

The Journal of Healthcare Engineering

Advancements of Medical Image Enhancement in Healthcare Applications

Lead Guest Editor: Shujun Fu

Guest Editors: Ming Zhang, Chengpo Mu, and Xiaohong Shen





Advancements of Medical Image Enhancement in Healthcare Applications

Journal of Healthcare Engineering

Advancements of Medical Image Enhancement in Healthcare Applications

Lead Guest Editor: Shujun Fu

Guest Editors: Ming Zhang and Chengpo Mu



Copyright © 2018 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
William Bertucci, France
Olivier Beuf, France
Daniel H.K. Chow, Hong Kong
Tiago H. Falk, Canada
Mostafa Fatemi, USA
Jesus Favela, Mexico
Joseph Finkelstein, USA
Jesus Fontecha, Spain
Antonio Gloria, Italy
Philippe Gorce, France
Valentina Hartwig, Italy
Andreas H. Hielscher, USA

Norio Iriguchi, Japan
Zhongwei Jiang, Japan
John S. Katsanis, Greece
Terry K.K. Koo, USA
Panagiotis Kosmas, UK
Michel Labrosse, Canada
Jui-Yang Lai, Taiwan
Feng-Huei Lin, Taiwan
Maria Lindén, Sweden
Francisco Lopez-Valdes, Spain
Andreas Maier, Germany
Mehran Moazen, UK
Rafael Morales, Spain

David Moratal, Spain
Vincenzo Positano, Italy
Alessandro Reali, Italy
Jose Joaquin Rieta, Spain
Sébastien Roth, France
Hélder A. Santos, Finland
Emiliano Schena, Italy
Maurizio Schmid, Italy
Jiann-Shing Shieh, Taiwan
Jinshan Tang, USA
Vinoy Thomas, USA
Ioannis G. Tollis, Greece

Contents

Advancements of Medical Image Enhancement in Healthcare Applications

Shujun Fu , Ming Zhang, Chengpo Mu , and Xiaohong Shen

Volume 2018, Article ID 7035264, 2 pages

Efficient OCT Image Enhancement Based on Collaborative Shock Filtering

Guohua Liu , Ziyu Wang , Guoying Mu , and Peijin Li 

Volume 2018, Article ID 7329548, 7 pages

Medical Imaging Lesion Detection Based on Unified Gravitational Fuzzy Clustering

Jean Marie Vianney Kinani, Alberto Jorge Rosales Silva, Francisco Gallegos Funes, Dante Mújica Vargas, Eduardo Ramos Díaz, and Alfonso Arellano

Volume 2017, Article ID 8536206, 14 pages

Semiautomatic Epicardial Fat Segmentation Based on Fuzzy c-Means Clustering and Geometric Ellipse Fitting

Vladimir Zlokolica, Lidija Krstanović, Lazar Velicki, Branislav Popović, Marko Janev,

Ratko Obradović, Nebojsa M. Ralević, Ljubomir Jovanov, and Danilo Babin

Volume 2017, Article ID 5817970, 12 pages

Frame-Based CT Image Reconstruction via the Balanced Approach

Weifeng Zhou and Hua Xiang

Volume 2017, Article ID 1417270, 7 pages

Automatic Radiographic Position Recognition from Image Frequency and Intensity

Ning-ning Ren, An-ran Ma, Li-bo Han, Yong Sun, Yan Shao, and Jian-feng Qiu

Volume 2017, Article ID 2727686, 9 pages

The Edge Detectors Suitable for Retinal OCT Image Segmentation

Su Luo, Jing Yang, Qian Gao, Sheng Zhou, and Chang'an A. Zhan

Volume 2017, Article ID 3978410, 13 pages

Deep Learning- and Transfer Learning-Based Super Resolution Reconstruction from Single Medical Image

YiNan Zhang and MingQiang An

Volume 2017, Article ID 5859727, 20 pages

Editorial

Advancements of Medical Image Enhancement in Healthcare Applications

Shujun Fu ¹, Ming Zhang,² Chengpo Mu ³, and Xiaohong Shen⁴

¹*School of Mathematics, Shandong University, Jinan, China*

²*Tufts Medical Center, Tufts University, Boston, MA, USA*

³*School of Mechatronics Engineering, Beijing Institute of Technology, Beijing, China*

⁴*Department of Biomedical Engineering, Emory University, Atlanta, GA, USA*

Correspondence should be addressed to Shujun Fu; shujunfu@163.com

Received 24 December 2017; Accepted 25 December 2017; Published 29 March 2018

Copyright © 2018 Shujun Fu 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.

1. Introduction

As one of the world's largest and fastest-growing industries, healthcare engineering refers to all aspects of the prevention, diagnosis, treatment, and management of illness, as well as the preservation and improvement of physical and mental health and well-being through medical services [1, 2].

Medical imaging technologies play more and more important roles not only in the diagnosis and treatment of diseases but also in disease prevention, health checkup, major disease screening, health management, early diagnosis, disease severity evaluation, choice of treatment methods, treatment effect evaluation, and rehabilitation. The status of medical imaging technologies has increased continuously in healthcare applications [3–6].

Due to its ability to make the diagnosis and treatment of disease, image-guided surgery, and other medical links more timely, accurate, and efficient, medical image enhancement has become a routine task. Through producing excellent tissue uniformity, optimized contrast, edge enhancement, artifact elimination, intelligent noise reduction, and so forth, cutting-edge image enhancement helps doctors accurately interpret medical images, a crucial foundation for better diagnosis and treatment [3–6].

2. The Special Issue

Although the special issue focuses on medical image enhancement, other important image processing methods

are reported from submitted papers, including image reconstruction, enhancement, segmentation, feature extraction, and recognition.

Next, we will introduce seven important works in the order of the above full-image processing procedure [5].

