358, no. 9289, pp. 1257–1264, 2001.
- [4] P. Palumbo and L. J. Melton, "Peripheral vascular disease and diabetes," in *Diabetes in America*, R. Aubert, Ed., pp. 401–408, Diane Publishing Company, Collingdale, PA, USA, 2nd edition, 1995.
- [5] B. F. Jones and P. Plassmann, "Digital infrared thermal imaging of human skin," *IEEE Engineering in Medicine and Biology Magazine*, vol. 21, no. 6, pp. 41–48, 2002.
- [6] J. Steketee, "Spectral emissivity of skin and pericardium," *Physics in Medicine and Biology*, vol. 18, no. 5, pp. 686–694, 1973.
- [7] B. F. Jones, "A reappraisal of the use of infrared thermal image analysis in medicine," *IEEE Transactions on Medical Imaging*, vol. 17, no. 6, pp. 1019–1027, 1998.
- [8] E. F. J. Ring and K. Ammer, "Infrared thermal imaging in medicine," *Physiological Measurement*, vol. 33, no. 3, pp. 33–46, 2012.
- [9] M. Etehadtavakol and E. Y. Ng, "Assessment of foot complications in diabetic patients using thermography: a review," in *Application of Infrared to Biomedical Sciences*, E. Y. K. Ng and M. Etehadtavakol, Eds., pp. 33–43, Springer, Berlin, Germany, 2017.
- [10] D. G. Armstrong, L. A. Lavery, P. J. Liswood, W. F. Todd, and J. A. Tredwell, "Infrared dermal thermometry for the high-risk diabetic foot," *Physical Therapy*, vol. 77, no. 2, pp. 169–175, 1997.
- [11] S. J. Benbow, A. W. Chan, D. R. Bowsher, G. Williams, and I. A. Macfarlane, "The prediction of diabetic neuropathic plantar foot ulceration by liquid-crystal contact thermography," *Diabetes Care*, vol. 17, no. 8, pp. 835–839, 1994.
- [12] M. Oe, R. R. Yotsu, H. Sanada, T. Nagase, and T. Tamaki, "Thermographic findings in a case of type 2 diabetes with foot ulcer and osteomyelitis," *Journal of Wound Care*, vol. 21, no. 6, pp. 274–278, 2012.
- [13] T. Tamaki, "Screening for osteomyelitis using thermography in patients with diabetic foot," *Ulcers*, vol. 2013, Article ID 284294, 6 pages, 2013.
- [14] T. Nagase, H. Sanada, K. Takehara et al., "Variations of plantar thermographic patterns in normal controls and non-ulcer diabetic patients: novel classification using angiosome concept," *Journal of Plastic, Reconstructive and Aesthetic Surgery*, vol. 64, no. 7, pp. 860–866, 2011.
- [15] C. Liu, J. J. van Netten, J. G. van Baal, S. A. Bus, and F. van der Heijden, "Automatic detection of diabetic foot complications with infrared thermography by asymmetric analysis," *Journal of Biomedical Optics*, vol. 20, no. 2, article 026003, 2015.
- [16] N. Kaabouch, Y. Chen, W.-C. Hu, J. W. Anderson, F. Ames, and R. Paulson, "Enhancement of the asymmetry-based overlapping analysis through features extraction," *Journal of Electronic Imaging*, vol. 20, no. 1, article 013012, 2011.
- [17] N. Kaabouch, W.-C. Hu, and Y. Chen, "Alternative technique to asymmetry analysis-based overlapping for foot ulcer

- examination: scalable scanning,” *Journal of Diabetes and Metabolism*, vol. 1, no. 5, pp. 1–6, 2016.
- [18] S. Araki, H. Nomura, and N. Wakami, “Segmentation of thermal images using the fuzzy c-means algorithm,” in *Proceedings of Second IEEE International Conference on Fuzzy Systems*, pp. 719–724, San Francisco, CA, USA, December 1993.
- [19] T.-H. Yoon, K.-S. Kim, J.-W. Lee, K. Dong-Jun, and S. Chul-Gyu, “Image segmentation of human forearms in infrared image,” in *Proceedings of International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 2762–2765, New York City, NY, USA, August 2006.
- [20] C. L. Herry, “Segmentation and extraction of regions of interest for automated detection of anomalies in clinical thermal infrared images,” Ph.D. thesis, Carleton University Ottawa, Ottawa, ON, Canada, 2008.
- [21] M. Blank and C. Kargel, “Infrared imaging to measure temperature changes of the extremities caused by cigarette smoke and nicotine gums,” in *Proceedings of IEEE Instrumentation and Measurement Technology Conference*, pp. 794–799, Sorrento, Italy, April 2006.
- [22] J. Gauci, O. Fazlon, K. P. Camilleri et al., “Automated segmentation of regions of interest from thermal images of hands,” in *Proceedings of 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 3822–3826, Jeju Island, South Korea, July 2017.
- [23] E. Ring and K. Ammer, “The technique of infrared imaging in medicine,” *Thermology International*, vol. 10, no. 1, pp. 7–14, 2000.
- [24] D. Cardone and A. Merla, “New frontiers for applications of thermal infrared imaging devices: computational psychophysiology in the neurosciences,” *Sensors*, vol. 17, no. 5, p. 1042, 2017.
- [25] L. R. Dice, “Measures of the amount of ecologic association between species,” *Ecology*, vol. 26, no. 3, pp. 297–302, 1945.

Research Article

Is Intensity Inhomogeneity Correction Useful for Classification of Breast Cancer in Sonograms Using Deep Neural Network?

Chia-Yen Lee ¹, Guan-Lin Chen,¹ Zhong-Xuan Zhang,¹ Yi-Hong Chou,²
and Chih-Chung Hsu ³

¹Department of Electrical Engineering, National United University, Miaoli, Taiwan

²Department of Radiology, Taipei Veterans General Hospital and National Yang Ming University, Taipei, Taiwan

³Department of Management Information Systems, National Pingtung University of Science and Technology, Neipu, Taiwan

Correspondence should be addressed to Chih-Chung Hsu; cchsu@mail.npust.edu.tw

Received 29 June 2018; Revised 25 October 2018; Accepted 18 November 2018; Published 4 December 2018

Guest Editor: Cesare Valenti

Copyright © 2018 Chia-Yen Lee et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The sonogram is currently an effective cancer screening and diagnosis way due to the convenience and harmlessness in humans. Traditionally, lesion boundary segmentation is first adopted and then classification is conducted, to reach the judgment of benign or malignant tumor. In addition, sonograms often contain much speckle noise and intensity inhomogeneity. This study proposes a novel benign or malignant tumor classification system, which comprises intensity inhomogeneity correction and stacked denoising autoencoder (SDAE), and it is suitable for small-size dataset. A classifier is established by extracting features in the multilayer training of SDAE; automatic analysis of imaging features by the deep learning algorithm is applied on image classification, thus allowing the system to have high efficiency and robust distinguishing. In this study, two kinds of dataset (private data and public data) are used for deep learning models training. For each dataset, two groups of test images are compared: the original images and the images after intensity inhomogeneity correction, respectively. The results show that when deep learning algorithm is applied on the sonograms after intensity inhomogeneity correction, there is a significant increase of the tumor distinguishing accuracy. This study demonstrated that it is important to use preprocessing to highlight the image features and further give these features for deep learning models. In this way, the classification accuracy will be better to just use the original images for deep learning.

1. Introduction

In recent years, with dietary habit and the change of lifestyle, a quick rise has been seen in female's breast cancer. In the developed and developing countries, breast cancer has, respectively, become the first and second causes of women's cancer death. According to the article published in International Journal of Cancer, 1993, one in every 8 women in European and American countries has breast cancer [1]. American Cancer Society pointed out as early as in 1999 that breast cancer is the first cause of women's death in the developed countries [2]. In the various breast cancer diagnosis procedures, ultrasound diagnosis is regarded as a kind of "highly accepted" tool because of low price, convenience, universality, and nonradiation. Ultrasound has

become a necessary diagnosis tool in all medical centers and a good tool for doctors to initially diagnose the breast cancer clinically.

However, ultrasonic images often contain lots of speckles, noises, and textures so that it is unable to clearly discriminate the detail changes of tissues, e.g., the tumor size and boundary. Therefore, many studies in the past have put forward ultrasound tumor segmentation algorithm, including the well-known deformation model [3, 4], clustering [5], overzero [6], thresholding [7], watershed method [8], and level set method [9, 10]. The above methods can be roughly divided into two categories, respectively, based on pixel or region information [11]. However, in order to make the segmentation results well [12–14], many experiment steps and parameter adjustments are involved. It may have

been affected by the developer's subjective ideas and habits so that these results may be not sufficiently objective, affecting the accuracy in doctor's judgment with naked eyes, and it is likely to cause misdiagnosis. Therefore, the traditional computer-aided diagnosis tool is limited.

Recently, deep learning has been widely used in various applications such as facial recognition [15, 16], object detection, and car identification. In medical imaging, deep learning still offers excellent performance on different fields [17, 18]. With the development of deep learning, the system architecture of computer-aided diagnosis has been changed, and the advantages unavailable in traditional computer-aided diagnosis system have been added. In the deep learning method, it is only needed to specify the training data, and then, the features in the images can be automatically extracted, and more abstract feature descriptions can be extracted according to the nerve cell depth of each layer, including the features from point to edge, contour, and even a higher level so as to gradually reduce the nerve cells at the same time of improving the features. In the deep learning method, selection and extraction of features, as well as data classification, are established under the same structure, which has a higher accuracy than traditional feature extraction.

To address the problem of the breast sonogram detection, there are several approaches that can be used such as a convolutional neural network (CNN) [19–21], deep autoencoder (DAE), or stack denoised autoencoder (SDAE). Zhou et al. demonstrated the importance of feature extraction and selection for tumor classification and confirmed that the classification effect based on the Shearlet-based texture feature was better than those of the other four feature extraction methods by AdaBoost and SVM. In the classification of 200 benign and malignant tumors, the accuracy is up to 90%, indicating the effectiveness of using machine learning for breast tumor classification [22]. In 2017, Moon added adaptive filtering to the CAD system (A-CAD) to emphasize the characteristics of tumor size. CAD was more robust when classifying tumors larger than or equal to 1 cm. The results showed that the accuracy increased from 73.1% to 81.4% after adding the adaptive filter. It is suggested that the tumor classification effect is better after using adaptive filtering [23]. Due to the speckle noise of ultrasonic images, Raha used four kinds of filters (median, high boost, Sobel, and average filter) for image preprocessing, and watershed was used for segmentation. Finally, the k-nearest neighbors algorithm was performed for the classification of benign and malignant tumors with a high accuracy of 96.4% [24]. Abdel-Nasser extracted a high-resolution (HR) image from a set of low-resolution (LR) images using super-resolution to overcome the speckle noise and artifacts problems. The 31 benign and 28 malignant tumors were divided into training and test data by leave-one-out cross validation (LOOCV) and classified by random forests [25]. The classification results with high accuracy were obtained. It was also emphasized that if the texture features were able to accurately represent the tumor information, it is very important to perform preprocessing. Cheng et al. conducted research on deep learning-based

CAD. To avoid using inaccurate image processing results to achieve tumor classification, the team used SDAE for US breast cancer, pulmonary nodules in CT scans classification. Relying on the advantages of SDAE well equipped with the automatic feature exploration mechanism and noise tolerance advantage, the classification result can reach 82.4% accuracy. Cheng et al. reported that the research is the first CAD study based on deep learning, which can further study other image features for more accurate classification results [26]. Overall, it is important to do image preprocessing to highlight the image features and further give these features for deep learning models.

In general, CNN is an end-to-end architecture and suitable for various image classification tasks. However, it is relatively hard to find the useful and meaningful features in a constrained-scale training set. In breast sonogram classification task, the training set is usually smaller than the size of the general training images. For example, the number of training images in ILSVRC12 [19] is 1.2 million, while that of our training set is 170. Instead, the autoencoder families deal with feature extraction task as a data reconstruction task. By minimizing the reconstruction error, the primary feature of the training images should be detectable. In this study, therefore, we tend to adopt the autoencoder network to extract the useful features and follow by concatenating a classifier to distinguish the type of breast sonogram. SDAE [27] is one of the state-of-the-art autoencoder architectures. Xing et al. adopted SDAE for the extraction of tumor features and classification of benign and malignant tumors [27]. SDAE is about to find out many typical patterns in the input training data, to deal with the high change problem of tumor margin, and moreover, it is provided with multilayer training automatic feature extraction and noise reduction ability so that the problems likely to occur under traditional methods may be avoided.

The accuracy of classifier is highly related to the quality of test images. As for ultrasonic image, due to the imaging principle and the properties of object to be tested, the research scholars are often encountered with two major problems, much noise and intensity inhomogeneity. Noise-resistant function has been added to SDAE when it is designed, to reduce the noise effect. In terms of intensity inhomogeneity, studies have found that after the preprocess of intensity inhomogeneity correction [28–31] of medical images, the imaging quality is effectively improved so as to reach a better image processing result. There are several intensity inhomogeneity correction methods. For example, the filter method [32, 33] is to take intensity inhomogeneity as low frequency signal and eliminate the intensity inhomogeneity with the filter method, but such a method has the elimination risk of important low-frequency information; the surface fitting method [34] is generally used to establish a curve with polynomial or spline, and the characteristics in the image are fit on such a curve to serve as image intensity inhomogeneity; the histogram method is to take the observed image as the convolution of original image and intensity inhomogeneity, and the original image can be obtained only through deconvolution; the image segmentation method [35, 36] is able to get the segmentation result

and estimated intensity inhomogeneity; or the method combining with many concepts above is adopted [37]. In this study, intensity inhomogeneity correction is conducted before the input of ultrasonic testing images. Then, the corrected images are input into the classifier system established on the basis of SDAE, hoping to overcome the effect of intensity inhomogeneity in ultrasonic imaging for improving the system resolution and quality.

Main aims of the paper are as follows:

- (i) A classification algorithm in breast sonogram without subjective influences is proposed
- (ii) Up to 85% accuracy can be obtained when deep learning was applied to the intensity inhomogeneity-corrected image
- (iii) The study overcomes drawbacks of sonogram and has reliable discrimination ability
- (iv) It has the potential to obtain the better classification results after image preprocessing

2. Materials and Methods

2.1. Input Data. This study was approved by the Institutional Review Boards of the Taipei Veterans General Hospital. All the experimental methods were carried out in accordance with the approved guidelines. In this study, we used two kinds of dataset. One is the data collected by our team (a database of 96 malignant and 74 benign images) and the other is the public dataset on the website, Rodrigues, Paulo Sergio (2017), "Breast Ultrasound Image," Mendeley Data, v1 (a database of 150 malignant and 100 benign images) [38].

2.2. SDAE. Artificial neural network is a kind of machine learning concept to simulate the learning of human's brain. As for each nerve cell, the nerve cell signals connected to it will influence the output of such nerve cell. The transmission of nerve signals is to add the received signals and then conduct nonlinear transformation, to get new output. The mutual connection in neural network decides the calculation method to make the actual output approach the expected output as much as possible through neural network adjustment.

The training of one neural network is to limit the output value to be equal to the input value, and this indicates that the output layer has the same quantity of nerve elements as the input layer, and the error between the two layers is used to adjust the weight of each layer. The training of autoencoder (AE) is unsupervised, so no label information is needed. Denoising autoencoder (DAE) is improved on the basis of AE. It is assumed that the input data include noise, so DAE is suitable to learn the features from the data including noise. SDAE is stacked by DAE, to obtain higher level features. The network training of SDAE is layer-wise, for it is of independent training between each DAE. As for the SDAE network after training, the decoding layer is eliminated, while the encoding layer generating features is reserved. To classify the data, logistic regression (LR) layer is added to be taken as output layer, and LR is of supervised

type, adding the expected output volume label information through backpropagation algorithm, and based on the error between actual output and expected output, the network weight between layers is fine-tuned. Therefore, the feature learning of system is the result of combination between SDAE pretraining and LR adjustment.

It is supposed that the input data are x and DAE anti-noise method is used to firstly set some input as 0 or add Gaussian noise to generate \bar{x} as shown in Figure 1. The way of thinking is to make the input not perfect in the very beginning so as to get better features after training. Therefore, a good result still can be achieved after the input is polluted. DAE encode is obtained according to one non-linear transformation equation:

$$Z = f_{\theta}(w\bar{x} + b), \quad (1)$$

where Z is the output of the hidden layer, called as feature description, w is the weight of input and hidden layers, b is the offset, and f_{θ} is the activation function of the hidden layer. DAE decoding and reconstruction are realized through the mapping function:

$$x' = g_{\theta'}(w'y + b'), \quad (2)$$

where x' denotes the transpose operation, x the output of DAE (the reconstruction), and w the corresponding weight. DAE may generate a depth network with many hidden layers through stacking. A standard SDAE structure includes two encoding layers and two decoding layers. In the encoding stage, the output of the first encoding layer is taken as the input of the second encoding layer. It is supposed that there are L hidden layers in the encoding stage, and we can obtain the activation function of number k encoding layer as follows:

$$Z^{(k+1)} = f_{\theta}(w^{k+1}Z^k + b^{k+1}), \quad k = 0, \dots, L-1, \quad (3)$$

where Z^0 is the original input and Z^L is the feature description of the highest layer. The decoding stage is the same as the encoding stage. The output of the first decoding layer is taken as the input of the second decoding layer, and in this way, the activation function of number k encoding layer can be obtained as follows:

$$x'^{(k+1)} = g_{\theta'}\left(w^{(L-k)T}x'^k + b'^{(k+1)}\right), \quad k = 0, \dots, L-1, \quad (4)$$

where x'^0 is the output of the last layer in the encoding stage and x'^L is the reconstruction of x . The weight of SDAE after pretraining will be the initial weight of adjustment and classification stages. Once the import and discriminative features have learned by SDAE, the next step is to find a suitable classifier to correctly detect the type of the extracted feature of the input image. Since there are many different classifiers such as linear classifier, perceptron, SVM, random forest, and so on, it is necessary to select one of them for the best performance purpose. Most of them are different from SDAE, while the perceptron is also based on the neural network. Inspired by perceptron learning, we

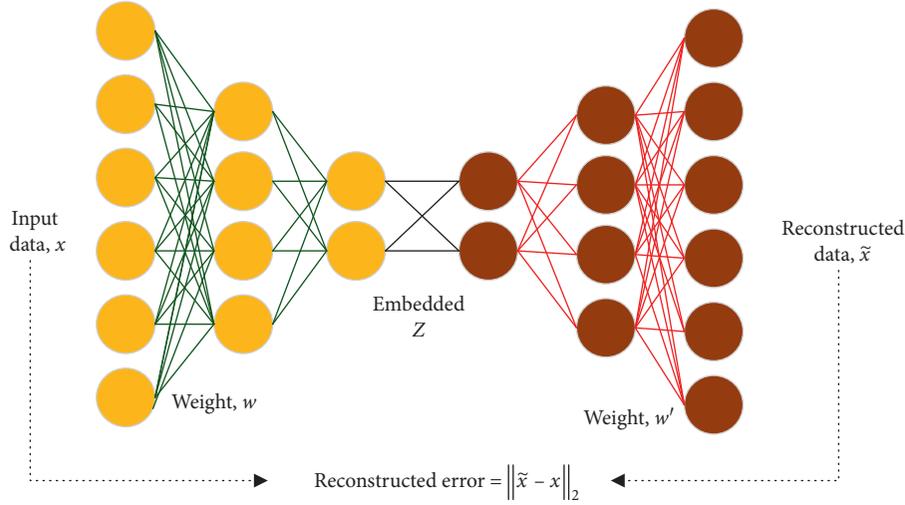


FIGURE 1: The flowchart of the reconstructed error of the DAE used in this paper.

concatenate a sigmoid layer to SDAE and formulate it as a cross-entropy loss function to keep it as an end-to-end architecture.

In the LR stage, the expected output is added, and the weight after pretraining is fine-tuned through the supervised learning method, and the activation function of LR layer is S (sigmoid):

$$S(t) = \frac{1}{1 + \exp(-t)}, \quad (5)$$

where t is the output Z^L of the final encoding layer and also the depth feature of pretraining of the SDAE method, and the output of S is the classification result, with the value ranging from 0 to 1.

The backpropagation algorithm is one kind of gradient descent methods. After the gradient direction is worked out, the weight of the classifier system will advance towards the direction with the quickest gradient descent. In addition, it is a kind of the greedy method, for it always advances towards the steepest direction, seeking for the biggest decent extent.

To realize the slight adjustment of weight, the error calculation method shall be defined. Different error calculation methods have different weight upgrading rules. Here, the common square error will be adopted:

$$E = \frac{1}{2} \sum_{d \in D} (t - y)^2, \quad (6)$$

where D indicates all input data, d indicates one of the input data, t is the expected output, and y is the actual output.

The target is to find out a set of weights, to realize the smallest calculation error:

$$\begin{aligned} w &= w + \Delta w, \\ \Delta w &= -\alpha \frac{\partial E}{\partial w}, \end{aligned} \quad (7)$$

where α indicates the learning rate. The weight upgraded mode is the batch mode. After all inputs are executed, the connecting weight will be changed.

2.3. Intensity Inhomogeneity Correction. The algorithm was composed of two parts, constrained fuzzy cell-based bipartition and the intensity inhomogeneity modelling [39]. When correcting the intensity inhomogeneity, the polynomial surface fitting method is adopted to estimate the image intensity inhomogeneity. Curve fitting is a kind of concept to represent the existing data through the mathematical method. Fitting is to obtain the discrete information through sampling and testing in engineering. Based on the data obtained in this way, it is hoped to obtain a continuous function (polynomial) or the gathering (spline) of many discrete equations to be identical to the data. Surface fitting is the popularization of curve fitting. It is supposed that the region of interest (ROI) image is composed of foreground (F) and background (B), and each area is regarded as the homogeneous area.

The observed gray level of image can be expressed as follows:

$$O_i = \mu_\psi + P_i + n_i, \quad (8)$$

where O_i is the gray level of number i pixel, μ_ψ is the average value of area where the pixel is located, $\psi \in \{F, B\}$ represents the area where the pixel is located, and the latter two items indicate intensity inhomogeneity and noise. Here, intensity inhomogeneity is supposed as the normal distribution of spatial changes, which is composed of a fixed variance and the average value P changing with the space, and it is represented with polynomial surface form. P_i represents the value of polynomial in number i pixel, that is, the average value of intensity inhomogeneity of number i pixel in the image, and n_i is the composition of change and noise in the polynomial surface indicating intensity inhomogeneity.

Adopt the least square method to minimize the cost function and estimate the polynomial surface as follows:

$$\varepsilon^2 = \frac{1}{N} \sum_{\forall i} (O_i - \mu_\psi - P_i)^2. \quad (9)$$

When the pixel is in the foreground area, $\mu_\psi = \mu_F$, and when the pixel is in the background area, $\mu_\psi = \mu_B$. $P_i = P(x_i, y_i)$, in which P is N time polynomial. After the polynomial surface is obtained, the image intensity will be adjusted based on results, to complete the image correction.

The image training and classification steps are shown as follows:

- (1) Input training image and initialize the neural network weight by use of SDAE
- (2) Remove decoding part and add LR structure, to establish SDAE-LR system
- (3) Add expected input volume label and slightly adjust the network weight, to complete the classifier

We use the original images of the two different databases, i.e., private dataset and BUSIS dataset, and conduct intensity inhomogeneity correction for the original images. Therefore, each database will generate two groups of images (the original image and the corrected image) and their labels as training data. These materials are trained in the deep learning model and then performed testing to predict the benign/malignant lesions in the image. In this study, we use five different models to compare the results. Figure 2 shows the flowchart for deep learning.

3. Results and Discussion

Intensity inhomogeneity often causes unclear contour of tumors in sonogram so that it is uneasy to judge the type of tumors clinically. The correction method is proposed for improving image quality and increasing the accuracy in classification. Figure 3 shows the comparison images of one malignant tumor with/without intensity inhomogeneity. Before correction, it is uneasy to find the whole edge of the tumor. There is a missing boundary part at the left bottom of the tumor so that a wrong judgment may be caused. However, after correction, the missing part can be clearly seen, so it could be correctly judged as a malignant tumor.

Certainly, intensity inhomogeneity correction also works well in benign tumor images. Figure 4 shows the comparison images of benign tumor with/without intensity inhomogeneity. The original image with intensity inhomogeneity has unclear boundary at the right and left of the tumor that would cause wrong classification result. After correction, the unclear boundary part can be obviously distinguished so as to classify the type of the tumor correctly.

In this study, we used two kinds of dataset. One is the data collected by our team, named private dataset (96 malignant and 74 benign images), and the other is the public dataset on the website, named BUSIS dataset (150 malignant and 100 benign images). In the private dataset, 143 images and 27 images are selected randomly as training data and test data, respectively. In the BUSIS dataset, 210 images and 40 images are selected randomly as training data and test data, respectively.

First, we feed ultrasound tumor ROI images into SDAE network sized of 28×28 . In the private dataset, 143 images and 27 images are selected randomly as training data and test data, respectively. In the BUSIS dataset, 210 images and 40 images are selected randomly as training data and test data, respectively. Second, a backpropagation algorithm is used on SDAE to learn the weights and the feature representation. Third, the images are labeled manually and conducted the second supervised training to classify the malignant/benign tumor images. The experiments of benign/malignant classification are grouped into two groups: (1) original images and (2) contrast-enhanced images by intensity inhomogeneity correction, which are shown in Figures 5–8, respectively.

For the private dataset, the classification accuracy values of testing results of two groups of images are, respectively, 63% and 82%. For the BUSIS dataset, the classification accuracy values of testing results of two groups of images are, respectively, 75% and 83%. The results indicate the improvement of image quality (contrast-enhanced images) actually achieve better classification performance. The results also show that because the image complexity in the private database is higher, the accuracy increases noticeably after intensity inhomogeneity correction.

As we stated previously, CNN may fail when the number of the training images is relatively small. To verify this, we also conduct an experiment of CNN for detecting benign and malignant tumors. The labeled training images are directly fed into AlexNet [19] to learn the prediction results (i.e., benign or malignant tumors). Similarly, the backpropagation algorithm is used to train the weights of AlexNet. Finally, the classification accuracy values of testing results of two groups of private dataset images are 44% and 37%, while the training accuracy values are 99.8% and 99.4%, respectively. Apparently, such CNN architecture is not suitable for such small-scale training images, leading to a serious overfitting problem. It also verified that the proposed SDAE is a proper choice for analyzing malignant/benign tumor images classification. In addition, three models (Inception v3 [40], ResNet [41], and DenseNet [42]) are also used for comparison. Table 1 shows the detection results of the original images (without correction) and the corrected images (with correction) of the private database, in which the displayed value means with/without intensity inhomogeneity correction, respectively. Table 2 shows the results of the BUSIS dataset. To sum up, because the complexity of the images in the BUSIS database is lower, the classification accuracy is higher than that of the private database before or after intensity inhomogeneity correction.

It is clear that the DenseNet achieves the best performance. However, it is also shown that the performance of the different CNN methods cannot achieve good performance, compared with SDAE approach. Although the number of the parameters of AlexNet is relatively low, the total number of parameters is still greatly large than that of SDAE. By fair comparison, we have chosen Resnet with 54 layers and Inception v3 with 34 layers to learn the classifiers with the default settings suggested by their approach. However, such lot parameters in these two networks lead to serious

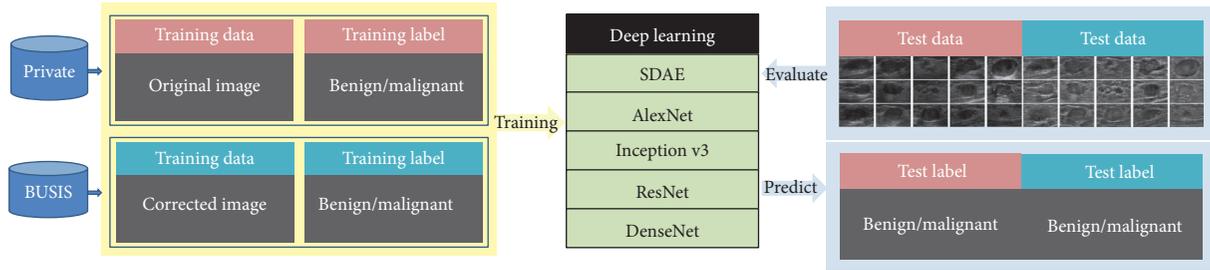


FIGURE 2: The flowchart for deep learning.

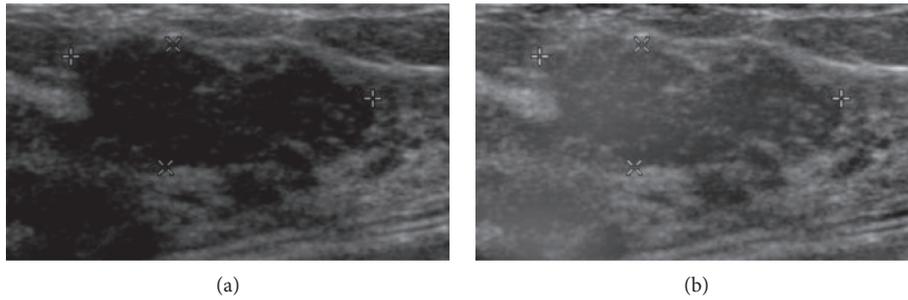


FIGURE 3: Malignant: (a) original image and (b) image after correction.

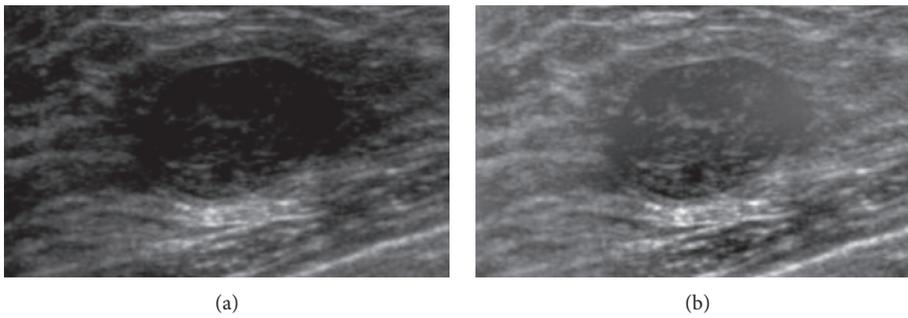


FIGURE 4: Benign: (a) original image and (b) image after correction.

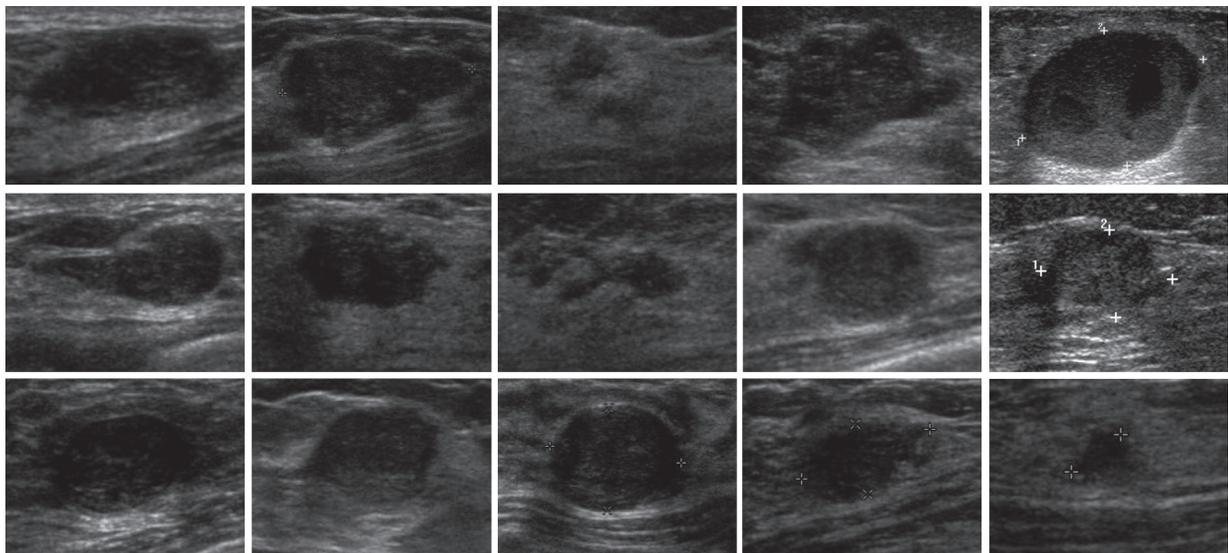


FIGURE 5: The examples of original test images in the private dataset.

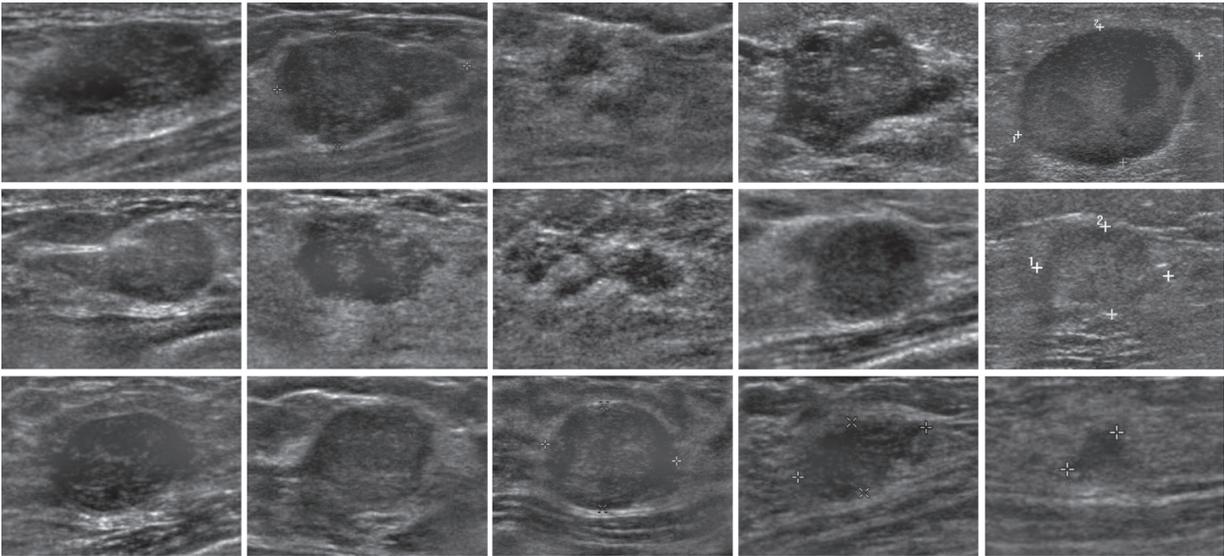


FIGURE 6: The examples of test images after correction in the private dataset.

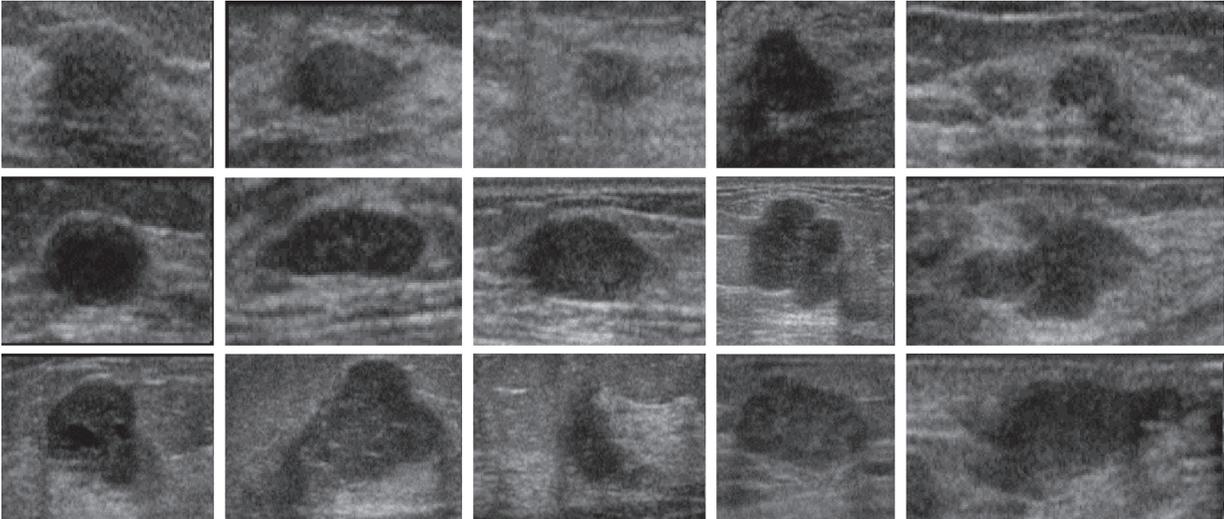


FIGURE 7: The examples of original test images in the BUSIS dataset.

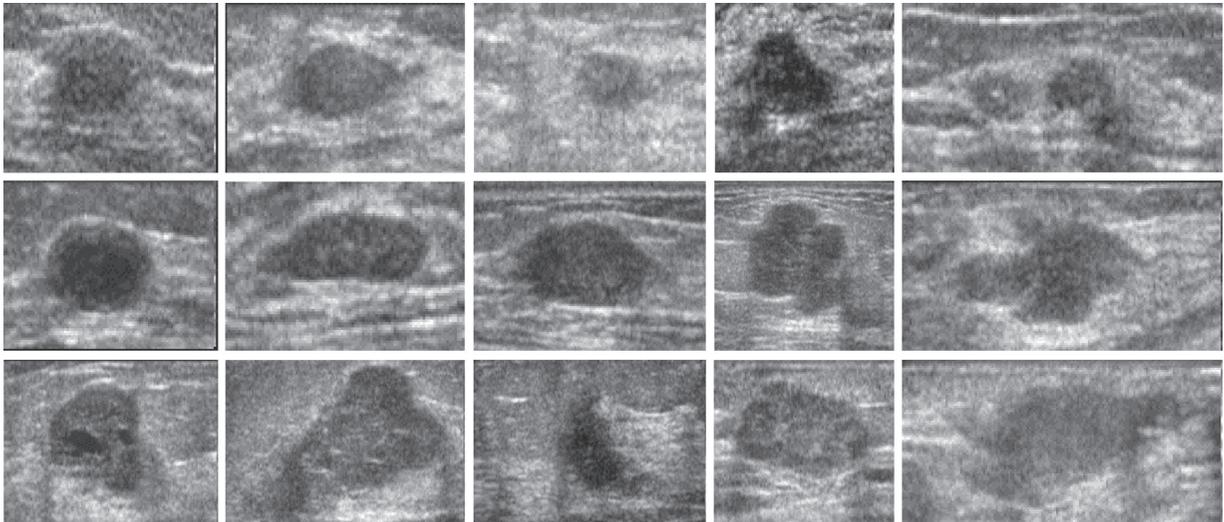


FIGURE 8: The examples of test images after correction in the BUSIS dataset.

TABLE 1: Classification results of the private database (the displayed value: original/corrected).

	SDAE	AlexNet	Inception v3	ResNet	DenseNet
TP	8/11	5/6	6/7	7/7	6/6
TN	9/11	7/4	6/5	6/7	8/7
FP	4/2	7/10	8/9	8/7	6/7
FN	6/3	8/7	7/6	6/6	7/7
Precision	0.67/ 0.85	0.42/ 0.38	0.43/0.44	0.5/0.5	0.5/0.46
Recall	0.57/ 0.79	0.38/ 0.46	0.46/0.54	0.46/ 0.54	0.46/0.46
Specificity	0.69/ 0.85	0.5/0.28	0.43/0.36	0.57/0.5	0.57/0.5
Accuracy	0.63/ 0.82	0.44/ 0.37	0.44/0.44	0.51/ 0.52	0.52/0.48
F-measure	0.62/ 0.85	0.4/0.41	0.44/0.48	0.48/ 0.52	0.48/0.46

TABLE 2: Classification results of the BUSIS database (the displayed value: original/corrected).

	SDAE	AlexNet	Inception v3	ResNet	DenseNet
TP	17/18	10/16	13/15	16/16	15/17
TN	13/15	12/14	14/16	13/10	14/13
FP	7/5	8/6	6/4	7/10	6/7
FN	3/2	10/4	7/5	4/4	5/3
Precision	0.71/0/ 78	0.56/ 0.73	0.68/0.79	0.69/ 0.61	0.71/0.71
Recall	0.85/0.9	0.5/0.8	0.65/0.75	0.8/0.8	0.75/0.85
Specificity	0.65/ 0.75	0.6/0.7	0.7/0.8	0.65/0.5	0.7/0.65
Accuracy	0.75/ 0.83	0.55/ 0.75	0.68/0.78	0.73/ 0.65	0.73/0.75
F-measure	0.77/ 0.83	0.53/ 0.76	0.67/0.77	0.74/ 0.70	0.73/0.77

overfitting problem and bad test performance. On the contrary, the most advanced network architecture, DenseNet, achieves good performance among four CNN methods; it still fails to obtain promising performance at all.

In [43], this study discriminates benign cysts from malignant masses in breast ultrasound by transferred deep neural network, Inception V3, with an accuracy of 89.44%. However, this study uses a total of 2058 breast ultrasound masses, comprising 1370 benign and 688 malignant lesions, so better results can be obtained. In [44], Han et al. use GoogLeNet with preprocessing for supporting the classification of breast lesions in ultrasound images. The networks showed an accuracy of about 90%. This study uses large data, 7408 images (4254 benign and 3154 malignant lesions), to get a good classification result. Apparently, CNN architecture is not suitable for such small-scale training images, leading to a serious overfitting problem. It also verified that the proposed SDAE is a proper choice for analyzing malignant/benign tumor images classification. In addition, Han et al. also proposed that the preprocessing can get better classification results.

4. Conclusion

This study reports that the proposed algorithm overcomes the problem of intensity inhomogeneity in the sonogram, and in combination with the deep learning method for tumor classification. This study compared five deep learning models, and SDAE achieved the best identification accuracy. Since this is a small amount of data, SDAE is a suitable choice for analyzing the small dataset.

For the original image group and the corrected image group in the private database, the accuracy values of identifying the benign and malignant tumors are 63% and 82%, respectively. For the same group in the public database, the accuracy values are 75% and 83%, respectively. The results refer that increasing the image quality will help improve the accuracy of classification. The image complexity in the private database is higher. From the results, it can be inferred that for images with high complexity, the classification accuracy increases noticeably after image preprocessing.

This work is different from the traditional algorithm to classify the type of lesions, i.e., benign or malignant in the sonogram. In the traditional way, the tumor contours must be delineated first by the segmentation technique before they can be classified. However, many segmentation algorithms need to adjust many parameters or depend on the developer's subjective ideas. In this way, the result of the segmentation in turn affects the classification result. The proposed method reduced the impact of segmentation processing steps to lead an objective classification result via deep learning. Moreover, this paper combines an intensity inhomogeneity correction method to make the classification result better. This way is advantageous, and it has the potential to obtain the better classification results after image preprocessing.

Abbreviations

AE:	Autoencoder
DAE:	Denoising autoencoder
SDAE:	Stack denoising autoencoder
LR:	Logistic regression
ROI:	Region of interest.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

Chia-Yen Lee and Chih-Chung Hsu conceptualized the study. Chia-Yen Lee was responsible for funding acquisition. Chia-Yen Lee, Chih-Chung Hsu, and Guan-Lin Chen performed the methodology. Chia-Yen Lee and Yi-Hong Chou administered the project. Yi-Hong Chou collected the

resources; Guan-Lin Chen performed software simulation. Chia-Yen Lee and Chih-Chung Hsu supervised the study. Chia-Yen Lee wrote the original draft. Chih-Chung Hsu reviewed and edited the manuscript.

Acknowledgments

This study was supported by the Ministry of Science and Technology, R.O.C. (Grant number: MOST 102-2218-E-239-004-MY2).

References

- [1] D. M. Parkin, P. Pisani, and J. Ferlay, "Estimates of the worldwide incidence of eighteen major cancers in 1985," *International Journal of Cancer*, vol. 54, no. 4, pp. 594–606, 1993.
- [2] A. C. Society, "Breast cancer facts & figures 1997-1998," 1999, <https://www.nrc.gov/docs/ml0716/ml071640135.pdf>.
- [3] V. Chalana, D. T. Linker, D. R. Haynor, and Y. Kim, "A multiple active contour model for cardiac boundary detection on echocardiographic sequences," *IEEE Transactions on Medical Imaging*, vol. 15, no. 3, pp. 290–298, 1996.
- [4] D. R. Chen, R. F. Chang, W. J. Wu, W. K. Moon, and W. L. Wu, "3-D breast ultrasound segmentation using active contour model," *Ultrasound in Medicine & Biology*, vol. 29, no. 7, pp. 1017–1026, 2003.
- [5] B. Cigale and D. Zazula, "Segmentation of ovarian ultrasound images using cellular neural networks," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 18, no. 4, pp. 563–581, 2004.
- [6] J. J. Clark, "Authenticating edges produced by zero-crossing algorithms," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 1, pp. 43–57, 1989.
- [7] M. S. Drew, J. Wei, and Z. N. Li, "Illumination-invariant image retrieval and video segmentation," *Pattern Recognition*, vol. 32, no. 8, pp. 1369–1388, 1999.
- [8] P. R. Hill, C. N. Canagarajah, and D. R. Bull, "Texture gradient based watershed segmentation," in *Proceedings of 2002 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 4, p. IV-3381, IEEE, Orlando, FL, USA, May 2002.
- [9] S. H. Lee and J. K. Seo, "Level set-based bimodal segmentation with stationary global minimum," *IEEE Transactions on Image Processing*, vol. 15, no. 9, pp. 2843–2852, 2006.
- [10] T. Brox and J. Weickert, "Level set segmentation with multiple regions," *IEEE Transactions on Image Processing*, vol. 15, no. 10, pp. 3213–3218, 2005.
- [11] J. F. Haddon and J. F. Boyce, "Image segmentation by unifying region and boundary information," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 10, pp. 929–948, 1990.
- [12] S. Joo, Y. S. Yang, W. K. Moon, and H. C. Kim, "Computer-aided diagnosis of solid breast nodules: use of an artificial neural network based on multiple sonographic features," *IEEE Transactions on Medical Imaging*, vol. 23, no. 10, pp. 1292–1300, 2004.
- [13] J. Z. Cheng, Y. H. Chou, C. S. Huang et al., "ACCOMP: augmented cell competition algorithm for breast lesion demarcation in sonography," *Medical Physics*, vol. 37, no. 12, pp. 6240–6252, 2010.
- [14] R. F. Chang, W. J. Wu, W. K. Moon, and D. R. Chen, "Automatic ultrasound segmentation and morphology based diagnosis of solid breast tumors," *Breast Cancer Research and Treatment*, vol. 89, no. 2, pp. 179–185, 2005.
- [15] Q. V. Le, "Building high-level features using large scale unsupervised learning," in *Proceedings of 2013 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 8595–8598, IEEE, Vancouver, Canada, May 2013.
- [16] G. B. Huang, H. Lee, and E. Learned-Miller, "Learning hierarchical representations for face verification with convolutional deep belief networks," in *Proceedings of 2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2518–2525, IEEE, Providence, RI, USA, June 2012.
- [17] H. C. Shin, M. R. Orton, D. J. Collins, S. J. Doran, and M. O. Leach, "Stacked autoencoders for unsupervised feature learning and multiple organ detection in a pilot study using 4D patient data," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 8, pp. 1930–1943, 2012.
- [18] Y. Xu, T. Mo, Q. Feng et al., "Deep learning of feature representation with multiple instance learning for medical image analysis," in *Proceedings of 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 1626–1630, IEEE, Florence, Italy, May 2014.
- [19] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems*, pp. 1097–1105, MIT Press, Cambridge, MA, USA, 2012.
- [20] M. H. Yap, G. Pons, J. Marti et al., "Automated breast ultrasound lesions detection using convolutional neural networks," *IEEE Journal of Biomedical and Health Informatics*, vol. 22, no. 4, pp. 1218–1226, 2017.
- [21] X. Xie, F. Shi, J. Niu, and X. Tang, "Breast ultrasound image classification and segmentation using convolutional neural networks," in *Proceedings of Pacific Rim Conference on Multimedia*, pp. 200–211, Springer, Hefei, China, September 2018.
- [22] S. Zhou, J. Shi, J. Zhu, Y. Cai, and R. Wang, "Shearlet-based texture feature extraction for classification of breast tumor in ultrasound image," *Biomedical Signal Processing and Control*, vol. 8, no. 6, pp. 688–696, 2013.
- [23] W. K. Moon, I. L. Chen, J. M. Chang, S. U. Shin, C. M. Lo, and R. F. Chang, "The adaptive computer-aided diagnosis system based on tumor sizes for the classification of breast tumors detected at screening ultrasound," *Ultrasonics*, vol. 76, pp. 70–77, 2017.
- [24] P. Raha, R. V. Menon, and I. Chakrabarti, "Fully automated computer aided diagnosis system for classification of breast mass from ultrasound images," in *Proceedings of 2017 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)*, pp. 48–51, IEEE, Chennai, India, March 2017.
- [25] M. Abdel-Nasser, J. Melendez, A. Moreno, O. A. Omer, and D. Puig, "Breast tumor classification in ultrasound images using texture analysis and super-resolution methods," *Engineering Applications of Artificial Intelligence*, vol. 59, pp. 84–92, 2017.
- [26] J. Z. Cheng, D. Ni, Y. H. Chou et al., "Computer-aided diagnosis with deep learning architecture: applications to breast lesions in US images and pulmonary nodules in CT scans," *Scientific Reports*, vol. 6, no. 1, p. 24454, 2016.
- [27] C. Xing, L. Ma, and X. Yang, "Stacked denoise autoencoder based feature extraction and classification for hyperspectral images," *Journal of Sensors*, vol. 2016, Article ID 3632943, 10 pages, 2016.

- [28] B. M. Dawant, A. P. Zijdenbos, and R. A. Margolin, "Correction of intensity variations in MR images for computer-aided tissue classification," *IEEE Transactions on Medical Imaging*, vol. 12, no. 4, pp. 770–781, 1993.
- [29] P. Vemuri, E. G. Kholmovski, D. L. Parker, and B. E. Chapman, "Coil sensitivity estimation for optimal SNR reconstruction and intensity inhomogeneity correction in phased array MR imaging," in *Proceedings of Biennial International Conference on Information Processing in Medical Imaging*, pp. 603–614, Springer, Glenwood Springs, CO, USA, July 2005.
- [30] G. Gerig, M. Prastawa, W. Lin, and J. Gilmore, "Assessing early brain development in neonates by segmentation of high-resolution 3T MRI," in *Proceedings of International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 979–980, Springer, Montréal, Canada, November 2003.
- [31] J. Ashburner and K. J. Friston, "Unified segmentation," *NeuroImage*, vol. 26, no. 3, pp. 839–851, 2005.
- [32] E. Ardizzone, R. Pirrone, and O. Gambino, "Exponential entropy driven HUM on knee MR images," in *Proceedings of 27th Annual International Conference of the Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005*, pp. 1769–1772, IEEE, Shanghai, China, September 2005.
- [33] J. G. Sled, A. P. Zijdenbos, and A. C. Evans, "A nonparametric method for automatic correction of intensity nonuniformity in MRI data," *IEEE Transactions on Medical Imaging*, vol. 17, no. 1, pp. 87–97, 1998.
- [34] J. A. Hernández, M. L. Mora, E. Schiavi, and P. Toharia, "RF inhomogeneity correction algorithm in magnetic resonance imaging," in *Proceedings of International Symposium on Biological and Medical Data Analysis*, pp. 1–8, Springer, Barcelona, Spain, November 2004.
- [35] A. W. C. Liew and H. Yan, "An adaptive spatial fuzzy clustering algorithm for 3-D MR image segmentation," *IEEE Transactions on Medical Imaging*, vol. 22, no. 9, pp. 1063–1075, 2003.
- [36] C. Li, R. Huang, Z. Ding, C. Gatenby, D. Metaxas, and J. Gore, "A variational level set approach to segmentation and bias correction of images with intensity inhomogeneity," in *Proceedings of International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 1083–1091, Springer, New York, NY, USA, September 2008.
- [37] J. Milles, Y. M. Zhu, G. Gimenez, C. R. Guttmann, and I. E. Magnin, "MRI intensity nonuniformity correction using simultaneously spatial and gray-level histogram information," *Computerized Medical Imaging and Graphics*, vol. 31, no. 2, pp. 81–90, 2007.
- [38] M. Xian, Y. Zhang, H. D. Cheng et al., "A benchmark for breast ultrasound image segmentation (BUSIS)," 2018, <http://arxiv.org/abs/1801.03182>.
- [39] C. Y. Lee, Y. H. Chou, C. S. Huang, Y. C. Chang, C. M. Tiu, and C. M. Chen, "Intensity inhomogeneity correction for the breast sonogram: constrained fuzzy cell-based bipartitioning and polynomial surface modeling," *Medical Physics*, vol. 37, no. 11, pp. 5645–5654, 2010.
- [40] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2818–2826, Las Vegas, NV, USA, June 2016.
- [41] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778, Las Vegas, NV, USA, June 2016.
- [42] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, vol. 1, no. 2, p. 3, Honolulu, HI, USA, July 2017.
- [43] T. Xiao, L. Liu, K. Li, W. Qin, S. Yu, and Z. Li, "Comparison of transferred deep neural networks in ultrasonic breast masses discrimination," *BioMed Research International*, vol. 2018, Article ID 4605191, 9 pages, 2018.
- [44] S. Han, H. K. Kang, J. Y. Jeong et al., "A deep learning framework for supporting the classification of breast lesions in ultrasound images," *Physics in Medicine & Biology*, vol. 62, no. 19, pp. 7714–7728, 2017.