For the sparse reconstruction of X-ray computed tomography (CT) image with a low-dose projection data, W. Zhou and H. Xiang proposed a wavelet frame-based regularization method, where a novel balanced hybrid model was considered with two sparse regularization terms related to the latent solution and the frame coefficient. A fast alternative direction method was devised to solve the proposed model so that four subproblems can be easily solved. Numerical experiments verified the efficiency and the accuracy of the proposed model.

Efficient enhancement of noisy optical coherence tomography (OCT) images is a key task for interpreting them correctly in the ophthalmology for vision-related diseases. G. Liu et al. proposed a collaborative shock filtering to better enhance details and layered structures of human retina image. Noisy OCT image is first denoised by a 3D collaborative filtering method with a novel similarity measure; then, the denoised image is sharpened by a shock-type filtering for edge and detail enhancement. In order to improve image contrast for the detection of tiny lesions, a gamma transformation is also used to enhance the images within proper gray levels.

To better interpret medical images plays an important role in medical diagnosis and treatment. Y. Zhang and

M.Q. An proposed a super-resolution reconstruction method based on transfer learning and deep learning, which contains one bicubic interpolation template layer and two convolutional layers. A SIFT feature-based transfer learning method is used to selectively add other types of images into the training dataset. The experiments on eight distinctive medical images showed the improvement of image quality with slightly sharper edges than other deep learning approaches in less time.

Image segmentation is a classic difficult problem linking low-level processing to high-level recognition. Three important works on three types of imaging modes are reported in this field. A fast and robust segmentation algorithm was presented by J. Marie et al. integrating the fuzzy clustering and the gravitational law. After the fuzzy enhancement of MRI images (T1/T2, FLAIR), the healthy structures in these images were separated from the unhealthy ones. Finally, the lesion contour is automatically outlined through an initialization-free level set evolution method. An 84%–93% overlap performance was obtained using clinical and synthetic brain datasets with different and heterogeneous types of lesions.

The segmentation of epicardial fat is also an important task for indicating the risk level for developing various cardiovascular diseases and predicting the progression of certain diseases. V. Zlokolica et al. proposed a semiautomatic approach for segmentation and quantification of epicardial fat from 3D CT images. After the user marks a patch corresponding to the epicardial fat, a 3D segmentation was done slice-by-slice integrating 2D processing using the fuzzy c-means clustering and ellipse fitting, followed by a filtering out of undesired parts of the target cluster.

Retinal layer thickness measurement offers important quantitative information for reliable diagnosis of retinal diseases and for the evaluation of disease development and medical treatment responses. This task critically depends on the accurate edge detection and segmentation of the retinal layers in OCT images. S. Luo et al. identified three most promising edge detection algorithms: Canny edge detector, the two-pass method, and the EdgeFlow technique. The quantitative evaluation results showed that the two-pass method outperforms consistently the other two methods in delineating and localizing the retinal layer boundaries in the OCT images. They also found that the OCT images contain more intensity gradient information than texture changes along the retinal layer boundaries.

With the development of digital X-ray imaging and processing methods, automatic categorization and analysis of massive digital radiographic images are urgently needed. N. Ren et al. developed a recognition method of radiographic positions using the frequency curve classification and gray information matching. Compared with predefined curve types (radiographic position) from a whole-body phantom image, a radiographic image is classified as one of the six positions with the help of image intensity similarity: the head, lung, lumber, pelvis, joint, and limbs. An average 93.78% accuracy is reported using the proposed method.

Acknowledgments

The guest editors of this special issue would like to thank all authors of the submitted papers, as well as all reviewers for their hard work and detailed reviews.

Shujun Fu
Ming Zhang
Chengpo Mu
Xiaohong Shen

References

- [1] P. W. Khong and D. Ghista, "Healthcare engineering for an efficient medical care system," *International Journal of Healthcare Technology and Management*, vol. 7, no. 5, pp. 429–439, 2006.
- [2] M.-C. Chyu, T. Austin, F. Calisir et al., "Healthcare engineering defined: a white paper," *Journal of Healthcare Engineering*, vol. 6, no. 4, pp. 635–648, 2015.
- [3] J. S. Duncan and N. Ayache, "Medical image analysis: progress over two decades and the challenges ahead," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 1, pp. 85–106, 2000.
- [4] J. L. Semmlow, *Biosignal and Biomedical Image Processing*, Marcel Dekker Inc., 2004.
- [5] I. Bankman, *Handbook of Medical Image Processing and Analysis*, Academic Press, Elsevier, 2009.
- [6] A. P. Dhawan, *Medical Image Analysis*, John Wiley & Sons Press, 2nd edition, 2011.

Research Article

Efficient OCT Image Enhancement Based on Collaborative Shock Filtering

Guohua Liu ¹, Ziyu Wang ², Guoying Mu ³, and Peijin Li ⁴

¹Department of Ophthalmology, Qilu Children's Hospital of Shandong University, Jinan 250022, China

²Department of Radiology, Yidu Central Hospital of Weifang, Qingzhou 262500, China

³Department of Ophthalmology, Shandong Provincial Hospital, Jinan 250021, China

⁴School of Computer Science and Technology, Shandong University, Jinan 250101, China

Correspondence should be addressed to Guoying Mu; mgyeyes@163.com

Received 5 May 2017; Accepted 14 August 2017; Published 1 February 2018

Academic Editor: Chengpo Mu

Copyright © 2018 Guohua Liu 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.

Efficient enhancement of noisy optical coherence tomography (OCT) images is a key task for interpreting them correctly. In this paper, to better enhance details and layered structures of a human retina image, we propose a collaborative shock filtering for OCT image denoising and enhancement. Noisy OCT image is first denoised by a collaborative filtering method with new similarity measure, and then the denoised image is sharpened by a shock-type filtering for edge and detail enhancement. For dim OCT images, in order to improve image contrast for the detection of tiny lesions, a gamma transformation is first used to enhance the images within proper gray levels. The proposed method integrating image smoothing and sharpening simultaneously obtains better visual results in experiments.