Research Article

Fully Automatic Segmentation and Three-Dimensional Reconstruction of the Liver in CT Images

ZhenZhou Wang , Cunshan Zhang, Ticao Jiao, MingLiang Gao, and Guofeng Zou

College of Electrical and Electronic Engineering, Shandong University of Technology, Zibo City 255049, China

Correspondence should be addressed to ZhenZhou Wang; zzwangsia@yahoo.com

Received 14 June 2018; Revised 30 September 2018; Accepted 22 October 2018; Published 18 November 2018

Guest Editor: Cesare Valenti

Copyright © 2018 ZhenZhou Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Automatic segmentation and three-dimensional reconstruction of the liver is important for liver disease diagnosis and surgical treatment. However, the shape of the imaged 2D liver in each CT image changes dramatically across the slices. In all slices, the imaged 2D liver is connected with other organs, and the connected organs also vary across the slices. In many slices, the intensities of the connected organs are the same with that of the liver. All these facts make automatic segmentation of the liver in the CT image an extremely difficult task. In this paper, we propose a heuristic approach to segment the liver automatically based on multiple thresholds. The thresholds are computed based on the slope difference distribution that has been proposed and verified in the previous research. Different organs in the CT image are segmented with the automatically computed thresholds, respectively. Then, different segmentation results are combined to delineate the boundary of the liver robustly. After the boundaries of the 2D liver in all the slices are identified, they are combined to form the 3D shape of the liver with a global energy minimization function. Experimental results verified the effectiveness of all the proposed image processing algorithms in automatic and robust segmentation of the liver in CT images.

1. Introduction

Liver diseases have become one of the most common causes of deaths in the world. Researchers have focus on the prevention and treatment of liver diseases for many years. In recent years, computed tomography (CT) imaging has been widely used in liver disease diagnosis and surgical treatment because tumors or hepatic lesions could be observed easily from the CT image. For the captured CT images, the liver slices need to be examined in the two dimensions one by one. Consequently, it lacks an overall image of the 3D liver. Furthermore, it takes clinicians considerable time to view all the slices and diagnose the disease or evaluate the liver function based on the information divided and presented in different images. Therefore, it is desirable that 2D liver slice is segmented from CT images, and the 3D liver is reconstructed automatically and robustly beforehand. Thus, the clinicians could get the information of the 3D liver at a glance and diagnose the liver disease or evaluate the liver

function more conveniently. This desire has led researchers worldwide to devote themselves to the research of coming up with automatic and robust liver segmentation methods. After so many years of research, it remains an open problem because the liver is adjacent to many other organs, such as the kidneys, spleen, stomach, intestines, and bones. In many cases, the intensity of the liver and that of the adjacent organ is indistinguishable. In addition, the shape of the liver varies according to the individuals. As a result, automatic and robust segmentation of the liver from the CT images remains as one of the most challenging artificial intelligence task for many decades.

After so many years of study and research, quite a few liver segmentation methods have been proposed, though none of them have achieved adequate accuracy so far. Among them, methods based on statistical and probabilistic models became the most popular ones [1–4]. Yet, such kinds of methods require a large size of training samples, which decreases the segmentation efficiency tremendously. Even if

dictionary and sparse coding is used to reduce the redundant information, the segmentation efficiency is still not satisfactory. In recent years, deep learning and convolutional neural network have been used to segment the liver in the CT images, and the reported accuracy appears to be promising [5, 6]. Similarly, deep learning or convolutional neural network-based methods rely heavily on the training datasets to yield accurate segmentation results. In other words, the accuracy will not be acceptable if the training datasets are not similar enough to the tested case. In reality, the liver of the patients varies tremendously, and it is thus difficult to acquire a complete training datasets. In such situations, both statistical/probabilistic models-based methods and neural network-based methods might fail completely in identifying the livers of some individuals.

There are also many other methods that do not rely on training datasets. The popular segmentation methods used in identifying the boundary of the liver in CT images include active contour [7], threshold selection [8], level set [9], and graph cuts [10–12]. Since its surrounding tissues and connected organs in different slices are different, one segmentation method or several combined segmentation methods [7–12] usually could not segment the liver robustly and fully automatically. Hence, heuristic methods are proposed to segment the liver, its surrounding tissues, and its connected organs as a whole [13–15]. For instance, the surrounding tissues are segmented first and then serve as a constraint for the segmentation of the liver. In this way, the surrounding tissues are changed from interference factors to the markers of constraint. Thus, the liver could be segmented robustly and automatically [16].

In this paper, we also propose a heuristic method based on multiple thresholds selection and morphological operations. With a global threshold, we could robustly segment all the connected organs in each CT slice, such as the ribs, the spine, the heart, the kidney, and the stomach. With another global threshold, we could robustly segment the liver with the logic not operation of the segmentation result of the connected organs. Therefore, robust threshold selection becomes a critical step in the proposed method. We utilize the slope difference distribution-based threshold selection (SDDTS) method to calculate multiple thresholds in this research work. The robustness of the slope difference distribution-based threshold selection method and its advantage over state-of-the-art segmentation methods have been verified in our previous studies [17, 18]. At first, we calibrate the parameters of the slope difference distribution-based threshold selection method with several typical CT images and then use the calibrated parameters for all the CT slices. After segmentation by the global threshold, the segmentation result is filtered by minimizing the Gibbs energy function [19] to reduce inhomogeneity. Then, morphological operations [20] are used to merge divided parts of the same organ, and the merged organ is used as liver segmentation constraint. Since the intensities of the surrounding tissues and that of the liver might be completely the same, we use centroids of the segmented ribs and the spine to fit a curved line, which is then used to separate the

liver and the surrounding tissues. After liver segmentation, its boundary is further refined by spline filtering.

2. Methods

The proposed method is heuristic, and it contains a series of image processing algorithms that vary depending on the index of the CT slice. The core image processing algorithms include (1) threshold selection based on slope difference distribution [17, 18] and image segmentation with the selected threshold; (2) energy minimization of segmentation result to eliminate noise; (3) morphological filtering; (4) morphological merging; and (5) spline filtering. The flowchart of the proposed method is shown in Figure 1. The multiple thresholds are calculated from the inputted 2D CT slice. One threshold is used for liver segmentation, and the other thresholds are used for constraint segmentation. The core image processing algorithms are used during constraint segmentation and computing the boundary of the segmented liver.

2.1. Threshold Selection. The slope difference distribution is computed from the histogram of the image, and it reflects the global variation rate of the histogram. With the assumption that the thresholding point between two pixel classes varies greatest, the thresholding point could be computed based on the variation rate of the histogram, i.e., the slope difference distribution. The slope difference distribution is formed by a series of slope difference that is calculated at each sampled point of the smoothed histogram. At each sampled point of the smoothed histogram, N points on its left are selected to fit a line model and N points on its right are selected to fit another line model. The slopes of these two line models are calculated. The left slope is subtracted from the right slope, and the slope difference at this sampled point is obtained. First of all, the histogram of the image needs to be calculated and smoothed as follows.

The grey-scale values of the image are scaled to the interval [1, 255], and then the histogram distribution $P(x)$ is calculated as follows:

$$P(x = i) = \frac{F_i}{F_j}, \quad i = 1, 2, \dots, 255, \quad j \in [1, 255], \quad (1)$$

where F_i denotes the total number of pixels that equals to i and F_j denotes the maximum number of pixels that occurs at j in the interval [1, 255]. The histogram distribution, $P(x)$, is then transformed to the frequency domain by the discrete Fourier transform (DFT):

$$F(k) = \sum_{x=1}^{255} P(x)e^{-i(2\pi kx/255)}, \quad k = 1, 2, \dots, 255. \quad (2)$$

Only the low frequencies from 0 to W and the symmetric frequencies from $255 - W$ to 255 are kept:

$$F'(k) = \begin{cases} F(k), & k = 1, 2, \dots, W, \\ F(k), & k = 255 - W, \dots, 254, 255, \\ 0, & k = W + 1, \dots, 255 - W, \end{cases} \quad (3)$$

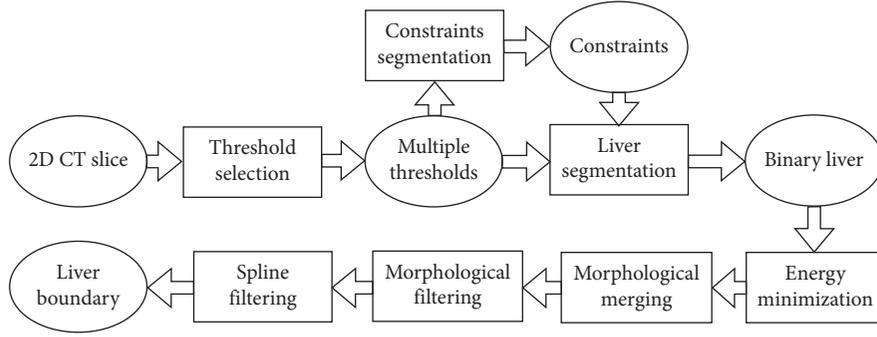


FIGURE 1: Flowchart of the proposed method.

where W is the bandwidth of the low-pass DFT filter. Its value and the value of N are obtained by parameter calibration [17, 18]. In general, there are always optimal values of W and N to yield the optimal threshold. However, it is still challenging to determine the optimal W and N by simple parameter calibration. When the values of W and N are not determined properly, the threshold will be determined inaccurately. The reason we choose the Fourier transform-based filter instead of other popular filters is based on the quantitative evaluation and comparison [19]. We found that the Fourier transformation-based filter outperforms other filters significantly in this conducted research work. Transforming from the frequency domain back to the spatial domain

$$P'(x) = \frac{1}{255} \sum_{k=1}^{255} F'(k) e^{i(2\pi xk/255)}, \quad x = 1, 2, \dots, 255, \quad (4)$$

where $P'(x)$ is the smoothed histogram distribution.

To compute the slope difference, we fit two-line models on both sides of the sampled point. The line model is formulated as follows:

$$y_i = ax_i + b, \quad (5)$$

where (x_i, y_i) , $i = 1 + N, \dots, 255 - N$, is the sampled point on the smoothed histogram distribution and a is the slope of the line model. The coefficient of the line model $[a, b]^T$ is computed as follows:

$$[a, b]^T = (B^T B)^{-1} B^T Y, \quad (6)$$

$$B = \begin{bmatrix} x_{i+1-N} & 1 \\ x_{i+2-N} & 1 \\ \vdots & \vdots \\ x_{i-1} & 1 \\ x_i & 1 \end{bmatrix} \text{ or } \begin{bmatrix} x_i & 1 \\ x_{i+1} & 1 \\ \vdots & \vdots \\ x_{i-2+N} & 1 \\ x_{i-1+N} & 1 \end{bmatrix}, \quad (7)$$

$$Y = [y_{i+1-N}, y_{i+2-N}, \dots, y_{i-1}, y_i]^T \text{ or } [y_i, y_{i+1}, \dots, y_{i-2+N}, y_{i-1+N}]^T, \quad (8)$$

where B is the design matrix of the least square fitting method and is the input data vector. Moreover, $[(x_{i+1-N}, y_{i+1-N}), (x_{i+2-N}, y_{i+2-N}), \dots, (x_{i-1}, y_{i-1}), (x_i, y_i)]$ are the N adjacent points at the left side of the point (x_i, y_i) , and

$[(x_i, y_i), (x_{i+1}, y_{i+1}), \dots, (x_{i-2+N}, y_{i-2+N}), (x_{i-1+N}, y_{i-1+N})]$ are the N adjacent points at the right side of the point (x_i, y_i) . The left slope and the right slope, a_l and a_r at point (x_i, y_i) , are then obtained from Equation (6). The slope difference at point (x_i, y_i) is then computed as follows:

$$s(i) = a_r(i) - a_l(i), \quad i = 1 + N, \dots, 255 - N. \quad (9)$$

The discrete version is denoted as $s(i)$, and its continuous version is denoted as that is named as slope difference distribution. Let the derivative of $s(x)$ equal zero and solve it, the valleys V_i , $i = 1, 2, \dots, M$, with greatest local variations are obtained. The positions where these valleys lie are the thresholds that separate different pixel classes. One fundamental property of the slope difference distribution is that the positions of the valleys change monotonically with the line model fitting parameter N . Hence, the parameter could be calibrated to yield the optimal threshold. After the optimum threshold T is selected, the image is segmented by the following equation:

$$I_b = \begin{cases} 1, & I(i, j) \geq T, \\ 0, & I(i, j) < T, \end{cases} \quad (10)$$

where I is the original image and I_b is the binarized image. (i, j) is the index of the pixel in the image. Since multiple thresholds are needed to segment different organs, Equation (9) is the basic format of segmenting a single object. More conditions need to be added in Equation (9), when several independent segmentation results are combined to segment the liver.

2.2. Energy Minimization. There are many popular noise reduction methods, e.g., the discrete Fourier transformation-based filter [20] and the wavelet image processing [21]. The noise reduction methods are usually applied beforehand in the preprocessing [20]. Since we did not reduce the noise beforehand, we apply a noise reduction procedure immediately after the segmentation. The noise would cause region inhomogeneity, and this inhomogeneity could be formulated by the Gibbs distribution:

$$P(X = x) = \frac{1}{\sum_{x \in \mathcal{L}} e^{-\sum_{c \in \mathcal{C}} V_c(x)}} e^{-\sum_{c \in \mathcal{C}} V_c(x)}, \quad (11)$$

where $V_c(x)$ is the potential function associated with clique c . The clique c is defined as a set of sites such that any two elements in the clique are neighbours of each other [19]. L is the total number of pixel classes. $V_c(x)$ is defined by the following equation:

$$V_c(x) = \begin{cases} -\beta, & \text{all values of } c \text{ are equal,} \\ \beta, & \text{else,} \end{cases} \quad (12)$$

where β is the constant and 1 is its default value. To reduce the noise in the binarized image, the total energy over the whole image is minimized as follows:

$$I_0 = \arg \min_{I_{\text{label}}(u,v) \in [1, M_{\text{in}}]} \prod_{l=1}^{M_{\text{in}}} \sum_{u=1}^U \sum_{v=1}^V \sum_{c \in C} V_c(x), \quad (13)$$

where I_0 is the filtered image with Gibbs energy minimization; u and v are the pixel indexes of the image in the vertical direction and in the horizontal direction, respectively; and U and V are the resolution of the image in the vertical direction and in the horizontal direction, respectively.

2.3. Morphological Filtering. Besides the noise, there are many small binarized blobs that do not belong to the liver or the segmented organ. These small blobs could not be removed by energy minimization. A popular way to remove these small blobs is to count the areas of all the blobs and remove some of the blobs based on their areas morphologically. However, there are also many situations in which these small blobs are connected with the liver or the segmented organ, which make the removal of them more difficult. To remove all these interference blobs, we propose a morphological filtering method that contains the following steps.

Step 1. Erode the segmentation result I_0 morphologically as follows:

$$I'_0 = I_0 \ominus B = \{z \mid (B)_z \subseteq I_0\}, \quad (14)$$

$$(B)_z = \{c \mid c = p + z, p \in B\}, \quad (15)$$

$$I_0 = I'_0, \quad (16)$$

where B is the 4-connected structure element with the disk shape and its radius is 1, p is the point in the structuring element B , and z is the translation vector.

Step 2. Repeat Step 1 N_F times. The default value of N_F is 8.

Step 3. Dilate the segmentation result I_0 morphologically as follows:

$$\begin{aligned} I'_0 &= I_0 \oplus B = \{z \mid (B^s)_z \cap I_0 \neq \emptyset\}, \\ I_0 &= I'_0, \end{aligned} \quad (17)$$

where B^s denotes the symmetric or supplement of B .

Step 4. Repeat Step 3 N_F times.

In summary, the proposed morphological filtering method removes the small blobs by a repeating morphological erosion process first. Then, it restores the eroded liver or other organs by a morphological dilation process with the same repeating times.

2.4. Morphological Merging. On the contrary, there are situations where the segmentation result of the organ specially the stomach is split into different parts. To utilize the segmentation result more effectively, it is required that the split parts are merged into a united one. To unite these split parts, we propose a morphological merging method that contains the following steps.

Step 1. Dilate the segmentation result I_0 morphologically as follows:

$$\begin{aligned} I'_0 &= I_0 \oplus B = \{z \mid (B^s)_z \cap I_0 \neq \emptyset\}, \\ I_0 &= I'_0. \end{aligned} \quad (18)$$

Step 2. Repeat Step 1 N_M times. The default value of N_M is 16.

Step 3. Erode the segmentation result I_0 morphologically as follows:

$$\begin{aligned} I'_0 &= I_0 \ominus B = \{z \mid (B)_z \subseteq I_0\}, \\ (B)_z &= \{c \mid c = p + z, p \in B\}, \\ I_0 &= I'_0. \end{aligned} \quad (19)$$

Step 4. Repeat Step 3 N_M times.

As can be seen, the morphological merging is the opposite process of the morphological filtering. It connects the split parts and merges them into one united one by a repeating morphological dilation process first. Then, it restores the dilated organ by a morphological erosion process with the same repeating times.

2.5. Spline Filtering. After the liver is segmented, its boundary B is extracted first. The final smooth boundary B_s is computed by minimizing the energy between the extracted boundary B and the fitted polynomial spline by the following equation:

$$E(B_s) = (1 - \alpha) \int |B_s(j) - B(j)|^2 dj + \alpha \int \left| \frac{d^2 B_s(t)}{dt^2} \right|^2 dt, \quad (20)$$

where α is the smoothing factor and its default value is 0.5.

2.6. Liver Segmentation. We use a typical CT image as shown in Figure 2(a) to illustrate the proposed segmentation method. First, the thresholds are computed with the slope difference distribution as shown in Figure 2(b). Four

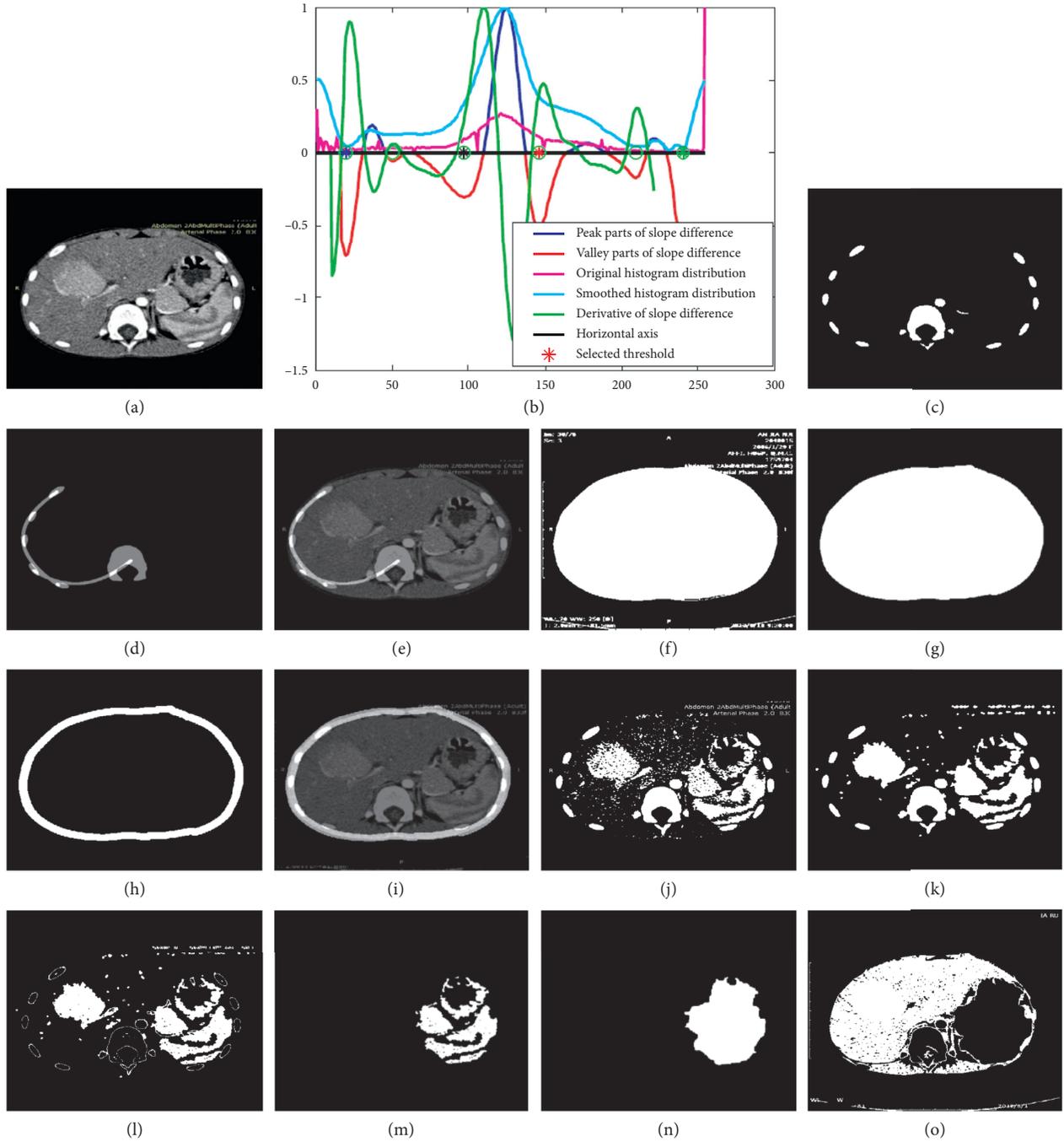


FIGURE 2: Demonstration of the proposed liver segmentation method with a typical image: (a) the original image; (b) threshold selection process by slope difference distribution; (c) the segmented ribs and spine; (d) the fitted curve based on the centroids of the ribs and the spine; (e) the fitted curve overlaying on the original image; (f) the segmented body; (g) the segmented body after morphological filtering; (h) the extracted circular part by erosion; (i) the circular part overlaying on the original image; (j) the segmented stomach; (k) the segmented stomach after energy minimization; (l) subtraction of the bones; (m) the extracted stomach after morphological filtering; (n) the merged stomach; (o) the segmented liver.

thresholds are calculated, and they are shown in Figure 2(b), respectively. The threshold T_1 to segment the spine and ribs is denoted by the green asterisk in the green circle, the threshold T_2 to segment the body is denoted by the blue asterisk in the green circle, the threshold T_3 to segment the stomach is denoted by the red asterisk in the green circle, and the threshold T_4 to segment the liver is denoted by the

black asterisk in the green circle. With the threshold T_1 , the spine and the ribs are segmented as follows:

$$I_{\text{bones}} = \begin{cases} 1, & I(i, j) \geq T_1, \\ 0, & I(i, j) < T_1, \end{cases} \quad (21)$$

where I_{bones} denotes the segmented spine and ribs, and it is shown in Figure 2(c). The centroids of these blobs are



FIGURE 3: Demonstration of the proposed liver segmentation method with a typical image: (a) the segmented liver after energy minimization; (b) the segmented liver after morphological filtering; (c) the segmented liver after morphological area filtering; (d) the calculated liver boundary overlaying on the original image.

calculated as the means of all the pixels they contain. Based on the computed centroids, the blobs are divided into two classes, the first class of blobs that include the blobs on the left and the second class that includes the blobs on the right and around the center. The blobs in the second class are filtered based on their areas, and only the largest one (the spine) is kept. The centroids of the blobs in the first class and the largest one in the second class are used to fit a second-order curve I_{curve} by the least squared method (Equations (6–8)) as shown in Figure 2(d). This fitted curve I_{curve} is used to separate the tissues with almost the same intensity from the liver. Figure 2(e) shows the fitted curve I_{curve} overlaying on the original image I . As can be seen, the liver and its neighboring tissue on the bottom are separated successfully.

With the threshold T_2 , the whole body is segmented as follows:

$$I_{\text{body}} = \begin{cases} 1, & I(i, j) \geq T_2, \\ 0, & I(i, j) < T_2, \end{cases} \quad (22)$$

where I_{body} denotes the segmented body, and it is shown in Figure 2(f). Figure 2(g) shows the segmentation result of the body after morphological filtering. As can be seen, all the small blobs are removed successfully. The filtered body is then eroded morphologically (Equations (14)–(16)) by the disk structuring element with the radius of N_c . N_c is calculated as the average width of the segmented ribs, and it is 16 in this specific example. The eroded body is then subtracted from the filtered part, and the constraint circular part I_{cir} is obtained as shown in Figure 2(h). This constraint circular part is also used to eliminate the surrounding tissues that have the same or similar intensities with that of the liver. The constraint circular part overlaying on the original image is shown in Figure 2(i). As can be seen, it is adjacent to the liver, but almost not covering any liver part.

With the threshold T_3 , the stomach is segmented as follows:

$$I_{\text{st}} = \begin{cases} 1, & I(i, j) \geq T_3, \\ 0, & I(i, j) < T_3, \end{cases} \quad (23)$$

where I_{st} denotes the segmented stomach, and it is shown in Figure 2(j). It is then filtered by energy minimization, and the result is shown in Figure 2(k). The filtered result is subtracted by the segmented bones I_{bones} shown in Figure 2(c), and the

subtraction result is shown in Figure 2(l). The morphological filtering is applied on the subtraction result, and then, only the parts on the right of the spine are kept as shown in Figure 2(m). The morphological merging is applied on the kept parts, and the merged stomach I_{stm} is shown in Figure 2(n). With the constraint of fitted curve I_{curve} , the circular part I_{cir} , and the merged stomach I_{stm} , the liver is segmented by the threshold T_4 as follows:

$$I_{\text{liver}} = \begin{cases} 1, & I(i, j) \geq T_4 \& I_{\text{curve}}(i, j) < 1 \& I_{\text{cir}}(i, j) \\ & < 1 \& I_{\text{stm}}(i, j) < 1, \\ 0, & \text{else,} \end{cases} \quad (24)$$

where I_{liver} denotes the segmented liver, and it is shown in Figure 2(o). It is then filtered by energy minimization, and the filtered result is shown in Figure 3(a). As can be seen, there are still significant interference blobs. Immediately following the energy minimization, the morphological filtering is applied, and the filtering result is shown in Figure 3(b). As can be seen, the interference blobs are reduced significantly, and the left blobs could be removed by the morphological area filtering. The largest blob (the liver) is kept after the morphological area filtering and it is shown in Figure 3(c). The boundary of the segmented liver is extracted and smoothed by spline filtering. Figure 3(d) shows the smoothed liver boundary overlaying on the original image. As can be seen, the extracted boundary matches the liver very well.

2.7. Three-Dimensional Reconstruction. After the two-dimensional livers in all the CT slices are segmented and their boundaries are extracted, each boundary is sampled evenly with the same number of points; i.e., the distance between any two adjacent sampled points on the same boundary is almost the same. The number of the sampling points is chosen as 200 in this study. Then, all the sampled points from all the slices are stacked together according to their practical pixel distances during the CT scanning. As a result, the three-dimensional coordinates of all the sampled points are obtained. The index of the sampled points in each slice is aligned with each other in the stacking direction (z direction). Then, we get 200 curved lines in the stacking direction, and the number of the points on the curved line equals the stacking number N_{stack} . We resample the curved lines by the following spline interpolation filter:

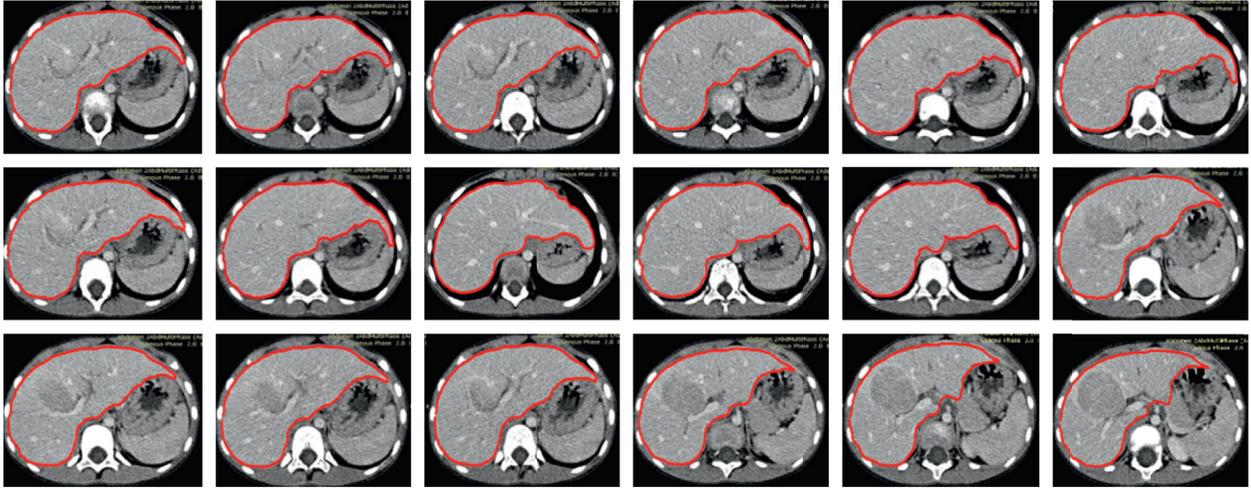


FIGURE 4: Segmentation results of some typical images in the first case, the proposed method.

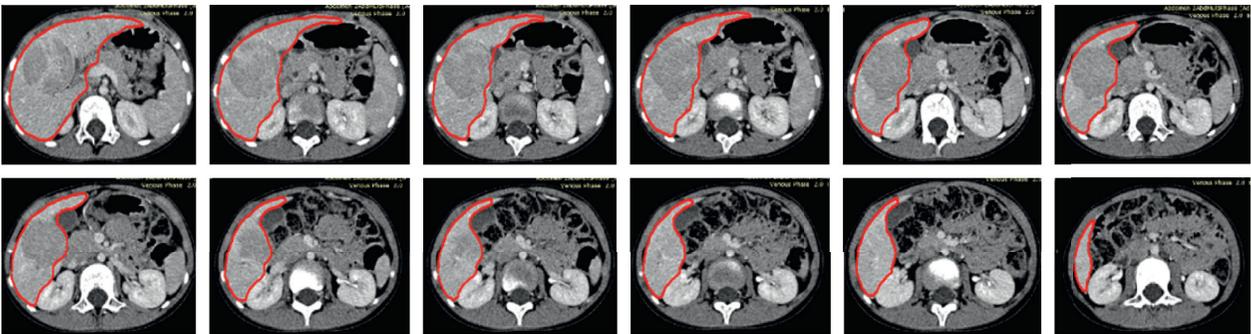


FIGURE 5: Segmentation results of some typical images in the second case, the proposed method.

$$P_f^f(m, n) = \underset{f}{\operatorname{argmin}} \left((1 - \alpha) \times \sum_{j=1}^{200} |P(m, j) - f(j)|^2 + \alpha \times \int \left| \frac{d^2 f(t)}{dt^2} \right|^2 dt \right), \quad (25)$$

where P_f^f denotes the smoothed point and P denotes the original sampled point. α is a smoothing factor and f is the fitted spline function ($m = 1, 2, \dots, N_{\text{stack}}$, $n = 1, 2, \dots, 200$).

3. Results and Discussion

Since the shapes of the liver and the surrounding organs also vary significantly in different slices, the number of the pixel classes in different slices varies accordingly. The organ that is adjacent to the liver is not fixed, and it might be the stomach, the kidney, or the heart. We thus define the liver segmentation into three cases depending on its adjacent organ. In the first case, there is only stomach adjacent to the liver. In the second case, there is kidney adjacent to the liver. In the third case, there is heart adjacent to the liver. In different cases, the number of the pixel classes is adjusted automatically based on the detected thresholds by the slope difference distribution. Then, the five core image processing algorithms are applied one by one. The optimal values of the input parameters are determined for different cases, respectively. After the optimal parameters are calibrated by trial and analysis, they will be

used for segmenting the livers in similar CT datasets. For the CT datasets with significant differences, the input parameters should be recalibrated again. As a result, it might require significant manual intervention to determine the optimal parameters for a specific-type CT dataset before the proposed method is run fully automatically for the whole CT dataset. The average time to process one CT image is 2.39 seconds in MATLAB with the i7-6700 CPU. Some typical liver segmentation results for the first case are shown in Figure 4. Some typical liver segmentation results for the second case are shown in Figure 5. Some typical liver segmentation results for the third case are shown in Figure 6. As can be seen, all the segmentation results are acceptable for clinical usage. With all the segmented two-dimensional liver from different slices, the whole three-dimensional liver is reconstructed by the method described in the above section. Figure 7 shows the reconstructed three-dimensional liver.

We use two public datasets to evaluate the proposed method further. The first public dataset (<https://eee.deu.edu.tr/moodle/mod/page/view.php?id=7872>) is liver transplantation donor database that provided by Emre Kavur from Dokuz Eylul University with the permission of Dokuz Eylul University Hospital. In this set, there are 20 upper abdominal CT image series that belongs to different patients and six sets are training sets. Only ground truths of the

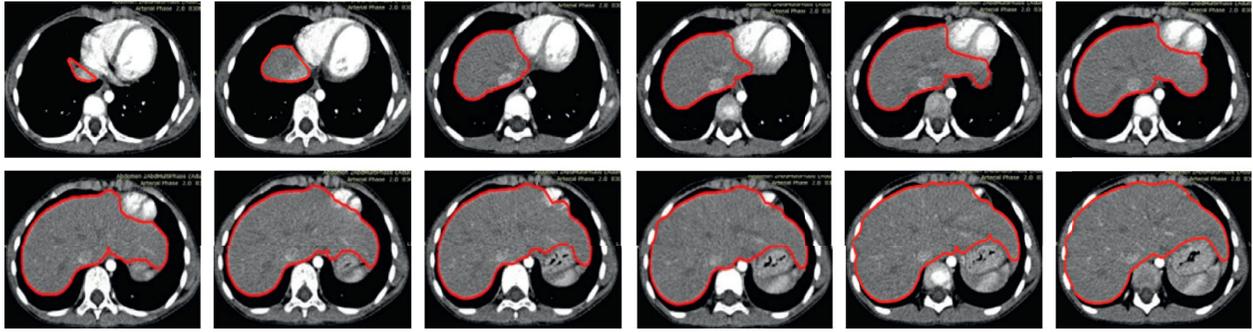


FIGURE 6: Segmentation results of some typical images in the third case, the proposed method.

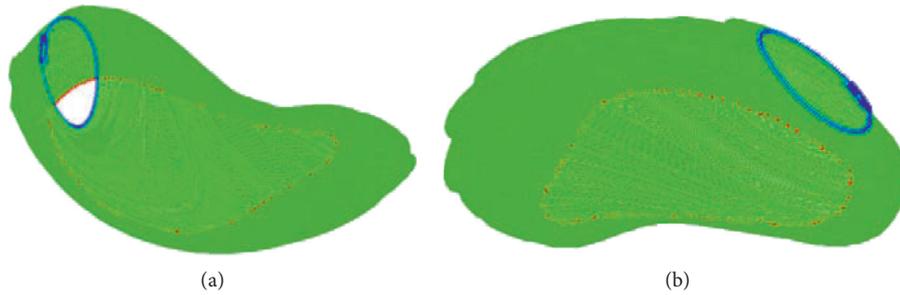


FIGURE 7: The reconstructed three-dimensional liver: (a) view 1; (b) view 2.

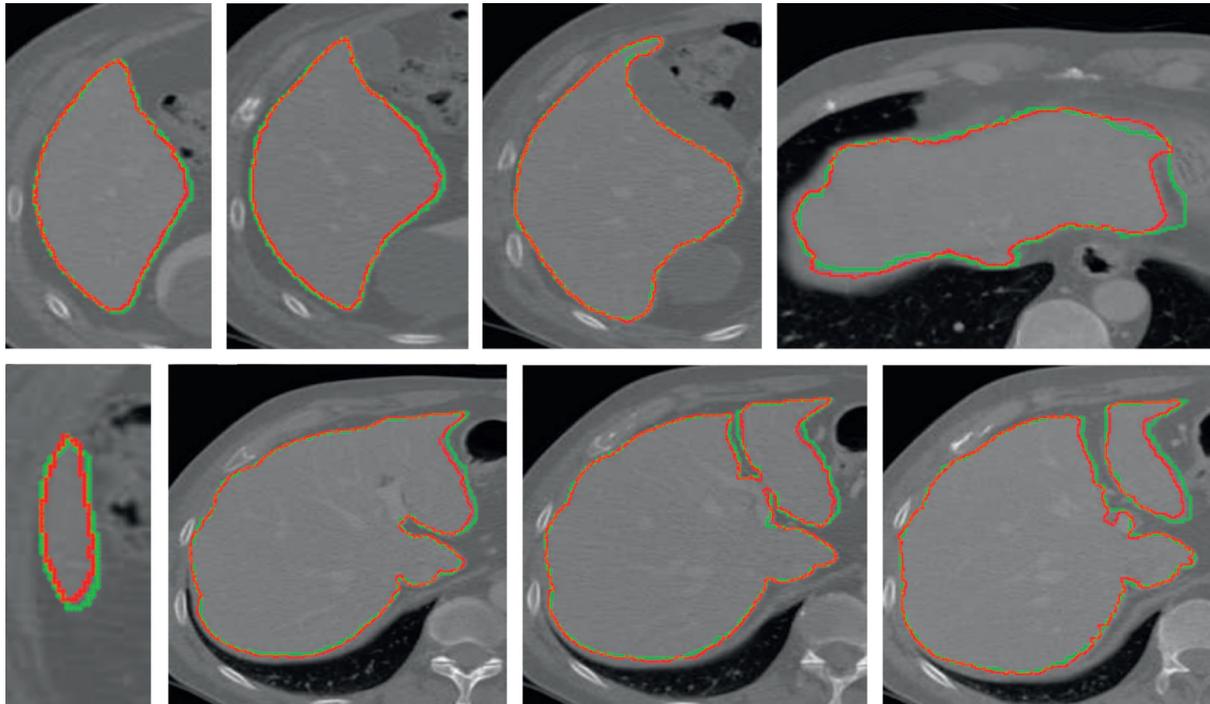


FIGURE 8: Segmentation results of some typical images from the first public dataset (the red contour denotes the automatically identified boundary, and the green contour denotes the manually identified boundary).

training sets are provided. Since the proposed method does not require training, we use the provided training sets to evaluate our method. The computed mean ratios of the volumetric overlap error (VOE), relative volume difference (RVD), average symmetric surface distance (ASD), root

mean square symmetric surface distance (RMSD), and maximum symmetric surface distance (MSSD) were $9.6 \pm 2.2\%$, $4.2 \pm 2.5\%$, 1.7 ± 0.9 mm, 2.4 ± 1.1 mm, and 9.2 ± 3.1 mm, respectively. We also show some qualitative results in Figure 8. As can be seen, the automatic identified

TABLE 1: Comparison of the proposed method with state-of-the-art methods on the 3Dircadb1 dataset.

Methods	VOE (%)	RVD (%)	ASD (mm)	RMSD (mm)	MSSD (mm)
[22]	12.99 ± 5.04%	5.66 ± 5.59%	2.24 ± 1.08	NA	25.74 ± 8.85
[23]	NA	3.62 ± 5.50%	1.94 ± 1.10	4.47 ± 3.30	34.60 ± 17.70
[24]	8.74 ± 2.37%	2.41 ± 1.71%	1.45 ± 0.37	2.55 ± 0.59	26.91 ± 7.72
[25]	15.6 ± 4.3%	5.8 ± 3.5%	2.0 ± 0.9	2.9 ± 1.5	7.1 ± 6.2
Proposed	10.25 ± 4.21%	4.7 ± 3.37%	1.84 ± 0.93	2.67 ± 1.1	15.23 ± 9.67

boundary of the liver matches the manually identified boundary well.

The second public dataset, 3D-IRCADB01 (<https://www.ircad.fr/research/3d-ircadb-01/>), consists of 20 CT scans with corresponding ground truth provided by IRCAD, the French Research Institute against Digestive Cancer. We show the quantitative comparisons with state-of-the-art methods [22–25] in Table 1. As can be seen, the comparisons are favorable. The proposed method does not require any work for training as state-of-the-art methods [22–25] do while the achieved accuracy is similar or better.

The accuracy of the proposed method differs significantly across different CT liver datasets because the proposed method is heavily relying on the accuracy of threshold selection. Yet, the optimal accuracy of slope difference-based threshold selection could not be guaranteed currently during unsupervised segmentation. In the near future, we will research on algorithms to make the slope difference-based threshold selection method be capable of always selecting optimal thresholds without supervision.

4. Conclusions

In this paper, a heuristic approach is proposed to segment the liver in CT images fully automatically. It calculates multiple thresholds simultaneously based on the slope difference distribution and then segment the CT image into different meaningful regions with these automatically calculated thresholds. The liver is segmented robustly with the constraint of the surrounding organs or tissues segmented beforehand. To reduce the noise, Gibbs distribution is utilized to minimize the global energy. A morphological filter is proposed to remove the small interference blobs. A morphological merging method is also proposed to unite the divided parts. Experimental results verified all the proposed image processing algorithms and the proposed approach. Since the proposed approach is very efficient and does not require any training datasets, it is thus very promising for clinical usage.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

The authors thank the hospital at Qingdao University for providing the dataset.

References

- [1] X. Wang, Y. Zheng, L. Gan, X. Wang, X. Sang, and X. Kong, “Liver segmentation from CT images using a sparse priori statistical shape model (SP-SSM),” *PLoS One*, vol. 12, no. 10, Article ID e0185249, 2017.
- [2] A. Saito, S. Yamamoto, S. Nawano, and A. Shimizu, “Automated liver segmentation from a postmortem CT scan based on a statistical shape model,” *International Journal of Computer Assisted Radiology and Surgery*, vol. 12, no. 2, pp. 205–221, 2016.
- [3] T. Okada, R. Shimada, M. Hori et al., “Automated segmentation of the liver from 3D CT images using probabilistic atlas and multi-level statistical shape model,” *Academic Radiology*, vol. 15, no. 11, pp. 1390–1403, 2008.
- [4] H. Ji, J. He, X. Yang, R. Deklerck, and J. Cornelis, “ACM-based automatic liver segmentation from 3-D CT images by combining multiple atlases and improved mean-shift techniques,” *IEEE Journal of Biomedical and Health Informatics*, vol. 17, no. 3, pp. 690–698, 2013.
- [5] P. Hu, F. Wu, J. Peng, P. Liang, and D. Kong, “Automatic 3D liver segmentation based on deep learning and globally optimized surface evolution,” *Physics in Medicine and Biology*, vol. 61, no. 24, pp. 8676–8698, 2016.
- [6] F. Lu, F. Wu, P. Hu, Z. Peng, and D. Kong, “Automatic 3D liver location and segmentation via convolutional neural network and graph cut,” *International Journal of Computer Assisted Radiology and Surgery*, vol. 12, no. 2, pp. 171–182, 2016.
- [7] A. Zareei and A. Karimi, “Liver segmentation with new supervised method to create initial curve for active contour,” *Computers in Biology and Medicine*, vol. 75, pp. 139–150, 2016.
- [8] T. S. Avşar and S. Arica, “Automatic segmentation of computed tomography images of liver using watershed and thresholding algorithms,” in *IFMBE Proceedings*, H. Eskola, O. Väisänen, J. Viik, and J. Hyttinen, Eds., Vol. 65, Springer, Singapore, 2017.
- [9] J. J. Lee, N. Kim, H. Lee et al., “Efficient liver segmentation using a level-set method with optimal detection of the initial liver boundary from level-set speed images,” *Computer Methods and Programs in Biomedicine*, vol. 88, no. 1, pp. 26–38, 2007.
- [10] Q. Huang, H. Ding, X. Wang, and G. Wang, “Fully automatic liver segmentation in CT images using modified graph cuts and feature detection,” *Computers in Biology and Medicine*, vol. 95, pp. 198–208, 2018.
- [11] R. Beichel, A. Bornik, C. Bauer, and E. Sorantin, “Liver segmentation in contrast enhanced CT data using graph cuts

- and interactive 3Dsegmentation refinement methods,” *Medical Physics*, vol. 39, no. 3, pp. 1361–1373, 2012.
- [12] M. Liao, Y. Q. Zhao, W. Wang et al., “Efficient liver segmentation in CT images based on graph cuts and bottleneck detection,” *Physica Medica*, vol. 32, no. 11, pp. 1383–1396, 2016.
- [13] Y. Chen, W. Zhao, Q. Wu, Z. Wang, and J. Hu, “Liver segmentation in CT images for intervention using a graph-cut based model,” in *Lecture Notes in Computer Science*, H. Yoshida, G. Sakas, and M. G. Linguraru, Eds., Vol. 7029, Springer, Berlin, Heidelberg, Germany, 2012.
- [14] A. H. Foruzan, Z. R. Aghaeizadeh, M. Hori, and Y. Sato, “Liver segmentation by intensity analysis and anatomical information in multi-slice CT images,” *International Journal of Computer Assisted Radiology and Surgery*, vol. 4, no. 3, pp. 287–297, 2009.
- [15] M. A. Selver, A. Kocaoğlu, G. K. Demir et al., “Patient oriented and robust automatic liver segmentation for pre-evaluation of liver transplantation,” *Computers in Biology and Medicine*, vol. 38, no. 7, pp. 765–784, 2008.
- [16] A. H. Foruzan, R. A. Zoroofi, M. Hori, and Y. Sato, “A knowledge-based technique for liver segmentation in CT data,” *Computerized Medical Imaging and Graphics*, vol. 33, no. 8, pp. 567–587, 2009.
- [17] Z. Z. Wang, J. J. Xiong, Y. M. Yang, and H. X. Li, “A flexible and robust threshold selection method,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 28, no. 9, pp. 2220–2232, 2018.
- [18] Z. Z. Wang, “A new approach for segmentation and quantification of cells or nanoparticles,” *IEEE Transactions on Industrial Informatics*, vol. 12, no. 3, pp. 962–971, 2016.
- [19] Z. Z. Wang, “Image segmentation by combining the global and local properties,” *Expert Systems with Applications*, vol. 87, pp. 30–40, 2017.
- [20] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, Prentice Hall, Upper Saddle River, NJ, USA, 3rd edition, 2008.
- [21] G. Sciortino, D. Tegolo, and C. Valenti, “Automatic detection and measurement of nuchal translucency,” *Computers in Biology and Medicine*, vol. 82, pp. 12–20, 2017.
- [22] F. Chung and H. Delingette, “Regional appearance modelling based on the clustering of intensity profiles,” *Computer Vision and Image Understanding*, vol. 117, no. 6, pp. 705–717, 2013.
- [23] M. Erdt and M. Kirschner, “Fast automatic liver segmentation combining learned shape priors with observed shape deviation,” in *Proceedings of IEEE 23rd International Symposium on Computer-Based Medical Systems (CBMS2010)*, pp. 249–254, Perth, Australia, October 2010.
- [24] C. F. Shi, Y. Z. Cheng, F. Liu, Y. D. Wang, J. Bai, and S. Tamura, “A hierarchical local region-based sparse shape composition for liver segmentation in CT scans,” *Pattern Recognition*, vol. 50, pp. 88–106, 2016.
- [25] C. Sun, S. Guo, H. Zhang et al., “Automatic segmentation of liver tumors from multiphase contrast-enhanced CT images based on FCNs,” *Artificial Intelligence in Medicine*, vol. 83, pp. 58–66, 2017.

Research Article

A Computer-Aided Pipeline for Automatic Lung Cancer Classification on Computed Tomography Scans

Emre Dandil 

Department of Computer Engineering, Faculty of Engineering, Bilecik Seyh Edebali University, Gulumbe Campus, 11210 Bilecik, Turkey

Correspondence should be addressed to Emre Dandil; emre.dandil@bilecik.edu.tr

Received 28 June 2018; Revised 24 September 2018; Accepted 8 October 2018; Published 1 November 2018

Guest Editor: Cesare Valenti

Copyright © 2018 Emre Dandil. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Lung cancer is one of the most common cancer types. For the survival of the patient, early detection of lung cancer with the best treatment method is crucial. In this study, we propose a novel computer-aided pipeline on computed tomography (CT) scans for early diagnosis of lung cancer thanks to the classification of benign and malignant nodules. The proposed pipeline is composed of four stages. In preprocessing steps, CT images are enhanced, and lung volumes are extracted from the image with the help of a novel method called lung volume extraction method (LUVEM). The significance of the proposed pipeline is using LUVEM for extracting lung region. In nodule detection stage, candidate nodules are determined according to the circular Hough transform-(CHT-) based method. Then, lung nodules are segmented with self-organizing maps (SOM). In feature computation stage, intensity, shape, texture, energy, and combined features are used for feature extraction, and principal component analysis (PCA) is used for feature reduction step. In the final stage, probabilistic neural network (PNN) classifies benign and malign nodules. According to the experiments performed on our dataset, the proposed pipeline system can classify benign and malign nodules with 95.91% accuracy, 97.42% sensitivity, and 94.24% specificity. Even in cases of small-sized nodules (3–10 mm), the proposed system can determine the nodule type with 94.68% accuracy.

1. Introduction

Nowadays, lung cancer is one of the ranking first causes of mortality worldwide among men and women [1, 2]. Although there are a lot of treatment options like surgery, radiotherapy, and chemotherapy, five year survival rate for patients is quite low [3]. However, survival rate may go up to 54% in case lung cancer is identified in an early stage [4]. Therefore, early detection of lung cancer is vital to decrease lung cancer mortality.

Medical imaging techniques have been important technology in screening of lung cancer recently. CT scan becomes a standard modality for detecting and assessing lung cancer [5]. Most of the lung nodules are usually benign. However, some nodules such as calcified, swollen, and hard can also be determined as benign. Similarly, a hard nodule generally is cancerous (malignant), but it may be considered as benign case in some cases [6]. Furthermore, medical CT images are needed to be diagnosed by radiologists.

Computer-aided detection (CAD) systems have been an important field in medical image processing. CAD systems also based on machine learning methods designed to diagnosis of cancer have become common in recent years. Radiologists and physicians may use findings of CAD systems as the second opinion before making their own final decisions. Therefore, CAD systems play an important role in CT scans to help radiologists for detection of lung cancer efficiently.

2. Related Work

Computer-aided detection (CAD) systems have been active research field for the pulmonary nodule detection and malign/benign nodule classification. Until now, many CAD systems have been proposed. For example, Ozekes and Camurcu proposed a method for pulmonary nodule detection method using template matching [7]. Schilham et al. presented a CAD system which consists of image

preprocessing, candidate nodule detection, feature extraction, and classification for nodule detection in chest radiographs [8]. Dehmeshki et al. detected lung nodules using shape-based genetic algorithm template matching [9]. Suarez-Cuenca et al. also designed a system which discriminates the nodules and non-nodule cases using iris filter in CT images [10]. Murphy et al. automatically performed lung nodule detection using k-nearest neighbours classifier [11]. Giger et al. realized CAD system to detect lung nodules on CT images using geometric features [12]. In addition, Hasegawa et al. proposed image processing methods for identification of lung nodules using CT scans [13]. In another study, Kanazawa et al. used a CAD system to identify pulmonary nodules with fuzzy features [14]. In 2005, Suzuki et al. proposed a method using ANN for classification of malignant and benign nodules on CT images [15]. Sun et al. compared support vector machines (SVM) with the some classification methods for detection of lung cancer on CT images [16]. Kuruvilla and Gunavathi proposed a system using ANNs for classification of lung cancer [17].