1. Introduction

Optical coherence tomography (OCT) is an important emerging noninvasive technique in medical imaging, which offers high-speed tomographic imaging of human tissues [1, 2]. Layered structures of the human retina provided by OCT imaging are commonly detected and measured for diagnosis and prognostic evaluation [3]. Thus, medical OCT imaging has a variety of clinical applications and is widely used in the ophthalmology for vision-related diseases. Since OCT imaging takes advantage of low-coherence interferometry, it inevitably suffers from speckle noise, visible as dark and bright spots on the imaged objects [4], which severely degrades OCT image contrast and quality and makes it particularly challenging to identify fine structures and features. Therefore, it is necessary to remove the speckle noise and to enhance OCT images in order to interpret OCT images correctly.

There are many approaches to reduce the speckle noise of OCT images for enhancing their image quality [5, 6]. For example, the geometry-oriented calculus of variations and

partial differential equations have been widely used in image processing [7, 8]: the total variation regularization [9] and the anisotropic diffusion [10] for image denoising and enhancement, the shock filters [11, 12] for image sharpening, and so forth. However, in the process of image denoising and enhancement, these classic geometrical regularization methods, based on operators in differential geometry such as gradient, divergence, and directional derivative, often tend to modify the image towards a piecewise constant function and blur fine features of the image, particularly the image's details and layered structures.

Other image enhancement methods include wavelet decomposition [13], local statistical [14], and self-similarity and sparsity-based methods [15–17]. These algorithms can reduce noise to a certain extent, but they also lead to the loss of image details and consequently the blur of important image features. Moreover, some algorithms are fairly time-consuming for them to enhance a noisy OCT image.

In 2005, Buades et al. proposed a nonlocal means (NLM) filter based on the self-similarity in the whole image, which can effectively preserve image details and textures [18, 19].

The nonlocal means filter produces impressive results in denoising textured patterns. A more complicated neighborhood filter based on the self-similarity of image blocks was introduced by Dabov et al. [20], which is actually also a weighted average algorithm, just after a transform-domain collaborative filtering. This algorithm achieves better denoising performance in terms of peak signal-to-noise ratio [21].

In this paper, to better preserve or enhance image details and layered structures, we propose an effective enhancement algorithm for OCT image denoising and enhancement based on collaborative shock filtering. The basic idea is that noisy OCT image is first denoised with a collaborative filtering of speckle noise following a gamma distribution, and then the denoised image is sharpened by a shock-type filtering for edge and detail enhancement. For dim OCT image, in order to improve the contrast of the retinal image for the detection of tiny lesions, a gamma transformation [22] is first used to enhance the image within proper gray levels.

Image smoothing and sharpening are two opposite operations in image processing. Generally speaking, image smoothing is to eliminate unnecessary and false discontinuous features (such as noise), while image sharpening is to produce or enhance some discontinuous features (such as edges and details) in proper positions of the image. In many cases, image denoising methods often inevitably blur image edges and details even if these methods are designed elaborately. For example, see following denoising experiments on noisy images. Thus, both image smoothing and sharpening are needed simultaneously even if only in an image denoising task. One will see that the methods integrating the above two operations simultaneously obtain better visual results in experiments [23].

The rest of this paper is organized as follows. In Section 2, we briefly review some classic works: shock filters and self-similarity filtering. Then, a collaborative shock-filtering algorithm integrating these classic methods is proposed in Section 3. Experiments on test images are described in Section 4. Finally, conclusions are drawn out in Section 5.

2. Shock Filters and Self-Similarity Filtering

In this section, we review some classic methods in image denoising and enhancement.

2.1. Shock Filter. In [24], some special ideas and techniques developed in numerical solutions of nonlinear hyperbolic equations were applied to feature-oriented image enhancement for the first time. Osher and Rudin introduced a novel image sharpening technique called the shock filter (SF) [11], which is based on a modification of the nonlinear Burgers' equation and simulates the shock wave calculation in the computational fluid mechanics [25]. Different from the nonlinear parabolic equation of diffusion-type process, they proposed a hyperbolic one:

$$\frac{\partial u}{\partial t} = -\text{sign}(u_{NN})|\nabla u|, \quad (1)$$

where u is the observed noisy image, sign is a sign function, and u_{NN} is the second directional derivative of the image

along a local normal direction to the isophote line. It detects an image edge using the zero crossing of u_{NN} , where a shock is formed at the speed of $|\nabla u|$.

Considering the image noise in the estimation of edges, Alvarez and Mazorra added a smoothing kernel and coupled the anisotropic diffusion with the shock filter (ADSF) [12] for noise elimination and edge sharpening:

$$\frac{\partial u}{\partial t} = -\text{sign}(G_\sigma * u_{NN})|\nabla u| + cu_{TT}, \quad (2)$$

where G_σ is a Gaussian kernel with standard deviation σ , u_{TT} is the second directional derivative of the image along a local tangent direction, and c is a constant to balance the anisotropic diffusion and the shock filter.

2.2. Self-Similarity Filtering. Another effective image denoising technique is neighborhood filters [8], which consider self-similarity between two pixels or blocks of the image both in spatial location and in gray level. Buades et al. proposed the following nonlocal means (NLM) filter [18, 19]:

$$NL_h u(x) = \frac{1}{c(x)} \int_{\Omega} \exp \left(- \frac{(G_\sigma * (u(x + \cdot) - u(y + \cdot)))^2(0)}{h^2} \right) u(y) dy, \quad (3)$$

where $c(x) = \int_{\Omega} \exp(-(G_\sigma * (u(x + \cdot) - u(y + \cdot)))^2(0)/h^2) dy$ is a normalization factor, and h is a filtering parameter related to noise level, and $\{x, y \in \Omega \subset \mathbf{R}^2\}$ are 2D spatial coordinates. The nonlocal means filter gives better results in denoising textured patterns.