In a recent study on lung nodule detection, Javaid et al. proposed a computer aided nodule detection method for the segmentation and detection of challenging in different type nodules [18]. ur Rehman et al. presented a systematic analysis of nodules detection techniques with the current trends and future challenges [19]. Wang et al. proposed a pulmonary nodule CAD based on semisupervised extreme learning machine [20]. Xie et al. proposed an automated pulmonary nodule detection system with 2D convolutional neural network (CNN) on LUNA16 dataset [21].

In this study, we have proposed fully automated computer-aided pipeline for the detection of pulmonary nodules and classification of benign/malign nodules in early stage. The contributions of this paper are (1) to review the systematic literature review; (2) to present the state of the art detection of pulmonary nodules and classification of lung cancer; (3) to propose the novel preprocessing method (LUVEM) for the lung volume extraction; (4) to suggest a novel candidate lung nodule detection method using CHT; (5) to design a holistic pipeline for the detection of pulmonary nodules as well as lung cancer; (6) the detailed comparison of feature extraction methods for lung nodule detection; and (7) to perform the detailed performance evaluation, high true detection rate, and low false detection rate for nodule detection and classification.

3. Architecture of the Computer-Aided Pipeline

Designed pipeline consists of four main stages such as image preprocessing (Stage I), lung nodule detection (Stage II), nodule feature computation (Stage III), and nodule classification (Stage IV). The work flow of the pipeline is presented in Figure 1.

3.1. Lung Image Preprocessing. In the first step of the image preprocessing stage, reading of CT images is performed. The CT scans obtained for the work are stored as DICOM

(Digital Imaging and Communications in Medicine) files [22].

The goal of image enhancement step is to prevent misleading results that may occur in subsequent processes. Thus, we firstly implemented the median filter to remove unnecessary noises and enhance the images. Moreover, the sharpening of nodule contours is an important step for the detection of nodules. Laplacian filter was used in our study. So, nodules on lung region were able to be detected more accurately. Furthermore, histogram equalization was also used in enhancement step in order to minimize contrast differences which occur due to scanning errors and to remove unnecessary grains.

In lung volume extraction step of image preprocessing stage, extracting of the lung region from CT image is performed. There are many methods for extracting lung volume from a lung CT scan [9, 10, 23–25]. However, these methods are complex, and they require more processing overhead. In some cases, these methods may lead to losses of information about lung regions or cause noise. The purpose of this step is to extract the lung region completely from the full lung CT image. Therefore, a simple but effective and novel method has been proposed in this study for lung volume extraction named as lung volume extraction method (LUVEM). The pseudocode of LUVEM is shown in Algorithm 1.

In LUVEM method, lung lobes are extracted from CT images with the help of morphological operations. LUVEM removes unrelated segments on the sides and edges of the preprocessed image and obtain the lung region successful. In the algorithm, input image is firstly converted to double-formatted image. Afterwards, 1 or 0 values are assigned to each pixel of double-formatted image according to low and high threshold values. The low and high threshold values are determined 0.25 and 0.65 in this algorithm, respectively. The method removes the bright areas on the edges of the lung CT image since their average values change between low and high values. After this process, the image is converted to binary format and performed morphological operations which are eroding, dilating, and filling, respectively. Finally, the image is again converted to gray-scale format. The segmentation examples of LUVEM can be seen in Figure 2. It is clearly seen that LUVEM can successfully extract the lung volume. In addition, quantitative evaluation of LUVEM will be reported below.

3.2. Lung Nodule Detection. The first step of lung nodule detection is candidate nodule detection. The nodule candidates in volume should be detected before nodule segmentation. The lung volume includes vessels and nodules. Moreover, the density of nodules, vessels, and lungs is different from each other [26]. Since the lung nodules have a circular and helical structure, they can be differentiated by means of circularity determination. Many methods have been suggested for identifying the round objects. Circular Hough transform (CHT), which proposed by Duda et al. [27], is one of the most successful method [28] for detection of round objects on the images. In this study, CHT

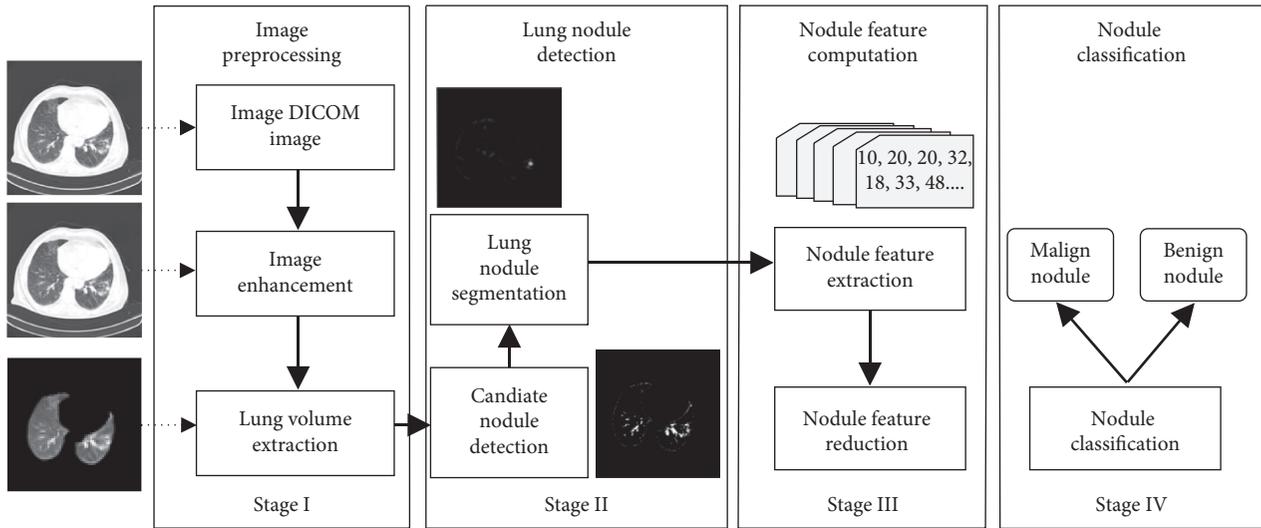


FIGURE 1: Work flow of the designed pipeline to detect lung cancer. The system consists of four stages: Stage I—enhancement of lung CT image and a novel lung volume segmentation method (LUVEM), Stage II—candidate nodule detection using CHT and segmentation of lung nodules using SOM, Stage III—computing of lung nodule features and reduction features using PCA, and Stage IV—classification of malign and benign lung nodule using PNN.

```

(1) procedure LungVolumeExtraction(input image, low, high) //beginning lung vol. extraction algorithm
(2)   I ← input image //input image
(3)   O ← ∅, Lo ← low, Hi ← high //output image, low and high threshold
(4)   DIm ← Converting (I, double) //converting image to double image
(5)   for (i ∈ row length of DIm) //starting loop of extraction step
(6)     for (j ∈ column length of DIm) //starting of inner loop
(7)       if DIm(i, j) ≥ Lo and DIm(i, j) ≤ Hi then //starting of if
(8)         DIm(i, j) ← 1; //assignment of binary 1
(9)       else
(10)        DIm(i, j) ← 0; //assignment of binary 0
(11)      end if //ending of if
(12)    end for //ending of inner loop
(13)  end for //ending loop of extraction step
(14)  BI ← Converting(DIm, binary) //converting image to binary image
(15)  FI ← Morphological eroding, dilating, filling (BI) //morphological operations
(16)  GIm ← Converting(FI, gray-scale) //converting image to gray-scale image
(17)  O ← GIm //return output image
(18) end procedure //ending of LUVEM method

```

ALGORITHM 1: The pseudocode of lung volume extraction method (LUVEM).

operations are used for candidate nodule detection. CHT can detect the round object in the image; moreover, it can also detect the noncircular object by means of some operations. The image dataset is divided into 3 categories according to the nodule size such as <10 mm, 10–20 mm, and >20 mm. In order to detect the nodules in different size by CHT, three minimum and maximum radiuses such as 3–12 mm, 10–20 mm, and 15–45 mm are determined. In Figure 3, it is shown that the examples of determination of candidate nodules on CT images.

The second step of lung nodule detection is nodule segmentation. In this study, SOM [29] is proposed to segment nodules on CT images. SOM is an unsupervised neural network learning [30] method. It can perform on

large/complex datasets [31, 32]. Furthermore, it designs data maps that can be interpreted easily. In addition to these advantages, SOM can easily segment very small nodules on the lung CT images [3]. The examples of segmented lung nodule images using SOM are shown in Figure 4.

3.3. Nodule Feature Computation. Generally, CAD systems segment lung nodules for the determination of nodule candidates, and then features extract from the candidate nodules. The popular features are geometric feature, gray level features, gradient features, and energy level features. Therefore, we extracted 2D significant features from lung

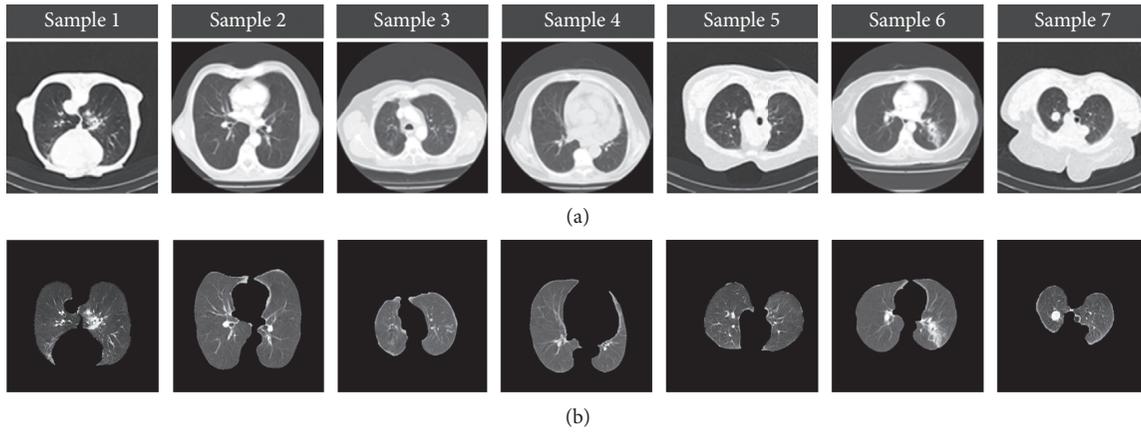


FIGURE 2: Extraction examples of lung regions: (a) preprocessed images and (b) lung volume extraction using LUVEM.

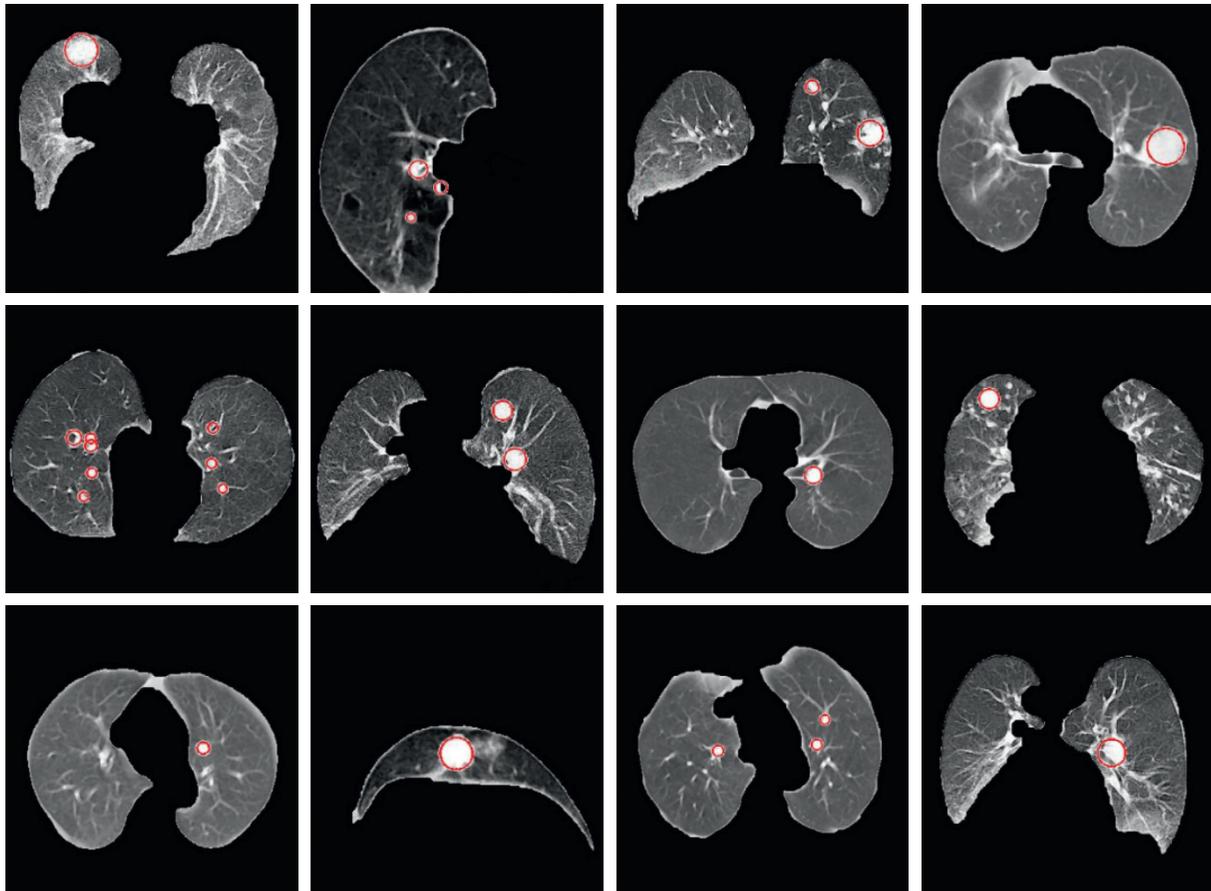


FIGURE 3: Determination of candidate lung nodule using CHT.

CT images to discriminate benign or malign nodule. Firstly, we used shape-based features for analyzing nodule geometry. We used first-order statistical features to obtain global statistic about nodule region. Moreover, we utilized gray level co-occurrence matrix (GLCM) texture features for gray level statistic of nodules. Finally, we extracted wavelet decomposition transform features to obtain the energy feature of nodules. All computed features are extracted from the slice of the segmented object.

First-order statistical features (SSF) of an image are calculated from the gray level histogram values of an image [33]. In this study, 6 basic features such as standard deviation, entropy, means, skewness, kurtosis, and variance were extracted by SSF using the histogram values of a gray level lung CT image. Shape-based features (SBF) allow feature extraction from the image by using geometric parameters [34]. Shape features give some information about an image such as sharpness, circularity, and convexity. In

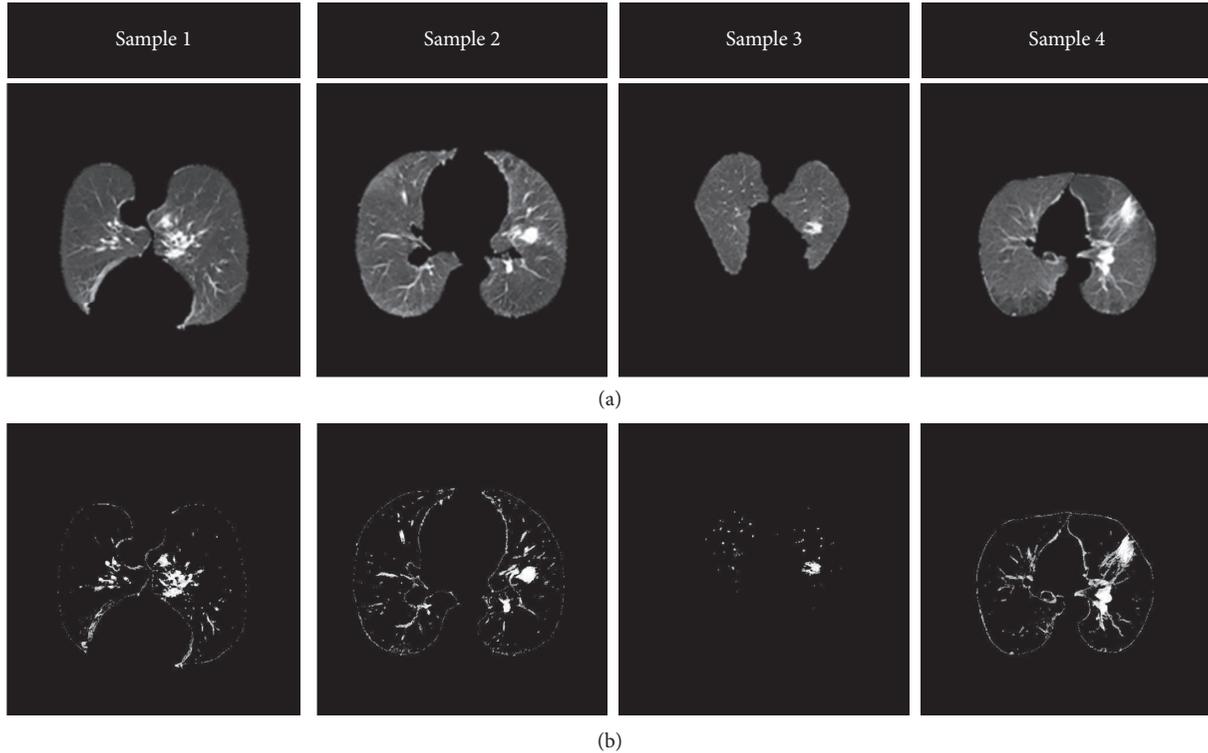


FIGURE 4: Examples of segmented lung nodules: (a) images of extracted lung volume and (b) segmentation of lung nodule using SOM.

this study, a total of 16 shape features were extracted to facilitate the determination of nodule type from CT lung images. Statistical features of a gray level image (GTF) of a texture are first derived with the help of GLCM texture features proposed by Haralick [35, 36]. GLCM method shows the relationship between pixels of different gray level and is widely used in applications of medical image processing. In this study, a total of 88 features were extracted with GLCM from 0° , 45° , 90° , and 135° angle directions in $d = 2$ distance. Wavelet decomposition transform can denote distribution of energy features of different regions (TEF). ROI of the CT image is divided into four subbands with 2D wavelet decomposition. Three images are created in low frequencies, and an image is created in high frequencies with wavelet decomposition transform from an image [37]. In this study, 13 energy features of an image are extracted with wavelet decomposition. The number of features extracted by each feature extraction method used in this study is presented in Table 1.

On lung CT images, malign nodules are generally more complex and irregular, while benign nodules are rounder with certain borders. Most of the benign nodules have small variance values. However, malign nodules show relatively higher variance values [3]. Figure 5 shows the examples of benign and malign lung nodules on CT images.

Since 123 features extracted are rather large in size, they may negatively affect accuracy during classification step. Thus, selecting the most appropriate features instead of using all features will be a more efficient method. We used PCA method for dimension reduction of feature vector.

TABLE 1: The number of extracted 2D features from lung CT images.

Feature extraction method	Number of feature	Order
SSF	6	0–6
SBF	16	7–22
GTF	$22 * 4 = 88$	23–110
TEF	13	111–123

PCA is used to reduce dimensionality of large dataset [38, 39]. We can select a number of features only up to one-third of the number of data (patterns) in the smaller of the two classes. Thus, for our work, the smallest class has 104 patterns (benign nodules), and since we split the data to half, one-third of 52 is around 17. Therefore, we selected with PCA the most appropriate 17 features (components) from 123 features. Figure 6 denotes principal component analysis of extracted features with cumulative variance. As can be seen from the chart in Figure 6, it is seen that the variance of the first 20 components is more selective.

3.4. Nodule Classification. In the proposed pipeline, we have used a probabilistic neural network (PNN) model to make automated decision about the nodule types (benign or malignant). PNN is an effective tool for many classification implementations and can easily make classifications [40, 41]. Figure 7 presents the architecture of the PNN designed for this study. Neuron number in the input layer is selected 17 according to the number of inputs.

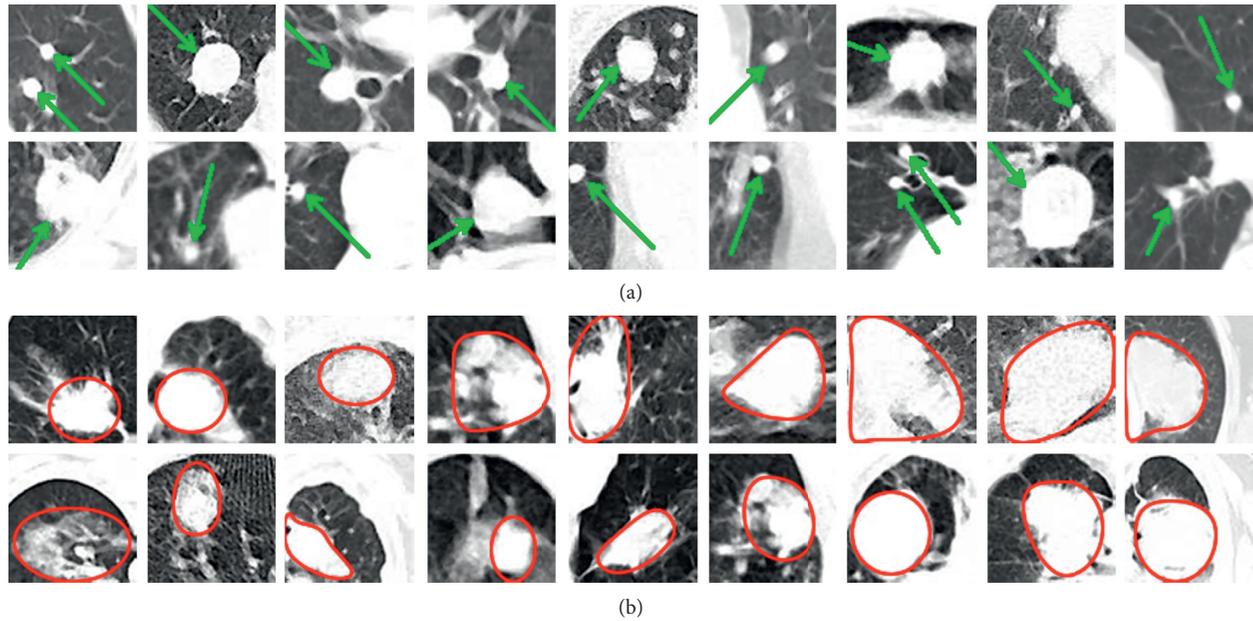


FIGURE 5: Examples of benign (a) and malign (b) lung nodules.

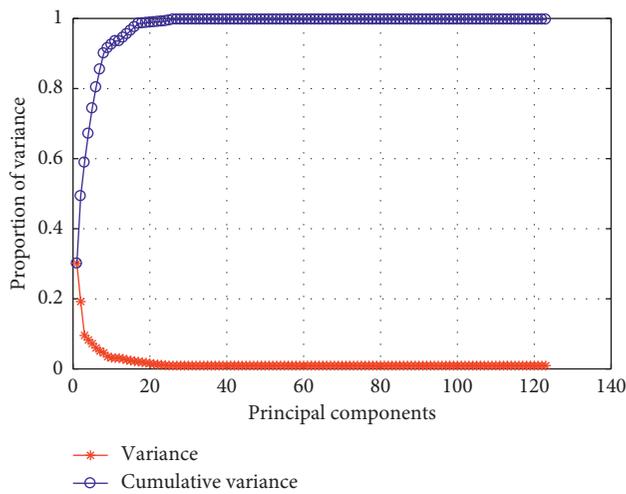


FIGURE 6: Principal component analysis of extracted features with cumulative variance.

4. Experimental Results

In this work, we realized all experiments using a PC with i7 processor, 16 GB memory, and Windows 10. Moreover, MATLAB software was used for performance evaluation of the proposed pipeline. In all experiments, leave-one-out cross validation technique was run at the level of nodule. So, all of 220 nodules (104 benign and 116 malignant) were used for both trainings and tests. Figure 8 summarizes the processing steps of the proposed pipeline.

4.1. CT Lung Dataset. In this study, an image dataset was prepared for the proposed pipeline. CT examinations were realized by using a helical CT scanner from Sincan Nafiz

Korez Hospital. Its acquisition parameters are slice collimation 1.0 mm and slice width 1 mm. Scans were acquired in 130 kV and 75 mAs. The size of the images was 512×512 pixels. The images were stored as DICOM format files. The database consists of 47 CT scans from 47 different patients. 35 of volunteer patients are male and remaining of them are female. Their ages are between 30 and 79 (mean 58.7 ± 10.5 years). All patients agreed that they have a legal and moral right to accurate and reliable information for the scan. These patients should be given clearly the diagnosis and prognosis with a simple language. There are a total of 9504 CT modality images in the database, and the number of CT slices per scan varies between 116 and 283. After the CT scan, the physician provided the selection of the slice where the nodule is fully visible. 1128 ROI, which includes a total of 220 nodules (104 benign and 116 malignant), were selected from 9504 CT images with the help of a lung physician and three experienced radiologists in the lung parenchyma. This process has been conducted by means of an annotation tool. The nodules were also approved by biopsy. Sizes of nodules change from 3 to 65 mm in diameter. The size distribution of the nodules is shown in Figure 9.

4.2. Validation of LUVEM with Evaluation Metrics. Proposed lung volume extraction method (LUVEM) in this study is compared with the standard manual segmentations using measurement metrics. We evaluate manual segmentations of the expert and automated segmentations of LUVEM using two popular overlap measures. We used a segmentation software tool developed by us for manual segmentation on the dataset. The software tool outlines edges automatically, presenting us to obtain contours of the nodule boundaries. The metrics evaluate the overlapping between the two sets. The first overlap metric, represents the

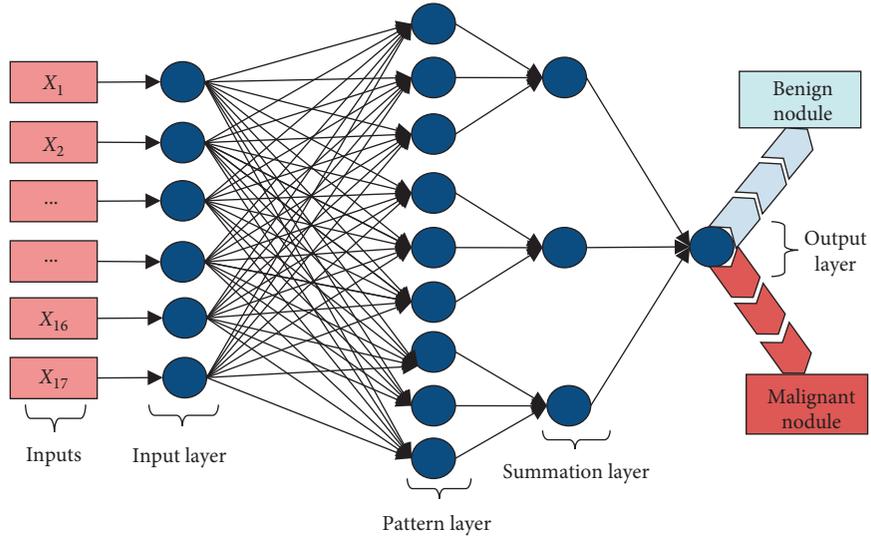


FIGURE 7: Probabilistic neural network architecture used in the proposed method for nodule classification.

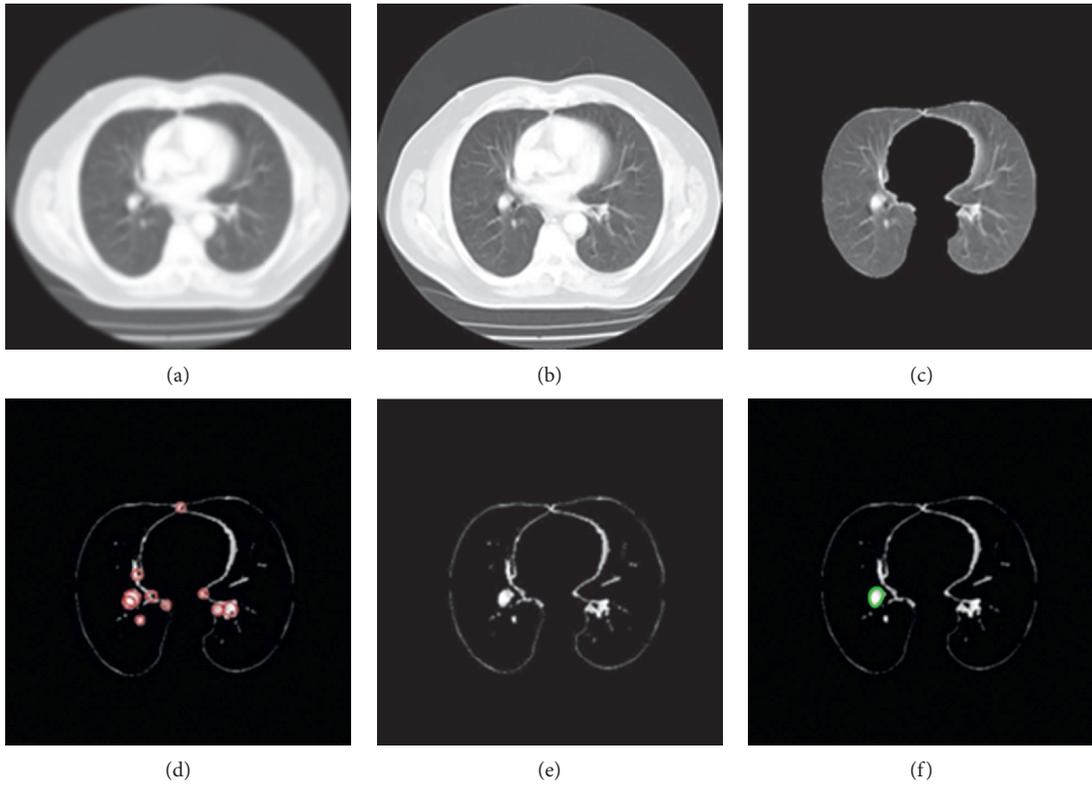


FIGURE 8: Processing steps of the proposed pipeline: (a) original DICOM image; (b) image preprocessing and enhancement; (c) lung volume extraction from CT scan; (d) detection of candidate nodules; (e) segmentation of nodules; (f) classification of nodules.

Jaccard coefficient (union overlap), defined as intersection over manual and automatic segmentations and measures the similarity of the S_1 and S_2 sets [42]. Our second overlap metric, the Dice coefficient (mean overlap), gives double the weight to agreements between the two sets [43]. Jaccard and Dice metrics are denoted in the following equations:

$$\begin{aligned} \text{Jaccard} &= \frac{|S_1 \cap S_2|}{|S_1 \cup S_2|}, \\ \text{Dice} &= \frac{2|S_1 \cap S_2|}{|S_1| + |S_2|}. \end{aligned} \tag{1}$$

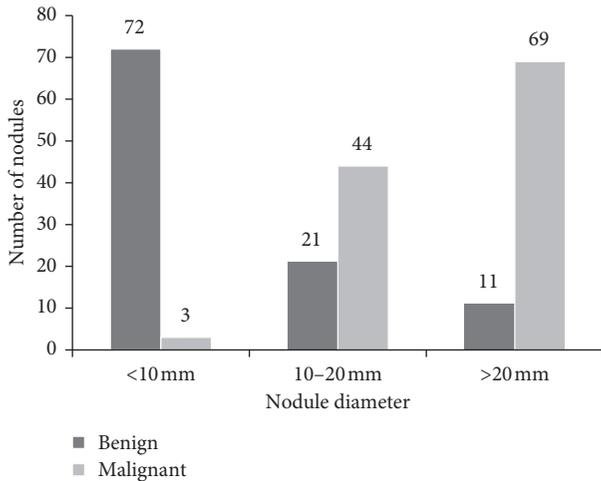


FIGURE 9: Size distribution of benign and malignant nodules in the image dataset.

We show the overlap metrics (Jaccard and Dice) that result from both LUVEM and Otsu's method. These results are the comparison of automated segmentations of LUVEM and Otsu methods with manual segmentation on 254 lung CT image in our database. The results in Table 2 showed LUVEM is higher in Jaccard overlap (0.867) and Dice overlap (0.938) than Otsu's method.

4.3. Detection Rates. Confusion matrixes of classification results with PNN according to each feature extraction and PCA method are presented in Table 3. As shown in Table 3, the usage of PCA affects the detection performance of the pipeline positively. Moreover, the usage of combined features extraction methods with PCA gives best success rate.

Table 4 presents the values of performance criteria obtained in the classification results of the proposed pipeline when feature extraction methods were used separately and together. According to the table, performance values are more successful when all feature extraction methods are used together. Accuracy (Acc) was found to be 92.27% when 123 features were used in classification without feature selection through PCA, and this rate was found to be 95.91% with the use of PCA. Similarly, more successful results were obtained in sensitivity (Sen), specificity (Spc), positive decision value (PDV), negative decision value (NDV), and F1 score criteria as presented in Table 4.

Since our CT image database was divided into 3 groups according to the size of nodules such as <10, 10-20, and >20 mm, we also realized a performance evaluation according to size group of the nodules. These experiments were realized with all together feature extraction method using PCA. Table 5 presents the result of detection performance depending upon nodule size. As shown in Table 5, the proposed pipeline can classify even small nodules with high success rates. Overall detection result of proposed pipeline according to the nodule size is 95.91%.

Receiver operator characteristic (ROC) curve is another popular performance evaluation criteria used in detection

TABLE 2: The Jaccard and Dice metrics measures for LUVEM.

	Jaccard overlap	Dice overlap
Otsu's method	0.587 ± 0.093	0.786 ± 0.088
LUVEM	0.867 ± 0.051	0.938 ± 0.032

TABLE 3: Confusion matrixes for feature extraction methods.

FE method	Classification results without PCA				Classification results with PCA			
	TP	FP	FN	TN	TP	FP	FN	TN
SSF	90	22	26	82	90	22	26	82
SBF	105	17	11	87	109	12	7	92
GTF	101	19	15	85	110	11	6	93
TEF	109	15	7	89	111	11	5	93
All FE methods (combined)	111	12	5	92	113	6	3	98

systems. Area under an ROC curve is measured according to sensitivity and specificity values of system. This area shows how the system is successful. Therefore, we also present ROC curve of our proposed detection system. Figure 10 shows ROC curve of the system obtained classification results of each lung nodule group and overall system. As seen in this graphic, area under ROC curve and true positive rate of small size nodules are lower than big size nodules. Here, as can be seen from this figure, if the nodule size is too large and too small, the success rate decreases.

Processing time is another performance criterion that we have used for the evaluation of the proposed pipeline. Longest time is needed for nodule detection step due to the use of SOM method for segmentation. Since SOM is an ANN model, it has a lot of time-consuming mathematical operations. In average, classification of a CT image as benign or malignant takes 2-3 seconds approximately. It can be accepted as a reasonable time period when it is compared with the time it needs for a physician to make decisions.

5. Conclusions and Discussion

In this study, a fully automated pipeline was proposed to classify benign and malign lung nodules on CT images. By means of the designed pipeline, nodule detection as well as benign/malign distinction was performed with high accuracy, sensitivity and specificity rates. Moreover, it was designed a preprocessing method called LUVEM for extracting the lung volume from CT images. SOM method was used to allow successful detection of lung nodules in early stages. According to the detailed experiment performed on large dataset with combined features, the proposed pipeline can differentiate benign/malign nodules with high accuracy rates such as 94.68% (3-10 mm), 96.92% (10-20 mm), and 96.25% (>20 mm) using PNN. The proposed pipeline can be used by the physicians as a supplementary tool for benign and malign nodule classification.

We evaluated the performance of the pipeline on Lung Imaging Database Consortium-Image Database Resource

TABLE 4: Overall performance results of proposed pipeline.

Performance criteria	Classification results without PCA					Classification results with PCA				
	SSF	SBF	GTF	TEF	All	SSF	SBF	GTF	TEF	All
Acc	78.18	87.27	84.55	90.00	92.27	78.18	91.36	92.27	90.00	95.91
Sen	77.57	90.52	87.07	93.97	95.67	77.57	93.97	94.83	95.67	97.42
Spc	78.85	83.65	81.73	85.58	88.46	78.85	88.46	89.42	89.42	94.24
PDV	80.36	86.07	84.17	87.90	90.24	80.36	90.08	90.91	90.98	94.96
NDV	79.93	88.78	85.00	92.71	94.85	79.93	92.93	93.94	94.90	97.03
F1	0.79	0.88	0.85	0.92	0.93	0.79	0.92	0.93	0.94	0.96

TABLE 5: Assessment of performance measurement criteria according to nodule size.

Nodule size (mm)	The number of nodule	Confusion matrix				Performance criteria					
		TP	FP	FN	TN	Acc	Sen	Spc	PDV	NDV	F1
<10	75	3	4	0	68	94.67	100	94.45	42.86	100	0.60
10–20	65	43	1	1	20	96.92	97.73	95.24	97.73	95.24	0.98
>20	80	67	1	2	10	96.25	97.10	90.91	98.53	83.34	0.98
Overall	220	113	6	3	98	95.91	97.42	94.24	94.96	97.03	0.96

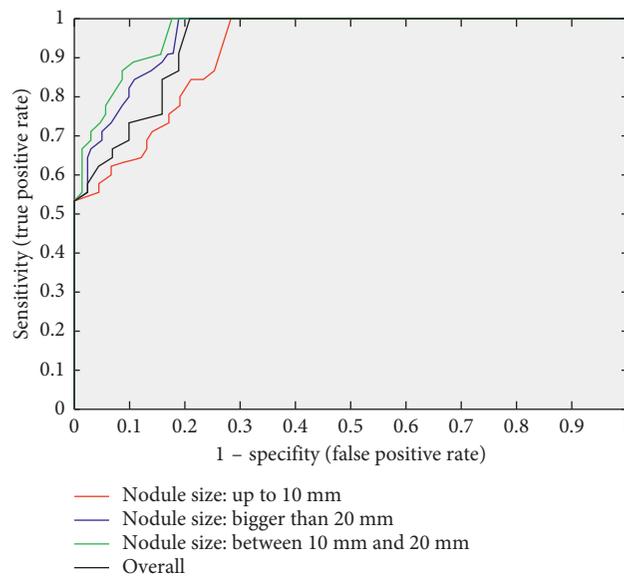


FIGURE 10: ROC curve of classification precision in proposed pipeline in different nodule diameter.

TABLE 6: The performance evaluation of proposed pipeline on LIDC-IDRI.

The number of nodules	Confusion matrix				Performance criteria					
	TP	FP	FN	TN	Acc	Sen	Spc	PDV	NDV	F1
38	22	2	4	10	84.21	84.62	83.33	91.67	71.43	0.88

Initiative (LIDC-IDRI) as well [50]. LIDC-IDRI dataset is the largest publicly available reference database for detection of lung nodules. We choose LIDC-IDRI dataset since it contains almost all the related information for lung CT including annotations on nodule sizes, locations, diagnosis results, and other information. We collected a total of 38 lung nodules from the dataset, including 26 malignant and 12 benign nodules. According to the evolution results on the proposed pipeline, accuracy obtained 84.21% using all FE

methods and PCA. In this test, F1 score result was found as 0.88. The obtained performance evaluation values of proposed pipeline on LIDC-IDRI dataset are denoted in Table 6.

There are some advantages of the proposed pipeline compared to the state-of-the-art systems. Firstly, the proposed pipeline has two diagnosis possibilities. It can perform nodule detection together with nodule classification. Second advantage of the proposed pipeline is to provide the detection of small nodules in the lung with the use of SOM

TABLE 7: The comparison of our pipeline with previously published CADs.

CAD system	CT image database	Number of cases	Nodule size(mm)	Sensitivity (%)	Average FPR
Dehmenski et al. [9]	Their own database	70	3–20	90.0	14.6
Suarez-Cuenca et al. [10]	Their own database	22	4–27	80.0	7.7
Opfer and Wiemeker [46]	LIDC database [47, 48, 50]	93	≥4	74.0	4
Rubin et al. [51]	Their own database	20	≥3	76	3
Sahiner et al. [49]	LIDC database [47, 48, 50]	48	3–36.4	79	4.9
Messay et al. [24]	LIDC database [47, 48, 50]	84	3–30	82.66	3
Suzuki et al. [52]	Their own database	101	8–20	80.3	16.1
Park et al. [53]	Their own database	38	Indefinite	80	–
Choi and Choi [23]	LIDC database [47, 48, 50]	32	3–30	94.1	5.45
Choi and Choi [44]	LIDC database [47, 48, 50]	58	3–30	95.28	2.27
Proposed method	Our database	47	3–35	97.42	4.54

method during segmentation step. This is remarkable in terms of early detection of lung cancer. Third advantage of the proposed pipeline is to have relatively high detection performance. Accuracy, sensitivity, and specificity of the system were calculated as 95.91%, 97.42%, and 94.24%, respectively. It is fairly difficult to compare formerly reported CAD systems due to different datasets, nodule types, sizes, and validation methods. We picked out some CAD systems to compare their performances. Some of them [23, 24, 44–46] used the LIDC database [47–49], and the other used their own databases. Table 7 denotes the comparison of the proposed pipeline with some CAD systems. When the results are analyzed, our pipeline has high sensitivity on our CT image dataset.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The author declares that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

The author would like to thank representatives of Sincan Nafiz Korez Hospital for creating the dataset.

References

- [1] R. Siegel, D. Naishadham, and A. Jemal, “Cancer statistics,” *CA: A Cancer Journal for Clinicians*, vol. 62, no. 1, pp. 10–29, 2012.
- [2] C. Jacobs, E. M. Van Rikxoort, T. Twellmann et al., “Automatic detection of subsolid pulmonary nodules in thoracic computed tomography images,” *Medical Image Analysis*, vol. 18, no. 2, pp. 374–84, 2014.
- [3] E. Dandil, M. Cakiroglu, Z. Eksi, M. Ozkan, O. K. Kurt, and A. Canan, “Artificial neural network-based classification system for lung nodules on computed tomography scans,” in *Proceedings of 2014 6th International Conference of Soft Computing and Pattern Recognition (SoCPaR)*, pp. 382–86, Tunis, Tunisia, August 2014.
- [4] N. Howlader, A. Noone, M. Krapcho et al., *SEER Cancer Statistics Review, 1975-2011*, National Cancer Institute, Bethesda, MD, USA, 2014.
- [5] H. Han, L. Li, F. Han, B. Song, W. Moore, and Z. Liang, “Fast and adaptive detection of pulmonary nodules in thoracic CT images using a hierarchical vector quantization scheme,” *IEEE Journal of Biomedical and Health Informatics*, vol. 19, no. 2, pp. 648–659, 2015.
- [6] Y. J. Jeong, C. A. Yi, and K. S. Lee, “Solitary pulmonary nodules: detection, characterization, and guidance for further diagnostic workup and treatment,” *American Journal of Roentgenology*, vol. 188, no. 1, pp. 57–68, 2007.
- [7] S. Ozekes and A. Y. Camurcu, *Automatic Lung Nodule Detection Using Template Matching*, Springer, Berlin, Germany, 2006.
- [8] A. M. Schilham, B. Van Ginneken, and M. Loog, “A computer-aided diagnosis system for detection of lung nodules in chest radiographs with an evaluation on a public database,” *Medical Image Analysis*, vol. 10, no. 2, pp. 247–58, 2006.
- [9] J. Dehmeshki, X. Ye, X. Lin, M. Valdivieso, and H. Amin, “Automated detection of lung nodules in CT images using shape-based genetic algorithm,” *Computerized Medical Imaging and Graphics*, vol. 31, no. 6, pp. 408–17, 2007.
- [10] J. J. Suarez-Cuenca, P. G. Tahoces, M. Souto et al., “Application of the iris filter for automatic detection of pulmonary nodules on computed tomography images,” *Computers in Biology and Medicine*, vol. 39, no. 10, pp. 921–33, 2009.
- [11] K. Murphy, B. Van Ginneken, A. M. Schilham, B. J. De Hoop, H. A. Gietema, and M. Prokop, “A large-scale evaluation of automatic pulmonary nodule detection in chest CT using local image features and k-nearest-neighbour classification,” *Medical Image Analysis*, vol. 13, no. 5, pp. 757–70, 2009.
- [12] M. L. Giger, K. T. Bae, and H. Macmahon, “Computerized detection of pulmonary nodules in computed tomography images,” *Investigative Radiology*, vol. 29, no. 4, pp. 459–65, 1994.
- [13] J. Hasegawa, K. Mori, J. Toriwaki, H. Anno, and K. Katada, “Automated extraction of lung cancer lesions from multi-slice chest CT images by using three-dimensional image processing,” *Systems and Computers in Japan*, vol. 25, no. 11, pp. 68–77, 1994.
- [14] K. Kanazawa, Y. Kawata, N. Niki et al., “Computer-aided diagnosis for pulmonary nodules based on helical CT images,” *Computerized Medical Imaging and Graphics*, vol. 22, no. 2, pp. 157–67, 1998.
- [15] K. Suzuki, F. Li, S. Sone, and K. Doi, “Computer-aided diagnostic scheme for distinction between benign and

- malignant nodules in thoracic low-dose CT by use of massive training artificial neural network," *IEEE Transactions on Medical Imaging*, vol. 24, no. 9, pp. 1138–50, 2005.
- [16] T. Sun, J. Wang, X. Li et al., "Comparative evaluation of support vector machines for computer aided diagnosis of lung cancer in CT based on a multi-dimensional data set," *Computer Methods and Programs in Biomedicine*, vol. 111, no. 2, pp. 519–24, 2013.
- [17] J. Kuruvilla and K. Gunavathi, "Lung cancer classification using neural networks for CT images," *Computer Methods and Programs in Biomedicine*, vol. 113, no. 1, pp. 202–9, 2014.
- [18] M. Javaid, M. Javid, M. Z. U. Rehman, and S. I. A. Shah, "A novel approach to CAD system for the detection of lung nodules in CT images," *Computer Methods and Programs in Biomedicine*, vol. 135, pp. 125–139, 2016.
- [19] M. Z. ur Rehman, M. Javaid, S. I. A. Shah, S. O. Gilani, M. Jamil, and S. I. Butt, "An appraisal of nodules detection techniques for lung cancer in CT images," *Biomedical Signal Processing and Control*, vol. 41, pp. 140–151, 2018.
- [20] Z. Wang, J. Xin, P. Sun, Z. Lin, Y. Yao, and X. Gao, "Improved lung nodule diagnosis accuracy using lung CT images with uncertain class," *Computer Methods and Programs in Biomedicine*, vol. 162, pp. 197–209, 2018.
- [21] H. Xie, D. Yang, N. Sun, Z. Chen, and Y. Zhang, "Automated pulmonary nodule detection in CT images using deep convolutional neural networks," *Pattern Recognition*, vol. 85, pp. 109–119, 2019.
- [22] DICOM (Digital Imaging and Communications in Medicine), <http://medical.nema.org>.
- [23] W. J. Choi and T. S. Choi, "Genetic programming-based feature transform and classification for the automatic detection of pulmonary nodules on computed tomography images," *Information Sciences*, vol. 212, pp. 57–78, 2012.
- [24] T. Messay, R. C. Hardie, and S. K. Rogers, "A new computationally efficient CAD system for pulmonary nodule detection in CT imagery," *Medical Image Analysis*, vol. 14, no. 3, pp. 390–406, 2010.
- [25] X. Ye, X. Lin, J. Dehmeshki, G. Slabaugh, and G. Beddoe, "Shape-based computer-aided detection of lung nodules in thoracic CT images," *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 7, pp. 1810–1820, 2009.
- [26] S. Akram, M. Y. Javed, A. Hussain, F. Riaz, and M. Usman Akram, "Intensity-based statistical features for classification of lungs CT scan nodules using artificial intelligence techniques," *Journal of Experimental and Theoretical Artificial Intelligence*, vol. 27, no. 6, pp. 1–15, 2015.
- [27] R. O. Duda and P. E. Hart, "Use of Hough transformation to detect lines and curves in pictures," *Communications of the ACM*, vol. 15, no. 1, pp. 11–15, 1972.
- [28] H. Yazid, H. Yazid, M. Harun et al., "Circular discontinuities detection in welded joints using Circular Hough Transform," *NDT and E International*, vol. 40, no. 8, pp. 594–601, 2007.
- [29] T. Kohonen, "Self-organizing maps of massive databases," *Engineering Intelligent Systems for Electrical Engineering and Communications*, vol. 9, no. 4, pp. 179–185, 2001.
- [30] S. Haykin, *Neural Networks, A Comprehensive Foundation*, Hamilton, ON, Canada, 1999.
- [31] D. Chi, "Self-organizing map-based color image segmentation with k-means clustering and saliency map," *ISRN Signal Processing*, vol. 2011, Article ID 393891, 18 pages, 2011.
- [32] C. Hollitt, "A convolution approach to the circle Hough transform for arbitrary radius," *Machine Vision and Applications*, vol. 24, no. 4, pp. 683–694, 2013.
- [33] U. Akilandeswari, R. Nithya, and B. Santhi, "Review on feature extraction methods in pattern classification," *European Journal of Scientific Research*, vol. 71, no. 2, pp. 265–272, 2012.
- [34] M. Yang, K. Kpalma, and J. Ronsin, "A survey of shape feature extraction techniques," *Pattern Recognition*, pp. 43–90, 2008.
- [35] D. A. Clausi, "An analysis of co-occurrence texture statistics as a function of grey level quantization," *Canadian Journal of Remote Sensing*, vol. 28, no. 1, pp. 45–62, 2002.
- [36] R. M. Haralick, K. Shanmuga, and I. Dinstein, "Textural features for image classification," *IEEE Transactions on Systems Man and Cybernetics*, vol. SMC-3, no. 6, pp. 610–621, 1973.
- [37] G. Van, P. Wouwer, D. Scheunders, and D. Van, "Statistical texture characterization from discrete wavelet representation," *IEEE Transactions on Image Processing*, vol. 8, no. 4, pp. 592–598, 1998.
- [38] H. Camdevyren, N. Demyr, A. Kanik, and S. Keskin, "Use of principal component scores in multiple linear regression models for prediction of Chlorophyll-a in reservoirs," *Ecological Modelling*, vol. 181, no. 4, pp. 581–589, 2005.
- [39] L. H. Chen and S. Y. Chang, "An adaptive learning algorithm for principal component analysis," *IEEE Transactions on Neural Networks*, vol. 6, no. 5, pp. 1255–1263, 1995.
- [40] K. Z. Mao, K. C. Tan, and W. Ser, "Probabilistic neural-network structure determination for pattern classification," *IEEE Transactions on Neural Networks*, vol. 11, no. 4, pp. 1009–1016, 2000.
- [41] D. F. Specht, "Probabilistic neural networks," *Neural Networks*, vol. 3, no. 1, pp. 109–118, 1990.
- [42] P. Jaccard, "The distribution of the flora in the alpine zone," *New Phytologist*, vol. 11, no. 2, pp. 37–50, 1912.
- [43] L. R. Dice, "Measures of the amount of ecologic association between species," *Ecology*, vol. 26, no. 3, pp. 297–302, 1945.
- [44] W.-J. Choi and T.-S. Choi, "Automated pulmonary nodule detection system in computed tomography images: a hierarchical block classification approach," *Entropy*, vol. 15, no. 2, pp. 507–523, 2013.
- [45] J. J. Erasmus, J. E. Connolly, H. P. McAdams, and V. L. Roggli, "Solitary pulmonary nodules: Part I. Morphologic evaluation for differentiation of benign and malignant lesions 1," *Radiographics*, vol. 20, no. 1, pp. 43–58, 2000.
- [46] R. Opfer and R. Wiemker, "Performance analysis for computer aided lung nodule detection on LIDC data," art. no. 65151C, *Medical Imaging 2007: Image Perception, Observer Performance, and Technology Assessment*, vol. 6515, pp. C5151–C5151, 2007.
- [47] M. F. McNitt-Gray, S. G. Armato, C. R. Meyer et al., "The Lung Image Database Consortium (LIDC) data collection process for nodule detection and annotation," *Academic Radiology*, vol. 14, no. 12, pp. 1464–1474, 2007.
- [48] A. P. Reeves, A. M. Biancardi, T. V. Apanasovich et al., "The Lung Image Database Consortium (LIDC): a comparison of different size metrics for pulmonary nodule measurements," *Academic Radiology*, vol. 14, no. 12, pp. 1475–1485, 2007.
- [49] B. Sahiner, L. M. Hadjiiski, H. P. Chan et al., "Effect of CAD on radiologists' detection of lung nodules on thoracic CT scans: observer performance study," art. no. 65151D, *Medical Imaging 2007: Image Perception, Observer Performance, and Technology Assessment*, vol. 6515, pp. D5151–D5151, 2007.
- [50] S. G. Armato III, G. McLennan, M. F. McNitt-Gray et al., "Lung image database consortium: developing a resource for the medical imaging research community 1," *Radiology*, vol. 232, no. 3, pp. 739–748, 2004.