A more complicated neighborhood filter called the Block-matching and 3D filtering (BM3D) algorithm was introduced by Dabov et al. [20], which has three main steps: grouping of similar image blocks, 3D collaborative filtering of these blocks in the spectrum domain, and aggregation of all local estimates. This algorithm includes basic and final estimations using the following weighted average:

$$\hat{u}(x) = \frac{\sum_R \sum_S w_{R,S} \hat{U}_{R,S}(x)}{\sum_R \sum_S w_{R,S} \chi_S(x)}, \quad x \in \Omega, \quad (4)$$

where R and S denote the reference block and the similar block, respectively, w is the weight of the corresponding block, $\hat{U}(x)$ is the local block-wise estimate, and $\chi(x)$ is a characteristic function. As they claim, this algorithm is currently one of best denoising methods in terms of peak signal-to-noise ratio [20, 21].

Finally, as one will see, the neighborhood filters are mainly used for image denoising, while the shock filters can be used for both image denoising and enhancement simultaneously, because of their inherent local backward diffusion [25].

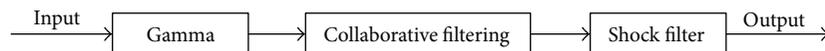


FIGURE 1: A flowchart of the proposed collaborative shock-filtering algorithm. A noisy OCT image is enhanced through three steps in sequence: gamma manipulation (gamma), collaborative filtering, and shock filter, respectively.

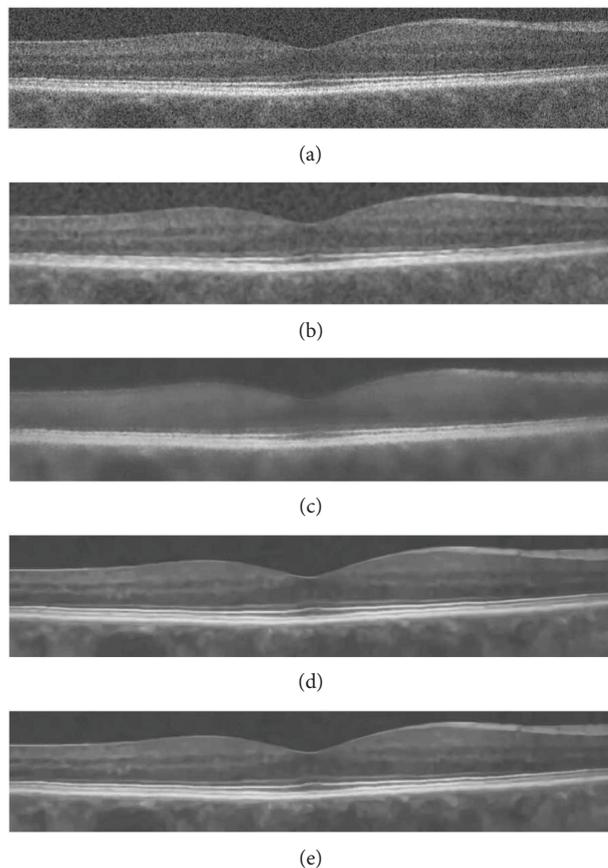


FIGURE 2: Enhancement of noisy OCT image: (a–e) original image, ADSF, NLM, BM3D, and the proposed collaborative shock filtering, respectively.

3. Collaborative Shock Filtering

The main information of an image is encoded in its edges, details, and structures. These components need to be processed in special ways to enhance the image as well as possible.

In order to better preserve or enhance image details and layered structures, we propose an efficient collaborative shock-filtering (CSF) algorithm for image denoising and enhancement, which fuses different advantages of previously mentioned classic methods: powerful capabilities of the self-similarity filtering in image edge denoising and the shock filters in edge sharpening.

To be specific, as illustrated in Figure 1, for a given degraded OCT image, it is first denoised with a collaborative filtering method; then, the denoised image is sharpened by a shock-type filtering for edge and detail enhancement [11]. For dim OCT images, in order to improve the contrast of the retinal image for the detection of tiny lesions, a gamma transformation [22] is first used to enhance the image within proper gray levels.

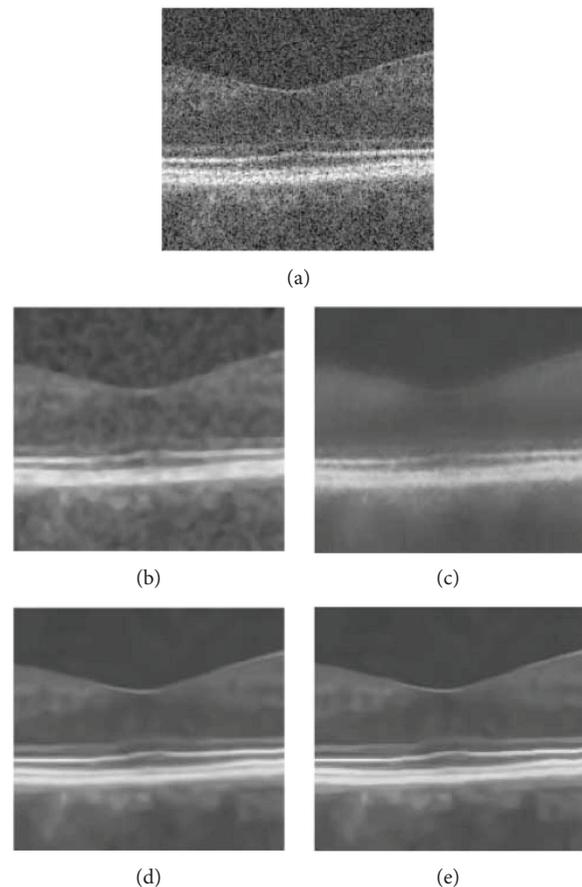


FIGURE 3: Zoomed parts in Figure 2: (a–e) original image, ADSF, NLM, BM3D, and the proposed collaborative shock filtering, respectively.