- [51] G. D. Rubin, J. K. Lyo, D. S. Paik et al., “Pulmonary nodules on multi-detector row CT scans: performance comparison of radiologists and computer-aided detection 1,” *Radiology*, vol. 234, no. 1, pp. 274–283, 2005.
- [52] K. Suzuki, S. G. Armato III, F. Li, S. Sone, and K. Doi, “Massive training artificial neural network (MTANN) for reduction of false positives in computerized detection of lung nodules in low-dose computed tomography,” *Medical Physics*, vol. 30, no. 7, pp. 1602–1617, 2003.
- [53] S. C. Park, J. Tan, X. Wang et al., “Computer-aided detection of early interstitial lung diseases using low-dose CT images,” *Physics in Medicine and Biology*, vol. 56, no. 4, pp. 1139–53, 2011.

Research Article

An Improved Fuzzy Connectedness Method for Automatic Three-Dimensional Liver Vessel Segmentation in CT Images

Rui Zhang,¹ Zhuhuang Zhou ,¹ Weiwei Wu ,² Chung-Chih Lin,³ Po-Hsiang Tsui ,^{4,5,6} and Shuicai Wu ¹

¹College of Life Science and Bioengineering, Beijing University of Technology, Beijing 100124, China

²College of Biomedical Engineering, Capital Medical University, Beijing 100069, China

³Department of Computer Science and Information Engineering, Chang Gung University, Taoyuan 33302, Taiwan

⁴Department of Medical Imaging and Radiological Sciences, College of Medicine, Chang Gung University, Taoyuan 33302, Taiwan

⁵Department of Medical Imaging and Intervention, Chang Gung Memorial Hospital at Linkou, Taoyuan 33302, Taiwan

⁶Medical Imaging Research Center, Institute for Radiological Research, Chang Gung University and Chang Gung Memorial Hospital at Linkou, Taoyuan 33302, Taiwan

Correspondence should be addressed to Shuicai Wu; wushuicai@bjut.edu.cn

Received 17 April 2018; Revised 22 September 2018; Accepted 4 October 2018; Published 29 October 2018

Guest Editor: Cesare Valenti

Copyright © 2018 Rui Zhang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In this paper, an improved fuzzy connectedness (FC) method was proposed for automatic three-dimensional (3D) liver vessel segmentation in computed tomography (CT) images. The vessel-enhanced image (i.e., vesselness image) was incorporated into the fuzzy affinity function of FC, rather than the intensity image used by traditional FC. An improved vesselness filter was proposed by incorporating adaptive sigmoid filtering and a background-suppressing item. The fuzzy scene of FC was automatically initialized by using the Otsu segmentation algorithm and one single seed generated adaptively, while traditional FC required multiple seeds. The improved FC method was evaluated on 40 cases of clinical CT volumetric images from the 3Dircadb ($n = 20$) and Sliver07 ($n = 20$) datasets. Experimental results showed that the proposed liver vessel segmentation strategy could achieve better segmentation performance than traditional FC, region growing, and threshold level set. Average accuracy, sensitivity, specificity, and Dice coefficient of the improved FC method were, respectively, $(96.4 \pm 1.1)\%$, $(73.7 \pm 7.6)\%$, $(97.4 \pm 1.3)\%$, and $(67.3 \pm 5.7)\%$ for the 3Dircadb dataset and $(96.8 \pm 0.6)\%$, $(89.1 \pm 6.8)\%$, $(97.6 \pm 1.1)\%$, and $(71.4 \pm 7.6)\%$ for the Sliver07 dataset. It was concluded that the improved FC may be used as a new method for automatic 3D segmentation of liver vessel from CT images.

1. Introduction

Hepatocellular carcinoma (HCC) is one of the most common malignancies in the world, especially in China with the fifth morbidity and the third mortality [1, 2]. Nowadays, main treatments to HCC include surgical resection, liver transplantation, and local thermal ablation [3]. Treatment planning and navigation based on medical imaging are essential for these procedures. Among different medical imaging modalities, computed tomography (CT) is commonly used for the guidance of liver tumor treatment. Three-dimensional (3D) segmentation of liver vessel is critical in

CT-based liver tumor treatment planning and navigation. However, manual segmentation of liver vessel in CT images is time consuming and tedious. Thus, there is a demand for computerized 3D segmentation of liver vessel in CT images [4, 5].

Currently, computerized liver vessel segmentation techniques can be classified into region growing [6–8], active contour models or level sets [9], graph cuts [10–12], extreme learning [13], deep learning [14], and fuzzy logic [15, 16]. However, it is still challenging to extract liver vessel in CT images, especially in those with low contrast [4]. Region growing methods [6–8] are simple with low computational

cost, but they are sensitive to seed point location and threshold. Active contour models or level sets [9] are among mainstream vessel segmentation algorithms, but they have limitations of complex initialization and typically use speed function which implicitly assumes that images are of good contrast. Graph cuts [10–12] are segmentation methods based on graph theory, but they are partly limited by the shrinking bias problem. Machine learning methods [13, 14] can take into account the diversity of liver vessel topologies and features, but they generally require plenty of training samples or long training time. Fuzzy connectedness (FC) methods [15, 16] are based on fuzzy logic. FC describes spatial connectedness between each voxel, rather than just focusing on intensity. Recently, Guo et al. [15] and Wang et al. [16] have demonstrated the potential of FC in liver vessel segmentation. However, for FC-based liver vessel segmentation in CT images, there are still issues to be addressed, including unsatisfying segmentation accuracy (especially for low-contrast CT images), requirement on multiple seeds, and sensitivity to initialization.

In this paper, an improved FC method was proposed for automatic 3D liver vessel segmentation in CT images. The vessel-enhanced image (i.e., vesselness image) was incorporated into the fuzzy affinity function of FC, rather than the intensity image used by traditional FC. An improved vesselness filter was also proposed based on the Jerman's vesselness filter [17] introduced recently. The fuzzy scene of FC was initialized by using the Otsu segmentation algorithm, and the quantity of seeds required was reduced to one which was generated automatically. The proposed method was evaluated on 40 cases of clinical CT image volumes, including low-contrast images. Experimental results demonstrate that the improved FC method can overcome the drawbacks of traditional FC and yield more satisfying segmentation performance.

2. Materials and Methods

Figure 1 shows the flow chart of the improved FC method. First, the liver volume of interest (VOI) image was obtained by using the liver mask, which could be obtained by using liver segmentation approaches [18]. The liver VOI image was then contrast enhanced by an adaptive sigmoid filtering which was initialized by K -means clustering and isotropically resampled. Subsequently, the improved vesselness filter was used to enhance the liver vessel and suppress the background, and a 3D vesselness image was obtained. Then, a 3D fuzzy scene was constructed with the 3D vesselness image by (1) incorporating the improved vesselness into the fuzzy affinity function of FC, (2) initializing the fuzzy scene by the Otsu algorithm, and (3) generating automatically one single seed. Finally, the 3D liver vessel was segmented on the basis of the 3D fuzzy scene and anisotropically resampled.

2.1. Dataset. Both simulated data ($n = 60$) and clinical CT data ($n = 40$) were used. The synthetic dataset VascuSynth [19, 20] was provided by the Medical Image Analysis Lab, School of Computing Science, Simon Fraser University,

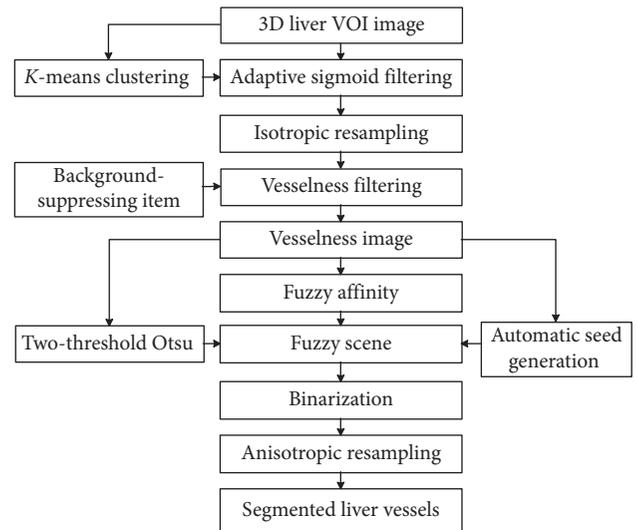


FIGURE 1: Flowchart of the improved fuzzy connectedness method for automatic 3D segmentation of liver vessel from CT images.

Canada. VascuSynth contains 10 groups of data, which are publically available at <http://vascusynth.cs.sfu.ca>. Each group consists of 12 randomly generated images with different quantity of bifurcations. Six groups of data were randomly selected; among them, the images with bifurcations ≥ 11 were included in this study. Gaussian white noise was also added to the raw data. The level of the Gaussian white noise was indicated by σ^2 , the variance of the noise. In this study, Gaussian white noise with $\sigma^2 = 30, 45,$ and 60 were added.

The clinical CT image datasets, 3Dircadb and Sliver07, were used. 3Dircadb contains 20 cases of contrast-enhanced CT (CE-CT) images. 3Dircadb was provided by the Research Institute against Digestive Cancer, France, and is publically available at <http://www.ircad.fr/research/3dircadb>. The pixel spacing is 0.56–0.86 mm, and the slice thickness is 1–4 mm. The number of slices ranges from 64 to 502, and the in-plane resolution is 512×512 pixels. The gold standard of liver vessel was provided by 3Dircadb, which was manually delineated by radiologists. Sliver07 contains 30 cases of CE-CT images, including 20 training sets and 10 testing sets. The 20 cases of training data are publically available at <http://www.sliver07.org> and were included in this study. However, Sliver07 did not provide the gold standard of liver vessel. Therefore, radiologists were asked to manually delineate the liver vessel to serve as the gold standard for the 20 cases of training data of Sliver07. The number of slices, in-plane resolution, and interslice resolution range from 64 to 394, from 0.58 to 0.81 mm, and from 0.7 to 5.0 mm, respectively.

2.2. Improved Vesselness Filter. The multiscale Hessian matrix-based filter (vesselness filter) is commonly used for vessel enhancement [4, 17]. Classical vesselness filters were proposed by Sato et al. [21] and Frangi et al. [22]. Since then, Li et al. [23], Erdt et al. [24], and Xiao et al. [25] proposed

improved methods for enhancing the vasculature. Recently, Jerman et al. [17] proposed a novel vesselness filter and demonstrated that it outperformed traditional vesselness filters. For completeness, the Jerman's filter was introduced briefly as below. Let λ_i , $i = 1, 2, 3$ denotes the Hessian eigenvalues of a 3D image at each coordinate \mathbf{x} . Considering the ideal eigenvalues' relationship $\lambda_2 \approx \lambda_3 \wedge |\lambda_{2,3}| \gg |\lambda_1|$ in vasculature, Jerman et al. [17] constructed a novel Hessian eigenvalues function to improve the enhancement performance by using a two-step piecewise compensation. In CT images, the magnitudes of λ_2 and λ_3 were lower at the vascular boundary or in the low-scale vessel ($|\lambda_3| \geq |\lambda_2| \approx |\lambda_1| \rightarrow Low$), which did not match the ideal Hessian eigenvalues relationship in vasculature, resulting in significant attenuation of the vesselness response. Therefore, Jerman et al. [17] performed a piecewise compensation on the eigenvalue λ_3 :

$$\lambda_p(\sigma) = \begin{cases} \lambda_3, & \text{if } \lambda_3 > \tau \max_{\mathbf{x}} \lambda_3(\mathbf{x}, \sigma), \\ \tau \max_{\mathbf{x}} \lambda_3(\mathbf{x}, \sigma), & \text{if } 0 < \lambda_3 \leq \tau \max_{\mathbf{x}} \lambda_3(\mathbf{x}, \sigma), \\ 0, & \text{otherwise,} \end{cases} \quad (1)$$

where σ is the vessel scale and τ is a threshold between 0 and 1. In addition, traditional vesselness filters would suppress blob-like structures and obtain poor response at vascular nodes ($|\lambda_1| \approx |\lambda_2| \approx |\lambda_3| \rightarrow High$). Thus, Jerman et al. [17] compensated the ellipsoid structure conforming to the condition $\lambda_2 \geq \lambda_p/2 > 0$ to construct the final vesselness function:

$$v_k = \begin{cases} 0, & \text{if } \lambda_2 \leq 0 \vee \lambda_p \leq 0, \\ 1, & \text{if } \lambda_2 \geq \lambda_p/2 > 0, \\ \lambda_2^2(\lambda_p - \lambda_2) [3/(\lambda_2 + \lambda_p)]^3, & \text{otherwise.} \end{cases} \quad (2)$$

Jerman et al. [17] evaluated their method on clinical image datasets of lung, cerebral, and fundus vasculatures. However, for the task of liver vessel enhancement in CT images, the Jerman's filter would enhance the liver contour, liver parenchyma, and noise. Therefore, the Jerman's filter was improved by incorporating adaptive sigmoid filter for contrast enhancement and by incorporating a background-suppressing item into the vesselness function of the Jerman's filter (Equation (2)).

The adaptive sigmoid filter is defined as

$$I_{\text{sigmoid}} = \left(1 + \exp\left(-\frac{I_{\text{VOI}} - \beta}{\alpha}\right) \right)^{-1}, \quad (3)$$

where I_{sigmoid} is the filtered image, I_{VOI} is the liver VOI image, and β and α represented the intensity center and the intensity range of the vasculature. In this study, β and α were obtained adaptively by the K -means clustering ($K = 5$). The internal structure of I_{VOI} was clustered into five regions with corresponding cluster centers. With the value of the intensity centers ranking from low to high, the five regions corresponded to the background, liver tumor, liver parenchyma, low-intensity vessel mixed with liver parenchyma, and high-intensity vessel, respectively. With the intensity means of the last two regions (m_1 and m_2), parameters β and α are calculated by

$$\begin{cases} \alpha = (m_2 - m_1)/2, \\ \beta = (m_2 + m_1)/2. \end{cases} \quad (4)$$

The background-suppressing item, $1 - e^{-R_s^2/2\gamma}$, was incorporated into Equation (2), yielding

$$v = \begin{cases} 0, & \text{if } \lambda_2 \leq 0 \vee \lambda_p \leq 0, \\ 1, & \text{if } \lambda_2 \geq \lambda_p/2 > 0, \\ \lambda_2^2(\lambda_p - \lambda_2) [3/(\lambda_2 + \lambda_p)]^3 (1 - e^{-R_s^2/2\gamma}), & \text{otherwise,} \end{cases} \quad (5)$$

where $R_s = \sqrt{\lambda_1^2 + \lambda_2^2 + \lambda_p^2}$ and γ is the background suppression coefficient, which was optimally set at $\lambda_p/3$.

Finally, the vesselness response was combined by calculating the maximum response of v in each scale σ , $\sigma \in [\sigma_{\min}, \sigma_{\max}]$:

$$I_{\text{vesselness}} = \sup\{v : \sigma_{\min} \leq \sigma \leq \sigma_{\max}\}, \quad (6)$$

where $I_{\text{vesselness}}$ is the final vessel-enhanced image (vesselness image). The improved vesselness filter algorithm is summarized in Algorithm 1.

2.3. Improved Fuzzy Connectedness. FC involved three kinds of fuzzy relationships: fuzzy adjacency, fuzzy affinity, and fuzzy connectivity. Fuzzy affinity represented the local similarity of the voxel pair (c, d) in the entire image scene C , denoted by $\mu_{\kappa}(c, d) \in [0, 1]$:

$$\mu_{\kappa}(c, d) = \mu_{\alpha}(c, d) [\omega_1 h_1(f(c), f(d)) + \omega_2 h_2(f(c), f(d))], \quad (7)$$

where $\mu_{\alpha}(c, d)$ is the fuzzy adjacency (a monotonic increasing function), and h_1 and h_2 are computed by

$$\begin{aligned} h_1(f(c), f(d)) &= e^{-1/2[|f(c)+f(d)/2-m/s|^2]}, \\ h_2(f(c), f(d)) &= e^{-1/2[|f(c)-f(d)-m/s|^2]}, \end{aligned} \quad (8)$$

where $f(\cdot)$ is the intensity of voxels; m and s are mean and standard deviation of $f(\cdot)$ in the VOI, respectively; and ω_1 and ω_2 are weight parameters, $\omega_1 + \omega_2 = 1$.

In this paper, the vesselness image obtained by using the improved vesselness filter was used as the input of the fuzzy affinity function, rather than the intensity image used by traditional FC. The improved fuzzy affinity function, $\mu'_{\kappa}(c, d)$, is defined as

$$\begin{aligned} \mu'_{\kappa}(c, d) &= \mu_{\alpha}(c, d) [\omega_1 h_1(I_{\text{vesselness}}(c), I_{\text{vesselness}}(d)) \\ &\quad + \omega_2 h_2(I_{\text{vesselness}}(c), I_{\text{vesselness}}(d))]. \end{aligned} \quad (9)$$

To adaptively set parameters m and s , the Otsu segmentation algorithm was adopted to the vesselness image. Two-threshold Otsu was used to yield a binary liver vessel mask. Parameters m and s are, respectively, set at the mean and standard deviation of the vesselness voxels belonging to the foreground of the vessel mask.

The weight parameters ω_1 and ω_2 are adaptively selected by using the method proposed by Pednekar et al. [26]:

Input: the liver VOI CT image I_{VOI} .

Output: the vesselness image $I_{vesselness}$.

- (1) Set the vessel scales $\sigma_{\min} \leftarrow 1$, $\sigma_{\max} \leftarrow 4$; threshold $\tau \leftarrow 0.6$.
- (2) Perform adaptive sigmoid filtering to I_{VOI} by using Equations (3) and (4) to obtain the filtered image $I_{sigmoid}$.
- (3) Perform isotropic resampling to $I_{sigmoid}$.
- (4) **For** $\sigma \leftarrow \sigma_{\min}$ **do**
- (5) Compute the Hessian matrix elements at each coordinate $H_{ij}(\mathbf{x}, \sigma)$;
- (6) Make eigenvalue decomposition $eigH(\mathbf{x}, \sigma) \rightarrow \lambda_i, i = 1, 2, \dots, D$;
- (7) Rank $\lambda_i \geq \lambda_{i+1}, i = 1, 2, \dots, D-1$;
- (8) Compensate λ_3 by using Equation (1);
- (9) Perform vesselness filtering by using Equation (5);
- (10) $\sigma \leftarrow \sigma_{\text{next}}$ until $\sigma = \sigma_{\max}$.
- (11) **End for**
- (12) Compute the vesselness image by using Equation (6).

ALGORITHM 1: Improved vesselness filter.

$$\omega_1 = \frac{h_1}{h_1 + h_2}, \quad (10)$$

$$\omega_2 = 1 - \omega_1.$$

The fuzzy scene of liver vessel was initialized with one single seed generated automatically on the vesselness image, binarized by a threshold T , and anisotropically resampled to yield the final liver vessel segmentation. Figure 2 illustrates automatic selection of one single seed. In Figure 2(a), the 3D image $I_{vesselness}$ was divided into several regions R_{sub} of $5 * 5 * 3$ voxels. The maximum vesselness voxels at each R_{sub} region were selected as potential seeds (denoted by the blue points in Figure 2(b)). Then, the regions of $5 * 5 * 3$ voxels around the potential seeds were constructed and denoted as R_{seed} , with each potential seed being the center of each R_{seed} region. The mean of the vesselness of each R_{seed} region was calculated. The potential seed having the largest vesselness mean in its R_{seed} region was automatically selected as the final single seed, which was indicated by the red point in Figure 2(b). The improved FC algorithm is summarized in Algorithm 2.

2.4. Evaluation. To analyze quantitatively the performance of the proposed vessel segmentation method, evaluation metrics including accuracy, sensitivity, specificity, and Dice coefficient were used:

$$\begin{aligned} \text{Accuracy} &= \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FN} + \text{TN} + \text{FP}}, \\ \text{Sensitivity} &= \frac{\text{TP}}{\text{FN} + \text{TP}}, \\ \text{Specificity} &= \frac{\text{TN}}{\text{TN} + \text{FP}}, \\ \text{Dice} &= \frac{2\text{TP}}{2\text{TP} + \text{FN} + \text{FP}}, \end{aligned} \quad (11)$$

where TP and TN are the numbers of voxels correctly segmented as vessel and background (i.e., nonvessel),

respectively; FP and FN are the numbers of voxels incorrectly segmented as vessel and background, respectively.

3. Results

Figure 3 shows the vessel segmented from the simulated data by using the improved FC method. Figure 3(a) represents the ground truth; Figure 3(b) shows the segmented vessel on the synthetic data; and Figures 3(c)–3(e) show the segmented vessel on the synthetic data added with Gaussian white noise $\sigma^2 = 30, 45, \text{ and } 60$, respectively. The segmentation performance of the improved FC method on the synthetic dataset ($n = 60$) is shown in Table 1, in terms of accuracy, sensitivity, specificity, and Dice coefficient. Although the sensitivity and Dice coefficient were decreased to some extent with increasing the level of Gaussian white noise, the segmentation performance was generally kept stable. It is thus indicated that the improved FC method is insensitive to Gaussian white noise.

Figure 4 shows the vessel-enhanced image by using the improved vesselness filter. Figure 4(a) shows the original CT image; Figure 4(b) shows the adaptive sigmoid filtered image; Figure 4(c) shows the isotropic resampled image; and Figure 4(d) shows the improved vesselness filtered image. It can be seen that the improved vesselness filter can effectively enhance the vessel while suppressing the background. The vesselness images obtained by using the Jerman's vesselness filter and the improved vesselness filter are shown in Figure 5. The intensity of the vesselness images ranged from 0 to 1. The contrast of vessel in CT images increased from Figures 5(a) to 5(c). Note that the Jerman's vesselness filter enhanced the liver contour and almost neglected the liver vessel for the low-contrast image (Figure 5(d)). As the image contrast increased, there was still undesired enhancement at the liver contour (Figures 5(e) and 5(f)). In addition, the Jerman's vesselness filter could not effectively suppress the background (Figure 5(f)). By contrast, the improved vesselness filter successfully enhanced the vessel while suppressing the background, with little enhancement at the liver contour (Figures 5(g)–5(i)).

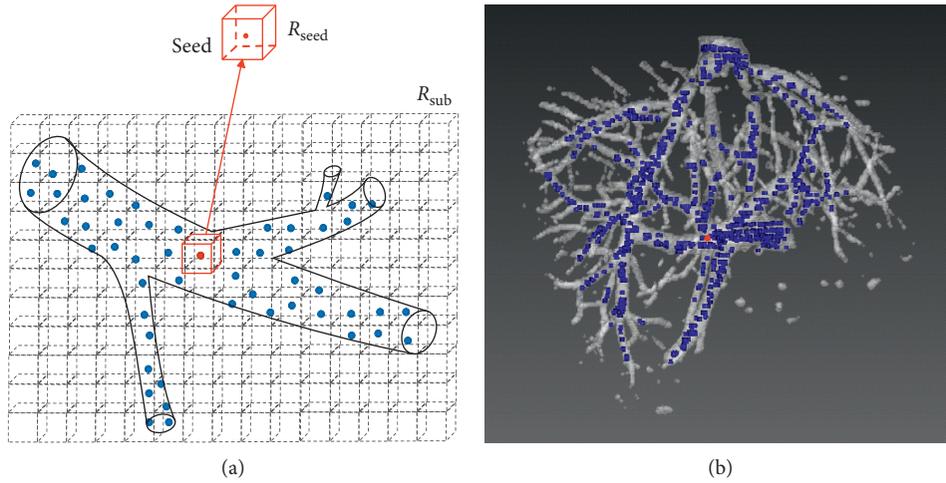


FIGURE 2: Multiple potential seeds and one single seed generated automatically on the vesselness image. (a) Illustration for automatic seed selection. (b) Multiple potential seeds (blue) and one single seed (red) indicated on the vesselness image.

Input: the vesselness image $I_{\text{vesselness}}$

Output: the liver vessel segmentation

- (1) Initialize the parameters, and set threshold $T \leftarrow 0.05$
- (2) Perform two-threshold Otsu to $I_{\text{vesselness}}$ to obtain the parameters m and s
- (3) Calculate the fuzzy affinity by using Equation (9)
- (4) Calculate adaptively the weights ω_1 and ω_2 by using Equation (10)
- (5) Compute potential seeds in R_{sub} regions, and select automatically the single seed with the largest vesselness mean in its R_{seed} region
- (6) Binarize the fuzzy scene and perform anisotropic resampling to obtain the final liver vessel segmentation

ALGORITHM 2: Improved fuzzy connectedness.

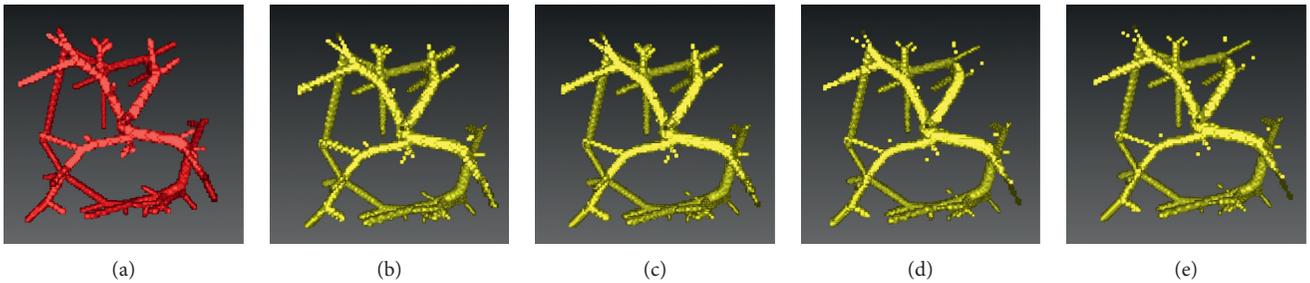


FIGURE 3: Segmentation of vessel from synthetic data by using the improved fuzzy connectedness method. (a) The gold standard of vessel. (b–e) The vessel segmented from the synthetic data added with Gaussian white noise $\sigma^2 = 0, 30, 45, \text{ and } 60$, respectively.

Figure 6 shows the liver vessel segmented by using the improved FC, depicted by yellow contour or surface. The gold standard of liver vessel is indicated by red contour or surface. The axial slices, sagittal slices, coronal slices, and 3D view are shown in Figures 6(a)–6(c), 6(d)–6(f), 6(g)–6(i), and 6(j)–6(k), respectively. It can be seen that the proposed method yielded satisfying segmentation performance. Figure 7 shows typical CT images from the 3Dircadb (Figures 7(a)–7(i)) and Sliver07 (Figures 7(j)–7(r)) datasets. The original CT images in Figures 7(a)–7(i) and 7(j)–7(r) are used in Figures 8 and 9, respectively. For vessel segmentation, Figures 7(a)–7(c)

and 7(j)–7(l) were of high contrast, while Figures 7(d)–7(i) and 7(m)–7(r) were of low contrast.

Figures 8 and 9 show the comparison of the improved FC method with traditional segmentation algorithms, including traditional FC [27], region growing [27], and threshold level set [27]. Figures 8(a)–8(c) and 9(a)–9(c) show the gold standard of liver vessel. Figures 8(d)–8(f) and 9(d)–9(f) show the vessel segmented by using the improved FC method. Figures 8(g)–8(i) and 9(g)–9(i) show the vessel segmented by using traditional FC with multiple potential seeds indicated by the blue points in Figure 2(b).

TABLE 1: Segmentation performance of the improved fuzzy connectedness method on the synthetic dataset.

	Group 1 ($n = 10$)			Group 2 ($n = 10$)			Group 3 ($n = 10$)			Group 4 ($n = 10$)			Group 5 ($n = 10$)			Group 6 ($n = 10$)								
	ACC (%)	SEN (%)	SPE (%)	DICE (%)	ACC (%)	SEN (%)	SPE (%)	DICE (%)	ACC (%)	SEN (%)	SPE (%)	DICE (%)	ACC (%)	SEN (%)	SPE (%)	DICE (%)	ACC (%)	SEN (%)	SPE (%)	DICE (%)				
$\sigma^2 = 0$	99.7 ± 0.1	96.6 ± 5.1	99.7 ± 0.2	84.3 ± 3.6	99.8 ± 0.1	91.2 ± 4.9	99.8 ± 0.1	86.6 ± 2.6	99.8 ± 0.1	91.5 ± 3.5	99.8 ± 0.1	85.4 ± 3.1	99.8 ± 0.1	91.5 ± 3.5	99.8 ± 0.1	86.0 ± 3.6	99.8 ± 0.1	91.5 ± 4.0	99.9 ± 0.1	85.4 ± 2.4	99.7 ± 0.3	91.0 ± 3.9	99.9 ± 0.1	85.6 ± 3.0
$\sigma^2 = 30$	99.7 ± 0.2	86.7 ± 4.9	99.7 ± 0.2	83.8 ± 4.1	99.7 ± 0.1	87.5 ± 7.3	99.8 ± 0.1	82.2 ± 3.3	99.7 ± 0.1	89.4 ± 4.7	99.8 ± 0.1	80.7 ± 3.1	99.7 ± 0.1	89.9 ± 5.2	99.8 ± 0.1	80.7 ± 3.6	99.7 ± 0.1	90.4 ± 5.2	99.8 ± 0.1	81.0 ± 2.5	99.7 ± 0.1	88.6 ± 6.6	99.8 ± 0.1	80.2 ± 2.5
$\sigma^2 = 45$	99.5 ± 6.3	91.9 ± 4.6	99.7 ± 0.1	81.3 ± 3.0	99.7 ± 0.1	88.3 ± 6.1	99.8 ± 0.1	81.8 ± 4.1	99.7 ± 0.1	90.4 ± 4.5	99.8 ± 0.1	80.7 ± 3.1	99.7 ± 0.1	89.7 ± 5.3	99.8 ± 0.1	80.8 ± 3.5	99.7 ± 0.1	89.1 ± 5.3	99.8 ± 0.1	80.9 ± 2.8	99.7 ± 0.1	87.8 ± 6.6	99.8 ± 0.1	80.6 ± 2.7
$\sigma^2 = 60$	99.6 ± 0.1	91.4 ± 4.3	99.7 ± 0.1	80.6 ± 2.4	99.7 ± 0.1	88.2 ± 5.5	99.8 ± 0.1	82.2 ± 4.1	99.7 ± 0.1	90.3 ± 4.5	99.8 ± 0.1	81.0 ± 3.1	99.7 ± 0.1	90.2 ± 5.1	99.8 ± 0.1	81.1 ± 3.3	99.7 ± 0.1	88.6 ± 5.1	99.7 ± 0.1	81.4 ± 2.6	99.7 ± 0.1	87.2 ± 6.4	99.8 ± 0.1	81.1 ± 2.7

ACC = accuracy; SEN = sensitivity; SPE = specificity; DICE = Dice coefficient.

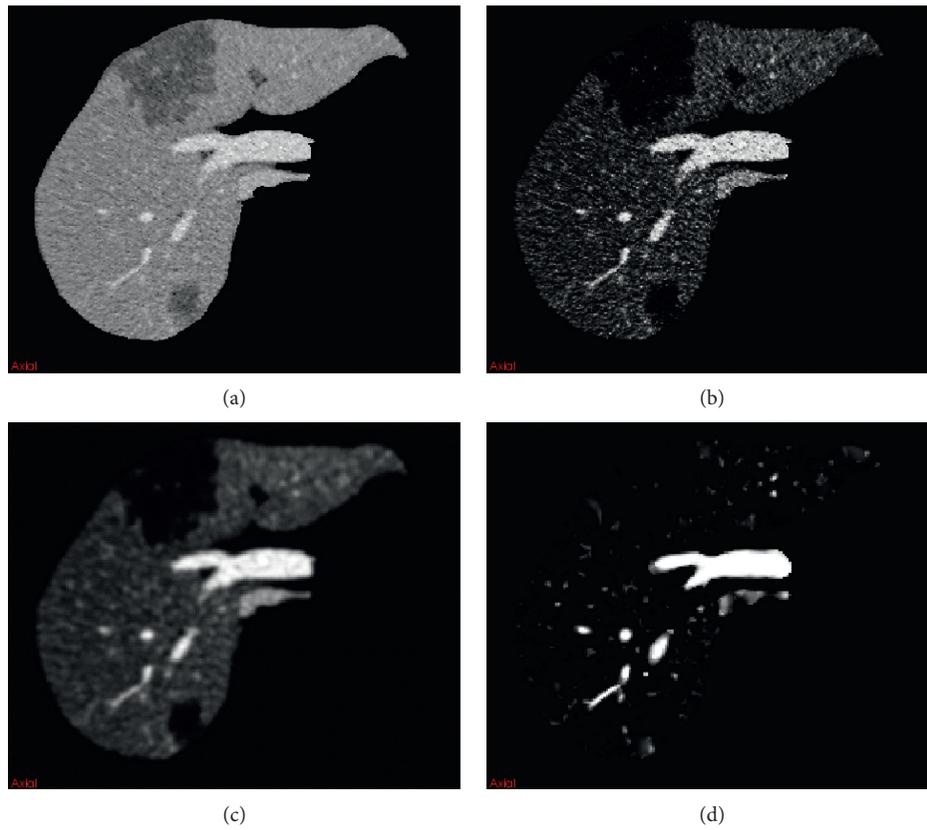


FIGURE 4: Illustration of the improved vesselness filter. (a) The original CT image. (b) The adaptive sigmoid filtered image of (a). (c) The isotropic resampled image of (b). (d) The improved vesselness filtered image of (c).

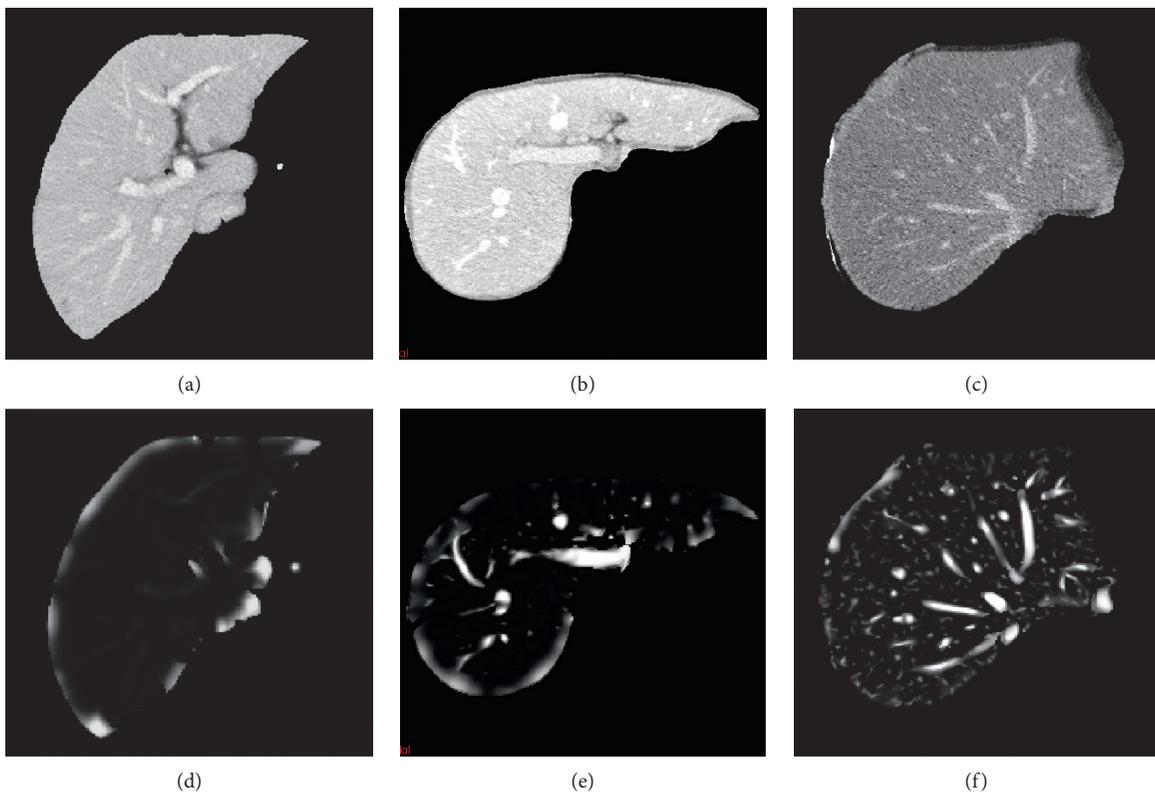


FIGURE 5: Continued.

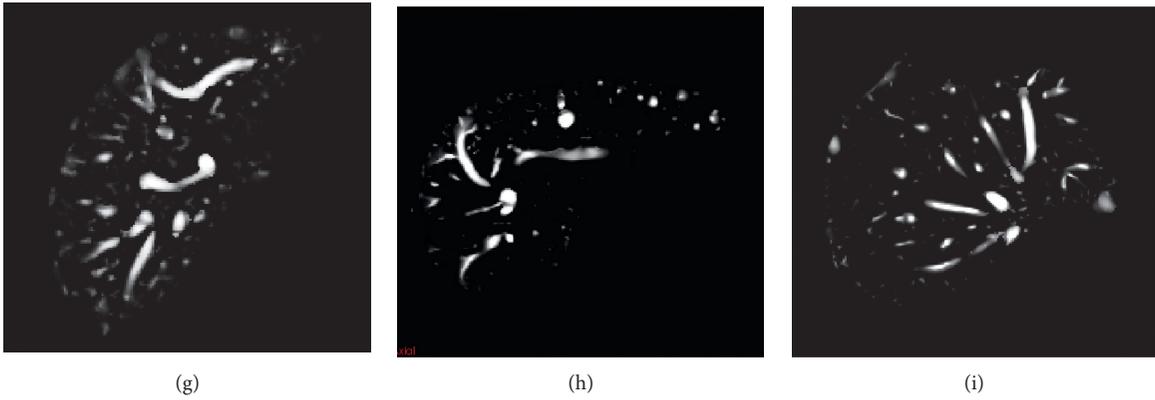


FIGURE 5: Comparison of the improved vesselness filter with the Jerman's vesselness filter. (a)–(c) The original CT images. (d)–(f) The vesselness images by using the Jerman's vesselness filtering to (a)–(c), respectively. (g)–(i) The vesselness images by using the improved vesselness filtering to (a)–(c), respectively.

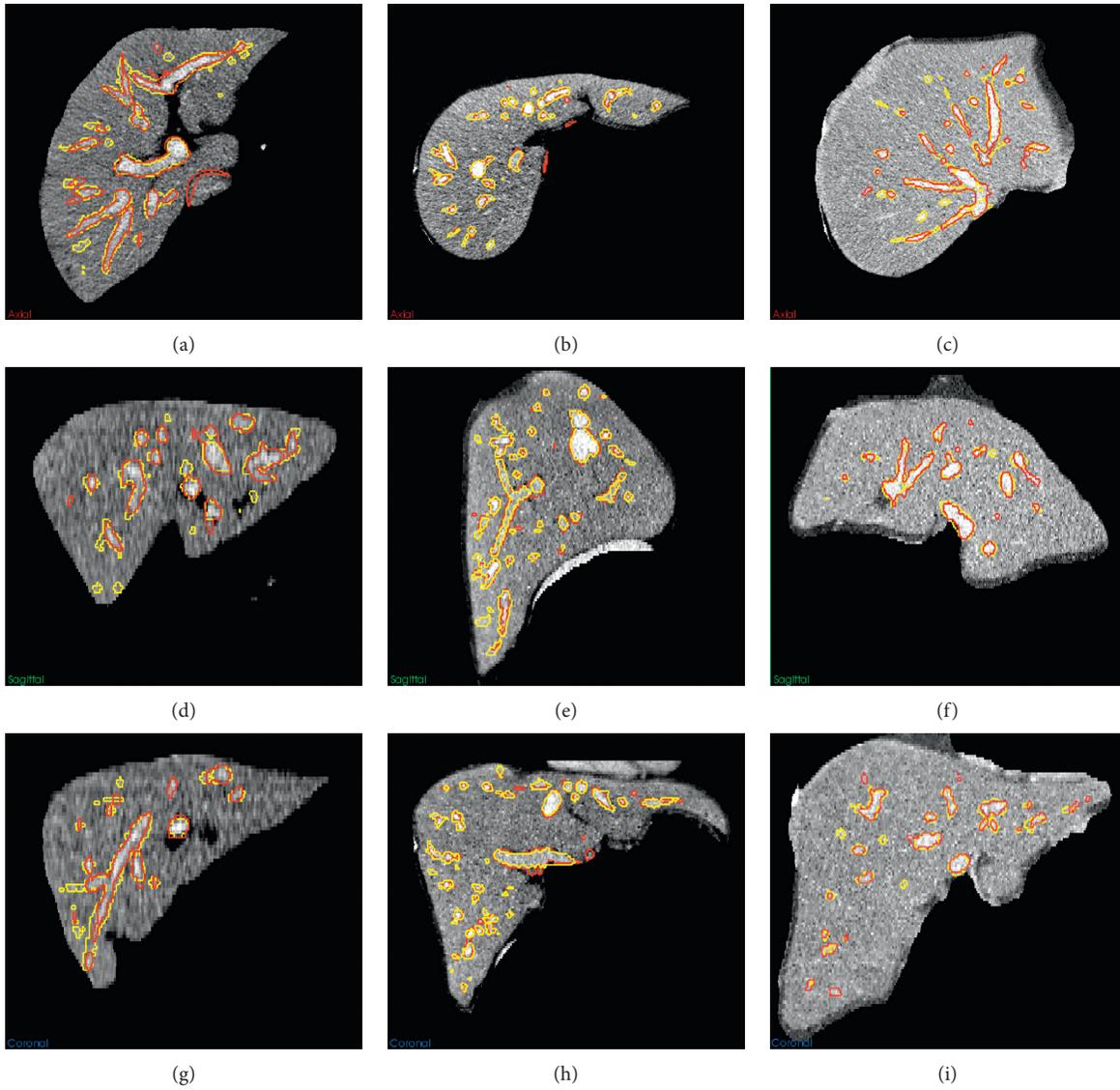


FIGURE 6: Continued.

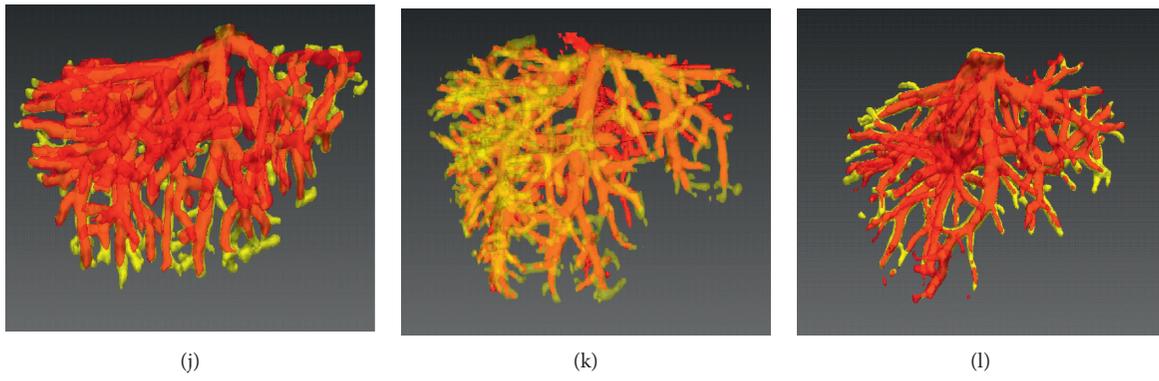


FIGURE 6: Liver vessel segmented by using the proposed method (yellow). The gold standard of vessel is depicted in red. Each column corresponds to one case. (a)–(c) The axial slices. (d)–(f) The sagittal slices. (g)–(i) The coronal slices. (j)–(k) The 3D view.

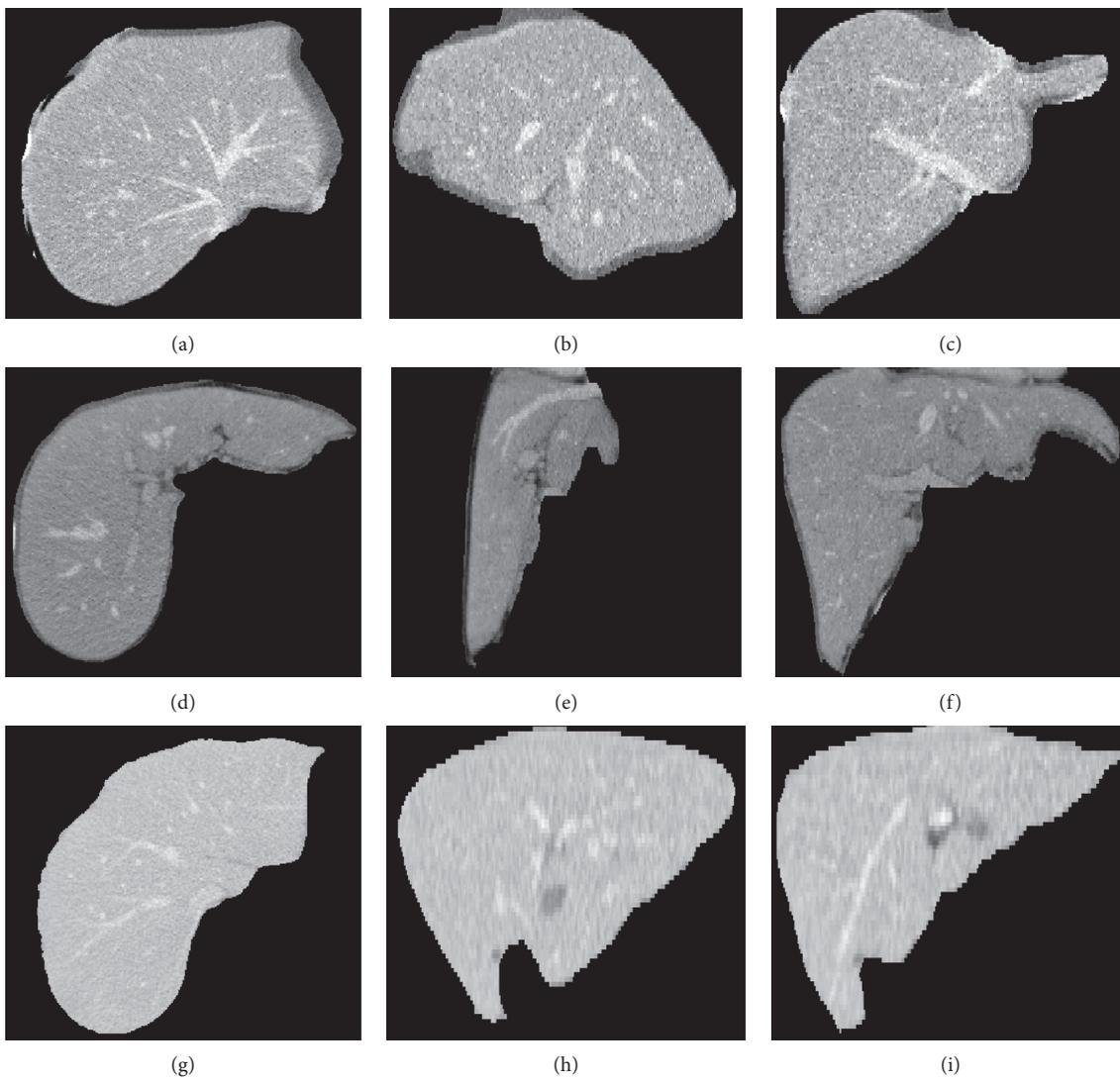


FIGURE 7: Continued.

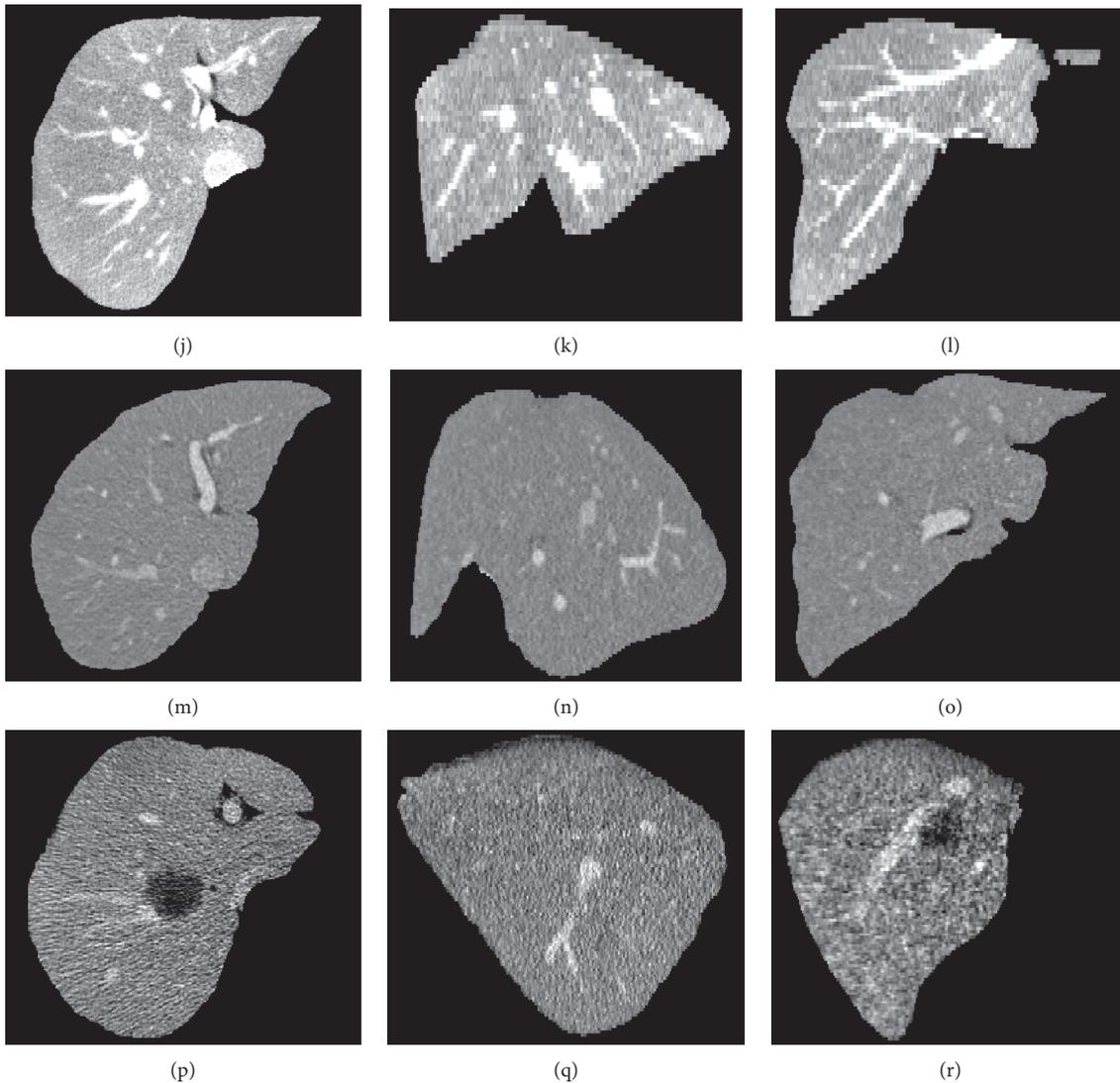


FIGURE 7: Typical CT images from the 3Dircadb (a)–(i) and Sliver07 (j)–(r) datasets. Each row corresponds to one case. The first, second, and third columns are axial slices, sagittal slices, and coronal slices, respectively. The original CT images in (a)–(i) and (j)–(r) are used in Figures 8 and 9, respectively. For vessel segmentation, (a)–(c) and (j)–(l) are of high contrast, while (d)–(i) and (m)–(r) are of low contrast.