Different from the BM3D algorithm for Gaussian noise removal [20], the block similarity measure is elaborated in the proposed collaborative filtering of speckle noise following a gamma distribution [26]. Based on a generalized likelihood ratio, the similarity criteria is defined as follows [27]:

$$S(A_i, A_j) = \frac{\text{mean}(A_i \odot A_j)}{\text{mean}((A_i + A_j) \odot (A_i + A_j))}. \quad (5)$$

Here, \odot is the Hadamard product of two image matrixes (blocks) and $\text{mean}(\cdot)$ is the function taking the mean of all elements of matrices. The larger value of S implies that image block A_i is more similar to A_j .

4. Experimental Results

In this section, we perform a lot of enhancement experiments to verify our method on various degraded OCT images. For

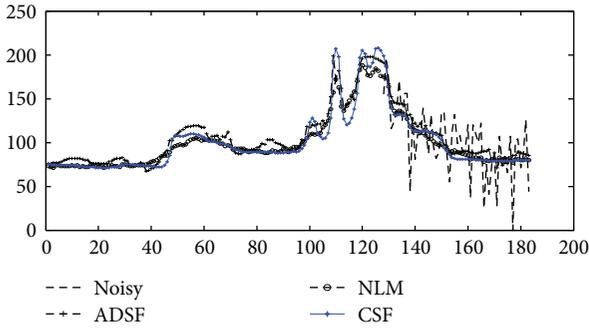
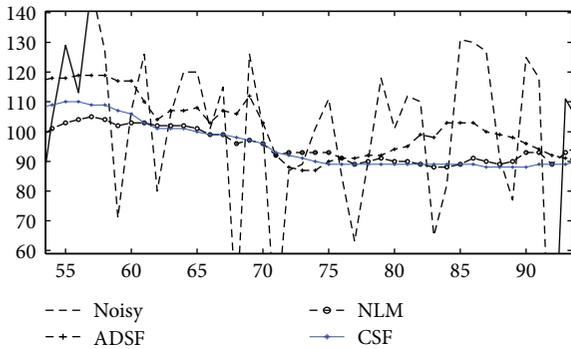
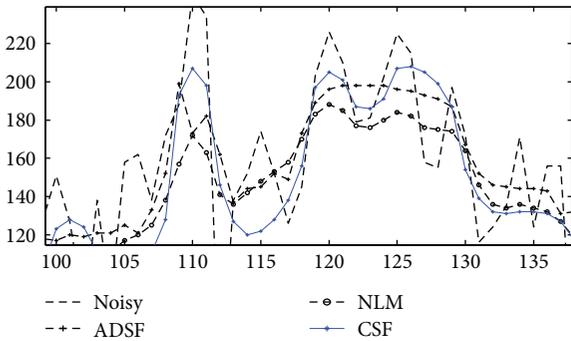


FIGURE 4: Enhancement of noisy OCT image: comparison of profiles (500th column) of original and enhanced images by ADSF, NLM, and the proposed collaborative shock filtering (CSF), respectively.



(a)

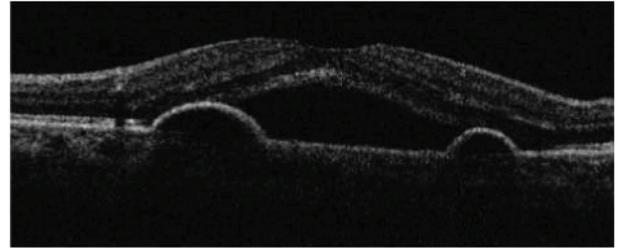


(b)

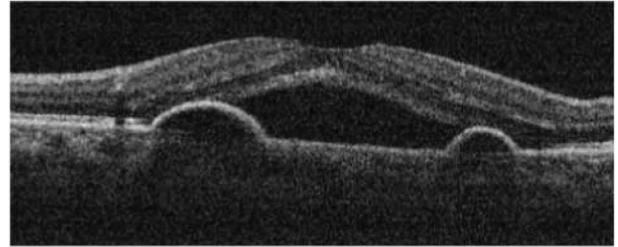
FIGURE 5: Enhancement of noisy OCT image: comparison of profiles of different regions of original and enhanced images by ADSF, NLM, and the proposed collaborative shock filtering (CSF): (a) flat region (line 55–90, 500th column) and (b) edge region (line 105–135, 500th column), respectively.

the above methods, numerical schemes presented in original works are used.

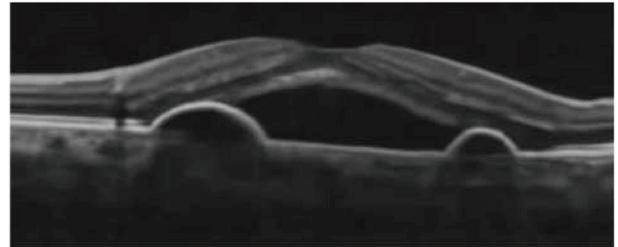
All methods are implemented using the MATLAB programming on OCT images with grayscale ranging from 0 to 1. In the comparison of different methods, parameters in each method are selected such that optimal visual results are obtained. Through three sequential processings of the gamma manipulation ($\gamma = 0.6$), the collaborative filtering ($\sigma = 50$) and the shock filter ($\Delta t = 0.05, n = 10$), where n



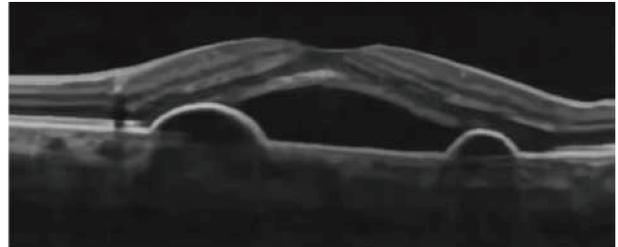
(a)



(b)



(c)



(d)

FIGURE 6: Enhancement of dim noisy OCT image: (a–d) original image, gamma transformation, BM3D, and the proposed collaborative shock filtering, respectively.

and Δt denote the iteration times and the temporal step size, tiny details and layered structures are shown clearly due to image contrast improvement and noise removal. Moreover, our method produces fewer overshoot artifacts while avoiding noise magnification.

In Figure 2, a severely noisy OCT image of size 183×894 is enhanced without image gray adjustment (gamma transformation). Its better effects with clearer layered structures and effective noise removal can be observed more clearly in zoomed parts of results in Figure 3, compared with other methods such as ADSF, NLM, and BM3D. The ADSF method does not efficiently remove speckle noise and produces blurry edges and layered structures with some annoying artifacts in flat regions. The NLM method produces very blurred layered structures with

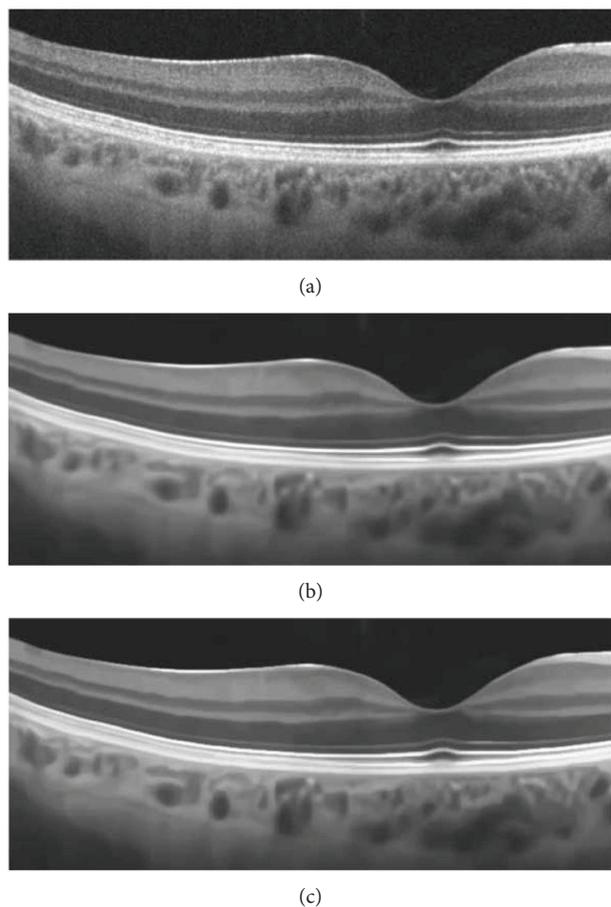


FIGURE 7: Enhancement of noisy OCT image: (a–c): original image, BM3D, and proposed collaborative shock filtering, respectively.

broken nonsmooth linear edges. The BM3D method produces better enhancement results with a few slightly blurry edges and layered structures. Blurry and nonsmooth-layered structures are bad for one to measure the thickness of layered structures to predict and evaluate related disorders in ophthalmology.

In order to observe enhancement effects by these methods more clearly, local profiles (500th column) of different results are shown in Figure 4. One can see that, compared with related methods, the proposed method removes speckle noise more effectively preserving layered structures without producing annoying artifacts, providing a chance to detect and evaluate retinal lesions faithfully by image enhancement. In Figure 5, local profiles of a flat region (line 55–90, 500th column) and an edge region (line 105–135, 500th column) of different results are shown. Superior denoising and edge sharpening by the proposed CSF method can be clearly observed: sharper edges with a smaller edge width are helpful for an accurate measure of layered structures.

In Figure 6, a very dim and noisy OCT image of size 194×477 is enhanced with image gray adjustment. Clearer layered structures can be easily observed with effective noise removal.

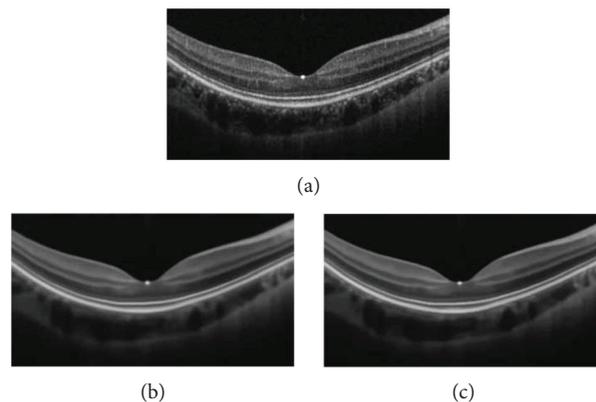


FIGURE 8: Enhancement of dim noisy OCT image (a–c): original image, BM3D, and the proposed collaborative shock filtering, respectively.