Figures 8(j)–8(l) and 9(j)–9(l) show the vessel segmented by using threshold level set with multiple potential seeds. Traditional FC with one single seed could not segment completely the liver vessel (Figures 8(j)–8(l) and 9(j)–9(l)). When multiple potential seeds were used for traditional FC, the segmentation performance show the vessel segmented by using traditional FC with one single seed indicated by the red point in Figure 2(b). Figures 8(m)–8(o) and 9(m)–9(o) show vessel segmented by using region growing with multiple potential seeds. Figures 8(p)–8(r) and 9(p)–9(r) shows improved, but it was still unsatisfying (Figures 8(g)–8(i) and 9(g)–9(i)). For region growing and threshold level set with multiple potential seeds, both undersegmentation and oversegmentation of vessel occurred (Figures 8(m)–8(r) and 9(m)–9(r)). It is interesting to discuss the segmentation performance on low-contrast cases shown in Figures 7(d)–7(i) and 7(m)–7(r). If a part of the main

vessel was low-contrast, it would be totally unsegmented, as indicated by the black arrows in Figures 8(h), 8(n), 8(q) and 9(h), 9(n), 9(q). When the peripheral vessel was low-contrast, it would be merged (Figures 8(h) and 8(q)) or missed (Figures 9(h), 9(n), 9(q)), as indicated by blue arrows. Even in the high-contrast images shown in Figures 7(a)–7(c) and 7(j)–7(l), part of the vessel was segmented falsely (Figures 8(m) and 8(p)) and the periphery vessel was not segmented (Figures 8(g) and 9(m), 9(p)), as indicated by green arrows. By contrast, the proposed method was capable to segment completely the liver vessel, even for the low-contrast images.

To compare further the improved vesselness filter with the Jerman's vesselness filter, the liver vessel segmented by using the improved FC method on the basis of the Jerman's vesselness filtering, rather than the improved vesselness filtering, is shown in Figure 10. Figures 10(a) and 10(b)

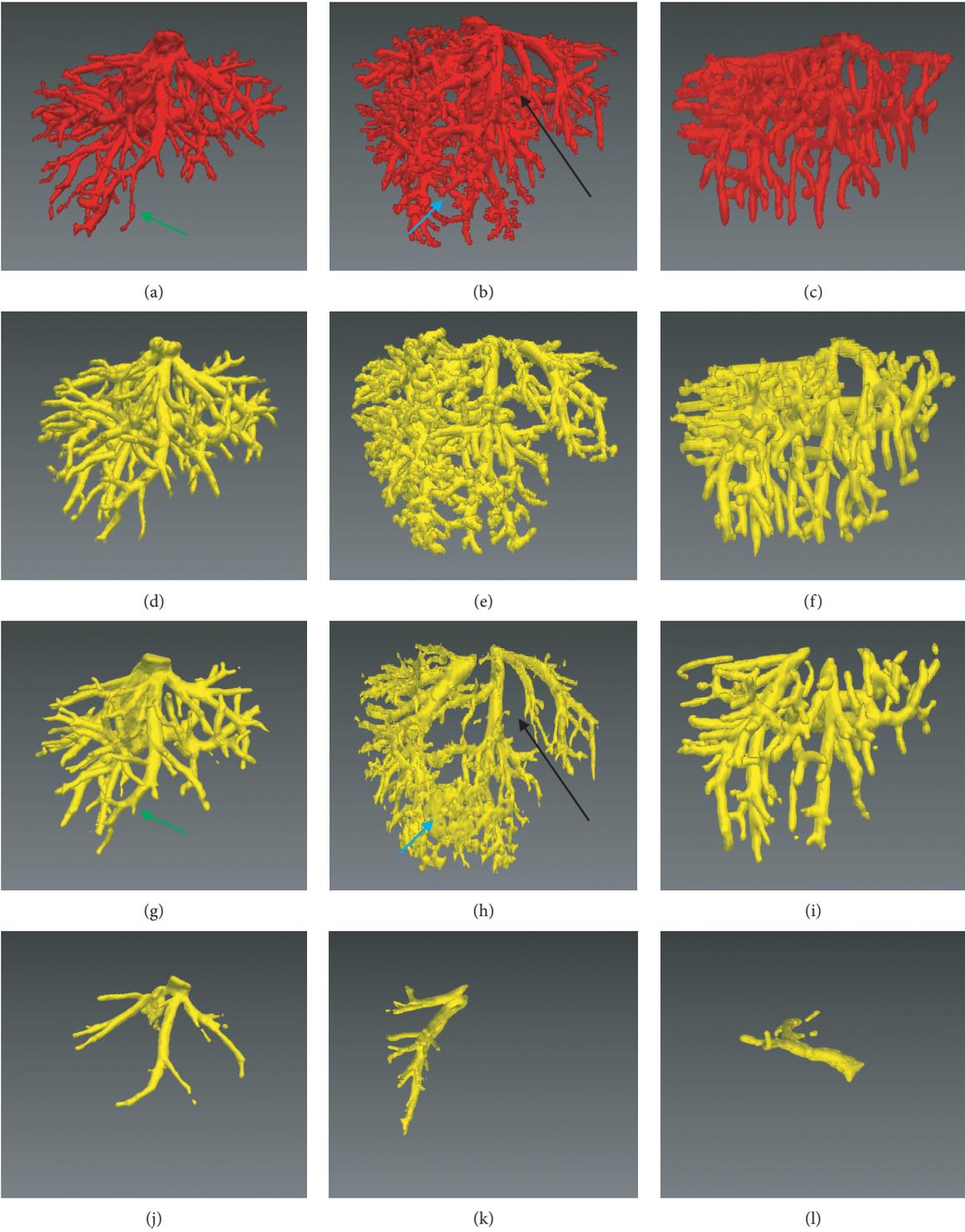


FIGURE 8: Continued.

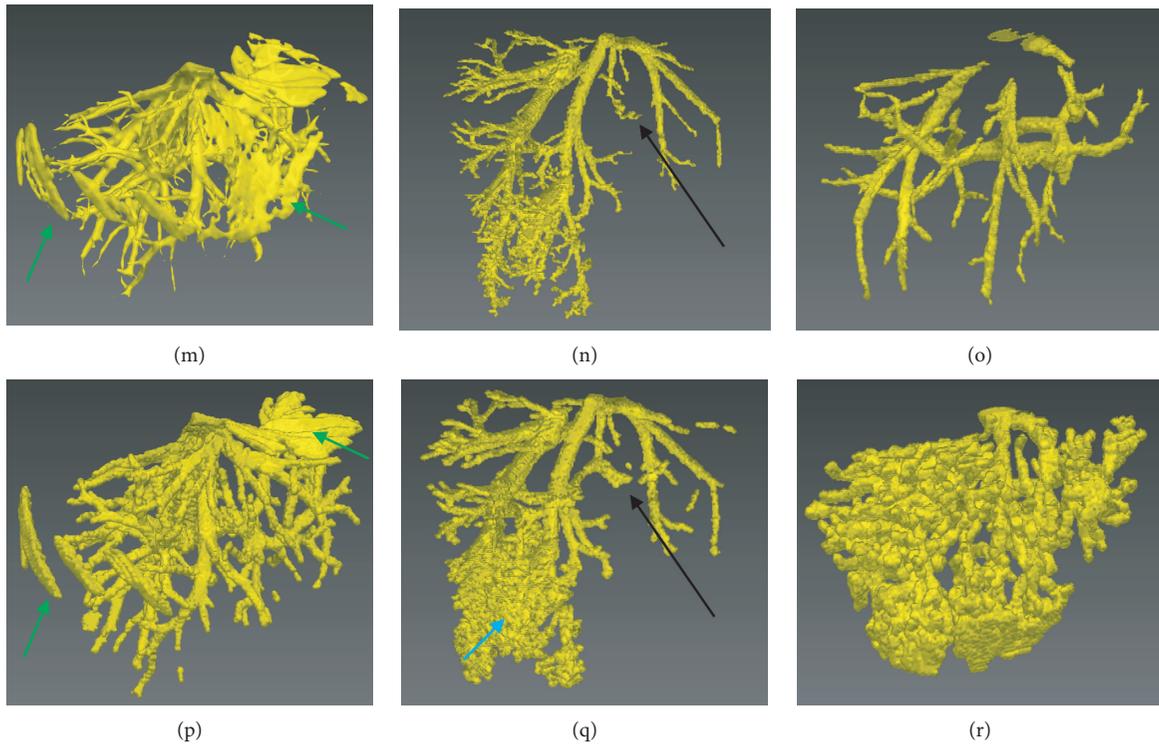


FIGURE 8: Comparison of the improved fuzzy connectedness (FC) method with traditional segmentation algorithms. Each column corresponds to one case. The original CT images of each case are shown in Figures 7(a)–7(i). (a)–(c) The gold standard of liver vessel. (d)–(f) The vessel segmented by using the improved FC. (g)–(i) The vessel segmented by using traditional FC with multiple potential seeds indicated by the blue points in Figure 2(b). (j)–(l) The vessel segmented by using traditional FC with one single seed indicated by the red point in Figure 2(b). (m)–(o) The vessel segmented by using region growing with multiple potential seeds. (p)–(r) The vessel segmented by using threshold level set with multiple potential seeds.

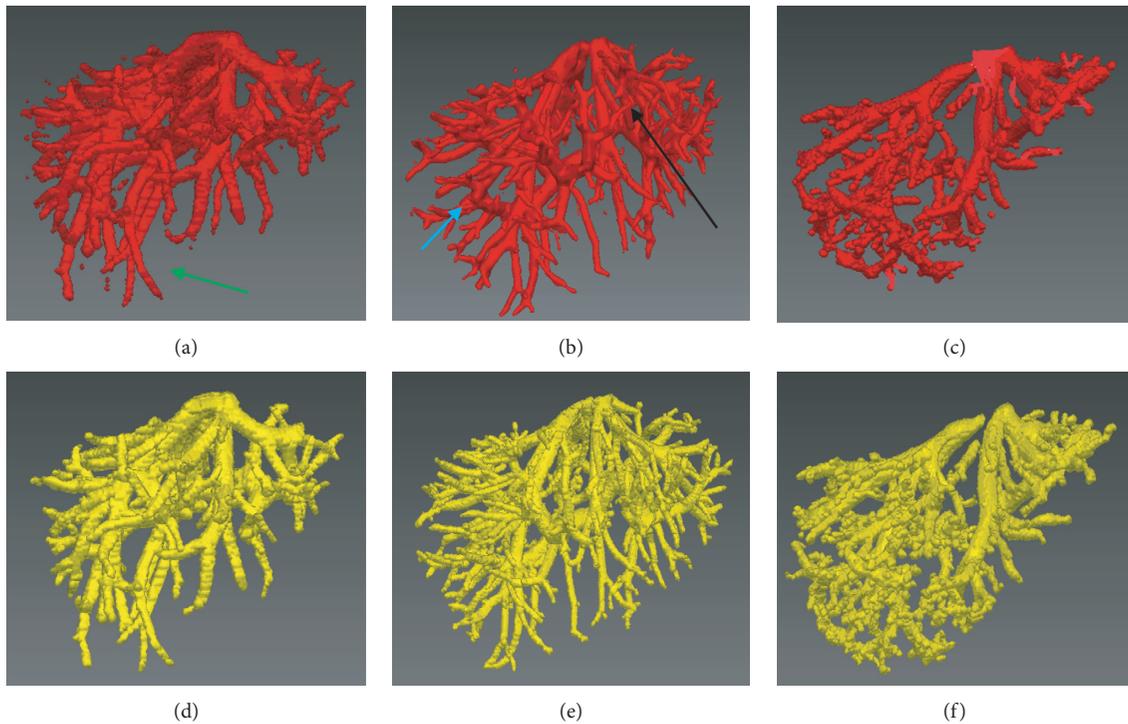


FIGURE 9: Continued.

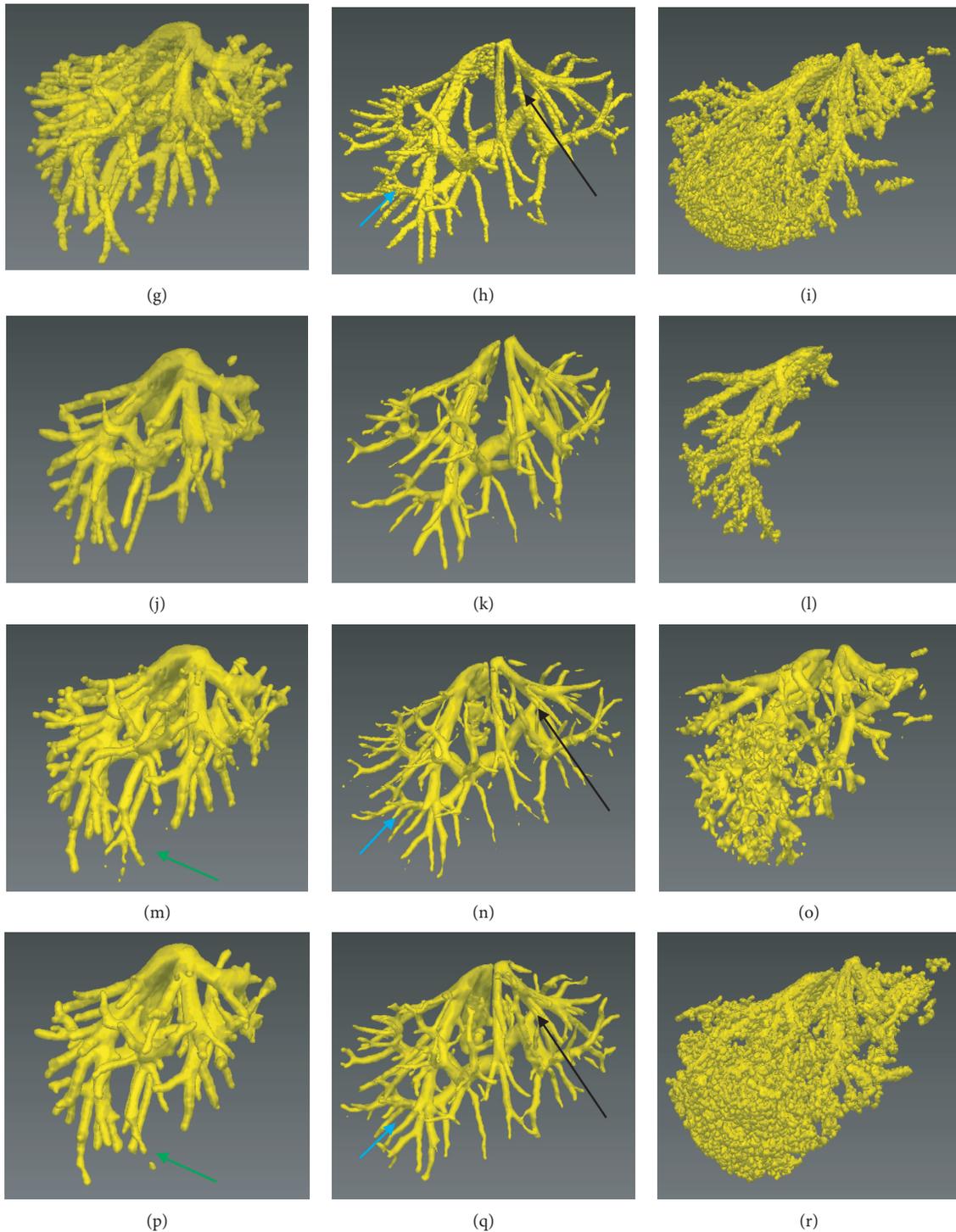


FIGURE 9: Comparison of the improved fuzzy connectedness (FC) method with traditional segmentation algorithms. Each column corresponds to one case. The original CT images of each case are shown in Figures 7(j)–7(r). (a)–(c) The gold standard of liver vessel. (d)–(f) The vessel segmented by using the improved FC. (g)–(i) The vessel segmented by using traditional FC with multiple potential seeds indicated by the blue points in Figure 2(b). (j)–(l) The vessel segmented by using traditional FC with one single seed indicated by the red point in Figure 2(b). (m)–(o) The vessel segmented by using region growing with multiple potential seeds. (p)–(r) The vessel segmented by using threshold level set with multiple potential seeds.

show that the Jerman's vesselness filter falsely enhanced the liver contour. Figure 10(c) shows that the Jerman's vesselness filter could not effectively suppress the background

(nonvessel) tissues. Quantitative comparison of the improved FC (with one single seed) with traditional FC, region growing, and threshold level set (with multiple

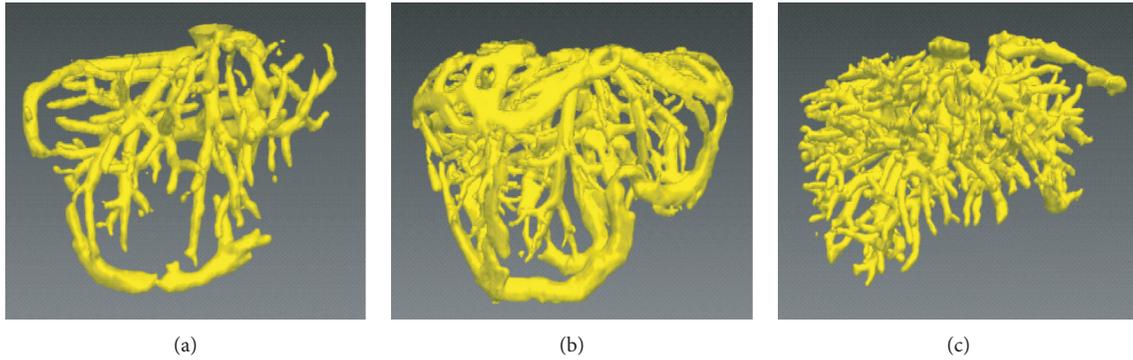


FIGURE 10: The liver vessel segmented by using the improved fuzzy connectedness method on the basis of the Jerman's vesselness filtering, rather than the improved vesselness filtering. (a) and (b) The Jerman's vesselness filter falsely enhances the liver contour. (c) The Jerman's vesselness filter fails to effectively suppress the background (nonvessel) tissues.

TABLE 2: Segmentation performance of the improved fuzzy connectedness (with one single seed), traditional fuzzy connectedness, region growing, and threshold level set (with multiple potential seeds).

	Improved fuzzy connectedness				Fuzzy connectedness				Region growing				Threshold level set			
	ACC (%)	SEN (%)	SPE (%)	DICE (%)	ACC (%)	SEN (%)	SPE (%)	DICE (%)	ACC (%)	SEN (%)	SPE (%)	DICE (%)	ACC (%)	SEN (%)	SPE (%)	DICE (%)
3Dircadb ($n = 20$)	96.4 ± 1.1	73.7 ± 7.6	97.4 ± 1.3	67.3 ± 5.7	82.2 ± 21.9	67.1 ± 25.5	82.5 ± 24.7	40.4 ± 16.2	84.1 ± 21.1	66.1 ± 25.0	86.4 ± 23.4	42.4 ± 15.6	94.3 ± 4.5	51.5 ± 20.3	97.0 ± 5.2	54.0 ± 13.9
Sliver07 ($n = 20$)	96.8 ± 0.6	89.1 ± 6.8	97.6 ± 1.1	71.4 ± 7.6	82.9 ± 25.5	85.0 ± 17.2	83.1 ± 26.7	45.1 ± 18.8	90.6 ± 12.6	84.1 ± 19.2	90.7 ± 13.6	56.2 ± 19.4	95.3 ± 2.9	82.5 ± 6.9	95.3 ± 3.2	67.5 ± 6.7

ACC = accuracy; SEN = sensitivity; SPE = specificity; DICE = Dice coefficient.

seeds) on the 3Dircadb ($n = 20$) and Sliver07 ($n = 20$) datasets are listed in Table 2 and shown in Figure 11, in terms of accuracy, sensitivity, specificity, and Dice coefficient. It can be observed that the improved FC outperformed traditional FC, region growing, and threshold level set. The average accuracy, sensitivity, specificity, and Dice coefficient of the improved FC method were, respectively, $(96.4 \pm 1.1)\%$, $(73.7 \pm 7.6)\%$, $(97.4 \pm 1.3)\%$, and $(67.3 \pm 5.7)\%$ for the 3Dircadb dataset and $(96.8 \pm 0.6)\%$, $(89.1 \pm 6.8)\%$, $(97.6 \pm 1.1)\%$, and $(71.4 \pm 7.6)\%$ for the Sliver07 dataset.

4. Discussion

4.1. Significance of This Study. 3D liver vessel segmentation is critical in computer-assisted liver tumor treatment planning and navigation. FC is an emerging method for image segmentation. However, traditional FC obtained unsatisfying performance for liver vessel segmentation in CT images, and it required multiple seeds and was sensitive to initialization. To address these issues, an improved FC method was proposed in this paper. Our method was fully automatic. The main contributions of this study were as follows. The Jerman's vesselness filter was improved by incorporating adaptive sigmoid filtering and a background-suppressing item. The improved vesselness filter effectively enhanced the vessel and suppressed the background. The improved vesselness response was incorporated into the fuzzy affinity function, increasing the

segmentation performance of FC. The fuzzy scene was initialized by two-threshold Otsu with one single seed, reducing the number of seeds and the sensitivity to initialization in traditional FC.

4.2. Implementation Details of the Algorithms. The algorithms described in this paper were implemented by using C++ and the Insight Segmentation and Registration Toolkit (ITK) (<http://itk.org>) [27]. The following ITK classes were mainly used:

- (1) The improved vesselness filter was implemented on the basis of the class `itk::HessianToObjectnessMeasureImageFilter`.
- (2) The improved FC method was implemented on the basis of the class `itk::SimpleFuzzyConnectednessScalarImageFilter`. This class was also used for the traditional FC segmentation.
- (3) The K -means clustering was implemented by using the class `itk::Statistics::ScalarImageKmeansImageFilter`.
- (4) The class `itk::SigmoidImageFilter` was used for sigmoid filtering.
- (5) The classes `itk::ResampleImageFilter` and `itk::IdentityTransform` were used for isotropic resampling.
- (6) The class `itk::ConfidenceConnectedImageFilter` was used for region growing segmentation.

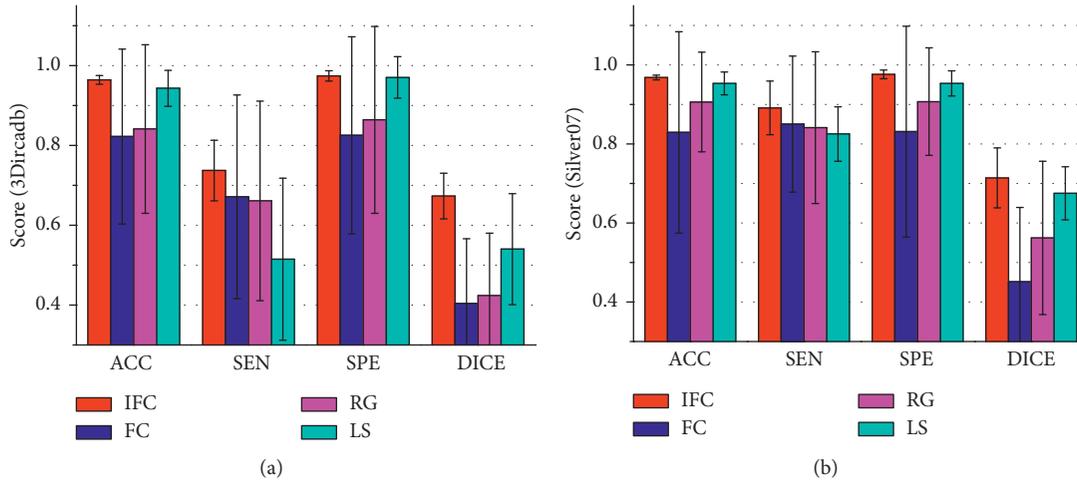


FIGURE 11: Comparison of the improved fuzzy connectedness (with one single seed) with traditional fuzzy connectedness, region growing, and threshold level set (with multiple potential seeds). (a) and (b) show the accuracy, sensitivity, specificity, and Dice coefficient for the 3Dircadb and Sliver07 datasets, respectively. ACC = accuracy; SEN = sensitivity; SPE = specificity; DICE = Dice coefficient; IFC = improved fuzzy connectedness; FC = fuzzy connectedness; RG = region growing; LS = threshold level set.

(7) The class `itk::ThresholdSegmentationLevelSetImageFilter` was used for threshold level set segmentation.

Average run time of the proposed algorithm was 200 s for 3Dircadb and 210 s for Sliver07. The improved vesselness filtering took approximately 30 s. The improved FC segmentation also took nearly 30 s. Each of the isotropic resampling and anisotropic resampling took around 60 s.

4.3. Sensitivity of the Proposed Algorithm to Key Algorithmic Parameters. Sensitivity analysis of key algorithmic parameters in Algorithms 1 and 2 was performed. The vessel scales σ_{\min} and σ_{\max} were set on the basis of the findings of Luu et al. [4]. Here, two key algorithmic parameters were analyzed: the threshold τ in the improved vesselness filter and the threshold T in the improved FC. The threshold τ in vesselness filter determined the degree of piecewise compensation on the eigenvalue λ_3 . In theory, the smaller the threshold τ is, the more enhancement on the vessel boundary would be obtained; however, a too small τ is prone to cause undersegmentation. The threshold T in FC determined the degree of undersegmentation or oversegmentation. A too small T caused undersegmentation, while a too large T resulted in oversegmentation. The value of T from 0.01 to 0.09 was tested, as the segmented vasculature would be incomplete when $T > 0.1$. For the compromise between undersegmentation and oversegmentation, the value of T was firstly fixed to 0.05 to analyze the sensitivity of the proposed algorithm to the threshold τ . Figure 12 shows the average accuracy, sensitivity, specificity, and Dice coefficient of the proposed method on 10 cases randomly selected from the 3Dircadb dataset. The threshold τ ranged from 0.1 to 0.9 ($T = 0.05$). The accuracy and Dice coefficient reached peak when τ was optimally set at 0.6. Then, the value of τ was fixed to 0.6 to analyze the sensitivity of the proposed algorithm to the threshold

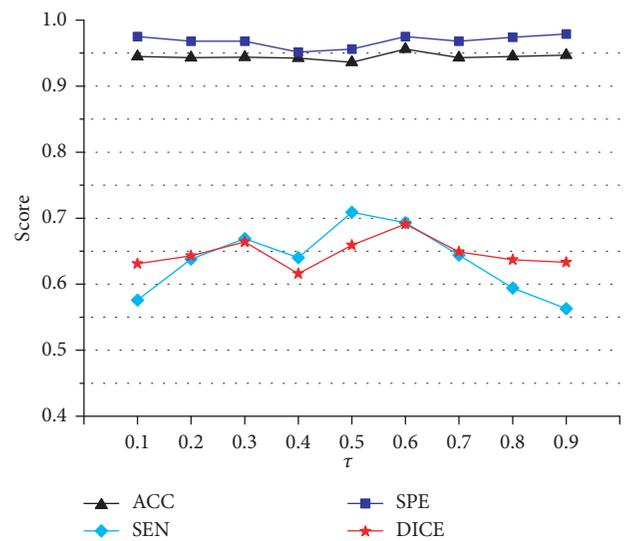


FIGURE 12: Evaluation of the segmentation performance of the improved fuzzy connectedness method on 10 cases randomly selected from the 3Dircadb dataset for the values of T ranging from 0.1 to 0.9 ($T = 0.05$). The value of T is optimally set at 0.6. ACC = accuracy; SEN = sensitivity; SPE = specificity; DICE = Dice coefficient.

T . Figure 13 shows the segmentation performance of the proposed method on the 10 randomly selected cases, with T ranging from 0.01 to 0.09 ($\tau = 0.6$). Based on the maximum value of the accuracy and Dice coefficient, the parameter T was optimally set at 0.05.

4.4. Comparison with Related Work. Table 3 shows a comparison of the improved FC method with related work in terms of segmentation method, dataset, number of cases, automation, precision, and run time. For the run time of the proposed method, it should be noted that each of the

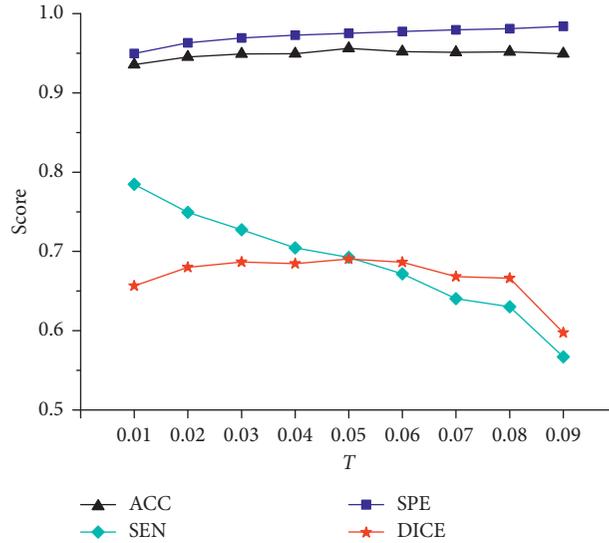


FIGURE 13: Evaluation of the segmentation performance of the improved fuzzy connectedness method on 10 cases randomly selected from the 3Dircadb dataset for the values of T ranging from 0.01 to 0.09 ($T = 0.6$). The value of T is optimally set at 0.05. ACC = accuracy; SEN = sensitivity; SPE = specificity; DICE = Dice coefficient.

TABLE 3: Comparison of the proposed method with related work.

Author	Year	Method	Dataset	#	Automation	Precision (%)	Run time (s)
Oliveira et al. [7]	2011	RG	Sliver07	20	Auto	—	—
Luu et al. [4]	2015	RG	Clinical CTA	51	Auto	ACC = 86.2; SEN = 85.1; SPE = 92.3	—
Esneault et al. [10]	2010	GC	Clinical CTA	1	Auto	—	10–100
Zeng et al. [12]	2017	GC	Clinical CTA	6	Auto	ACC = 97.7; SEN = 79.8; SPE = 98.6	390
Sangsefidi et al. [11]	2018	GC	3Dircadb/Clinical CTA	7	Auto	DICE = 74.0	560
Shang et al. [9]	2011	LS	Clinical CTA	20	Auto	SEN* = 91.0	480
Ahmadi et al. [28]	2016	FCC	Sliver07	20	Auto	ACC = 91.0; SEN = 94.1; SPE = 83.6	27.1
Zeng et al. [13]	2016	ML	Clinical CTA	6	Auto	ACC = 98.1; SEN = 74.2; SPE = 99.3	0.05–0.1
Guo et al. [15]	2015	FC	Clinical CTA	4	Semi	—	112.5
Wang et al. [16]	2016	FC	Clinical CTA	3	Semi	—	22
Huang et al. [14]	2018	DL	3Dircadb	20	Auto	ACC = 97.1; SEN = 74.3; SPE = 98.3; DICE = 67.5	230
Ours	2018	IFC	3Dircadb	20	Auto	ACC = 96.4; SEN = 73.7; SPE = 97.4; DICE = 67.3	200
			Sliver07	20	Auto	ACC = 96.8; SEN = 84.4; SPE = 97.6; DICE = 71.4	210

* Evaluation by the number of vascular nodes; CTA = computed tomography angiography; RG = region growing; GC = graph cuts; LS = level set; FCC = fuzzy C-means clustering; ML = machine learning; FC = fuzzy connectedness; DL = deep learning; IFC = improved fuzzy connectedness; ACC = accuracy; SEN = sensitivity; SPE = specificity; DICE = Dice coefficient; Auto = automatic; Semi = semiautomatic.

improved vesselness filtering and the improved FC segmentation only took around 30 s.

Firstly, the proposed method was compared with related work that used the 20 cases of the Sliver07 training dataset. Oliveira et al. [7] used region growing for liver vessel segmentation, but they only performed visual assessment for the segmentation. Ahmadi et al. [28] segmented liver vessel by using fuzzy C-means clustering and initialized the parameters by the genetic algorithm. Though the run time was shorter, the training process was more complex, and the accuracy and specificity of Ahmadi et al. [28] were lower than those of the proposed method. Then, the proposed method was compared with related work that used the 3Dircadb dataset. Huang et al. [14] segmented liver vessel on

the 20 cases of 3Dircadb by using the 3D U-Net network. Their method reduced the need for the quantity of training data, but it required long training time (48 h). The accuracy, sensitivity, specificity, and Dice coefficient of Huang et al. [14] were slightly higher than those of the proposed method. Sangsefidi et al. [11] employed graph cuts for segmenting liver vessel, but they evaluated their method on only few cases of 3Dircadb.

Finally, the proposed method was compared with related work that used clinical data other than Sliver07 and 3Dircadb. These studies mostly used CT angiography (CTA) images, which were specific CT for vasculature with clear vascular boundary. However, in the context of computer-assisted liver tumor treatment planning and

navigation, CE-CT images may be used more commonly, as liver tumors could be observed in CE-CT images. Though region growing methods had relatively higher operation efficiency, they are depended on the number and distribution of seeds, resulting in unsatisfied segmentation performance even in high-contrast CTA images [4]. Graph cuts and level set methods would take long time to segment liver vessel [9–12]. Esneault et al. [10] just showed the segmentation on one case of data, and they reported that the segmented vascular branches needed to be registered, which would take more time. Zeng et al. [12] reported that their method only achieved good performance on high-contrast CTA images, so their method might be restricted in practical applications when only low-contrast CT images are available. The similar issue existed in Shang et al. [9] and Zeng et al. [13]. Shang et al. [9] evaluated the sensitivity by the number of vascular nodes (denoted as SEN* in Table 3), but this evaluation metric may not be rigorous. In comparison with Guo et al. [15] and Wang et al. [16] which increased the time efficiency of traditional FC, this study focused on improving the segmentation performance and reducing the number of seeds and the sensitivity to initialization. In addition, our method did not require manual interaction to select the seed.

4.5. Limitations and Future Work. One limitation of this study is the small number of clinical data with the gold standard (40 cases). More clinical data may be used in future (if possible) to further verify the performance of the proposed method. In addition, the algorithmic steps of isotropic resampling and anisotropic resampling are time consuming, each taking around 60 s. This limitation may be overcome in future work.

5. Conclusions

An improved FC method was presented for automatic liver vessel segmentation in CT volumetric images. The Jerman's vesselness filter was improved by incorporating adaptive sigmoid filtering and a background-suppressing item. The improved vesselness filter effectively enhanced the liver vessel while suppressing the background. The improved vesselness response was incorporated into the fuzzy affinity function of FC. The fuzzy scene was initialized by two-threshold Otsu with one single seed generated automatically, reducing the number of seeds and the sensitivity to initialization in traditional FC. The improved FC method was evaluated on 40 cases of clinical CT volumetric images. Experimental results showed that the proposed liver vessel segmentation strategy could achieve better segmentation performance than traditional FC, region growing, and threshold level set. It is concluded that the proposed algorithm may be used as a new method for automatic 3D liver vessel segmentation in CT images.

Data Availability

The VasuSynth dataset is publically available at <http://vasusynth.cs.sfu.ca/>. The 3DirCADb dataset is publically

available at <http://www.ircad.fr/research/3dircadb>. The training data of the Sliver07 dataset are publically available at <http://www.sliver07.org/>.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant Nos. 61801312 and 71661167001), Beijing Natural Science Foundation (Grant No. 4184081), China Postdoctoral Science Foundation (Grant No. 2017M620566), Postdoctoral Research Fund of Chaoyang District, Beijing (Grant No. 2017ZZ-01-03), and Basic Research Fund of Beijing University of Technology.

References

- [1] J. Ferlay, I. Soerjomataram, R. Dikshit et al., "Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012," *International Journal of Cancer*, vol. 136, no. 5, pp. E359–E386, 2015.
- [2] W. Chen, R. Zheng, P. D. Baade et al., "Cancer statistics in China," *CA: A Cancer Journal for Clinicians*, vol. 66, no. 2, pp. 115–132, 2016.
- [3] M. W. Lee, S. S. Raman, N. H. Asvadi et al., "Radiofrequency ablation of hepatocellular carcinoma as bridge therapy to liver transplantation: a 10-year intention-to-treat analysis," *Hepatology*, vol. 65, no. 6, pp. 1979–1990, 2017.
- [4] H. M. Luu, C. Klink, A. Moelker et al., "Quantitative evaluation of noise reduction and vesselness filters for liver vessel segmentation on abdominal CTA images," *Physics in Medicine and Biology*, vol. 60, no. 10, pp. 3905–3926, 2015.
- [5] M. Moghbel, S. Mashohor, R. Mahmud, and M. I. B. Saripan, "Review of liver segmentation and computer assisted detection/diagnosis methods in computed tomography," *Artificial Intelligence Review*, 2018, In press.
- [6] D. Selle, B. Preim, A. Schenk, and H. O. Peitgen, "Analysis of vasculature for liver surgical planning," *IEEE Transactions on Medical Imaging*, vol. 21, no. 11, pp. 1344–1357, 2002.
- [7] D. A. Oliveira, R. Q. Feitosa, and M. M. Correia, "Segmentation of liver, its vessel and lesions from CT images for surgical planning," *BioMedical Engineering OnLine*, vol. 10, no. 1, p. 30, 2011.
- [8] Y. Wang, B. Fang, J. Pi et al., "Automatic multi-scale segmentation of intrahepatic vessel in CT images for liver surgery planning," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 27, no. 1, article 1357001, 2013.
- [9] Y. Shang, R. Deklerck, E. Nyssen et al., "Vascular active contour for vessel tree segmentation," *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 4, pp. 1023–1032, 2011.
- [10] S. Esneault, C. Lafon, and J. L. Dillenseger, "Liver vessel segmentation using a hybrid geometrical moments/graph cuts method," *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 2, pp. 276–283, 2010.
- [11] N. Sangsefidi, A. H. Foruzan, and A. Dolati, "Balancing the data term of graph-cuts algorithm to improve segmentation of hepatic vascular structures," *Computers in Biology and Medicine*, vol. 93, pp. 117–126, 2017.
- [12] Y. Z. Zeng, Y. Q. Zhao, P. Tang et al., "Liver vessel segmentation and identification based on oriented flux symmetry

- and graph cuts,” *Computer Methods and Programs in Biomedicine*, vol. 150, pp. 31–39, 2017.
- [13] Y. Z. Zeng, Y. Q. Zhao, M. Liao et al., “Liver vessel segmentation based on extreme learning machine,” *Physica Medica*, vol. 32, no. 5, pp. 709–716, 2016.
- [14] Q. Huang, J. Sun, H. Ding, X. Wang, and G. Wang, “Robust liver vessel extraction using 3D U-Net with variant dice loss function,” *Computers in Biology and Medicine*, vol. 101, pp. 153–162, 2018.
- [15] X. Guo, S. Huang, X. Fu, and X. Huang, “Vascular segmentation in hepatic CT images using adaptive threshold fuzzy connectedness method,” *Biomedical Engineering Online*, vol. 14, no. 1, p. 57, 2015.
- [16] L. Wang, D. Li, and S. Huang, “An improved parallel fuzzy connected image segmentation method based on CUDA,” *Biomedical Engineering Online*, vol. 15, no. 1, p. 56, 2016.
- [17] T. Jerman, F. Pernuš, B. Likar, and Ž. Špiclin, “Enhancement of vascular structures in 3D and 2D angiographic images,” *IEEE Transactions on Medical Imaging*, vol. 35, no. 9, pp. 2107–2118, 2016.
- [18] W. Wu, Z. Zhou, S. Wu, and Y. Zhang, “Automatic liver segmentation on volumetric CT images using supervoxel-based graph cuts,” *Computational and Mathematical Methods in Medicine*, vol. 2016, Article ID 9093721, 2016.
- [19] P. Jassi and G. Hamarneh, “Vascusynth: vascular tree synthesis software,” *Insight Journal*, pp. 1–12, 2011, <http://hdl.handle.net/10380/3260>.
- [20] 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, 2010.
- [21] Y. Sato, S. Nakajima, N. Shiraga et al., “Three-dimensional multi-scale line filter for segmentation and visualization of curvilinear structures in medical images,” *Medical Image Analysis*, vol. 2, no. 2, pp. 143–168, 1998.
- [22] A. F. Frangi, W. J. Niessen, K. L. Vincken, and M. A. Viergever, “Multiscale vessel enhancement filtering,” in *Proceedings of International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 130–137, Springer, Cambridge, MA, USA, October 1998.
- [23] Q. Li and S. Sone, “Selective enhancement filters for nodules, vessels, and airway walls in two-and three-dimensional CT scans,” *Medical Physics*, vol. 30, no. 8, pp. 2040–2051, 2003.
- [24] M. Erdt, M. Raspe, and M. Suehling, “Automatic hepatic vessel segmentation using graphics hardware,” in *Medical Imaging and Augmented Reality*, pp. 403–412, Springer, Heidelberg, Germany, 2008.
- [25] C. Xiao, M. Staring, D. Shamonin et al., “A strain energy filter for 3D vessel enhancement with application to pulmonary CT images,” *Medical Image Analysis*, vol. 15, no. 1, pp. 112–124, 2011.
- [26] A. Pednekar, I. A. Kakadiaris, and U. Kurkure, “Adaptive fuzzy connectedness-based medical image segmentation,” in *Proceedings of Indian Conference on Computer Vision, Graphics & Image Processing*, Ahmedabad, India, December 2002.
- [27] L. Ibanez, W. Schroeder, L. Ng, and J. Cates, *The ITK Software Guide*, U.S. National Library of Medicine, Bethesda, Maryland, 2005.
- [28] K. Ahmadi, A. Karimi, and N. B. Fouladi, “New technique for automatic segmentation of blood vessels in CT scan images of liver based on optimized fuzzy-means method,” *Computational and Mathematical Methods in Medicine*, vol. 2016, Article ID 5237191, 2016.

Research Article

Pixel-Label-Based Segmentation of Cross-Sectional Brain MRI Using Simplified SegNet Architecture-Based CNN

Bijen Khagi and Goo-Rak Kwon 

Department of Information and Communication Engineering, Chosun University, 375 Seosuk-Dong, Dong-Gu, Gwangju 501-759, Republic of Korea

Correspondence should be addressed to Goo-Rak Kwon; grkwon@chosun.ac.kr

Received 19 July 2018; Accepted 25 September 2018; Published 28 October 2018

Guest Editor: Cesare Valenti

Copyright © 2018 Bijen Khagi and Goo-Rak Kwon. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Using deep neural networks for segmenting an MRI image of heterogeneously distributed pixels into a specific class assigning a label to each pixel is the concept of the proposed approach. This approach facilitates the application of the segmentation process on a preprocessed MRI image, with a trained network to be utilized for other test images. As labels are considered expensive assets in supervised training, fewer training images and training labels are used to obtain optimal accuracy. To validate the performance of the proposed approach, an experiment is conducted on other test images (available in the same database) that are not part of the training; the obtained result is of good visual quality in terms of segmentation and quite similar to the ground truth image. The average computed Dice similarity index for the test images is approximately 0.8, whereas the Jaccard similarity measure is approximately 0.6, which is better compared to other methods. This implies that the proposed method can be used to obtain reference images almost similar to the segmented ground truth images.

1. Introduction

Deep neural networks have been highly successful in segmenting outdoor scenes with high complexity, dissimilar patterns, variable texture, and wide pixel range. In the present study, this model is used for segmenting MRI images of the brain, which are relatively simpler than outdoor scenes. The precise segmentation of a 2D image has always been a challenging task, and various approaches have been proposed for better accuracy, such as supervised and unsupervised, manual and automatic, and standalone and neural network-based techniques. Similarly, deep convolutional neural networks (CNNs) have been effective in machine learning and have had impact on various industrial, medical, and commercial fields. Generally, image segmentation is the process of presenting and partitioning image content into distinguishable parts. Moreover, segmentation methods from edge detection as well as supervised and unsupervised methods have been proposed. Similarly, neural networks have been developed for medical image processing, particularly in MRI image segmentation and Alzheimer's disease classification [1, 2]. Brain MRI segmentation is

fundamental in several clinical applications and influences the outcome of the entire analysis because various processing operations rely on accurate segmentation of anatomical and structural regions. For instance, MRI segmentation is frequently used for calculating and imagining different brain structures, delineating lesions, analyzing brain development, and image-guided intrusions and surgical preparation. In MRI, tissue is heterogeneously concentrated in terms of intensity owing to the bias field and the partial volume effect that reflects the actual content of the brain, namely, white matter (WM), gray matter (GM), and cerebrospinal fluid (CSF). Therefore, accurate and selective methods should be chosen.

In contrary to existing methods [3], which have a certain way of feature extraction and criteria like thresholding, contours, and clustering, this method has been used extensively by many researchers and found to be excellent in case of MRI segmentation as well. But on the contrary, deep neural network are now proving to be better, highly computational for large data, and powerful because of encoder-decoder-based network or CNN architecture. The features are automatically investigated from low level features like

edge, blob, and line to high level features like color, shape, and detail in a hierarchical manner by each layer. The activation layer like ReLU helps to make those features more clear and computable. Hence, we can easily get our segmentation result using our model. The only problem will be to train the network as it requires a large amount of ground truth and design the network appropriately.

2. Background and Methodology

2.1. Semantic Segmentation. In semantic segmentation, the image is segmented on a pixel-label basis, that is, each pixel is associated with a certain defined class. Its applications include scene understanding, autonomous driving, object recognition, machine translation, and machine vision. Semantic segmentation has been improved by using full CNNs [4] and deep CNNs [5–8]. These neural networks are trained in an end-to-end, pixel-to-pixel manner on each layer for image segmentation.

2.2. SegNet Layer. The SegNet layer is a deep full CNN architecture adapted for semantic segmentation that was proposed by Vijay Badrinarayanan et al. [5]. Generally, the semantic segmentation approach is used for outdoor, indoor, and road scenes mostly for a large number of classes. SegNet was originally designed for scene understanding applications. Hence, it should be efficient in terms of memory, operation, and computational time. It is also considerably smaller in terms of the number of trainable parameters than other competing architectures, and it can be used in training end-to-end pixel-label classes using stochastic gradient descent and the cross-entropy loss function.

The encoder used in SegNet is identical to the convolutional layers in VGG16 [9]. The fully connected layers of VGG16 have been removed in SegNet, and thus the encoder network is considerably reduced and easier to train compared to other recent architectures [5, 6, 10, 11]. The most important constituent of SegNet is the encoder-decoder network, which consists of a hierarchy of downsampling encoders matching each upsampling decoder with associated feature vectors cycling inside them.

2.3. CNN and Architecture. CNNs have always been important in machine learning; by using various types of neural networks, systematic training and testing of image and pixel labels can be performed. The encoder network used here consists of convolution layers of 64 filters, each of size 3×3 , manually padded, followed by batch normalization and ReLU activation unit and repeatedly followed by same convolution, batch normalization, and ReLU for proper downsampling and robust feature extraction. Same is the case with decoder convolution network but firstly unpool layer and then convolution layer following batch normalization and ReLU.

Proposed CNN has an encoder network and a matching decoder network, which is followed by a final pixel-based classification layer. This architecture is shown in Figure 1. To simplify the architecture, two encoder and two decoder networks have been employed: *encoder1* is mapped to

decoder1, and *encoder2* is mapped to *decoder2*. *encoder1* consists of *encoder1_conv1*, *encoder1_bn_1*, *encoder1_relu_1*, and *encoder1_maxpool_1* in hierarchical order, whereas *dencoder1* consists of *dencoder1_unpool_1*, *dencoder1_conv1*, *dencoder1_bn_1*, and *dencoder1_relu_1*. *encoder2* and *decoder2* are similarly structured. Here, *encoder1* is followed by *encoder2*, and *dencoder2* is followed by *dencoder1*, as shown in Figure 2. The first 13 layers constitute an encoder network that performs the convolution with 64 filter banks of size 3×3 to obtain sets of features along with batch normalization in a minibatch set of 8 images. ReLU acts as an activation function $f(x) = \max(0, x)$, which can be used by neurons, as any other activation function, to eliminate negative values. Thereafter, the max pooling layer with a 2×2 window and stride size 2 (nonoverlapping window) is executed, so that the resulting output is downsampled by a factor of 2. Multiple layers of max pooling downsampling are used to achieve more translation invariance and robust pixel classification. Similarly, the decoder in the decoder network upsamples the input layer feature maps unpooling the memorized max pooling indices with the location of maximum feature values from the corresponding encoder feature maps. It is followed by the convolution and batch normalization layers to produce dense features that are similar in size to the input image. The details of the simplified architecture are tabulated in Table 1.

3. Experimental Setup

For the experiments, T1-weighted structural brain MRI data were used that are available on OASIS (open access series of imaging studies). OASIS is an open access website [12], created by the Alzheimer's Disease Research Center at Washington University. The dataset consists mainly of brain MRI images from Alzheimer's disease patients aged 18 to 96 and normal human brain MRI images for comparative study. All experiments were conducted using Matlab R2017b on an i3 4160, 4 GB RAM windows desktop. To reduce computation time, the neural network was trained by a single GeForce GTX 1050 Ti GPU using parallel computing.

3.1. Image Extraction and Preprocessing. The dataset consisted of several types of MRI scans with raw, processed, and segmented 3D raw files or analyze format file (.img .hdr). Cross-sectional averaged and coregistered scan images were used that were obtained in the native acquisition space resampled to 1 mm isotropic voxels [12] from 50 subjects (cross-sectional MRI brain scans of dimensions $208 \times 176 \times 160$). The MRIcon software package was used to extract slices from each mid cross-sectional MRI to generate images of size 208×176 pixels, each representing a single MRI scan.

Two disc images were selected from ID OAS1_0001_MR1 to OAS1_0080_MR1, consisting of 76 images originally. The skull stripped image was used as training image, and the segmented images (the image is already segmented into four parts) of each training image were used as training labels or ground truth. Later, the trained network was used to segment the test MRI images, and the result was compared with the

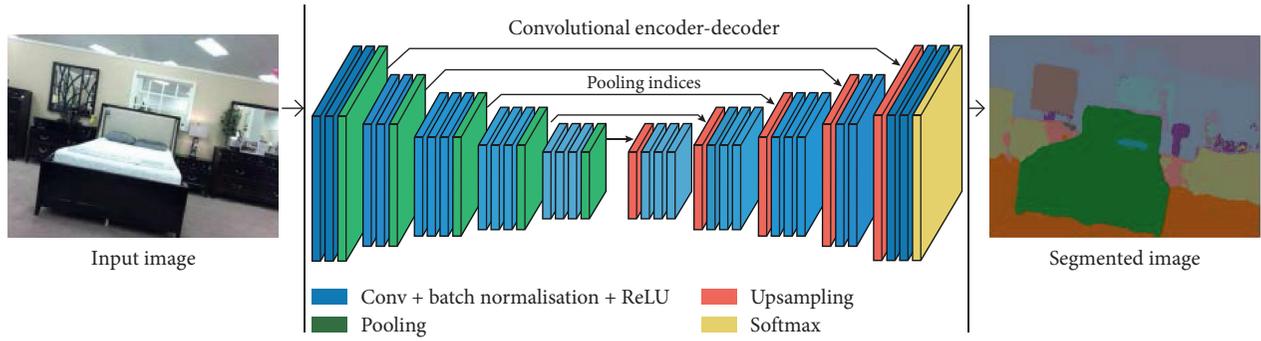


FIGURE 1: SegNet architecture pictorial representation as presented by Vijay Badrinarayanan et al. [5].