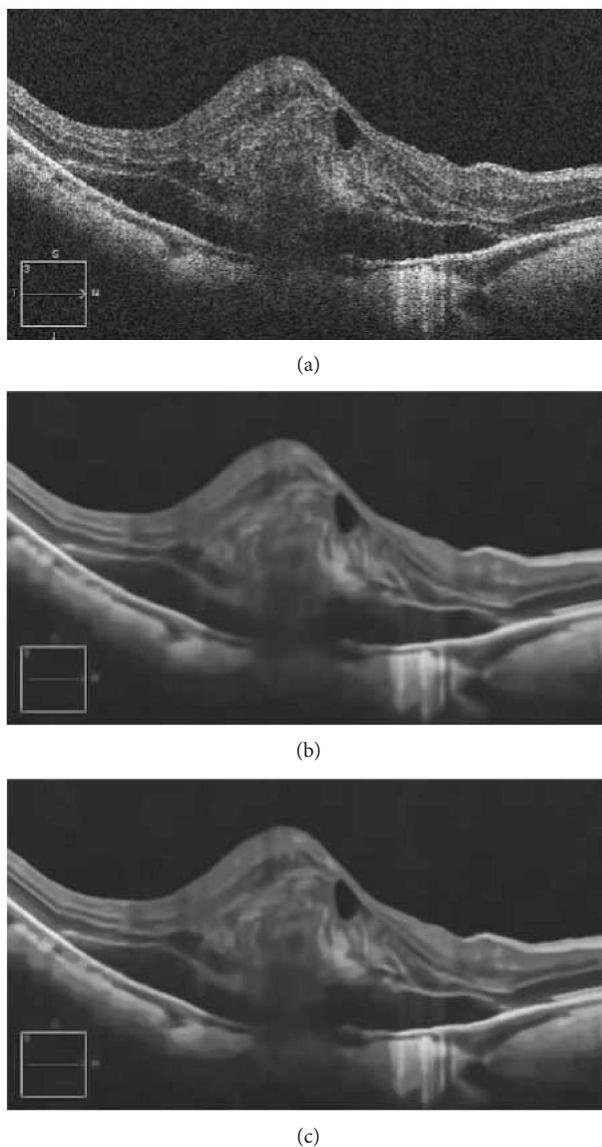


FIGURE 9: Enhancement of dim noisy OCT image: (a–c): original image, BM3D, and the proposed collaborative shock filtering, respectively.

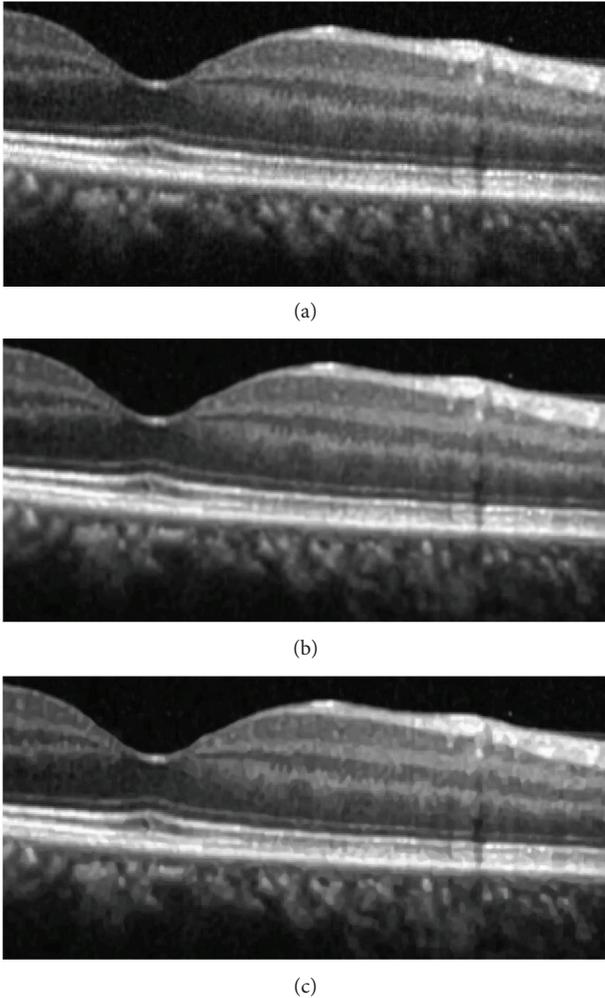


FIGURE 10: Enhancement of blurred noisy OCT image: (a–c) original image, BM3D, the and proposed collaborative shock filtering, respectively.

In Figures 7–9, three much noisy OCT images of sizes (245×586) , (187×341) , and (250×453) are enhanced without image gray adjustment. Clearer layered structures can be easily observed with effective noise removal.

In Figure 10, a blurred noisy OCT image of size 594×1263 is enhanced without image gray adjustment. In this case, more sharpening with the shock filter ($\Delta t = 0.1$, $n = 10$) is used to ensure clearer layered structures and effective noise removal.

Finally, we compare the proposed CSF method with the BM3D method for edge sharpening in OCT image enhancement. In Figure 11, typical results of previous experiments by two methods are shown to carefully observe the enhancement of layered structures in OCT images. Obviously, the CSF method produces sharper edges and clearer layered structures around the retinal fovea. In Figure 12, moreover, it is clearer for one to observe profiles of an edge region (line 105–135, 500th column) in enhancing the first test image in Figure 2: the proposed CSF method produces sharper edges with smaller edge width, which is consistent with the visual observation.

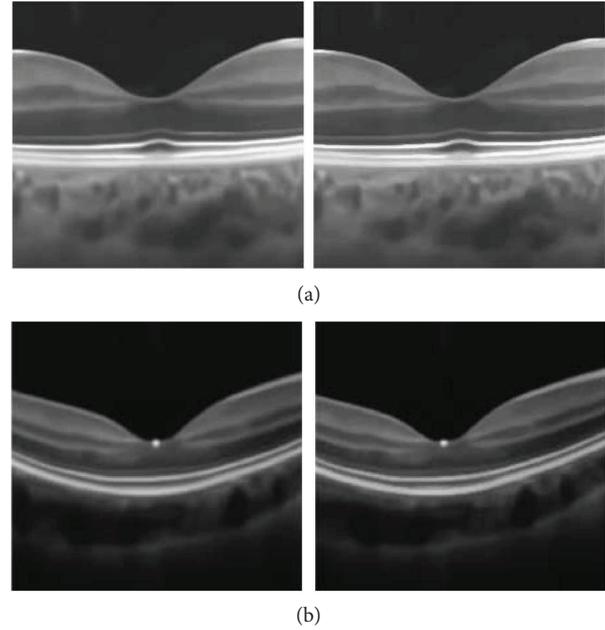


FIGURE 11: Enhancement of noisy OCT images: (a) results by BM3D and (b) results by the proposed collaborative shock filtering, respectively.