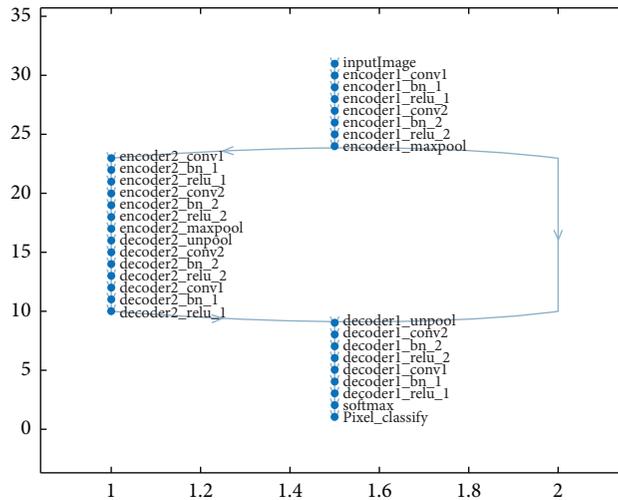


FIGURE 2: Schematic representation of 31 layers and 34 connections used in proposed CNN Network.

TABLE 1: Simplified SegNet network.

S. No.	Layer name	Type	Description
1	“Image input”	Image	208 × 1761 images with “zero center” normalization
2	“encoder1_conv1”	Convolution	64 3 × 3 × 1 convolutions with stride [1 1] and padding [1 1 1 1]
3	“encoder1_bn_1”	Batch normalization	Batch normalization with 64 channels
4	“encoder1_relu_1”	ReLU	ReLU
5	“encoder1_conv2”	Convolution	64 3 × 3 × 64 convolutions with stride [1 1] and padding [1 1 1 1]
6	“encoder1_bn_2”	Batch normalization	Batch normalization with 64 channels
7	“encoder1_relu_2”	ReLU	ReLU
8	“encoder1_maxpool”	Max pooling	2 × 2 max pooling with stride [2 2] and padding [0 0 0 0]
9	“encoder2_conv1”	Convolution	64 3 × 3 × 64 convolutions with stride [1 1] and padding [1 1 1 1]
10	“encoder2_bn_1”	Batch normalization	Batch normalization with 64 channels
11	“encoder2_relu_1”	ReLU	ReLU
12	“encoder2_conv2”	Convolution	64 3 × 3 × 64 convolutions with stride [1 1] and padding [1 1 1 1]
13	“encoder2_bn_2”	Batch normalization	Batch normalization with 64 channels
14	“encoder2_relu_2”	ReLU	ReLU
15	“encoder2_maxpool”	Max pooling	2 × 2 max pooling with stride [2 2] and padding [0 0 0 0]
16	“decoder2_unpool”	Max unpooling	Max unpooling
17	“decoder2_conv2”	Convolution	64 3 × 3 × 64 convolutions with stride [1 1] and padding [1 1 1 1]
18	“decoder2_bn_2”	Batch normalization	Batch normalization with 64 channels
19	“decoder2_relu_2”	ReLU	ReLU
20	“decoder2_conv1”	Convolution	64 3 × 3 × 64 convolutions with stride [1 1] and padding [1 1 1 1]
21	“decoder2_bn_1”	Batch normalization	Batch normalization with 64 channels
22	“decoder2_relu_1”	ReLU	ReLU
23	“decoder1_unpool”	Max unpooling	Max unpooling
24	“decoder1_conv2”	Convolution	64 3 × 3 × 64 convolutions with stride [1 1] and padding [1 1 1 1]
25	“decoder1_bn_2”	Batch normalization	Batch normalization with 64 channels

TABLE 1: Continued.

S. No.	Layer name	Type	Description
26	“decoder1_relu_2”	ReLU	ReLU
27	“decoder1_conv1”	Convolution	$4 \times 3 \times 3 \times 64$ convolutions with stride [1 1] and padding [1 1 1 1]
28	“decoder1_bn_1”	Batch normalization	Batch normalization with 4 channels
29	“decoder1_relu_1”	ReLU	ReLU
30	“Softmax”	Softmax	Softmax
31	“Pixel_classify”	Pixel classification layer	Class weighted cross-entropy loss with “background,” “CSF,” “GM,” and “WM” classes

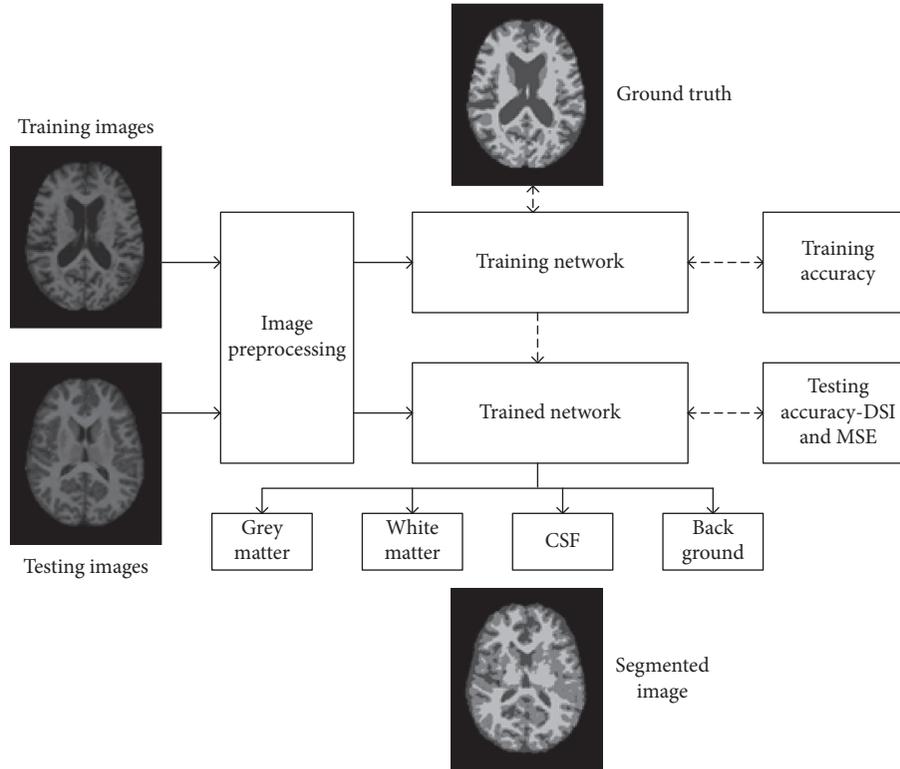


FIGURE 3: Representation of the proposed method, the training and testing image is separated, so that training images after preprocessing feeds into CNN network along with its respective ground truth, after reaching to convergence, the training is stopped and the network is now called trained. This trained network is used to test other test image separately to get segmented image, which is compared with its ground truth itself for performance analysis.

TABLE 2: Training accuracy, intersection over union (IoU), and MeanBFscore for each assigned class.

	Accuracy	IoU	MeanBFscore
Background	0.98877	0.9855	0.99456
CSF	0.92848	0.66983	0.8923
Gray	0.77666	0.66022	0.93897
White	0.83603	0.79073	0.90347

ground truth segmentation. Regarding the training environment, “Stochastic Gradient Descent with Momentum” was selected as the training optimization algorithm, with an initial learning rate of 0.001. To facilitate smooth training, the training was carried out in minibatches of 8 files per epoch, with data augmentation carried out at a random reflection in X-axis and rotation of ± 10 degrees from the original position of each image. The predesigned SegNet layers created a training network, which was to undergo

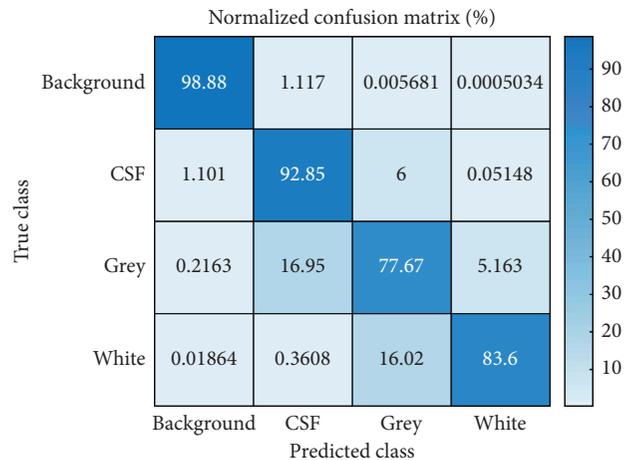


FIGURE 4: Confusion matrix of the experiment. The diagonal represents the accuracy of predicted class versus true class.

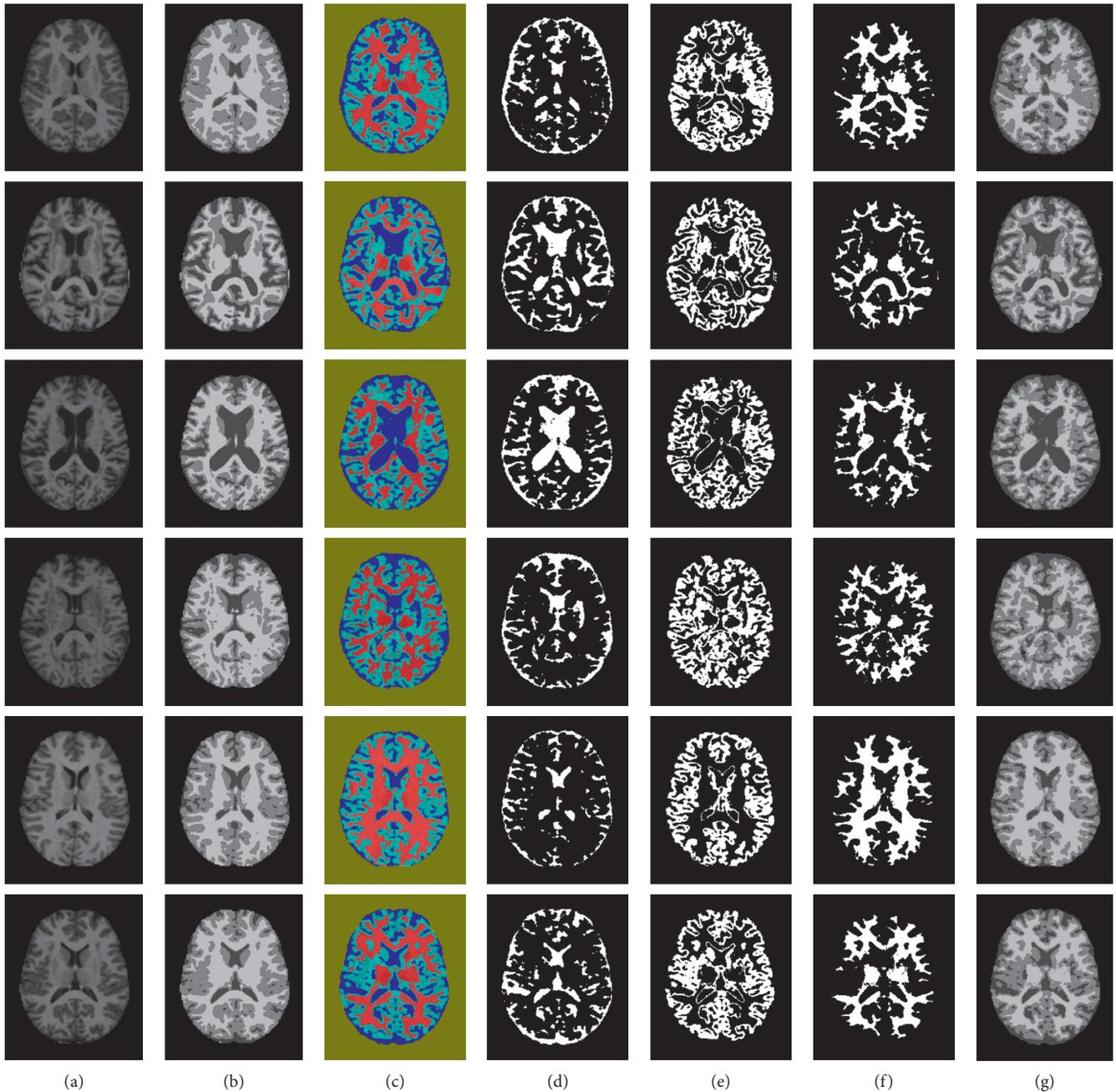


FIGURE 5: Original image and ground truth image presented along with other images, as a result of proposed segmentation: (a) original image, (b) ground truth, (c) segmented image by proposed method (color), (d) CSF part (binary image), (e) GM part (binary image), (f) WM part (binary image), and (g) segmented Image in gray (c).

a stepwise feature extraction process on each CNN layer. Additionally, the classification of pixels was facilitated by classweight–classname pairs and a cross-entropy loss function. The SegNet layer acted as the training framework, whereas the pixel classification layer acted as classification output. The overall workflow of proposed method is illustrated in Figure 3.

3.2. Training and Testing Accuracy. Seventy-six images were selected from ID OAS1_0001_MR1 to OAS1_0080_MR1 for

training (including the augmented images) excluding four missing MRI and six images from ID OAS1_0081_MR1 to OAS1_0087_MR1 for testing excluding OAS1_0082_MR1. The overall training accuracy was 91.47 with mean global accuracy 0.91, mean accuracy 0.88248, mean IoU 0.88248, and WeightedIoU 0.84. The clusterwise accuracy, IoU, and MeanBFScore are tabulated in Table 2. The intersection over union (IoU) for the best predicted image was approximately 0.8477, whereas IoU for the worst predicted image was approximately 0.625. The confusion matrix obtained from the classification is shown in Figure 4. The obtained result

TABLE 3: Comparison of performance parameters for each result image (Figure 5(g)), with respective ground truth image (Figure 5(b)).

Test image ID	Parameter	CSF part	Gray part	White part	Mean value
OAS1_0081_MR1	Dice similarity	0.54	0.75	0.85	0.71
	Jaccard similarity	0.37	0.59	0.74	0.57
	Mean squared error	—	—	—	29.47
OAS1_0083_MR1	Dice similarity	0.84	0.75	0.79	0.80
	Jaccard similarity	0.73	0.60	0.66	0.66
	Mean squared error	—	—	—	19.32
OAS1_0084_MR1	Dice similarity	0.85	0.71	0.78	0.78
	Jaccard similarity	0.74	0.55	0.64	0.64
	Mean squared error	—	—	—	25.02
OAS1_0085_MR1	Dice similarity	0.72	0.67	0.73	0.71
	Jaccard similarity	0.56	0.51	0.57	0.55
	Mean squared error	—	—	—	32.52
OAS1_0086_MR1	Dice similarity	0.74	0.85	0.92	0.84
	Jaccard similarity	0.59	0.74	0.85	0.73
	Mean squared error	—	—	—	9.52
OAS1_0087_MR1	Dice similarity	0.64	0.75	0.85	0.74
	Jaccard similarity	0.47	0.60	0.74	0.60
	Mean squared error	—	—	—	27.58

TABLE 4: Comparison of deep learning approaches for brain structure segmentation.

Authors	CNN style	Dimension	Accuracy	Data
Zhang et al. [13]	Patchwise	2D	DSC 83.5% (CSF), 85.2% (GM), 86.4% (WM)	Private data (10 healthy infants)
Nie et al. [14]	Semantic-pixelwise	2D	DSC 85.5% (CSF), 87.3% (GM), 88.7% (WM)	Private data (10 healthy infants)
de Brebisson et al. [15]	Patchwise	2D/3D	Overall DSC 72.5% \mp 16.3%	MICCAI 2012-multi-atlas labeling
Moeskops et al. [16]	Patchwise	2D/3D	Overall DSC 73.53%	MICCAI 2012-multi-atlas labeling
Proposed method	Pixel-label semantic (SegNet CNN)	2D	DSC 72.2% (CSF), 74.6% (GM), 81.9% (WM)	OASIS cross-sectional MRI

shows, out of total training image, the network could correctly classify around 99% of background pixel, 93% of CSF pixel, 78% of GM, and 83.6% of WM which indicates the network is trained and ready to perform segmentation in other testing image.

3.3. Dice and Jaccard Similarity Index. To assess the performance of the method, the Dice similarity index, the Jaccard coefficient, and the mean squared error (MSE) of each tested image were calculated with reference to the ground truth image available in the same database. For comparison, each image was converted into a label image as that of ground truth. From the experiment, it can be clearly seen that results of high visual quality were obtained, with almost 80% Dice similarity index in each test image.

The Dice similarity coefficient of two sets x and y is defined as

$$\text{Dice}(x, y) = 2 \times |\text{intersection}(x, y)| / (|x| + |y|), \quad (1)$$

where $|x|$ represents the cardinality of the set x and $|y|$ represents the cardinality of the set y .

Similarly, the Jaccard similarity coefficient is defined as

$$\text{Jaccard}(x, y) = |\text{intersection}(x, y)| / |\text{union}(x, y)|. \quad (2)$$

MSE is defined as

$$\text{MSE} = \frac{1}{M \times N} \sum_{i=1}^M \sum_{j=1}^N (I - I'), \quad (3)$$

where I and I' stand for the pixel intensity value for the ground truth reference image of size $M \times N$ and the simulated image pixel value of the same image size, that is, $M \times N$, respectively. Both the Dice similarity index and the Jaccard similarity index are important parameters for determining how closely the images I and I' are related, and IoU is used for determining how closely they are spatially matched, with no wrong mapping. Similarly, MSE was calculated to authenticate the similarity index and the resemblance of the simulated result I' to the ground truth I with minimum loss of information.

4. Results and Discussion

The results obtained appear satisfactory and visually distinguishable. Figures 5(a)–5(g) show the results of the

experiment. The first column (a) contains the original MRI images obtained from the OASIS database, which are cross-sectional T1 images; the next column (b) contains the ground truth or the segmented image of respective images in column (a). The third column (c) shows the main results, which are segmented using the proposed method, that is, segmentation based on pixel label. The remaining three columns (d), (e), and (f) show the extracted binary image as a classification result of (c). The segmented image (c) is represented by gray level intensity in (g), which is compared with the ground truth to evaluate the Dice similarity coefficient (DSC) of each class, namely, WM, GM, and CSF. Table 3 presents the performance parameter for each image presented in row (a) of the original image in Figure 5.

4.1. Comparison with Other State-of-the-Art Methods. The computed mean DSC was approximately 80% (highest 84% and lowest 71%) among 6 test images. To compare this result, similar previous approaches for brain image segmentation are presented in Table 4. Zhang et al. [12] used a patchwise CNN for private data of 10 healthy infants, and Nie et al. [13] used semantic approach for the same type of data. Our approach was superior to those by de Brebisson et al. [14] and Moeskops et al. [15] in terms of DSC, but the dataset used here is OASIS mid cross-sectional T1 MRI 2D images instead of MICCAI 2012 Atlas.

5. Conclusions

In conclusion, we have successfully applied deep learning technique for image segmentation with convincing results. Specifically, we are able to segment closely related brain MRI images on pixel-label basis using encoder-decoder network of SegNet layer, which is generally used in semantic segmentation of outdoor scene. This suggests us that, with certain modification and simplified architecture, deep neural network can be effective in medical MRI image segmentation as like natural outdoor images.

Data Availability

The Open Access Series of Imaging Studies (OASIS) data were acquired through grants: P50 AG05681, P01 AG03991, R01 AG021910, P20 MH071616, and U24 RR021382. The MRI analyze format file (.img, and .hdr) data used in preparation of this article are publicly available on the OASIS database (<http://www.oasis-brains.org/#data>).

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This research was supported by the Brain Research Program through the National Research Foundation of Korea funded by the Ministry of Science, ICT & Future Planning

(NRF-2014M3C7A1046050). And, this work was supported by the National Research Foundation of Korea Grant funded by the Korean Government (NRF-2017R1A2B4006533).

References

- [1] M. Tabaton, P. Odetti, S. Cammarata et al., "Artificial neural networks identify the predictive values of risk factors on the conversion of amnesic mild cognitive impairment," *Journal of Alzheimer's Disease*, vol. 19, no. 3, pp. 1035–1040, 2010.
- [2] D. Jha, J. I. Kim, and G. R. Kwon, "Diagnosis of Alzheimer's disease using dual-tree complex wavelet transform, PCA, and feed-forward neural network," *Journal of Healthcare Engineering*, vol. 2017, Article ID 9060124, 13 pages, 2017.
- [3] I. Despotović, B. Goossens, and W. Philips, "MRI segmentation of the human brain: challenges, methods, and applications," *Computational and Mathematical Methods in Medicine*, vol. 2015, Article ID 450341, 23 pages, 2015.
- [4] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3431–3440, China, 2015.
- [5] 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, 2017.
- [6] C. Liang-Chieh, G. Papandreou, I. Kokkinos, K. Murphy, and A. Yuille, "Semantic image Segmentation with deep convolutional nets and fully connected CRFs," in *Proceedings of International Conference on Learning Representations*, San Diego, CA, USA, May 2015.
- [7] H. Noh, S. Hong, and B. Han, "Learning Deconvolution network for semantic segmentation," in *Proceedings of the 2015 IEEE International Conference on Computer Vision*, pp. 1520–1528, 2015.
- [8] C. Szegedy, W. Liu, Y. Jia et al., "Going deeper with convolutions," in *The Conference on Computer Vision and Pattern Recognition*, pp. 1–9, 2015.
- [9] K. Simonyan and A. Zisserman, *Very Deep Convolutional Networks for Large-Scale Image Recognition*, arXiv preprint arXiv:1409.1556, 2014.
- [10] R. Socher, C. C. Lin, C. Manning, and A. Y. Ng, "Parsing natural scenes and natural language with recursive neural networks," in *Proceedings of International Conference on Machine Learning*, pp. 129–136, Lille, France, July 2011.
- [11] S. Hong, H. Noh, and B. Han, "Decoupled deep neural network for semi supervised semantic segmentation," in *Proceedings of Neural Information Processing Systems*, pp. 1495–1503, 2015.
- [12] <http://www.oasis-brains.org/app/template/Index.vm>.
- [13] W. Zhang, R. L. H. Deng, L. Wang et al., "Deep convolutional neural networks for multimodality isointense infant brain image segmentation," *NeuroImage*, vol. 108, pp. 214–224, 2015.
- [14] D. Nie, L. Wang, Y. Gao, and D. Sken, "Fully convolutional networks for multi-modality isointense infant brain image segmentation," in *Proceedings of IEEE 13th International Symposium on Biomedical Imaging (ISBI)*, pp. 1342–1345, Prague, Czech Republic, 2016.

- [15] A. de Brebisson and M. Giovanni, "Deep neural networks for anatomical brain segmentation," in *The Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Dublin, Ireland, 2015.
- [16] P. Moeskops, M. A. Viergever, A. M. Mendrik, L. S. de Vries, M. J. N. L. Benders, and I. Išgum, "Automatic segmentation of MR brain images with a convolutional neural network," *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1252–1261, 2016.

Research Article

Rank-Two NMF Clustering for Glioblastoma Characterization

Aymen Bougacha ,¹ Ines Njeh,¹ Jihene Boughariou,¹ Omar Kammoun,²
Kheireddine Ben Mahfoudh,² Mariem Dammak,³ Chokri Mhiri,³ and Ahmed Ben Hamida¹

¹ATMS-ENIS, Advanced Technologies for Medicine and Signals, Department of Electrical and Computer Engineering, National Engineers School, Sfax University, Sfax, Tunisia

²Department of Radiology, University Hospital Habib Bourguiba, Sfax, Tunisia

³Department of Neurology, University Hospital Habib Bourguiba, Sfax, Tunisia

Correspondence should be addressed to Aymen Bougacha; aymen.bougacha@gmail.com

Received 29 June 2018; Accepted 26 September 2018; Published 23 October 2018

Academic Editor: Vincenzo Conti

Copyright © 2018 Aymen Bougacha et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This study investigates a novel classification method for 3D multimodal MRI glioblastomas tumor characterization. We formulate our segmentation problem as a linear mixture model (LMM). Thus, we provide a nonnegative matrix M from every MRI slice in every segmentation process' step. This matrix will be used as an input for the first segmentation process to extract the edema region from T2 and FLAIR modalities. After that, in the rest of segmentation processes, we extract the edema region from T1c modality, generate the matrix M , and segment the necrosis, the enhanced tumor, and the nonenhanced tumor regions. In the segmentation process, we apply a rank-two NMF clustering. We have executed our tumor characterization method on BraTS 2015 challenge dataset. Quantitative and qualitative evaluations over the publicly training and testing dataset from the MICCAI 2015 multimodal brain segmentation challenge (BraTS 2015) attested that the proposed algorithm could yield a competitive performance for brain glioblastomas characterization (necrosis, tumor core, and edema) among several competing methods.

1. Introduction

Brain tumor represents 85% to 90% of all primary central nervous system tumors. It is one of the main sources for the increase in death rate among children and adults in the world. Bauer et al. [1] noted that glioma could be considered as the largest common brain tumor with the maximum death rate. According to its severity, such brain tumor could be classified as low-grade glioblastomas (LGG) and high-grade glioblastomas (HGG). The low-grade tumors keep developing for many years and could be designed as slow invaders of brain safety tissue. On the other hand, the high-grade tumors known as glioblastomas multiform (GBM) are incurable with an average life of one year after its revelation. Such invasive tumors are very heterogeneous due to their morphological, cytological, and molecular variability. It might have a variety of shapes, might be of any size, and might appear at any location and in different image intensities.

In behalf on their frequency and severity, glioblastomas continue to be the major therapeutic issue for neurosurgeons, neuro-oncologists, and radiation therapists.

The magnetic resonance imaging (MRI) could be considered as one of the main noninvasive modalities used to explore glioblastomas brain tumor for diagnosis, evaluation as well as for inspection of the addressed treatment effect. Such procedure offers the generation of different sequences by modifying the excitation and the repetition times during the acquisition of the image. Each sequence provides relevant structural information. The main four standard MRI modalities are the T1-weighted MRI (T1), the T2-weighted MRI (T2), the T1-weighted MRI with gadolinium contrast enhancement (T1-Gd), and the fluid-attenuated inversion recovery (FLAIR). Conventionally, T1 images are used to differentiate healthy tissue, while T2 images provide a light signal on the image which helps to delineate the region of the edema. In T1-Gd images, the hyperintense given by the accumulated contrast agent (gadolinium ions) in the active

cellular region of the tumor tissue allows us to facilitate the observation of the tumor boundary. The necrotic cells are observed by a hypointense part of the tumor core, as they do not interact with the contrast agent, which makes them easily distinguishable from the active cell region. In FLAIR images, the suppression of the signal of molecule water provides a good observation of edema region from cerebrospinal fluid (CSF).

Glioblastomas segmentation is a challenging task that could be considered as an essential preprocessing task in brain tumors diagnoses. Manual segmentation is a tedious and a time-consuming process for radiologists.

Consequently, automatic segmentation algorithms would be recommended in order to obtain accurate and reliable brain tumor delimitation but it remains a persistent challenge due to the structural complexity of glioblastomas tumors. Furthermore, such tumors present essentially four different zones: edema, which represents an excess accumulation of fluid in the intracellular or extracellular spaces of the brain, nonenhancing solid core, necrotic/cystic core, and enhancing core.

Several research studies have been investigated to segment different tumor zones in multiple MRI modalities (T1, T1-Gd, T2, and Flair) [1–3]. Dupont et al. [4] present, in their review, four main classes in order to segment glioblastomas tumor: region-based approach, edge-based approach, and classification-based algorithms approach.

In the region-based approach, we intend to implement segmentation by merging neighbourhood pixels that have similar characteristics. A region-based method presented by Franz et al. [5] is used to differentiate the enhanced tumor portion, the necrotic zone, and the edema zone. Only two modalities (T1-Gd and Flair) have been used as an input for this algorithm, and only the image intensity has been employed as a feature in order to delimit different tumor region's zone. As a consequence, coherent intensity pixels have been grouped into three classes: tumor enhancement zone, necrosis zone, and edema zone. Sachdeva et al. [6] introduced an edge-based method based on image texture's intensity and a specific active contour to achieve semi-automatic segmentation. The authors used multimodal MRI (T1-weighted, T2-weighted, and T1-Gd MRI) to test their algorithm. Essidike et al. [7] proposed a two-step brain tumor segmentation. For the first step, a numerical simulation of the optical correlation has been used to detect brain tumor, and an active contour model is used to detect region for the next step.

Healthy tissues extraction can help to provide GBM structure segmentation, and atlas-based approaches have been used in this way. Prastawa et al. [8] introduced an automatic brain tumor segmentation with edema's detection. This algorithm used only T2 MRI. Pixels classification of cerebrospinal fluid (CSF), white matter (WM), and gray matter (GM) was performed from atlas template. The unclassified pixels have been labelled as tumor or edema.

Classification approaches are widely used in image segmentation. These methods consist in clustering pixels depending on different features used as an input vector (intensity, texture, neighbours, and spatial distribution in

the image) of a clustering algorithm. There are two groups of classification approaches: supervised approaches or unsupervised approaches. Wu et al. [9] applied a supervised method. A multimodal MRI is segmented into superpixels to minimize the sampling problem. Then, features were extracted from the superpixels using multilevel Gabor wavelet filters. These features are used to power the support vector machine (SVM) model. The theory of conditional random fields has been applied to segment the tumor based on the output of the SVM models. Finally, the marking noise was removed using "structural knowledge." This system was applied with 20 GBM cases. Recent studies [10–12] have used the deep learning technique for the segmentation of GBM tumors. These methods of segmentation differ according to the training concept. Havaei et al. [10] performed a modified conventional neural network (CNN) and a two-phase training to touch on problems related to the unbalance of GBM labels. Zhao et al. [12] used a three-segmentation model based on fully convolutional neural networks (FCNNs), conditional random fields (CRFs), and recurrent neural networks (RNNs). These models are trained with 2D image patches and slices acquired in axial, coronal, and sagittal views, respectively, and mixed them to segment brain tumors. Hussain et al. [11] implemented a deep conventional neural network (DCNN) where two networks are piled over one another to construct a new linear nexus architecture. The first network holds in parallel placing of layers, whereas in the second network, layers are structured linearly.

Corso et al. [13] proposed a Bayesian model classification. This unsupervised method has used two concepts: class model and graph cuts. The objective was to fuse speed of graph cuts and statistical distribution efficiency of the class model. The proposed method was executed to twenty GBM cases with T1, T1-Gd, T2, and FLAIR previously investigated by experts. Presented as one of the popular unsupervised clustering methods, Cordova et al. [14] developed a fuzzy c-means GBM segmentation using T1-GD images. This method has been tested with thirty seven cases. In [15, 16], authors applied a hierarchical nonnegative matrix factorization (hNMF) on multiparametric MRI to provide tissue characterization. The specification of tissue's patterns was obtained, and the spatial distribution of each tissue type was visualized. Li et al. [17] also applied hNMF to brain MRSI data for GBM tissue's differentiation.

In this work, we propose a novel classification method for 3D multimodal MRI glioblastomas tumor characterization. We formulate our segmentation problem as a linear mixture model (LMM). Thus, we provide a nonnegative matrix M from every MRI slice in every segmentation process' step. This matrix will be used as an input for the first segmentation process to extract edema region from T2 and FLAIR modalities. After that, in the rest of segmentation processes, we extract the edema region from T1c modality, generate the matrix M and segment the necrosis, the enhanced tumor, and the nonenhanced tumor regions as described in the method's flowchart (see Figure 1). In the segmentation process, we apply a rank-two NMF clustering which could be defined as a blind source separation technique [18]. It consists in approximately the factorization of

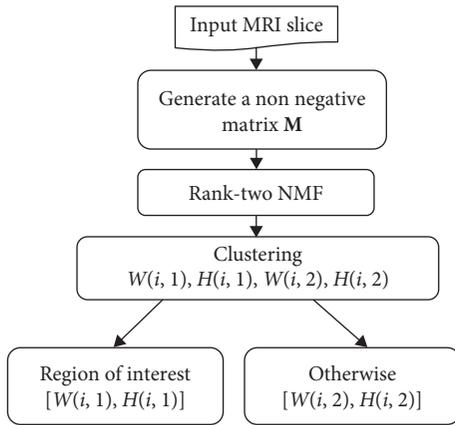


FIGURE 1: Segmentation process.

a matrix M into the product of a source matrix W and an abundance matrix H . This method has been used as a brain tumor segmentation with MRSI (magnetic resonance spectroscopy image) [17] and multiparametric MRI data [15, 16]. The main contribution of this study and the differences between our work and others mentioned previously lies on the application of the GLCM features for nonnegative matrix M and the use of a rank-two NMF instead of the hierarchical NMF. The proposed method does not require a training dataset, as is the case of the many existing methods. Quantitative assessment over the publicly existing training and testing dataset from the MICCAI Multimodal Brain Tumor Segmentation 2015 (BraTS 2015) challenge [19] confirm that the proposed method provides a competitive performance.

The remainder of this paper is arranged as follows: the materials and methods section where we define the Multimodal Brain Tumor Segmentation Benchmark (BraTS 2015) data and illustrate the segmentation methodology. The results and discussion shows the experimental results with a discussion. Finally, the conclusion section illustrates various perspectives of this work.

2. Materials and Methods

The obtained results were based on approved evaluations using the Multimodal Brain Tumor Segmentation Benchmark (BraTS 2015) [19]. In this section, we present in details the used dataset, the evaluation metrics, and the different steps of the proposed methodology: the preprocessing step, the feature extraction, and the rank-two NMF segmentation. The proposed approach could be outlined according to the flowchart (see Figure 2).

2.1. Multimodal Brain Tumor Segmentation Benchmark (BraTS 2015). The Multimodal Brain Tumor Segmentation dataset (BraTS 2015) is in continuation of BraTS 2012, BraTS 2013, and BraTS 2014. It has been organized by B. Menze, M. Reyes, K. Farahani, and J. Kalpathy-Cramer in conjunction with the MICCAI 2015 conference. This available publicly training and testing dataset could be considered as very

useful to compare the existing method and to gauge the current state of the art in brain tumor segmentation. It consists in comparing and evaluating 3D MRI brain tumor regions obtained by segmenting multimodal imaging dataset. Such task could be considered as a challenging task in medical image analysis due to the unpredictable appearance and shape of glioblastomas tumor. The coregistered, the skull-stripped, and the annotated training dataset are available via the Virtual Skeleton Database (VSD) [20].

Training dataset, testing dataset, and the ground truth are stored as signed 16-bit integers, but only positive values are used. Four MRI modalities are proposed for each case: T1 modality, T2 modality, T1c modality, and FLAIR modality. The manual segmentations (ground truth) of the patient images have the following five different labels: (1) for necrosis, (2) for edema, (3) for nonenhancing tumor, (4) for enhancing tumor, and (0) for everything else.

The evaluation is done for 3 different tumor subcompartments:

- (i) *Region 1.* Complete tumor (labels 1 + 2 + 3 + 4 for patient data and labels 1 + 2 for synthetic data)
- (ii) *Region 2.* Tumor core (labels 1 + 3 + 4 for patient data and label 2 for synthetic data)
- (iii) *Region 3.* Enhancing tumor (label 4 for patient data and n.a. for synthetic data)

The total case of training data is 274 patients (220 high-grade tumors and 54 low-grade tumors), while the testing dataset contains 110 subjects with low-grade glioma (LGG) and high-grade glioma (HGG).

2.2. Evaluation Metric. In this study, the dice (DM) [21] and the sensitivity metrics were used to evaluate the quantitative performance and the quality of segmentation. This requires computing the similarities between ground-truth segmentations provided with BraTS 2015 dataset and the obtained results. These metrics take values within the interval $[0..1]$, where 1 indicates a perfect match and 0 a complete mismatch. An automated segmentation should be uploaded directly to the evaluation page to obtain the dice metric score [20].

2.3. Proposed Algorithm. In this work, we present a performed algorithm in order to segment the different tumor regions (necrosis, edema, nonenhancing tumor, enhancing tumor, and everything else). The proposed methodology is tested and validated using BraTS 2015. We present the flowchart that describes the segmentation process (see Figure 1).

We formulate our segmentation problem as a linear mixture model (LMM). As depicted in Figure 1, from an input MRI scan, we generate a nonnegative matrix M , that is, matrix $M \in \mathbb{R}_+^{m \times n}$ with m features and n voxels as follows (see Figure 3): the (i, j) th entry $M(i, j)$ of the matrix M is the i th GLCM feature of the j th voxel. Hence, each column of M is equal to the feature signature of a voxel while each row is a vectorized image at a given feature. The linear mixing

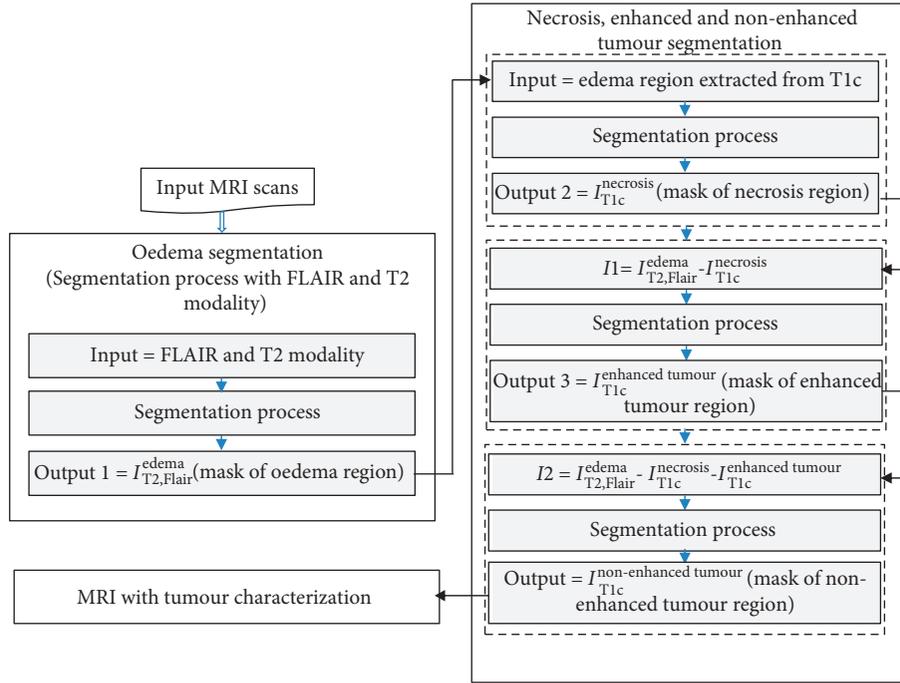
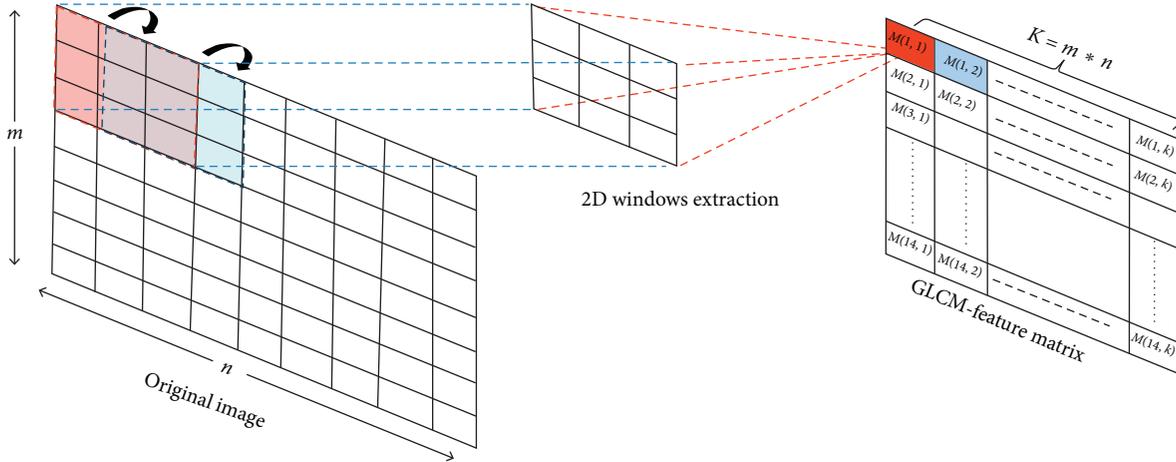


FIGURE 2: Flowchart of the proposed GBM characterization.

FIGURE 3: GLCM-feature matrix m generation.

model (LMM) assumes that the feature signature of each voxel is a linear combination of the feature signatures of the constitutive pattern in image (endmembers), where the weights in the linear combination are the abundances of each endmember in this voxel [22].

Supposing the image encloses r endmembers, and designating $W(:, k) \in \mathbb{R}^m$ ($1 \leq k \leq r$) the feature signatures of the endmembers, we can write the LMM as

$$M(:, j) = \sum_{k=1}^r W(:, k)H(k, j), \quad 1 \leq j \leq n, \quad (1)$$

where $H(k, j)$ is the abundance of the k th endmember in the j th voxel, so $\sum_{k=1}^r H(k, j) = 1$ for all j , which is specified to as

the abundance sum-to-one constraint. As all matrices concerned M , W and H are nonnegative; the LMM is corresponding to nonnegative matrix factorization (NMF). Having a nonnegative matrix $M \in \mathbb{R}_+^{m \times n}$ and a factorization rank r , discover two nonnegative matrices $W \in \mathbb{R}_+^{m \times r}$ and $H \in \mathbb{R}_+^{r \times n}$ such that $M \approx WH$. However, having an MRI slice with r endmembers, it consists in to cluster the pixels into r clusters, and each cluster is equivalent to one endmember. Mathematically, having a matrix $M \in \mathbb{R}_+^{m \times n}$, we aim to find r disjoint clusters $C_k \subset \{1, 2, \dots, n\}$ for $1 \leq k \leq r$ so that $\cup_{k=1,2,\dots,r} C_k = \{1, 2, \dots, n\}$ and so that all pixels in C_k are monopolized by the same endmember.

MRI imaging systems provide images with high resolution and high tissue contrast. These images are defined

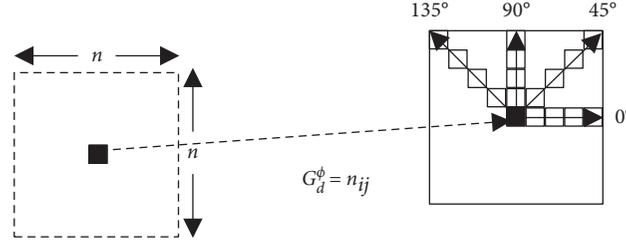


FIGURE 4: 2D GLCM computation for $n * n$ window. Main directions (0° , 45° , 90° , and 135°) and a distance d are used. Mean value is affected to central voxel.

with a depth of up to 16 bits corresponding to 65535 intensity levels. In order to simplify the calculation, all intensity values were rearranged to a gray level values with a maximum of 255, and texture features are computed using the grayscale co-occurrence matrix (GLCM). GLCM has been useful in various image processing fields. It is a squared matrix $G(N,N)$ where N represents the number of gray level existing in the window. This matrix is a structure that represents the co-occurring intensity values at a given offset. This is defined by the fact that the GLCM gives information on how often a gray level arise at different directions and a distance d . Usually, four directions are looked up in the 2D case: $\phi = 0^\circ$, $\phi = 45^\circ$, $\phi = 90^\circ$, and $\phi = 135^\circ$. The structure of the 2D GLCM is shown in Figure 4, where n_{ij} is the number of co-occurrences of gray levels i and j at a specific direction ϕ and a distance d . Haralick et al. [23] defined texture features calculated using the GLCM. Moreover, Haralick recommended utilizing the average value of the features calculated for the four directions to ensure rotation invariance.

The GLCM features (Table 1) used in this study were extracted at a distance $d = 1$ with mean value for four directions.

As illustrated in Figure 2 the MRI modalities are used as follows: from T2 and Flair modalities, we apply the segmentation process in order to obtain the edema mask region $I_{T2,Flair}^{edema}$. Then, we apply the obtained mask on T1c modality and we apply the segmentation process in order to obtain the necrosis region's mask $I_{T1c}^{necrosis}$. We calculate, after that, the intermediary image $I1 = I_{T2,Flair}^{edema} - I_{T1c}^{necrosis}$ which will be used as an input to the segmentation process in order to obtain the enhanced tumor region's mask $I_{T1c}^{enhanced\ tumor}$. We also calculate a second intermediary image $I2 = I_{T2,Flair}^{edema} - I_{T1c}^{necrosis} - I_{T1c}^{enhanced\ tumor}$, and we apply the segmentation process in order to obtain the non-enhanced tumor region's mask $I_{T1c}^{nonenhanced\ tumour}$.

As we can see in Figure 2 that in every segmentation process, we aim to cluster the input MRI slice into two clusters. However, we propose to use a rank-two NMF clustering. Having a nonnegative matrix $M \in \mathbb{R}_+^{m*n}$, rank-two NMF searches two nonnegative matrices $W \in \mathbb{R}_+^{m*2}$ and $H \in \mathbb{R}_+^{2*n}$, where $M \approx WH$. This factorization is a two-dimensional description of the data; more literally, it conceives the columns of M onto a two-dimensional pointed cone developed by the columns of W . Therefore, the approach to segment the MRI slice, in other words to cluster the columns of M , is to selecting the clusters like this: $C_1 = \{i | H(i, 1) \geq H(i, 2)\}$, which represents the region of interest and $C_2 = \{i | H(i, 1) < H(i, 2)\}$ represents the otherwise zone.

3. Results and Discussion

In this section, we report the segmentation result obtained by the proposed method over the publicly training dataset from BraTS 2015. We also present the quantitative evaluation by computing the dice metric and the sensitivity for, respectively, complete tumor, tumor core, and the enhancing tumor. This section is supported by illustration that depict typical example of the obtained results.

Figure 5 depicts the segmentation obtained on two high-grade gliomas from the training dataset. The green zone corresponds to the edema region, the yellow zone represents the enhanced tumor, the red zone is the necrosis, and the blue color represents the nonenhanced tumor. Table 2 attests the performance of the proposed algorithm by a great score for dice and sensitivity metric.

The segmentation methodology proposed in this paper can process an immense diversity of tumors because it does not depend on contrast enhancement. It segments the whole brain, including healthy tissue types, and automatically identifies edema, nonenhanced, enhanced tumor, and necrosis region. Delineating the edema region can be valuable for surgical planning and description of radiation therapy fields, and since the edema region demonstrates the volume over which the tumor applies obvious chemical effects, recognition of areas of interest to various investigators is involved in tumor growth and treatment. Delineating the edema region can also be valuable for surgical planning and radiation therapy. Often, edema regions need to be treated to minimize the risk of recurrence.

We have carried out the proposed method to MR data from patients with glioblastoma tumors. These images include tumors with different intensities, sizes, locations, and shapes. This authorizes us to demonstrate the large field of application of our algorithm.

We have executed our tumor characterization method on BRATS 2015 challenge dataset. Two cases have been selected randomly in this experiment. Definitely, there are four label types in this dataset, including necrosis, edema enhanced, and nonenhanced tumor. As pointed out in Table 2, the dice ratio is superior to 0.85, illustrating good overlap with ground truth. Moreover, the sensitivity is superior to 0.8 which means that the segmentation results are reliable enough.

The results of this table also illustrate that the quality of the segmentation for whole tumor is better than for core tumors because of their well-defined boundaries. Enhancement of the approach for segmenting core tumors could still be valuable.

TABLE 1: GLCM features.

Feature	Formula	Feature	Formula
Contrast	$\sum_{p=1}^N (p^2 \sum_{i=1}^N \sum_{j=1}^N G_d^\phi(i, j)), \quad p = i - j $	Sum average	$\sum_{i=1}^{2N-1} (i p_{x+y}(i))$
Energy	$\sum_{i=1}^N \sum_{j=1}^N G_d^\phi(i, j)^2$	Cluster shade	$\sum_{i=1}^N \sum_{j=1}^N (i + j - \mu_x - \mu_y)^3 G_d^\phi(i, j)$
Dissimilarity	$\sum_{i=1}^N \sum_{j=1}^N (i - j) G_d^\phi(i, j)$	Cluster prominence	$\sum_{i=1}^N \sum_{j=1}^N (i + j - \mu_x - \mu_y)^4 G_d^\phi(i, j)$
Entropy	$-\sum_{i=1}^N \sum_{j=1}^N G_d^\phi(i, j) \log(G_d^\phi(i, j))$	Maximum probability	$\sum_{i=1}^N \sum_{j=1}^N \max_{i,j} \{G_d^\phi(i, j)\}$
Correlation	$\sum_{i=1}^N \sum_{j=1}^N G_d^\phi(i, j) \frac{(i - \mu_x)(i - \mu_y)}{\sigma_x \sigma_y}$ $\mu_x = \sum_{i=1}^N \sum_{j=1}^N j G_d^\phi(i, j), \mu_y = \sum_{i=1}^N \sum_{j=1}^N i G_d^\phi(i, j)$ $\sigma_x = \sum_{i=1}^N \sum_{j=1}^N (j - \mu_x)^2 G_d^\phi(i, j)$ $\sigma_y = \sum_{i=1}^N \sum_{j=1}^N (i - \mu_y)^2 G_d^\phi(i, j)$	Difference variance	$\sum_{i=1}^{2N-1} (i^2 p_{x-y}(i))$
Homogeneity	$\sum_{i=1}^N \sum_{j=1}^N \frac{1}{1 + (i - j)^2} G_d^\phi(i, j)$	Autocorrelation	$\sum_{i=1}^N \sum_{j=1}^N (i * j) G_d^\phi(i, j)$
Variance	$\sum_{i=1}^N \sum_{j=1}^N (1 - \mu)^2 G_d^\phi(i, j)$	Sum entropy	$-\sum_{i=1}^{2N-1} (p_{x+y}(i) \log(p_{x+y}(i)))$
Difference entropy	$-\sum_{i=0}^{2N-1} (p_{x-y}(i) \log(p_{x-y}(i)))$	Sum variance	$-\sum_{i=1}^{2N-1} (1 - \text{SumEnt})^2 p_{x+y}(i)$

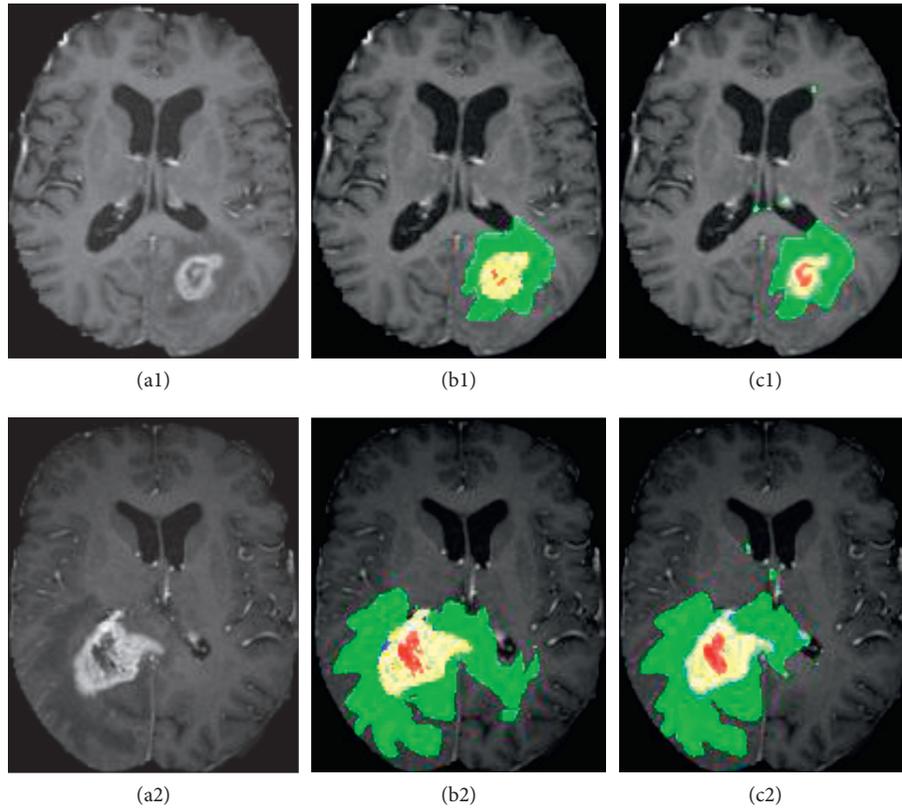


FIGURE 5: Examples of characterization results obtained from rank-two NMF methods on BRATS 2015 data. T1c images with high-grade tumor case HG-02 (a1), HG-03 (a2); b1-b2 ground truth; c1-c2 results using rank-two NMF; edema (green), necrosis (red), enhanced tumor (yellow), and nonenhanced tumor (blue).

4. Conclusion

In this paper, we define a novel methodology for 3D multimodal MRI GBM tumor characterization. Unlike from the classic tumor segmentation methods, in the proposed

method, we observe the brain tumor segmentation task as a four-class (tumor (including necrosis, enhanced, and nonenhanced tumor), edema, and normal tissue) classification problem regarding three modalities T2, FLAIR, and T1c. We formulate our segmentation problem as a linear

TABLE 2: Quantitative results for the BRATS 2015 MRI images.

Class	Dice	Sensitivity
Complete tumor	0.87	0.84
Tumor core	0.77	0.64
Enhancing tumor	0.74	0.61

mixture model (LMM). Thus, we provide a nonnegative matrix M from every MRI slice in every step of the segmentation process. This matrix will be used as an input for the first segmentation process to extract edema from T2 and FLAIR modality. After that, in the rest of segmentation processes, we extract the edema region from T1c modality, generate the matrix M from this modality, and segment necrosis, enhanced tumor, and nonenhanced tumor regions. In the segmentation process, we apply a rank-two NMF clustering. Compared to the traditional tumor segmentation methodologies, the proposed method is easy to achieve and quite robust to high-intensity inhomogeneity images. Comparison results on BRATS 2015 challenge dataset illustrate the superior achievements of the proposed method.

As a perspective, we will apply the proposed method through all training data and also the proposed testing data in order to attest the performance of the algorithm.

Data Availability

The BRATS 2015 data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] S. Bauer, R. Wiest, L.-P. Nolte, and M. Reyes, "A survey of MRI-based medical image analysis for brain tumor studies," *Physics in Medicine and Biology*, vol. 58, no. 13, pp. R97–R129, 2013.
- [2] E.-S. A. El-Dahshan, H. M. Mohsen, K. Revett, and A.-B. M. Salem, "Computer-aided diagnosis of human brain tumor through MRI: a survey and a new algorithm," *Expert Systems with Applications*, vol. 41, no. 11, pp. 5526–5545, 2014.
- [3] N. Gordillo, E. Montseny, and P. Sobrevilla, "State of the art survey on MRI brain tumor segmentation," *Magnetic Resonance Imaging*, vol. 31, no. 8, pp. 1426–1438, 2013.
- [4] C. Dupont, N. Betrouni, N. Reyns, and M. Vermandel, "On image segmentation methods applied to glioblastoma: state of art and new trends," *IRBM*, vol. 37, no. 3, pp. 131–143, 2016.
- [5] A. Franz, S. Remmele, and J. Keupp, "Brain tumour segmentation and tumour tissue classification based on multiple MR protocols," in *Proceedings of Medical Imaging 2011: Image Processing*, Lake Buena Vista, FL, USA, August 2011.
- [6] J. Sachdeva, V. Kumar, I. Gupta, N. Khandelwal, and C. K. Ahuja, "A novel content-based active contour model for brain tumor segmentation," *Magnetic Resonance Imaging*, vol. 30, no. 5, pp. 694–715, 2012.
- [7] A. Essadike, E. Ouabida, and A. Bouzid, "Brain tumor segmentation with Vander Lugt correlator based active contour," *Computer Methods and Programs in Biomedicine*, vol. 160, pp. 103–117, 2018.
- [8] M. Prastawa, E. Bullitt, S. Ho, and G. Gerig, "A brain tumor segmentation framework based on outlier detection," *Medical Image Analysis*, vol. 8, no. 3, pp. 275–283, 2004.
- [9] W. Wu, A. Chen, L. Zhao, and J. Corso, "Brain tumor detection and segmentation in a CRF (conditional random fields) framework with pixel-pairwise affinity and superpixel-level features," *International Journal of Computer Assisted Radiology and Surgery*, vol. 9, no. 2, pp. 241–253, 2013.
- [10] M. Havaei, A. Davy, D. Warde-Farley et al., "Brain tumor segmentation with deep neural networks," *Medical Image Analysis*, vol. 35, pp. 18–31, 2017.
- [11] S. Hussain, S. M. Anwar, and M. Majid, "Segmentation of glioma tumors in brain using deep convolutional neural network," *Neurocomputing*, vol. 282, pp. 248–261, 2018.
- [12] X. Zhao, Y. Wu, G. Song, Z. Li, Y. Zhang, and Y. Fan, "A deep learning model integrating FCNNs and CRFs for brain tumor segmentation," *Medical Image Analysis*, vol. 43, pp. 98–111, 2018.
- [13] J. J. Corso, E. Sharon, S. Dube, S. El-Saden, U. Sinha, and A. Yuille, "Efficient multilevel brain tumor segmentation with integrated bayesian model classification," *IEEE Transactions on Medical Imaging*, vol. 27, no. 5, pp. 629–640, 2008.
- [14] J. S. Cordova, E. Schreiber, C. G. Hadjipanayis et al., "Quantitative tumor segmentation for evaluation of extent of glioblastoma resection to facilitate multisite clinical trials," *Translational Oncology*, vol. 7, no. 1, pp. 40–W5, 2014.
- [15] N. Sauwen, D. M. Sima, S. Van Cauter et al., "Hierarchical non-negative matrix factorization to characterize brain tumor heterogeneity using multi-parametric MRI," *NMR in Biomedicine*, vol. 28, no. 12, pp. 1599–1624, 2015.
- [16] N. Sauwen, M. Acou, D. M. Sima et al., "Semi-automated brain tumor segmentation on multi-parametric MRI using regularized non-negative matrix factorization," *BMC Medical Imaging*, vol. 17, no. 1, p. 29, 2017.
- [17] Y. Li, D. M. Sima, S. V. Cauter, A. R. Croitor Sava et al., "Hierarchical non-negative matrix factorization (hNMF): a tissue pattern differentiation method for glioblastoma multiforme diagnosis using MRSI," *NMR in Biomedicine*, vol. 26, no. 3, pp. 307–319, 2013.
- [18] D. D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol. 401, no. 6755, pp. 788–791, 1999.
- [19] B. H. Menze, A. Jakab, S. Bauer et al., "The multimodal brain tumor image segmentation Benchmark (BRATS)," *IEEE Transactions on Medical Imaging*, vol. 34, no. 10, pp. 1993–2024, 2015.
- [20] M. Kistler, S. Bonaretti, M. Pfahrer, R. Niklaus, and P. Büchler, "The virtual Skeleton Database: an open access repository for biomedical research and collaboration," *Journal of Medical Internet Research*, vol. 15, no. 11, pp. e245–12, 2013.
- [21] L. R. Dice, "Measures of the amount of ecologic association between species," *Ecology*, vol. 26, no. 3, pp. 297–302, 1945.
- [22] N. Gillis, D. Kuang, and H. Park, "Hierarchical clustering of hyperspectral images using rank-two nonnegative matrix factorization," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 4, pp. 2066–2078, 2015.
- [23] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-3, no. 6, pp. 610–621, 1973.

Research Article

A Novel Method to Quantify Longitudinal Orthodontic Bone Changes with In Vivo Micro-CT Data

Chao Wang ^{1,2}, Li Cao,³ Chongshi Yang ^{4,5} and Yubo Fan ⁶

¹Assistant Professor, College of Stomatology, Chongqing Medical University, Chongqing, China

²Chongqing Municipal Key Laboratory of Oral Biomedical Engineering of Higher Education, Chongqing, China

³Attending Doctor, Department of Orthodontics, Stomatological Hospital of Chongqing Medical University, Chongqing, China

⁴Attending Doctor & Lecturer, Department of Orthodontics, Stomatological Hospital of Chongqing Medical University, Chongqing, China

⁵Chongqing Key Laboratory for Oral Diseases and Biomedical Sciences, Chongqing, China

⁶Professor, Laboratory for Biomechanics and Mechanobiology of Ministry of Education, School of Biological Science and Medical Engineering, Beihang University, Beijing, China

Correspondence should be addressed to Chongshi Yang; cqmudental@gmail.com

Received 26 May 2018; Accepted 29 July 2018; Published 1 October 2018

Academic Editor: Vincenzo Conti

Copyright © 2018 Chao Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Orthodontic tooth movement (OTM) is the result of region-specific bone modeling under a load. Quantification of this change in the alveolar bone around a tooth is a basic requirement to understand the mechanism of orthodontics. The purpose of this study was to quantify subregional alveolar bone changes during orthodontic tooth movement with a novel method. In this study, 12 Sprague-Dawley (SD) rats were used as an orthodontic model, and one side of the first upper molar was used to simulate OTM. The alveolar bone around the mesial root was reconstructed from in vivo micro-CT images and separated from other parts of the alveolar bone with two semicylinder filters. The amount and rate of OTM, bone mineral density (BMD), and bone volume (BV) around the root were calculated and compared at 5 time points. The results showed that the amount of tooth movement, BMD, and BV can be evaluated dynamically with this method. The molar moved fastest during the first 3 days, and the rate decreased after day 14. BMD decreased from day 0 to day 14 and returned from day 14 to day 28. BV decreased from day 0 to day 7 and from day 14 to day 28. The method created in this study can be used to accurately quantify dynamic alveolar bone changes during OTM.

1. Introduction

Orthodontic tooth movement (OTM) results from the modeling and remodeling of the alveolar bone under a prolonged and moderate load. With selective bone removal and apposition, the tooth moves through the alveolar bone, carrying its attachment apparatus with it. Then, the newly generated bone is remodeled to increase its bone density and mechanical strength to perform function in the new position [1].

Alveolar bone modeling has region-specific characteristics. Different processes of bone modeling can be observed at different sides under the same orthodontic load. The pressure side and tension side show bone resorption mediated by osteoclasts and bone apposition mediated by

osteoblasts, respectively [1, 2]. Therefore, partitioning the alveolar bone around the root is necessary when evaluating the amount and characteristics of bone changes in orthodontics.

Bone morphologic measures, such as bone mineral density (BMD), bone volume (BV), and other microstructure parameters, were calculated by definition of regions of interest (ROIs) in previous studies [3–5]. In these studies, a cube of the alveolar bone near the root was extracted as a ROI to evaluate bone resorption and apposition. Although this method is straightforward, the calculation is affected by the relatively arbitrary nature of ROI selection, especially in in vivo studies. Therefore, determining the same ROI at different time points is important to rule out this influence.

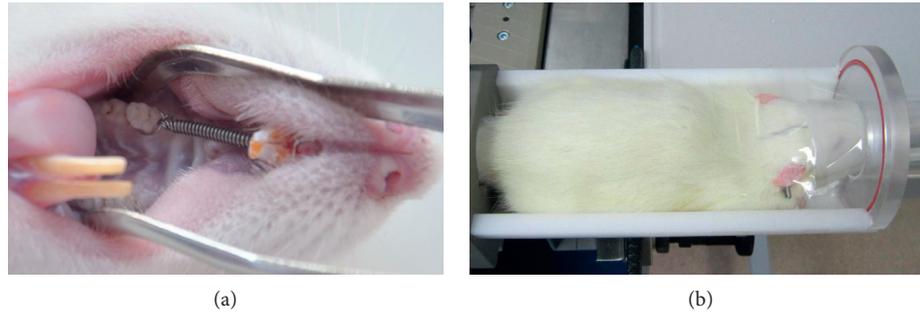


FIGURE 1: Animal study: (a) OTM of the SD rat; (b) fix and inhalation anesthesia.

In this study, sequential images of the Sprague-Dawley (SD) rats were acquired from in vivo micro-CT. The alveolar bone around the mesial root was reconstructed and separated from other parts of the alveolar bone with two semicylinder filters, which were defined as the ROI. BMD and BV were calculated and compared at different time points. The objective of this study was to accurately quantify dynamic bone resorption and apposition with this novel method and to evaluate the modeling of the alveolar bone on tooth movement.

2. Materials and Methods

2.1. Animal Study and Micro-CT Scan. This project was conducted with approval from the Ethics Committee of Chongqing Medical University. Twelve male SD rats (aged 6–8 weeks and weighing 180–220 g) were used as the experimental animals and were fed a standardized laboratory rat diet under conventional conditions ($25 \pm 2^\circ\text{C}$ and a 12-hour light/dark cycle). The right maxillary first molar of each rat served as the orthodontic side, and the left first molar served as a control. On the orthodontic side, the first molar and incisor were connected with a nickel-titanium coil spring (wire diameter: 0.008 mm; Protect, Zhejiang, China). The coil spring was activated to generate a about 25 g continuous force to move the first molar forward (Figure 1(a)). The maxillary left first molar served as the control without any orthodontic load. The animal study protocol has been previously described [6].

Each animal was scanned with an in vivo micro-CT (viva CT40, SCANCO Medical, Brüttisellen, Switzerland) under isoflurane anesthesia (2.3–2.5 v/v %) at day 0 (before the orthodontic load) and days 3, 7, 14, and 28 after orthodontic loads (Figure 1(b)). The scan settings were 70 kV, 114 μA , and 350 ms integration time with an isotropic voxel resolution of 10.5 μm and a calibration with 1200 mg HA/ccm. The scan started 3 mm before the mesial root of the first molar and included the entire first molar and the surrounding alveolar bone. The scan field was determined in 2D X-ray scout view. The coil springs were removed before scanning and reattached after scanning to prevent metal artifacts. In this study, each scan lasted 45 min, and approximately 600 cross-sectional images were generated. The rats were sacrificed by overinhalation of CO_2 after the last scan. These images acquired from scanning were exported in

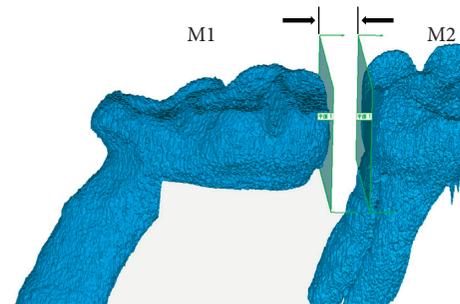


FIGURE 2: Measurement of OTM distance.

the Digital Imaging and Communications in Medicine (DICOM) format and processed by a graphic workstation.

2.2. Image Processing: Measurement of OTM Distance. The measurement of tooth movement was calculated as the distance between the proximal surfaces of the first and second molar. The images generated from micro-CT were imported into Mimics software (version 10.0, Materialise, Leuven, Belgium) to segment the crowns of the first and second molars in 2D sections by determining the Hounsfield value and manual mask segmentation, such as region growth, morphology, and edit operations. Three-dimensional objects of the crown were calculated from 2D image sequences and exported in the stereolithography (STL) format into Geomagic Studio (2012, Raindrop Inc., Rock Hill, South Carolina, USA). A plane was fitted according to the shape of the distal surface of the crown of the first molar and transferred to the second molar until tangent to the mesial surface of the second molar. The distance of the plane between the two positions can be determined as the tooth movement distance (Figure 2). The amount of tooth movement at 5 time points can be calculated with this method.

2.3. Image Processing: the Calculation of BV and BMD. The first molar of the SD rats has five roots; the mesial root is the largest and is far from the other four roots, thus enabling clear evaluation of the modeling of the alveolar process around the mesial root.

In this study, a novel method was created to quantify alveolar bone changes under an orthodontic load through the following steps: 3D reconstruction of the tooth,

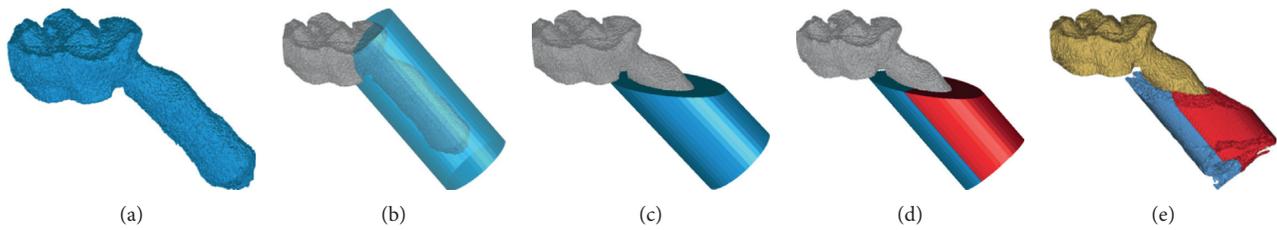


FIGURE 3: The reconstruction, partition, and extraction of the alveolar bone around the mesial root: (a) the reconstruction of the crown and mesial root of the upper first molar; (b) generation of the cylinder around the root; (c) partition of the cylinder with the alveolar crest plane; (d) division of the cylinder into the mesial and distal sides according to the direction of the orthodontic load; (e) reconstruction of the alveolar bone within semicylinder filters.

superimposition of 3D images at different time points, extraction of the alveolar bone around the root, and calculation of parameters in SCANCO.

2.3.1. Reconstruction of the Alveolar Bone, Crown, and Mesial Root of the Upper First Molar. The crown and mesial root of the first molar around the alveolar bone were reconstructed and exported in the STL format with the abovementioned method (Figure 3(a)).

2.3.2. Formation of Cylinder Filters at Different Time Points. 3D images of the molar and alveolar bone were imported into Geomagic Studio to generate a cylinder that can be used to define the ambit of the quantified bone. The height of the cylinder was determined by the following points: the start point was the intersecting point of the occlusal surface and the long axis of the mesial root; the end point was 0.4 mm past the apical according to the anatomic characteristics at this region. The diameter was set as 1.8 mm, taking into consideration the distance to the other roots (Figure 3(b)). The cylinder was converted into a CAD object and partitioned with the plane of the alveolar bone ridge (Figure 3(c)). Then, the cylinder was divided into the mesial part and distal part according to the direction of the orthodontic load (Figure 3(d)).

There were five image sequences corresponding to 5 scans for each animal. The positions of the upper first molar and surrounding structures varied at different time points as the rat cannot be fixed in the exact same position during scanning. Therefore, the image at one time point needs to be matched to the others through rotating and translating operations. In this study, the initial scan (day 0) was set as the baseline and superimposed onto the latter images (days 3, 7, 14, and 28) with its semicylinders. The buccal and lingual grooves on the crown were chosen as references of superimposition as these characteristics are obvious, easily recognized, and not subjected to abrasion caused by mastication. The cylinder generated at time point 1 was then allocated to the roots at other time points with the same size. The CAD objects of the two semicylinders were generated and saved in the STL format.

2.3.3. Alveolar Bone Extraction around the Mesial Root. The STL objects were projected to 2D sections generated by micro-CT scans by calculating the mask from an object

operation. The alveolar bone within semicylinders was extracted to form a new mask through a Boolean operation and exported in the DICOM format (Figure 3(e)).

2.3.4. Analysis in SCANCO Analysis Module. The DICOM files generated in the previous step were imported into the evaluation module of SCANCO viva CT40 (version 4.1). Bone modeling parameters, such as BMD and BV, were calculated at different time points after determination of the ROI, which included all the images within semicylinders and were rendered with a value of 3–5 HU (the images have reached the threshold in Mimics software and were used to generate dichromatic images: background < 3 HU, alveolar bone > 3 HU) (Figures 4(a)–4(c)).

2.4. Statistics. Statistical analysis was performed using SPSS software (version 13.0, SPSS Inc., Chicago, USA). Repeated measures analysis of variance was used for comparisons of the rate of tooth movement and bone morphology parameters, including BMD and BV, among different time points.

3. Results

The upper first molars of the SD rats on the orthodontic side moved mesial under the load of the Ni-Ti spring (Figures 5(a) and 5(b)). The blank control side showed no tooth movement. The distance at adjacent time points significantly increased ($P < 0.05$). The amount of OTM steadily increased over time within the first 2 weeks and slightly slowed down over the last 2 weeks (Figure 5(a)). The rate of OTM rapidly increased during the first 3 days ($P < 0.05$), slightly increased from day 3 to day 14, and slightly decreased from day 14 to day 28 (Figure 5(b)).

Figure 6 shows the BMD changes in the alveolar bone around the mesial root at different time points. The BMD value significantly changed between adjacent time points ($P < 0.05$); BMD decreased at day 3, reached its lowest point at day 14, and then markedly returned. The mesial and distal parts of the alveolar bone showed the same changes as the entire bone. However, the BMD of the alveolar bone on the mesial side was higher than that of the entire bone, and the BMD of the distal part was lower than that of the entire bone.

The volume of the alveolar bone around the mesial root is the sum of its mesial and distal parts. The three parts

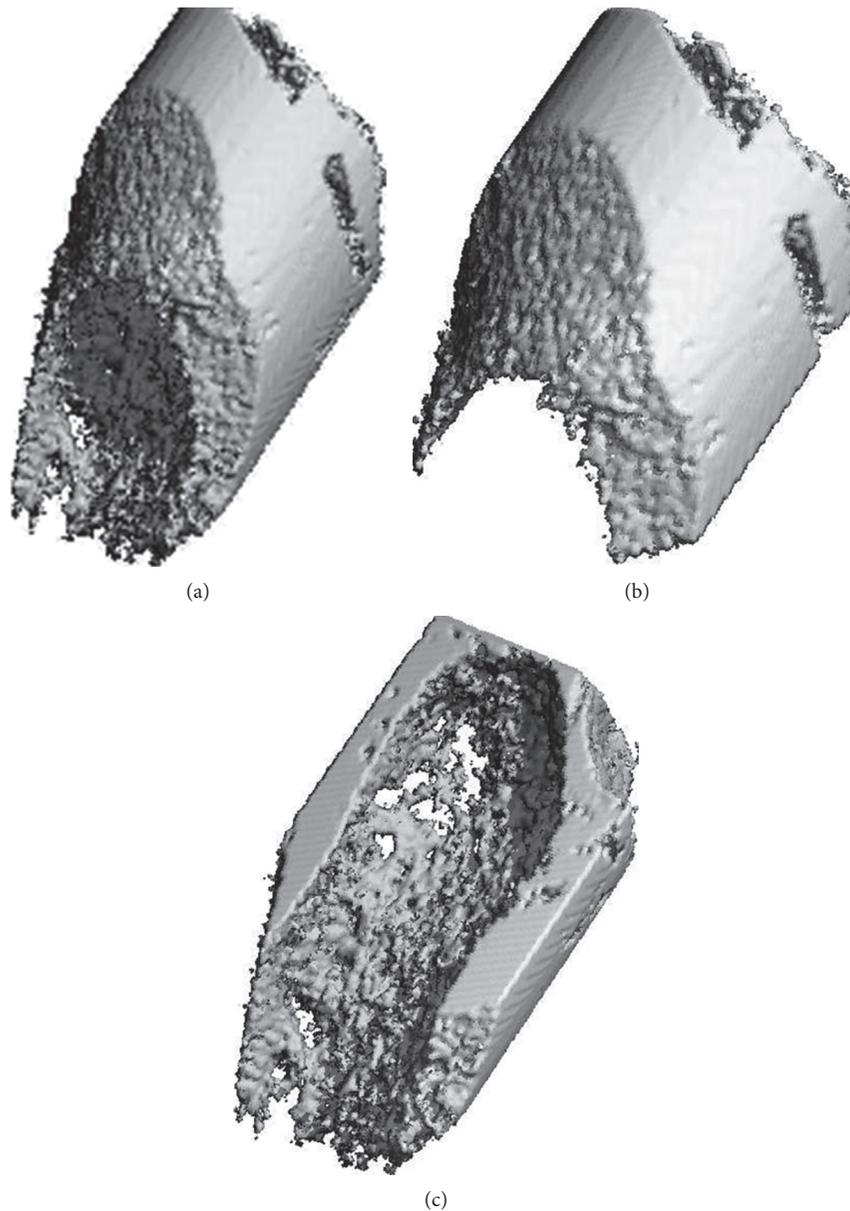


FIGURE 4: Reconstruction of the extracted alveolar bone around the mesial root (a), and mesial part (b) and distal part (c) in SCANCO viva CT40.

showed the same change tendency, decreasing from the initial time point to day 28 (Figure 7). The decrease in BV at the mesial part was more obvious than that at the distal part from day 14 to day 28 ($P < 0.05$).

4. Discussion

Direct measurement of tooth movement distance in the rat specimens is infeasible due to the smaller molar size (the length of the crown is approximately 3 mm), the tiny initial tooth movement (< 0.5 mm), and insufficient exposure of the distal space for the posterior location of the molar. Presently, the main approach is 2D measurement by determining the distance of the landmarks on the molars with X-ray films [7, 8]. However, this method suffers from more errors caused

by inconsistent X-ray magnification, projection angulation, and other factors, such as interference from artifacts caused by other anatomic structures. Currently, micro-CT is used to measure tooth movement distance in rodents [3, 9]. More precise linear and angulation measurements can be acquired in three-dimensional micro-CT films, which display the teeth and surrounding apparatus without the interference of other structures. However, the nature of this approach is a 2D measurement with greater arbitrariness when determining the direction of tooth movement, especially when the direction of tooth movement does not conform to the scan plane or vertical scan plane, which may cause greater errors. In this study, the tooth movement distance was determined by transferring a plane that was fitted according to the shape of the distal plane of the first molar, which was

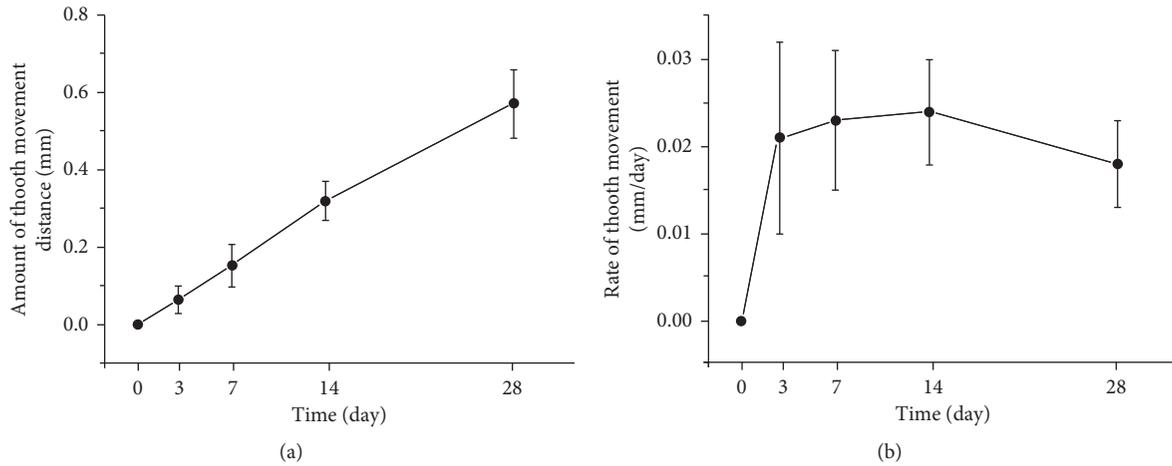


FIGURE 5: Amount and rate of tooth movement on the orthodontic side: (a) amount of orthodontic tooth movement; (b) rate of orthodontic tooth movement.

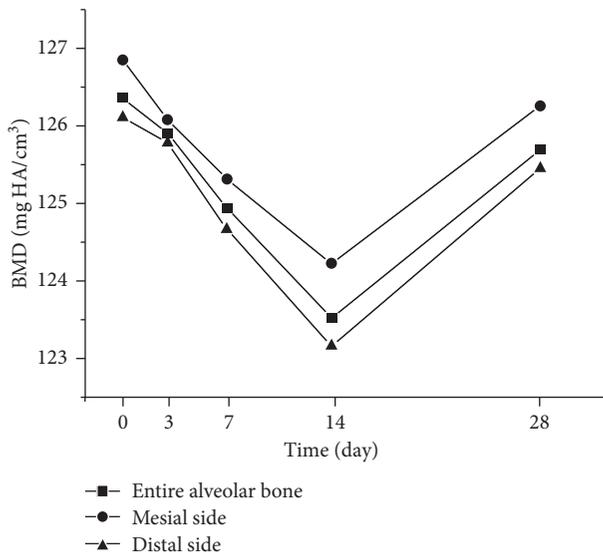


FIGURE 6: BMD changes in the alveolar bone at different time points.

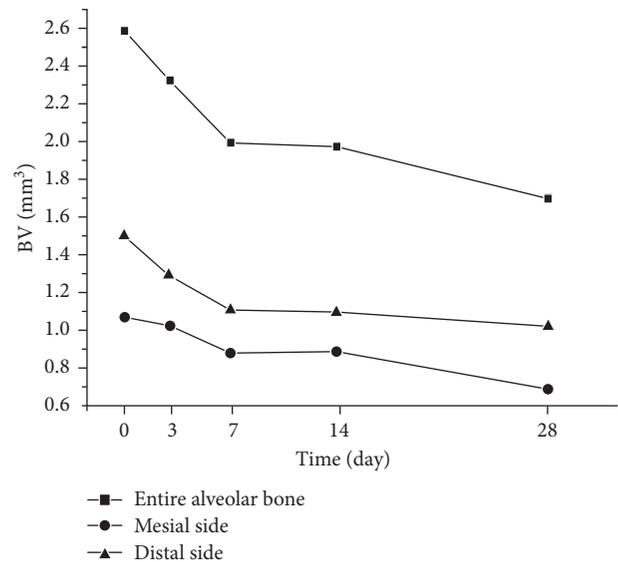


FIGURE 7: BV changes in the alveolar bone at different time points.

reconstructed from micro-CT images, to the distal plane of the second molar. This method can obtain better and more accurate tooth movement results based on micro-CT, excluding the interference of the animal position during scanning and artifacts caused by other structures.

Although OTM is a continuous process, it can be divided into the following stages: the initial stage, lag stage, and linear movement stage [10]. Tooth movement at the initial stage is caused by deformation of the PDL within 1-2 days; the lag stage shows slow tooth movement, which is blocked by hyalinization on the pressure side of the PDL and adjacent alveolar bone; and tooth movement at the linear movement stage is the fastest as the hyalinization is absorbed. The three-stage theory can clarify and explain OTM in the clinic. In this study, we did not observe marked, staged tooth movement. The amount of OTM steadily increased over time within the first 2 weeks and slightly slowed

down in the last 2 weeks. The most rapid tooth movement was observed during the initial 3 days; the same level was approximately maintained after this point, as shown in Figure 5(b). Von Böhl et al. [11, 12] demonstrated that the hyalinization caused by ischemic injury to the PDL will hinder tooth movement. The nontypical tooth movement without remarkable delayed and linear movement stages observed in this study may be caused by the low possibility of hyalinization in the PDL of the rats, which have a wider PDL relative to the size of the tooth compared with that of human patients.

Bone density changes with generation and functional regulation of osteoblasts and osteoclasts. Osteoclast absorbs the surface of the alveolar bone to form an acidic micro-environment, leading to demineralization of the adjacent bone, with degradation of the collagen matrix after secretion of various hydrolases. Osteoblasts mediate bone formation through secretion of collagen and glycoprotein to form the

osteoid and matrix vesicle on the osteoid to form the mineralization center. Although the mechanisms of osteoblasts and osteoclasts can be revealed in cell biology and molecular biology in histological studies, micro-CT scans can quantify their functional activity through analysis of BV and bone density. In this study, the BMD markedly decreased from the initial time point to day 14 and almost returned to the original level in the last 2 weeks (Figure 6). This result demonstrated that the entire alveolar bone modeling was dominated by bone resorption in the early stage and bone apposition in the later stage. As shown in Figure 6, the change tendency of BMD was almost the same on the mesial and distal sides. This result indicated that it cannot be considered simply that only bone resorption exists in the moving side and bone formation in the opposite side. More subregions of the alveolar bone, such as the cervix of the root or apex of the root, are needed to reflect the characteristic change of alveolar bone modeling.

The rate of OTM has a close relationship with bone density. Hashimoto et al. found that the rate of OTM increased remarkably with a decrease in bone mineral content (BMC) after ovariectomy in the Wistar rats [3, 13]. Conversely, OTM decreased with an increase in bone minerals after diphosphonate treatment [14–16]. These findings indicate that the rate of OTM is negatively correlated with bone mineral content. Our study verified this result. The molar showed the largest velocity when the bone density on the pressure side was the lowest and then exhibited slower OTM when bone density increased.

Bone loss caused by orthodontic treatment seems inevitable, especially when tooth movement occurs at the extraction space and lower incisors [17, 18]. The most common method to assess bone loss in orthodontic treatment is distance measurement from 2D section films reconstructed from CBCT [17–19]. Bone resorption can also be detected and evaluated with the BV parameter three-dimensionally. In this study, the entire alveolar bone around the mesial root decreased at the initial stage and final stage, and the decrease on the mesial side was more obvious after 14 days (Figure 7). This decrease is not only the result of alveolar bone resorption caused by OTM but also the result of the reduction in the alveolar ridge height when the tooth moved forward.

Traditionally, bone modeling and remodeling can be evaluated through osteoclast counts, visualization of absorption lacuna, and new bone apposition from histologic films or semiquantification of cytokines, such as PCNA and TGF beta-1, from immunohistochemical staining of sections [20]. Another option is quantification of RNA markers of osteoclastic cells and osteoclastic regulators, such as M-CSF, RANKL, OPG, OPN, BSP, and OCN [7]. These methods are more sensitive in detecting bone changes. However, these methods are invasive and have to be done via the tissue biopsy or autopsy after animals are sacrificed; thus, more experimental animals are needed in a longitudinal study. The micro-CT technique as a noninvasive method is more suitable for observing and quantifying bone changes in the same location over a longer period of time. Another advantage of micro-CT is volume measurement with a wider

scope versus a slice or thin layer with histological methods. Therefore, micro-CT and imaging process were utilized to quantify bone changes of subregional bone segments in this study.

In vivo micro-CT offers the possibility to monitor dynamic changes in the bone microstructure through repeated scanning. However, studies about the effect of repeated radiation from micro-CT on small animals are lacking. Several studies have illustrated that repeated radiation with lower doses of 400~600 mGy is unlikely to have an impact on any changes in bone microstructures [21]. However, more extensive research is still needed to clarify the effect of radiation on bone modeling and remodeling according to different scan protocols. Waarsing et al. [22] first quantified dynamic changes in the microstructure of the rat tibia after ovariectomy based on image registration of in vivo micro-CT films. Schulte et al. demonstrated the sensitivity and repeatability of bone formation and bone resorption characteristics, such as mineral surface (MS), mineral apposition rate (MAR), eroded surface (ES), and mineral resorption rate (MRR), acquired with in vivo micro-CT scans compared with those acquired with fluorescent staining [23]. Nishiyama et al., Liu et al., and Lukas et al. [24–26] demonstrated the broad application prospects in experiments. Based on these studies, we created a method to quantify alveolar bone changes through image processing based on in vivo micro-CT scans.

The main difference between the method in this study and the traditional method is the definition of ROI. In the traditional method, the ROI cannot be accurately located in the same place due to the position change of anatomic structures caused by the inability to maintain the same animal positions during several scans. Another disadvantage is the inflexible definition of the border and shape of the ROI, which should be defined according to research objectives and anatomic characteristics. By contrast, in this study, the ROIs were all the same and related to the root and crown in several scans. Moreover, the partition can be flexibly performed according to research demands, such as the direction of the orthodontic load. However, the determination and registration of ROI may increase the labor intensity. Moreover, high resolution images are needed to construct the anatomic structures used for registration at different time points.

Furthermore, with the strategy and technical route created in this study and the wide use of CBCT, we can obtain a more complicated, accurate, and flexible evaluation of the bone remodeling and modeling effects in the clinic treatments, such as space-closing, expansion, maxillary protraction, functional treatment, and corticotomy.

5. Conclusions

The present study introduced a nondestructive 3D technique used to evaluate longitudinal bone morphology changes that result from an orthodontic load based on micro-CT images. Special regions of the alveolar bone can be extracted through image processing and partitioned according to clinical or research requirements. In addition, rat bone modeling was

evaluated with this technique. The results indicated that the rate of OTM is negatively correlated with BMD. The BV of the alveolar bone decreased during the OTM procedure, and some differences were detected between the mesial and distal sides of the alveolar bone around the mesial root of the upper first molar of the SD rats.

Data Availability

The data used to support the findings of this study are an intact data chain including original micro-CT scan images, the generated CAD files, the STLs, and DICOM files, as well as the final results calculated in in-vivo micro-CT. These files are so large (more than 250 GB) to be accessed in the network storage system. These files are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors would like to acknowledge and thank Dr. Jinlin Song, Stomatological Hospital of Chongqing Medical University for his contribution to this study. This work was supported by the Science and Technology Research Program of Chongqing Municipal Education Commission (Grant No. KJ1702020), the Science and Technology Project of Yubei District, Chongqing (Grant No. 2017 (Agriculture and Society) 44).

References

- [1] W. R. Proffit, "Chapter 8: the biologic basis of orthodontic therapy," in *Contemporary Orthodontics*, W. R. Proffit, H. W. Fields, and D. M. Sarver, Eds., Mosby Elsevier, Maryland Heights, MO, USA, 5th edition, 2012.
- [2] V. Krishnan and Z. Davidovitch, "Cellular, molecular, and tissue-level reactions to orthodontic force," *American Journal of Orthodontics and Dentofacial Orthopedics*, vol. 129, no. 4, pp. 469.e1–469.e32, 2006.
- [3] Y. Xu, T. Zhao, W. Xu, and Y. Ding, "Periodontal microstructure change and tooth movement pattern under different force magnitudes in ovariectomized rats: an in-vivo micro-computed tomography study," *American Journal of Orthodontics and Dentofacial Orthopedics*, vol. 143, no. 6, pp. 828–836, 2013.
- [4] N. Ru, S. S. Liu, L. Zhuang, S. Li, and Y. Bai, "In vivo microcomputed tomography evaluation of rat alveolar bone and root resorption during orthodontic tooth movement," *Angle Orthodontists*, vol. 83, no. 3, pp. 402–409, 2013.
- [5] X. Xu, J. Zhou, F. Yang, S. Wei, and H. Dai, "Using micro-computed tomography to evaluate the dynamics of orthodontically induced root resorption repair in a rat model," *PLoS One*, vol. 11, no. 3, Article ID e0150135, 2016.
- [6] C. Yang, Y. Zhang, Y. Zhang, Y. Fan, and F. Deng, "Three-dimensional quantification of orthodontic root resorption with time-lapsed imaging of micro-computed tomography in a rodent model," *Journal of X-Ray Science and Technology*, vol. 23, no. 5, pp. 617–626, 2015.
- [7] S. S. Baloul, L. C. Gerstenfeld, E. F. Morgan, R. S. Carvalho, T. E. Van Dyke, and A. Kantarci, "Mechanism of action and morphologic changes in the alveolar bone in response to selective alveolar decortication-facilitated tooth movement," *American Journal of Orthodontics and Dentofacial Orthopedics*, vol. 139, no. 4, pp. S83–S101, 2011.
- [8] C. Gonzales, H. Hotokezaka, M. Yoshimatsu, J. H. Yozgatian, M. A. Darendeliler, and N. Yoshida, "Force magnitude and duration effects on amount of tooth movement and root resorption in the rat molar," *Angle Orthodontist*, vol. 78, no. 3, pp. 502–509, 2008.
- [9] C. Gonzales, H. Hotokezaka, Y. Arai et al., "An in vivo 3D micro-CT evaluation of tooth movement after the application of different force magnitudes in rat molar," *Angle Orthodontist*, vol. 79, no. 4, pp. 703–714, 2009.
- [10] C. J. Burstone, "The biomechanics of tooth movement," in *Vistas in Orthodontics*, B. S. Kraus and R. A. Riedel, Eds., Lea & Febiger, Philadelphia, PA, USA, 1962.
- [11] M. Von Böhl, J. C. Maltha, J. W. Von Den Hoff, and A. M. Kuijpers-Jagtman, "Focal hyalinization during experimental tooth movement in beagle dogs," *American Journal of Orthodontics and Dentofacial Orthopedics*, vol. 125, no. 5, pp. 615–623, 2004.
- [12] M. Von Böhl, J. C. Maltha, H. Von den Hoff, and A. M. Kuijpers-Jagtman, "Changes in the periodontal ligament after experimental tooth movement using high and low continuous forces in beagle dogs," *Angle Orthodontist*, vol. 74, pp. 16–25, 2004.
- [13] M. Hashimoto, H. Hotokezaka, I. Sirisoontorn et al., "The effect of bone morphometric changes on orthodontic tooth movement in an osteoporotic animal model," *Angle Orthodontist*, vol. 83, no. 5, pp. 766–773, 2013.
- [14] Y. Fujimura, H. Kitaura, M. Yoshimatsu et al., "Influence of bisphosphonates on orthodontic tooth movement in mice," *European Journal of Orthodontics*, vol. 31, no. 6, pp. 572–577, 2009.
- [15] J. Choi, S. H. Baek, J. I. Lee, and Y. I. Chang, "Effects of clodronate on early alveolar bone remodeling and root resorption related to orthodontic forces: a histomorphometric analysis," *American Journal of Orthodontics and Dentofacial Orthopedics*, vol. 138, no. 5, pp. 548.e1–548.e8, 2010.
- [16] N. R. Kaipatur, Y. Wu, S. Adeeb, T. R. Stevenson, P. W. Major, and M. R. Doschak, "Impact of bisphosphonate drug burden in alveolar bone during orthodontic tooth movement in a rat model: a pilot study," *American Journal of Orthodontics and Dentofacial Orthopedics*, vol. 144, no. 4, pp. 557–567, 2013.
- [17] H. Lund, K. Gröndahl, and H. G. Gröndahl, "Cone beam computed tomography evaluations of marginal alveolar bone before and after orthodontic treatment combined with premolar extractions," *European Journal of Oral Sciences*, vol. 120, no. 3, pp. 201–211, 2012.
- [18] D. T. Garlock, P. H. Buschang, E. A. Araujo, R. G. Behrents, and K. B. Kim, "Evaluation of marginal alveolar bone in the anterior mandible with pretreatment and posttreatment computed tomography in nonextraction patients," *American Journal of Orthodontics and Dentofacial Orthopedics*, vol. 149, no. 2, pp. 192–201, 2016.
- [19] L. O. Castro, I. O. Castro, A. H. De Alencar, J. Valladares-Neto, and C. Estrela, "Cone beam computed tomography evaluation of distance from cemento-enamel junction to alveolar crest before and after nonextraction orthodontic treatment," *Angle Orthodontist*, vol. 86, no. 4, pp. 543–549, 2016.

- [20] L. Wang, W. Lee, D. L. Lei, Y. P. Liu, D. D. Yamashita, and S. L. Yen, "Tissue responses in corticotomy- and osteotomy-assisted tooth movements in rats: histology and immunostaining," *American Journal of Orthodontics and Dentofacial Orthopedics*, vol. 136, no. 6, pp. 770.e1–770.e11, 2009.
- [21] A. Dare, R. Hachisu, A. Yamaguchi, S. Yokose, S. Yoshiki, and T. Okano, "Effects of ionizing radiation on proliferation and differentiation of osteoblast-like cells," *Journal of Dental Research*, vol. 76, no. 2, pp. 658–664, 1997.
- [22] J. H. Waarsing, J. S. Day, J. C. Van Der Linden et al., "Detecting and tracking local changes in the tibiae of individual rats: a novel method to analyse longitudinal in vivo micro-CT data," *Bone*, vol. 34, no. 1, pp. 163–169, 2004.
- [23] F. A. Schulte, F. M. Lambers, G. Kuhn, and R. Müller, "In vivo micro-computed tomography allows direct three-dimensional quantification of both bone formation and bone resorption parameters using time-lapsed imaging," *Bone*, vol. 48, no. 3, pp. 433–442, 2011.
- [24] K. K. Nishiyama, G. M. Campbell, R. J. Clinck, and S. K. Boyd, "Reproducibility of bone micro-architecture measurements in rodents by in vivo micro-computed tomography is maximized with three-dimensional image registration," *Bone*, vol. 46, no. 1, pp. 155–161, 2010.
- [25] X. S. Liu, L. Ardeshirpour, J. N. VanHouten, E. Shane, and J. J. Wysolmerski, "Site-specific changes in bone micro-architecture, mineralization, and stiffness during lactation and after weaning in mice," *Journal of Bone and Mineral Research*, vol. 27, no. 4, pp. 865–875, 2012.
- [26] C. Lukas, D. Ruffoni, F. M. Lambers et al., "Mineralization kinetics in murine trabecular bone quantified by time-lapsed in vivo micro-computed tomography," *Bone*, vol. 56, no. 1, pp. 55–60, 2013.

Research Article

Evaluation of Commonly Used Algorithms for Thyroid Ultrasound Images Segmentation and Improvement Using Machine Learning Approaches

Prabal Poudel ¹, Alfredo Illanes ¹, Debdoot Sheet,² and Michael Friebe¹

¹Department of Medical Engineering, Otto-von-Guericke University Magdeburg, 39106 Magdeburg, Germany

²Department of Mechanical Engineering, Indian Institute of Technology Kharagpur, 822113 Kharagpur, India

Correspondence should be addressed to Prabal Poudel; prabal.poudel@ovgu.de

Received 4 May 2018; Revised 13 August 2018; Accepted 29 August 2018; Published 23 September 2018

Academic Editor: Orazio Gambino

Copyright © 2018 Prabal Poudel et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The thyroid is one of the largest endocrine glands in the human body, which is involved in several body mechanisms like controlling protein synthesis and the body's sensitivity to other hormones and use of energy sources. Hence, it is of prime importance to track the shape and size of thyroid over time in order to evaluate its state. Thyroid segmentation and volume computation are important tools that can be used for thyroid state tracking assessment. Most of the proposed approaches are not automatic and require long time to correctly segment the thyroid. In this work, we compare three different nonautomatic segmentation algorithms (i.e., active contours without edges, graph cut, and pixel-based classifier) in freehand three-dimensional ultrasound imaging in terms of accuracy, robustness, ease of use, level of human interaction required, and computation time. We figured out that these methods lack automation and machine intelligence and are not highly accurate. Hence, we implemented two machine learning approaches (i.e., random forest and convolutional neural network) to improve the accuracy of segmentation as well as provide automation. This comparative study intends to discuss and analyse the advantages and disadvantages of different algorithms. In the last step, the volume of the thyroid is computed using the segmentation results, and the performance analysis of all the algorithms is carried out by comparing the segmentation results with the ground truth.

1. Introduction

The segmentation and volume computation of thyroid are of prime importance when it comes to the diagnosis and treatment of thyroid diseases. Thyroid is a butterfly-shaped gland located below the Adam's apple on the front of the neck. Most of the thyroid diseases like Graves' (excessive production of thyroid hormones), subacute thyroiditis (inflammation of thyroid), thyroid cancer, goitre (thyroid swelling), and thyroid nodule (small abnormal lump growths in thyroid) involve changes in the shape and size of thyroid [1]. Hence, it is essential to compute the volume of thyroid over time to identify whether the thyroid is healthy or not. We use ultrasound (US) imaging for data acquisition instead of other imaging modalities as it is much safer and painless when used on the patients.

Several research works have been proposed on how to segment the thyroid in individual 2D US images. Zhao et al.

[2] proposed several approaches (edge detection, method of threshold value, region splitting and merging, watershed segmentation, active contour, graph theory, US image segmentation based on Ncut, and segmentation based on improved normalized cut) based on 2D segmentation of thyroid in US images. Kaur and Jindal [3] segmented thyroid from 2D US and scintigraphy images using active contour without edges, localized region-based active contour, and distance regularized level set. Augustin et al. [4] tested and segmented thyroid US images using fuzzy c-means algorithm, histogram clustering, QUAD tree, region growing, and random walk [5]. A polynomial Support Vector Machine (SVM) was used [6] to segment the thyroid gland in US images. A local region-based active contour was proposed [7] to segment and compute the area of segmented thyroid in a 2D US image. Another region-based active-contour implementation to segment medical images was

carried out by Mylona et al. [8, 9] where they encoded the local geometry information (i.e., orientation coherence in the edges of the regions to be segmented) to control the evolution of the contour. Similarly, thyroid segmentation in US images using a novel boundary detection method and local binary patterns for texture analysis was proposed by Keramidas et al. [10]. Level-set active-contours models for thyroid segmentation in US images were used in [11, 12]. These methods were mainly based on variable background active contour and joint echogenicity texture. Garg and Jindal later worked on feed-forward neural network [13] to segment thyroid gland from US images. Recently, Narayan et al. [14] made use of the speckle-related pixels and imaging artefacts as source of information to perform multi-organ (thyroid, carotid artery, muscles, and trachea) segmentation in thyroid US images.

Similarly, several research works have been carried out to segment a full 3D thyroid image. Kollorz et al. [15] proposed a semi-automated approach to classify thyroid gland for volumetric quantification using geodesic active contour. Chang et al. [16] proposed radial basis function (RBF) neural network to segment the blocks of thyroid gland. 3D mass spring models for thyroid cartilage segmentation by creating a 3D deformable shape models were proposed by Dornheim et al. [17] but on computed tomography (CT) images. A complete segmentation and analysis of 3D thyroid images was carried out by Osman [18] by thresholding the voxel intensities and then connecting similar voxels to predict the segmenting regions.

The aforementioned approaches have limitations in the sense that they work either on a single 2D image or on a whole 3D image and they do not make use of the spatial relationship between the neighbouring US slices. Hence, we propose three widely used segmentation algorithms which usually work on a 2D image but can be extended to segment a sequence of freehand US images by making use of the spatial relationships between the corresponding image frames. These three approaches are based on active contours without edges (ACWE), graph cut (GC), and pixel-based classifier (PBC). In case of ACWE, the centroid of the segmented image is used as the priori information to find the location of contour initialization in the corresponding slices. GC allows the user to select the foreground and background areas in one image, and these information are transferred to the corresponding slices for further initialization. In PBC, the user clicks inside and outside of the thyroid region to extract the features for thyroid segmentation which are used to train the decision trees and later to classify thyroid and nonthyroid regions in the corresponding images. Our approach of segmenting individual slices and then reconstructing them to a volume possess greater advantages than segmenting directly on the volume itself as segmenting on the 3D image is very complex and is difficult to control. Also, segmentation in 2D allows to analyse the shape of the thyroid in detail as compared to segmenting directly on the thyroid volume.

The main purpose of this work is to compare three nonautomatic segmentation techniques, which are based on ACWE [19], GC [20], and PBC [16] to perform the segmentation in the thyroid images. They are compared based on their accuracy, robustness, ease of use, and computation

time. We also compare the results of these approaches to some of the existing methods [17] that use mass spring models. These algorithms were chosen over others as they can be used not only on one image but also on a sequence of US images in a dataset to produce a 3D segmented thyroid as the information from a segmented image could be transferred to the corresponding image slices to segment them. Additionally, when the segmentation is ongoing in different images in a dataset, the user could directly interact with the segmentation results and correct them if there are any under- or oversegmentation. After segmentation, the segmented images are later used for 3D reconstruction and volume computation using ImFusion [21] and MeVisLab [22]. We figured out that the nonautomatic methods pose several disadvantages and thus implemented two automatic machine learning based methods such as Random Forest Classifier (RFC) and Convolutional Neural Network (CNN) and compared their performance with the nonautomatic methods. We came to the conclusion that the commonly used algorithms could not segment a series of US images highly accurately as compared with these supervised learning techniques.

2. Materials and Methods

In this section, we will explain the three nonautomatic as well as the two automatic methods that are compared in this work to segment the thyroid glands in US images. The automatic methods use 3D thyroid images while the nonautomatic methods use 2D images. We will also present the 3D reconstruction (using segmented results from nonautomatic methods) as well as volume computation technique.

2.1. Active Contour without Edges

2.1.1. Preprocessing. ACWE segmentation was followed by a preprocessing step as the algorithm mainly worked on the gradient information for contour evolution and the preprocessing step improved the gradient visualization. US images mainly contain speckle noise [23] and have low contrast [16]. The speckle noise is produced by the interference of the returning ultrasound waves at the transducer aperture as the ultrasound images are produced when the reflected sound waves from different surfaces inside the body are picked up by the transducer. To enhance the contrast and suppress the speckle noise, a preprocessing step is carried out. Contrast enhancement [24] is used to improve the visibility of the thyroid region. In this work, we make use of Histogram equalization technique which is one of the methods used in contrast enhancement. It helps in recovering the lost contrast in the image by remapping the brightness values such that they are distributed over all the pixels. After histogram equalization, median filter is applied to suppress the speckle noise. It not only reduces speckles but also preserves the gradient/edge information.

2.1.2. Segmentation. After preprocessing, the segmentation process was carried out using the level-set approach developed by Chan and Vese [19]. It is based on the minimization of the Mumford-Shah functional and involves four main steps.

In the first step, the user starts by initializing a rectangle/square mask around the region to be segmented. The initialization of the mask (Figure 1) is a very important step in this algorithm as a wrong initialization can lead to the segmentation of unnecessary segments inside the image. The initial mask separates the image into two regions: foreground (=1) which is inside of the mask and the background (=0) which is outside of the mask.

In the second stage, a Signed Distance Function (SDF), \varnothing , is computed from the initial mask (C) by using the Euclidean distance. SDF is one of the methods of representing the level sets which are used to keep track of evolving curve over time. Our goal is to evolve $\varnothing(x, y)$ when the evolving contour (C) is the zero level set of $\varnothing(x, y, t)$ at each time t .

Thirdly, the forces that control the evolution of the initial contour are computed. These forces are force from the image and force from curvature. Hence, they are calculated as follows:

$$F_{\text{image}} = \int_{\text{inside } C} (I - \mu_{\text{in}})^2 + \int_{\text{outside } C} (I - \mu_{\text{out}})^2, \quad (1)$$

where I is the image, μ_{in} is the average inside the contour, and μ_{out} is the average outside the contour.

$$F_{\text{curvature}} = \frac{\varnothing_x^2 * \varnothing_{yy} + \varnothing_y^2 * \varnothing_{xx} - 2\varnothing_x \varnothing_y \varnothing_{xy}}{(\varnothing_x^2 + \varnothing_y^2)^{3/2}}. \quad (2)$$

All the derivatives are computed using central differentiation method. Using these two forces, the equation of the curve is computed using the Taylor expansion given by the following equation:

$$\varnothing((x, y), t + \Delta t) = \Delta t * \varnothing_t + \varnothing((x, y), t), \quad (3)$$

where

$$\varnothing_t = \alpha * F_{\text{curvature}} + \frac{F_{\text{image}}}{\max|F_{\text{image}}|}, \quad (4)$$

$$\Delta t = \frac{1}{\max(\varnothing_t) + \varepsilon}, \quad (5)$$

where α represents the smoothing term and ε represents the coefficient to satisfy Courant, Friedrichs, and Lewy (CFL) condition [25]. The evolution of the contour stops after the given number of iterations are complete, giving us the segmented thyroid image.

In the last stage, the result of the segmentation on the first image of the dataset is used to segment rest of the images in the dataset. After the segmented thyroid is obtained, its centre of mass is computed. This centre of mass is used to find probable centre of mass of the thyroid in next image slice.

It is computed by making use of the tracking matrices obtained during the data acquisition phase. Each image in the dataset has an associated tracking matrix which gives the transformation from the origin of electromagnetic (EM) tracking system to the centre of the image. Hence, the centre of each image can be computed using the transformation matrix which has the information about the centre of each image in the dataset.

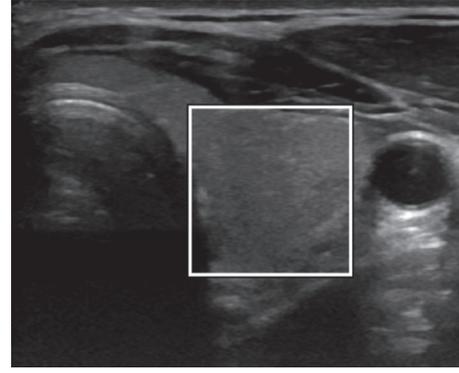


FIGURE 1: ACWE initialization of the mask by the user.

Using this information, the Euclidean distance between the image centres of the current and the next image is computed and the angle between the centres is computed. After computing the distance and angle between the two image centres, a probable centre of mass of the thyroid in the next image is computed by traversing the same distance and angle from the centre of mass of the current segmented thyroid [26]. Centre of mass computed this way will serve to be the centre of rectangle in the next image frame around which the new mask will be initialized automatically. The size of the rectangle will be the same as it was drawn by the user in the first image. In this way, the automatic initialization of segmentation mask is done in the consecutive image frames which will undergo the ACWE algorithm to produce segmented thyroids. The schematic description of the approach is presented in Figure 2.

A fixed number of iterations is set by the user for the contour evolution. By increasing the number of iterations, the computation time will be higher. Hence, a trade-off between the accuracy and computation time has to be maintained while running this algorithm.

2.2. Graph Cut. This approach makes use of the GrabCut algorithm from Rother et al. [20]. It is also a semi-automatic 2D segmentation algorithm just like the ACWE as the user needs to mark the regions as being thyroid and nonthyroid in the initialization phase. It starts with the user creating an initial trimap by marking the thyroid region to be segmented by using yellow scribbles and the surrounding (i.e., nonthyroid) regions by using violet scribbles as seen in Figure 3. The pixels outside of the violet scribble are marked as known background, pixels inside of the violet scribble are marked as unknown, and the yellow scribble areas are marked as definite foreground. The schematic description of the approach can be seen in Figure 4.

Then, an initial image segmentation is computed where all the unknown pixels are placed in the foreground class and all the known background pixels are placed in the background class. These initial foreground and background classes are used to construct foreground and background Gaussian Mixture Models (GMMs) using the Orchard-Bouman clustering algorithm [27]. Each pixel in the foreground class is assigned to the most likely Gaussian component in the foreground GMM, and similarly, each pixel in

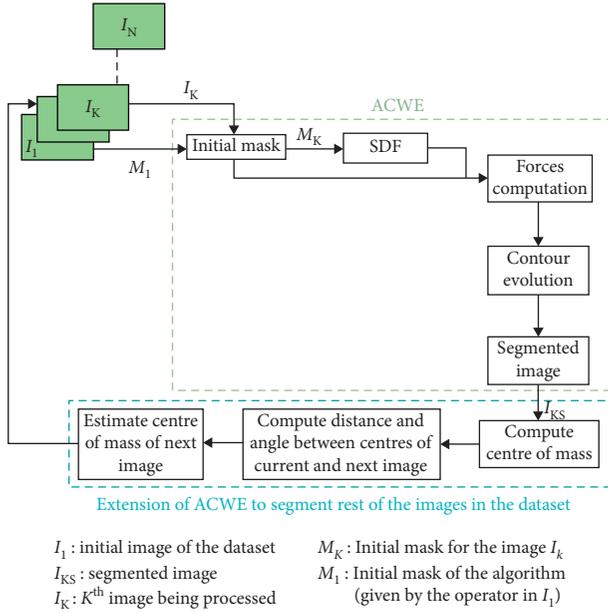


FIGURE 2: Schematic description of ACWE segmentation method.

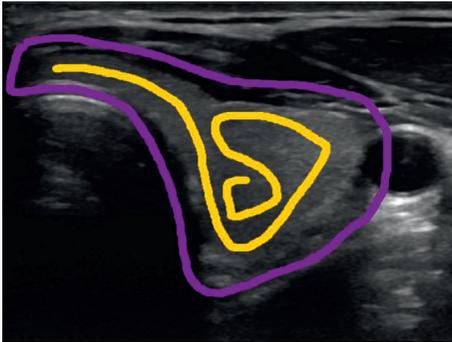


FIGURE 3: Graph cut initialization by the user.

the background class is assigned to the most likely background Gaussian component. With the new distribution of the pixels, the initial GMMs are disregarded and new GMMs are learned from the pixel distributions in each of the two classes.

Finally, a graph is built which consists of each pixel as node and two special nodes (i.e., foreground and background). All of these nodes are connected by two types of edges (also called as links). The first link (i.e., N -link) connects a pixel to its 8-neighbourhood pixels. These links describe the penalty for placing a segmentation boundary between the neighbouring pixels. The second link (i.e., T -link) connects each pixel to the foreground and background nodes. Each of these links has a weight which represents the probability of a pixel belonging to either a foreground or a background. These probabilities are computed in the GMM models and updated in each iteration until a convergence is reached to get a segmented thyroid. The weight of the N -links between pixel m and its 8-neighbourhood pixels, n , is computed as

$$N(m, n) = \frac{\gamma}{\text{dist}(m, n)} e^{-\beta \|z_m - z_n\|^2}, \quad (6)$$

where z_m is the color of pixel m , $\gamma = 50$ as suggested by Blake et al. [28], and β is given as follows by Boykov and Jolly [29]:

$$\beta = \frac{1}{2 \langle \|z_m - z_n\|^2 \rangle}. \quad (7)$$

The initial user initializations are interpolated in the corresponding slices to mark the thyroid and nonthyroid regions (i.e., foreground and background) and create corresponding GMMs. The aforementioned processes are then repeated in each individual images to segment all the thyroid in the dataset. The advantage of this algorithm over ACWE is that it is much faster than ACWE and the user can interact with the result of the segmentation (i.e., postsegmentation) and correct if any errors are present. The results of the segmentation from GC in all the 2D images are used to reconstruct the 3D thyroid by using MeVisLab [22]. The 3D model is updated as soon as the user tries to improve the segmentation results by further interaction in the segmented images. Thus, the accuracy of the algorithm is directly proportional to the number of user interactions on the segmented images.

The increased number of user interactions adds to the computation time of the algorithm. Hence, an optimum number of user interaction should be chosen to obtain the best segmentation results with minimum user interaction. For this purpose, the user interaction in every 10 slices or every 2 mm was proposed.

2.3. Pixel-Based Classifier. This approach is based on training the decision trees by using different features computed from the images. In this work, three image features are computed. The selection of the features is based on the work of Chang et al. [16]. The first feature that is computed is the coefficient of variation $C_{v=(\sigma/\mu)}$, where σ means the standard deviation and μ is the mean of the user selected region during the initialization process. This coefficient is computed in two different sized neighbourhoods (i.e., 4-neighbours and 8-neighbours) of every pixel, thus resulting in two features. The third feature that is computed is the mean of the smaller of the two neighbourhoods. So, the first two features are the coefficient of variation at two different sized neighbourhoods of every pixel, and the third feature is the mean of the smaller of the two neighbourhoods.

The algorithm starts by the user clicking on the inside and outside of the thyroid in several thyroid images from where the features are computed which are then passed as training input for the decision trees. The trained trees later classify the different regions in the image as thyroid or nonthyroid. After segmentation, the user can click in more regions to improve the segmentation results. However, selection of wrong thyroid regions for training the decision trees might result in oversegmentation. So, the user should carefully select the thyroid regions. The presented approach is shown in the schematic diagram in Figure 5.

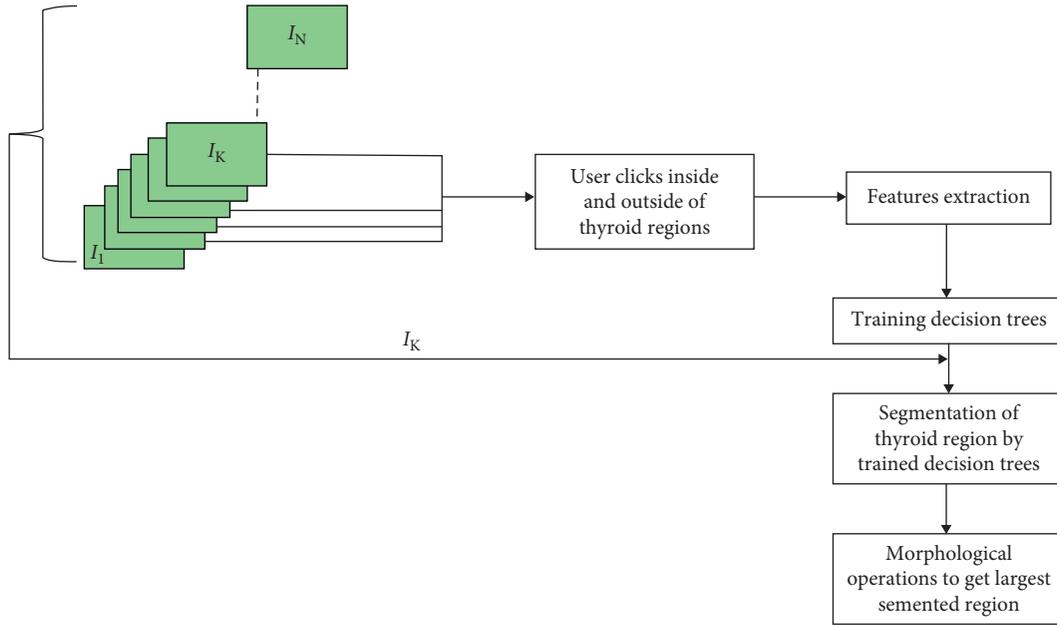


FIGURE 5: Schematic description of PBC segmentation method.

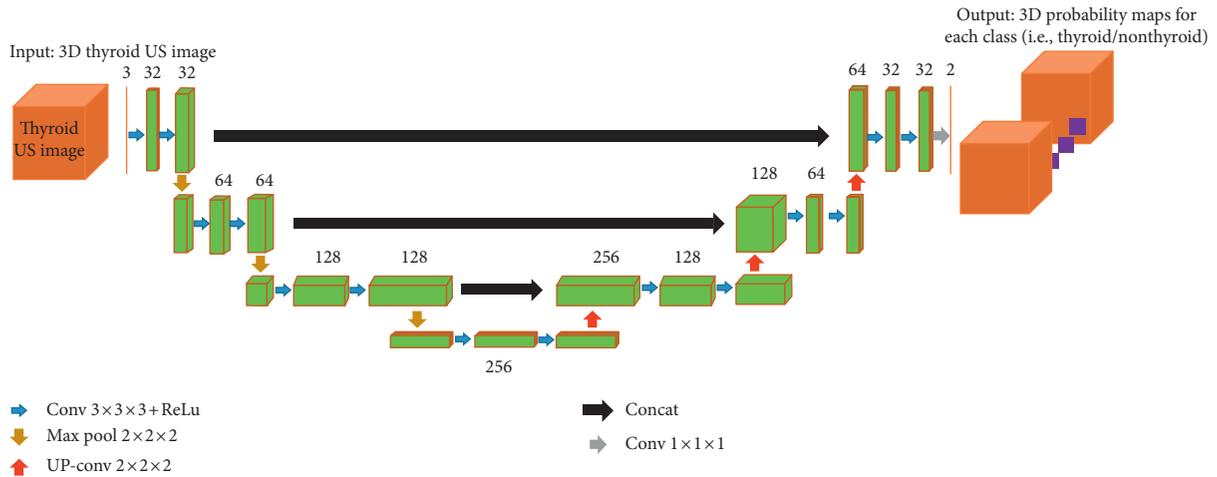


FIGURE 6: Architecture of our 3D U-net CNN. Each green box represents the feature maps.

consists of a 3D thyroid US images and its corresponding ground truth. The input can be represented as $D = (I_n, G_n)$, where I_n denotes one of the thyroid US image and G_n denotes its ground truth obtained from medical experts.

The network consists of two paths (i.e., downsampling/encoder/left side and upsampling/decoder/right side). The downsampling path consists of two $3 \times 3 \times 3$ convolutions followed by a rectified linear unit (ReLU) in each layer and then a $2 \times 2 \times 2$ max pooling with stride of 2 in each dimension. The number of feature channels is doubled in each downsampling step. The upsampling path remaps the lower resolution feature maps to the higher resolution space of the input images. It does this by upsampling the feature maps followed by a $2 \times 2 \times 2$ convolution (upconvolution) which halves the number of feature channels in each upsampling step, a concatenation with the corresponding feature map

from the downsampling path and two $3 \times 3 \times 3$ convolutions, each followed by a ReLU activation. The final convolution layer uses a $1 \times 1 \times 1$ convolution with a voxel-wise softmax activation function to compute a 3D probability map for each of the target label (i.e., thyroid or nonthyroid) as the output of our network.

Since the available datasets were only with 10 datasets consisting of 1416 images, we had to make sure that the network was not overfitting. We performed data augmentation by rotating the images at random angles between -10° and $+10^\circ$, translating between -20 and $+20$ voxels in each dimension, and scaling between -1.5 and 1.5 times from the original size, and since the thyroid are in the left and right sides in the human body, we also flipped the images. We added a dropout of 25% after each pooling layer so that the unnecessary neurons are discarded. Finally, we used Adam

optimizer with relatively low learning rate to make sure that the network was not overfitting. During the training, we observed that the validation accuracy was very close to that of the training accuracy which proves that our network was not overfitting.

2.5.1. 3D Reconstruction and Volume Computation. This step involves 3D reconstruction and volume computation of the segmented thyroid from ACWE and PBC using ImFusion [21] and GC using MeVisLab imaging tools. The segmented 2D images are stored as binary images which are processed to make a video. The video file is passed along with the tracking data to ImFusion and MeVisLab for 3D reconstruction of the thyroid as well as volume computation. The reconstruction is done by the interpolation between the corresponding image frames of the ultrasound sweep to fill the empty spaces between the image slices.

Volume computation of thyroid is particularly important for the medical doctors as this allows them to keep track of the size of the thyroid over time and diagnose whether the patients have any thyroid disorders or not.

3. Experimental Results

3.1. Data Collection. We acquired the thyroid datasets from different clinical university hospital-based sources. A total of 6 healthy human datasets were acquired using the General Electric (GE) Logiq E9 US system which was equipped with the Ascension driveBay EM tracking system. These dataset along with the ground truth are available at OpenCAS [33]. A ML6-15 linear probe was used to acquire the data. All the images were acquired along with a tracking matrix that gave the transformation from the origin of the EM tracking system to the centre of the image. These matrices are used for the 3D reconstruction of the segmented thyroid. The images for the evaluation of nonautomatic methods had a size of 760×500 pixels.

A total of 1416 2D images corresponding to 10 datasets were acquired and used for the evaluation of both the automatic and nonautomatic methods. The 3D models of all the 10 datasets were used for evaluating the automatic methods. All these datasets are stored in the DICOM format. To evaluate the accuracy of our segmentation approaches, we acquired the ground truth by manually tracing the thyroid contour with the help of two medical experts from Magdeburg university clinic using MeVisLab. The datasets are presented in Table 1. The results and discussion may be presented separately, or in one combined section, and may optionally be divided into headed subsections.

3.2. Evaluation Procedure. For evaluation of the segmented images, we compare all the segmentation algorithms using two performance measures. We compute Dice's coefficient (DC) to compare the segmentation accuracy between active contours, graph cut, and pixel-based classifier. Similarly, we compute Hausdorff distance (HD) to compare the accuracy of all the algorithms with the works of Dornheim et al. [17].

TABLE 1: Acquired datasets for the evaluation of nonautomatic and automatic methods.

Dataset	Number of images
D1	96
D2	50
D3	94
D4	55
D5	135
D6	167
D7	216
D8	211
D9	201
D10	191
Total	1416

These measures are computed by comparing the segmentation results with the ground truth images.

Dice's coefficient is a numerical estimate used for comparing the similarity of two samples. In our case, it is a measure to see how accurate our segmented results were by comparing the segmentation results with the ground truth obtained by manual segmentation of the thyroid by trained medical staff. It ranges from 0 to 1, 0 meaning the two datasets are completely different from each other and 1 meaning the datasets completely overlap with each other. It is computed by using the following formula:

$$\text{Dice's coefficient} = \frac{2|X \cap Y|}{|X| + |Y|}, \quad (9)$$

where X is the segmented image and Y is the ground truth.

Similarly, Hausdorff distance measures how far the two subsets of a metric space are from each other. In other words, it is the greatest of all the distances from a point in one set to the closest point in the other set; so the less the distance is, the closer the sets are. It is computed by using the following formula:

$$\text{Hausdorff distance} = \max_{x \in X} \left\{ \min_{y \in Y} \{d(x, Y)\} \right\}, \quad (10)$$

where x are the pixels in the segmented image X and y are the pixels in the ground truth image Y .

The results of segmentation are later used for 3D reconstruction and volume computation. We compare the volumes of the segmented thyroid obtained from all the five algorithms. The accuracy of volume computation is computed by comparing the volume of the 3D reconstructed segmented thyroid to that of the ground truth.

3.3. Analysis of Segmentation and 3D Reconstruction. This section is further divided into two subsections where the first subsection will present the visual analysis of the segmented images and the second subsection will present the quantitative comparison of accuracy, robustness, ease of use, and computation time of all the segmentation algorithms that we have discussed.

3.3.1. Visual Analysis. As mentioned earlier, a total of 1416 images in the ten datasets were taken for the evaluation

procedure. An example of segmented thyroid US image from each of the proposed algorithms will be presented in this section.

The result of segmentation in four thyroid images using ACWE along with the user-initialized mask is shown in Figure 7, GC is shown in Figure 8, PBC is shown in Figure 9, RFC is shown in Figure 10, and CNN is shown in Figure 11. These segmentation results show that the automatic methods produce better segmented thyroid as compared to the non-automatic methods. Within the nonautomatic methods, ACWE and GC give a better approximation of the segmented thyroid region compared to PBC as it has few oversegmented areas. In case of ACWE, the number of iterations of contour evolution is set by the user in order to optimize between the accuracy and the computation time, and because of this, the contour does not reach the narrow areas like the isthmus as shown in Figure 4. We also allow the user to stop the segmentation process where GC is more user-friendly as it allows the user to disregard the oversegmented areas in the post-segmentation stage. PBC works by computing the features from the areas the user select during the initialization process, and because of this, only those areas that have very similar features to that of initialized areas are segmented as thyroid region. This results in undersegmentation as well as oversegmentation most of the times.

Similarly, for the visualization, we performed the 3D reconstruction of the segmented thyroid using the whole set of 2D segmented images. The 3D reconstructed thyroid using ImFusion is shown in Figure 12 and MeVisLab is shown in Figure 13. With MeVisLab, we could even segment the neighbouring artery (i.e., arteria carotis) using a Hessian-based vesselness filter [34].

3.3.2. Quantitative Analysis. In this section, the comparison of the accuracy of segmentation in all the five algorithms (i.e., ACWE, GC, PBC, RFC, and CNN) in terms of DC is presented in Table 2. Also, the comparison of segmentation accuracy of all the five algorithms with two of the standard algorithms [17] in terms of HD is presented in Table 3.

The volume of the segmented thyroids from ACWE, GC, and PBC was computed after the 3D reconstruction using ImFusion and MeVisLab and is presented in Table 4. The results of the volume computation from segmentation results show a close correlation with the segmentation results as well as ground truth in terms of accuracy.

We compared the three nonautomatic algorithms not only based on their accuracy of segmentation but also on other factors like the computation time, robustness of the algorithm, number of user interactions required, etc. All the algorithms performed differently on average where ACWE performed the best with an average DC of 0.800, PBC performed the worst with an average DC of 0.670, and GC performed relatively well with an average DC of 0.765. Even though ACWE was found to be the best performer, it is not accurate enough to use for clinical practices as they require relatively higher accuracy.

ACWE produced undersegmented and oversegmented results in some cases as the contour evolution (set by the

user) does not reach all the regions of thyroid (e.g., isthmus of thyroid) as well as due to the wrong initialization of the contour (this happens when the segmentation results from one image frame are used to segment the corresponding image frames). In order to address these problems, the user could stop the ongoing segmentation at any image frame and change the number of iterations as well as re-initialize the initial mask. 7.7 re-initializations were required on average per dataset. Similarly, the average computation time for ACWE was around 369 seconds in average making it the slowest of all the algorithms, and the initialization determined the rest of the segmentation process. Hence, it is not very robust as compared to the other algorithms. All the methods were implemented in MATLAB in a Lenovo T430 ThinkPad notebook with Intel Core i5-3320M CPU, 2.60 GHz processor, and 8.00 GB RAM.

GC required the most number of user interactions (i.e., 36 scribbles on average) as the user could visualize the segmentation results instantly and improve it with more interactions. Hence, the quality of the results is directly proportional to the number of user interactions using this algorithm. The computation time was around 98 seconds on average per dataset. Graph cut is robust compared to the other two approaches as the user can control the results of the segmentation (i.e., during postsegmentation).

PBC required very few user interactions as the user had to click twice, one inside and one outside of the thyroid. However, the user could take more samples by additional clicks to improve the segmentation results. On average, 4.8 clicks were made while segmenting the images. In the same time, if wrong samples were taken, the user had to start the process from the beginning. This makes the algorithm less robust as compared to GC and ACWE. The computation time was around 10 seconds making it the fastest of all the algorithms. The comparison of the computation time and the number of user interactions required in all the three algorithms are shown in Table 5.

The RFC and CNN yielded an average DC of 0.862 and 0.876, respectively, in ten datasets when tested using a 9-fold cross validation. The computation time for the prediction of one volume was on average 15.62 seconds for the RFC and 34.45 seconds for the CNN. These approaches had higher accuracies of segmentation as compared with ACWE, GC, and PBC. Also, these methods are highly robust as the algorithm does not depend on user interaction. Both of these approaches were also implemented in the same workstation as mentioned before.

4. Discussions and Conclusions

As mentioned earlier, it is essential to keep track of thyroid shape and size over time as it helps to diagnose whether the thyroid is healthy or pathological. In this paper, we have worked on three thyroid segmentation techniques which attempted to extend the 2D segmentation algorithm to generate a 3D segmented thyroid. We have evaluated these algorithms on the basis of accuracy of segmentation, computation time, number of user interactions required, and the

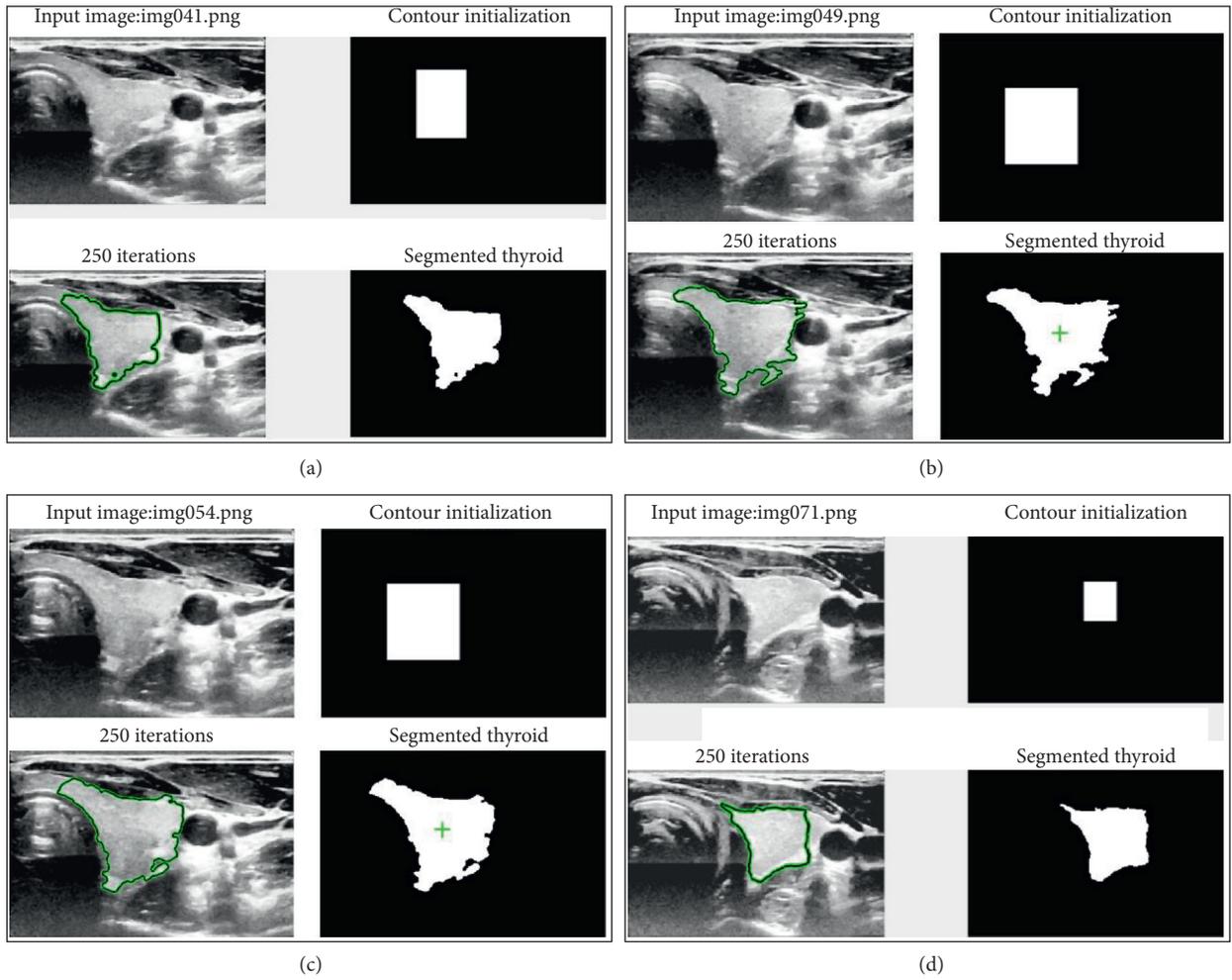


FIGURE 7: Segmentation of 4 different thyroid images using ACWE.

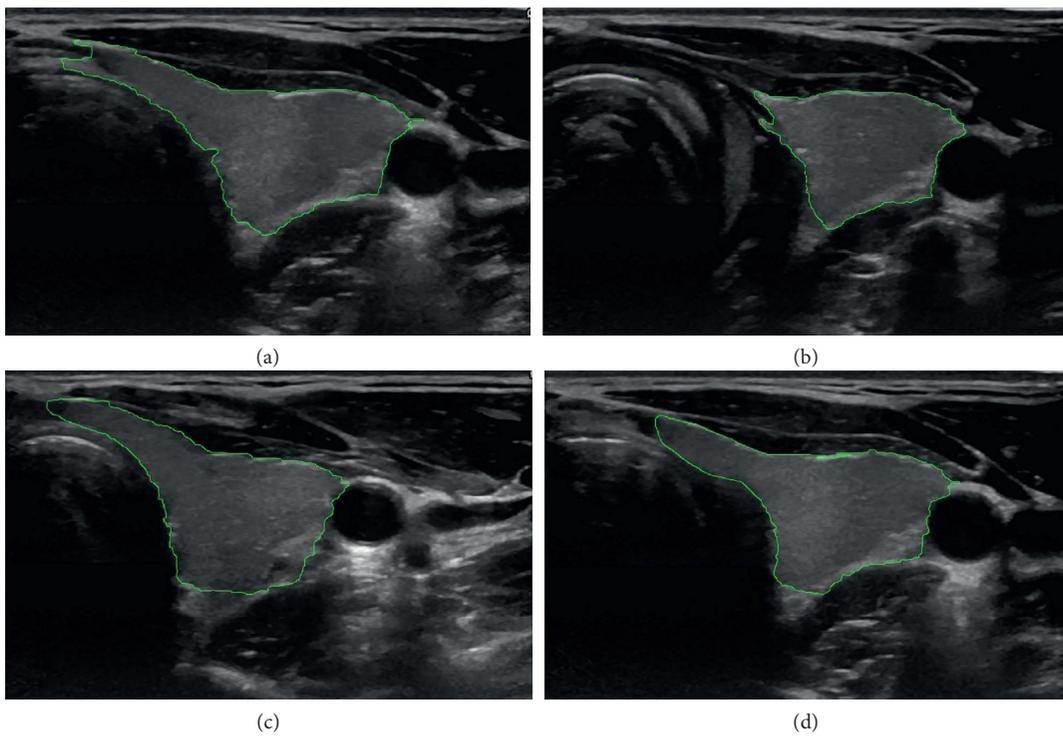


FIGURE 8: Segmentation of 4 different thyroid images using GC.

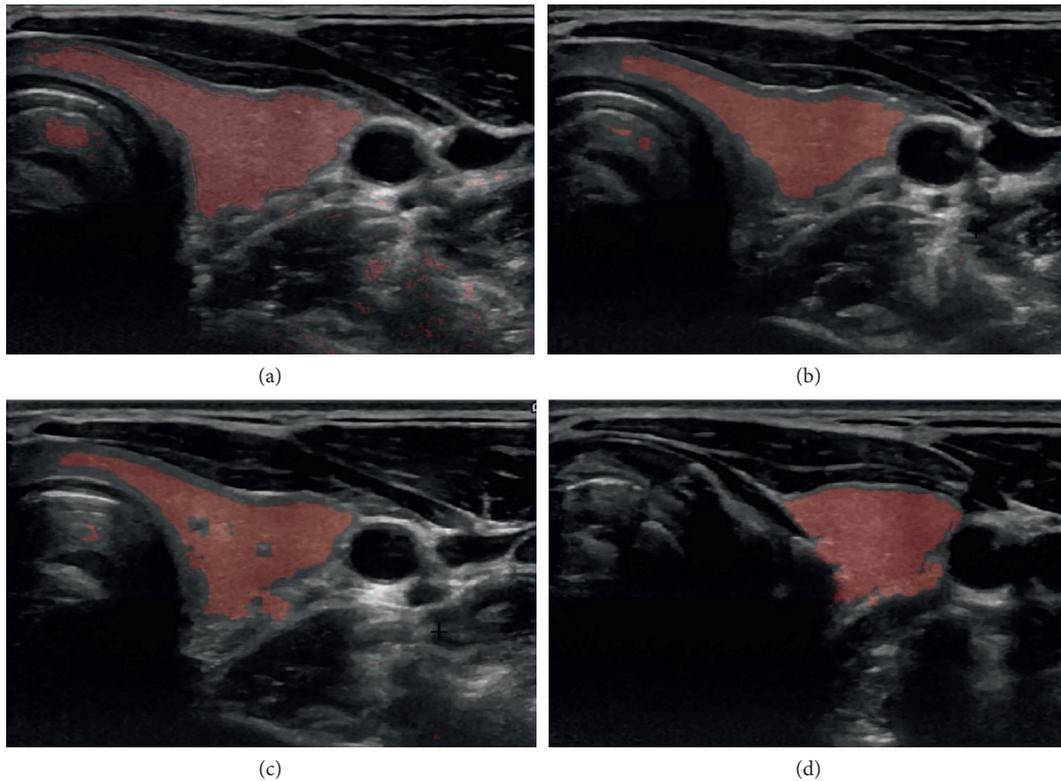


FIGURE 9: Segmentation of 4 different thyroid images using PBC.

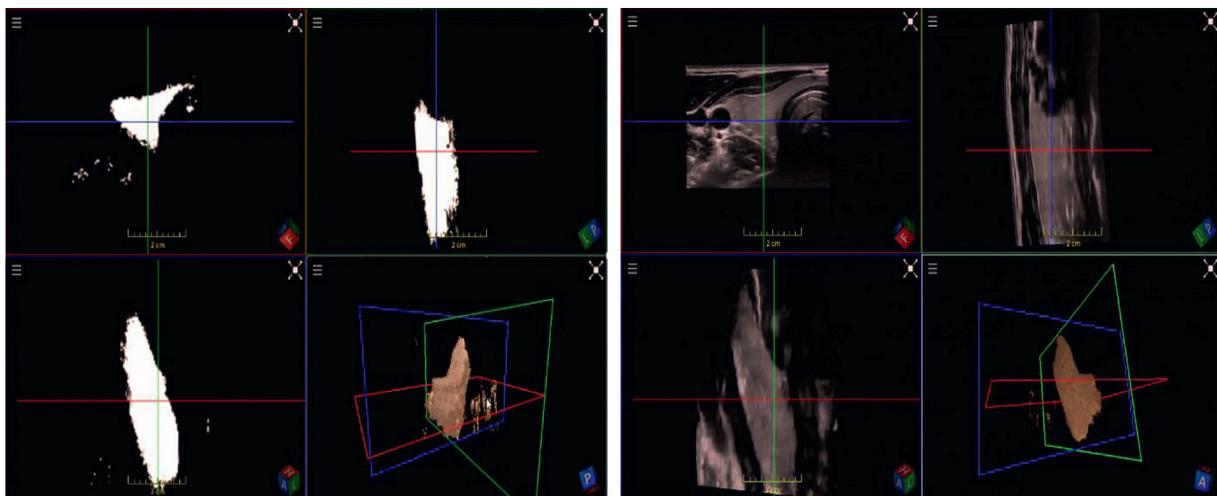


FIGURE 10: Segmentation of thyroid using RFC. Left: segmented thyroid images as binary images with three different viewing angles (top-left, top-right, and bottom-left, and 3D thyroid in bottom-right). Right: original thyroid images with three different viewing angles (top-left, top-right, and bottom-left, and segmented 3D thyroid in bottom right).

robustness. At the same time, a comparison analysis was carried out with the works of Dornheim et al. [17].

We found that all three nonautomatic algorithms performed to different levels. However, a specific approach can be chosen if faster results are required or the least human interaction is desired. The result of volume computation corresponds to the segmentation as well as to the ground truth results which shows that the volume-rendering process

was correct. The accuracy of the discussed algorithms could be further improved.

The computation time of ACWE could be accelerated by reducing the image resolution and using different initialization shapes (e.g., ellipse as thyroid is elliptical in shape). Similarly, the highly echogenic areas near the thyroid could be detected by preprocessing and later the evolution of the contour could be restricted to these areas which would

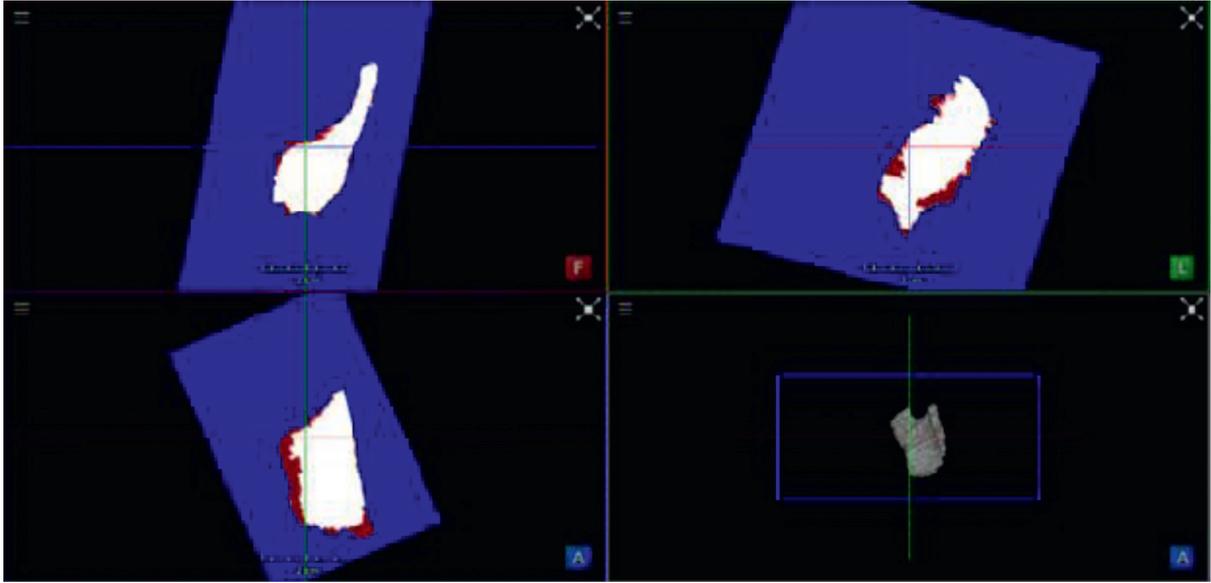


FIGURE 11: Comparison of the segmented thyroid (red) with the ground truth (white) using CNN. Top-left, top-right, and bottom-left: three different viewing angles of segmented thyroid and ground truth. Bottom-right: segmented thyroid in 3D.

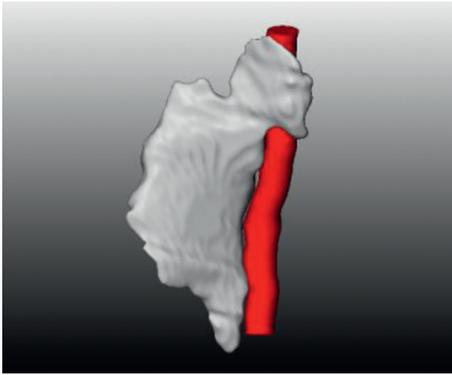


FIGURE 12: 3D reconstructed thyroid using ImFusion.

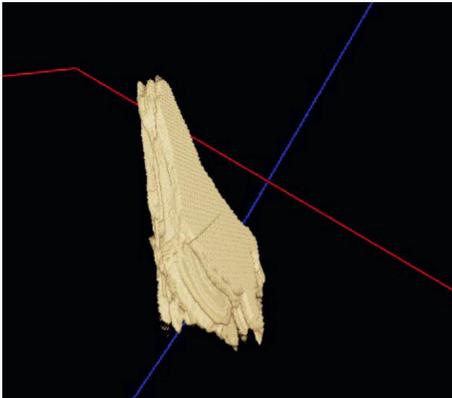


FIGURE 13: 3D reconstructed thyroid (white) along with the neighbouring artery (red) using MeVisLab.

reduce the oversegmentation. The preprocessing step can be further worked on with new contrast enhancement and filtering algorithms so that we generate a good quality ultrasound images before segmentation. In case of graph cut,

TABLE 2: Comparison of DC in five segmentation algorithms.

Dataset	ACWE	GC	PBC	RFC	CNN
D1	0.841	0.729	0.749	0.859	0.863
D2	0.819	0.636	0.666	0.864	0.876
D3	0.804	0.706	0.610	0.853	0.872
D4	0.816	0.841	0.680	0.872	0.869
D5	0.771	0.706	0.673	0.831	0.879
D6	0.781	0.853	0.623	0.853	0.874
D7	0.788	0.848	0.659	0.895	0.861
D8	0.746	0.746	0.547	0.877	0.888
D9	0.785	0.676	0.732	0.841	0.901
D10	0.852	0.912	0.761	0.875	0.877
Average	0.800	0.765	0.670	0.862	0.876

TABLE 3: Comparison of our approaches with other segmentation algorithms.

Dataset	Hausdorff distance (mm)
ACWE	8.1
GC	8.3
PBC	9.5
RFC	7.5
CNN	7.0
Volumetric mass spring model	11.1
Surface mass spring model	9.8

a postprocessing step could be added which could take the shape prior information of the thyroid and remove the oversegmented areas automatically. Segmentation by pixel-based classifier could be improved with more image features. It can be made fully automatic using machine learning approaches and a postprocessing step to remove the oversegmented areas just like in graph cut. Also, advanced thresholding and connected component analysis could be

TABLE 4: Volume comparison of 2D segmented and 3D reconstructed thyroid to ground truth in cm³.

Dataset	ACWE	GC	PBC	Ground truth
D1	10.15	8.79	9.04	12.07
D2	11.46	8.90	9.30	13.99
D3	12.45	10.97	9.42	15.51
D4	11.91	13.82	9.86	14.64
D5	10.83	9.78	9.41	13.93
D6	12.18	13.95	10.90	9.86
D7	10.85	11.68	9.07	13.77
D8	10.66	10.66	7.82	14.29
D9	11.91	10.25	11.10	15.16
D10	10.52	11.26	9.40	12.35
Average	11.29	11.01	9.53	13.56

TABLE 5: Comparison of average computation time and number of interactions.

Approach	Computation time (sec)	Number of user interaction
ACWE	369	7.7
GC	98	36
PBC	10	4.8
RFC	15.62	None
CNN	34.45	None

performed to get the largest connected component and subsequently remove any elements outside that component to get a better segmented thyroid. Furthermore, all the acquired datasets were from healthy patients, so pathological datasets have to be acquired and tested on the discussed algorithms to evaluate their practical usefulness.

We figured out that the first three methods lacked automation and machine intelligence, were not highly accurate, and required long computation time. Hence, we implemented an RFC and a CNN that predict for each voxel the probability of belonging to the thyroid. Both approaches were trained for each voxel the probability of belonging to the thyroid in the available ten datasets, and they show better results as compared to the nonautomatic approaches.

As next steps, we will investigate several other thyroid segmentation approaches based on machine learning that operate directly on the volumetric three-dimensional ultrasound data instead of the 2D frames volumetrically compounded with isotropic spacing to form a 3D volume [35]. Similarly, future steps towards these automatic approaches must include more training data especially those with thyroid diseases as we have carried out our tests on healthy thyroid images only.

Data Availability

The data used for the evaluation purpose are uploaded in Open-CAS and available publicly (<http://opencas.webarchiv.kit.edu/data/thyroid.zip>).

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

We would like to thank General Electrics, USA, for providing us with the LogiqE9 Ultrasound equipment to generate the Thyroid Ultrasound Data. Special thanks are due to our clinical partners at the University of Magdeburg (Prof. C. Arens) for helping us obtain the Thyroid Ultrasound Datasets. This work has been funded by Federal Ministry of Education and Research (BMBM) in the context of the INKA Project (Grant no. 03IPT7100X).

References

- [1] Understanding Thyroid Problems: The Basics, <https://www.webmd.com/women/picture-of-the-thyroid>.
- [2] J. Zhao, W. Zheng, L. Zhang, and H. Tian, "Segmentation of ultrasound images of thyroid nodule for assisting fine needle aspiration cytology," *Health Information Science and Systems*, vol. 1, no. 1, p. 5, 2013.
- [3] J. Kaur and A. Jindal, "Comparison of thyroid segmentation algorithms in ultrasound and scintigraphy images," *International Journal of Computer Applications*, vol. 50, no. 23, pp. 24–27, 2012.
- [4] A. S. Agustin, S. S. Babu, and K. T. Nadu, "Thyroid segmentation on US medical images: an overview," *International Journal of Emerging Technology and Advanced Engineering*, vol. 2, no. 12, 2012.
- [5] A. Stefano, S. Vitabile, G. Russo et al., "An enhanced random walk algorithm for delineation of head and neck cancers in PET studies," *Medical & Biological Engineering & Computing*, vol. 55, no. 6, pp. 897–908, 2017.
- [6] D. Selvathi and V. Sharnitha, "Thyroid classification and segmentation in ultrasound images using machinelearning algorithms," in *Proceedings of 2011 International Conference on Signal Processing, Communication, Computing and Networking Technologies*, pp. 836–841, IEEE, Piscataway, NJ, USA, 2011.
- [7] N. H. Mahmood and A. H. Rusli, "Segmentation and area measurement for thyroid ultrasound image," *International Journal of Scientific & Engineering Research*, vol. 2, no. 12, pp. 1–8, 2011.
- [8] E. A. Mylona, M. A. Savelonas, and D. Maroulis, "Automated adjustment of region-based activecontour parameters using local image geometry," *IEEE Transactions on Cybernetics*, vol. 44, no. 12, pp. 2757–2770, 2014.
- [9] E. A. Mylona, M.A. Savelonas, and D. Maroulis, "Self-parameterized active contours based on regional edge structure for medical image segmentation," *SpringerPlus*, vol. 3, no. 1, p. 424, 2014.
- [10] E. G. Keramidias, D. K. Iakovidis, D. Maroulis, and S. Karkanis, "Efficient and effective ultrasound image analysis scheme for thyroid nodule detection," in *Proceedings of International Conference Image Analysis and Recognition*, pp. 1052–1060, Springer, Montreal, Canada, August 2007.
- [11] D. E. Maroulis, M. A. Savelonas, D. K. Iakovidis, S. A. Karkanis, and N. Dimitropoulos, "Variable background active contour model for computer-aided delineation of nodules in thyroid ultrasound images," *IEEE Transactions on Information Technology in Biomedicine*, vol. 11, no. 5, pp. 537–543, 2007.
- [12] M. A. Savelonas, D. K. Iakovidis, I. Legakis, and D. Maroulis, "Active contours guided by echogenicity and texture for delineation of thyroid nodules in ultrasound images," *IEEE*

- Transactions on Information Technology in Biomedicine*, vol. 13, no. 4, pp. 519–527, 2009.
- [13] H. Garg and A. Jindal, “Segmentation of thyroid gland in ultrasound image using neural network,” in *Proceedings of 2013 Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT)*, pp. 1–5, Tiruchengode, Tamil Nadu, India, July 2013.
- [14] N. S. Narayan, P. Marziliano, J. Kanagalingam, and C. G. Hobbs, “Speckle patch similarity for echogenicity-based multiorgan segmentation in ultrasound images of the thyroid gland,” *IEEE Journal of Biomedical and Health Informatics*, vol. 21, no. 1, pp. 172–183, 2017.
- [15] E. N. Kollorz, D. A. Hahn, R. Linke, T. W. Goecke, J. Hornegger, and T. Kuwert, “Quantification of thyroid volume using 3-d ultrasound imaging,” *IEEE Transactions on Medical Imaging*, vol. 27, no. 4, pp. 457–466, 2008.
- [16] C.-Y. Chang, Y.-F. Lei, C.-H. Tseng, and S.-R. Shih, “Thyroid segmentation and volume estimation in ultrasound images,” *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 6, pp. 1348–1357, 2010.
- [17] J. Dornheim, L. Dornheim, B. Preim, K.D. Tonnies, I. Hertel, and G. Strauss, “Stable 3d mass-spring models for the segmentation of the thyroid cartilage,” in *Proceedings of the DAGM, Otto-von-Guericke-Universitat Magdeburg*, Vienna, Austria, August 2005.
- [18] A. Osman, *Automated evaluation of three dimensional ultrasonic datasets*, Ph.D. thesis, INSA de Lyon, Villeurbanne, France, 2013.
- [19] T. F. Chan and L. A. Vese, “Active contours without edges,” *IEEE Transactions on Image Processing*, vol. 10, no. 2, pp. 266–277, 2001.
- [20] C. Rother, V. Kolmogorov, and A. Blake, “Grabcut: interactive foreground extraction using iterated graph cuts,” *ACM*, vol. 23, no. 3, pp. 309–314, 2004.
- [21] ImFusion GmbH, <https://www.imfusion.de>.
- [22] MeVisLab, <https://www.mevislab.de/>.
- [23] R. Gir, L. Jain, and R. Rai, “Speckle reduction of synthetic aperture radar images using median filter and savitzky-golay filter,” *International Journal of Computer Applications*, vol. 113, no. 11, pp. 38–43, 2015.
- [24] P. C. Tay, C. D. Garson, S. T. Acton, and J. A. Hossack, “Ultrasound despeckling for contrast enhancement,” *IEEE Transactions on Image Processing*, vol. 19, no. 7, pp. 1847–1860, 2010.
- [25] R. Courant, K. Friedrichs, and H. Lewy, “On the partial difference equations of mathematical physics,” *IBM Journal of Research and Development*, vol. 11, no. 2, pp. 215–234, 1967.
- [26] P. Poudel, C. Hansen, J. Sprung, and M. Friebe, “3d segmentation of thyroid ultrasound images using active contours,” *Current Directions in Biomedical Engineering*, vol. 2, no. 1, pp. 467–470, 2016.
- [27] M. T. Orchard and C. A. Bouman, “Color quantization of images,” *IEEE Transactions on Signal Processing*, vol. 39, no. 12, pp. 2677–2458, 1991.
- [28] A. Blake, C. Rother, M. Brown, P. Perez, and P. Torr, “Interactive image segmentation using an adaptive GMMRF model,” in *Proceedings of European Conference on Computer Vision*, pp. 428–441, Springer, Prague, Czech Republic, May 2004.
- [29] Y. Y. Boykov and M. P. Jolly, “Interactive graph cuts for optimal boundary and region segmentation of objects in ND image,” in *Proceedings of Eighth IEEE International Conference on Computer Vision (ICCV’01)*, vol. 1, pp. 105–112, Vancouver, British Columbia, Canada, July 2001.
- [30] L. Breiman, “Random forests,” *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.
- [31] A. Crimini and J. Shotton, *Decision forests for Computer Vision and Medical Image Analysis*, Springer Science and Business Media, 2013.
- [32] O. Cicek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger, “3D U-Net: learning dense volumetric segmentation from sparse annotation,” in *Proceedings of International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 424–432, Springer, Athens, Greece, October 2016.
- [33] Open-CAS, <http://opencas.webarchiv.kit.edu/data/thyroid.zip>.
- [34] Y. Sato, S. Nakajima, H. Atsumi et al., “3d multi-scale line filter for segmentation and visualization of curvilinear structures in medical images,” in *Proceedings of First Joint Conference Computer Vision, Virtual Reality and Robotics in Medicine and Medical Robotics and Computer-Assisted Surgery, CVRMed-MRCAS’97*, pp. 213–222, Springer, Grenoble, France, March 1997.
- [35] A. Karamalis, W. Wein, O. Kutter, and N. Navab, “Fast hybrid freehand ultrasound volume reconstruction,” in *Proceedings of Medical Imaging 2009: Visualization, Image-Guided Procedures, and Modeling*, vol. 7261, International Society for Optics and Photonics, Vista, FL, USA, February 2009.