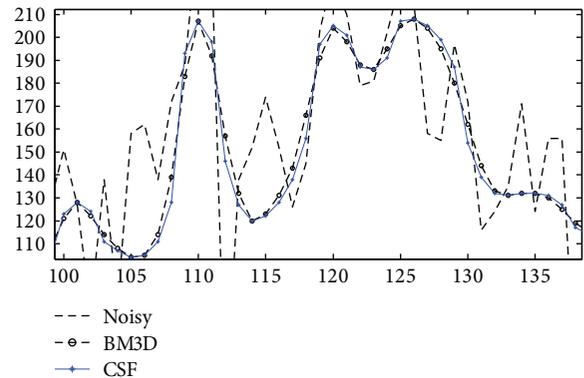


FIGURE 12: Enhancement of noisy OCT image (first test image in Figure 2): comparison of profiles (line 105–135, 500th column) of enhanced images by BM3D and the proposed collaborative shock filtering (CSF), respectively.

5. Conclusions and Future Work

To better enhance image details and layered structures of noisy optical coherence tomography (OCT) image, we propose a collaborative shock filtering for OCT image denoising and enhancement. Noisy OCT image is first denoised with a collaborative filtering method based on a new similarity measure, followed by a sharpening step by a shock-type filtering for edge and detail enhancement. Finally, in order to improve the contrast of dim OCT image, a gamma transformation is used to enhance the image within proper gray levels. Simultaneously, integrating the image smoothing and sharpening of the proposed method obtains better visual results in image enhancement experiments.

In future work, we will develop robust measurement tools for quantitative analysis of vision-related diseases based on this work.

Conflicts of Interest

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

Acknowledgments

The research has been supported in part by the National Natural Science Foundation of China (61272239 and 61671276) and the Science and Technology Development Project of Shandong Province of China (2014GGX101024).

References

- [1] D. Huang, E. A. Swanson, C. P. Lin et al., "Optical coherence tomography," *Science*, vol. 254, no. 5035, pp. 1178–1181, 1991.
- [2] J. M. Schmitt, "Optical coherence tomography (OCT): a review," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 5, no. 4, pp. 1205–1215, 1999.
- [3] J. G. Fujimoto, W. Drexler, J. S. Schuman, and C. K. Hitzenberger, "Optical coherence tomography (OCT) in ophthalmology: introduction," *Optics Express*, vol. 17, no. 5, pp. 3978–3979, 2009.
- [4] J. M. Schmitt, S. Xiang, and K. M. Yung, "Speckle in optical coherence tomography," *Journal of Biomedical Optics*, vol. 4, no. 1, pp. 95–105, 1999.
- [5] H. M. Salinas and D. C. Fernández, "Comparison of PDE-based nonlinear diffusion approaches for image enhancement and denoising in optical coherence tomography," *IEEE Transactions on Medical Imaging*, vol. 26, no. 6, pp. 761–771, 2007.
- [6] P. Sudeep, S. I. Niwas, P. Palanisamy et al., "Enhancement and bias removal of optical coherence tomography images: an iterative approach with adaptive bilateral filtering," *Computers in Biology and Medicine*, vol. 71, pp. 97–107, 2016.
- [7] T. Chan and J. Shen, *Image Processing and Analysis: Variational, PDE, Wavelet, and Stochastic Methods*, the Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2005.
- [8] G. Aubert and P. Kornprobst, *Mathematical Problems in Image Processing: Partial Differential Equations and the Calculus of Variations*, Springer-Verlag, New York, NY, USA, 2nd edition, 2006.
- [9] L. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D: Nonlinear Phenomena*, vol. 60, no. 1–4, pp. 259–268, 1992.
- [10] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 7, pp. 629–639, 1990.
- [11] S. Osher and L. Rudin, "Feature-oriented image enhancement using shock filters," *SIAM Journal on Numerical Analysis*, vol. 27, no. 4, pp. 919–940, 1990.
- [12] L. Alvarez and L. Mazorra, "Signal and image restoration using shock filters and anisotropic diffusion," *SIAM Journal on Numerical Analysis*, vol. 31, no. 2, pp. 590–605, 1994.
- [13] Z. Jian, Z. Yu, L. Yu, B. Rao, Z. Chen, and B. J. Tromberg, "Speckle attenuation in optical coherence tomography by curvelet shrinkage," *Optics Letters*, vol. 34, no. 10, pp. 1516–1518, 2009.
- [14] A. Wong, A. Mishra, K. Bizheva, and D. A. Clausi, "General bayesian estimation for speckle noise reduction in optical coherence tomography retinal imagery," *Optics Express*, vol. 18, no. 8, pp. 8338–8352, 2010.
- [15] L. Fang, S. Li, Q. Nie, J. A. Izatt, C. A. Toth, and S. Farsiu, "Sparsity based denoising of spectral domain optical coherence tomography images," *Biomedical Optics Express*, vol. 3, no. 5, pp. 927–942, 2012.
- [16] H. Yu, J. Gao, and A. Li, "Probability-based non-local means filter for speckle noise suppression in optical coherence tomography images," *Optics Letters*, vol. 41, no. 5, pp. 994–997, 2016.
- [17] D. Thapa, K. Raahemifar, and V. Lakshminarayanan, "Reduction of speckle noise from optical coherence tomography images using multi-frame weighted nuclear norm minimization method," *Journal of Modern Optics*, vol. 62, no. 21, pp. 1856–1864, 2015.
- [18] A. Buades, B. Coll, and J. Morel, "A review of image denoising algorithms, with a new one," *Multiscale Modeling & Simulation*, vol. 4, no. 2, pp. 490–530, 2005.
- [19] A. Buades, B. Coll, and J.-M. Morel, "Nonlocal image and movie denoising," *International Journal of Computer Vision*, vol. 76, no. 2, pp. 123–139, 2008.
- [20] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3D transform-domain collaborative filtering," *IEEE Transactions on Image Processing*, vol. 16, no. 8, pp. 2080–2095, 2007.
- [21] V. Katkovnik, A. Foi, K. Egiazarian, and J. Astola, "From local kernel to nonlocal multiple-model image denoising," *International Journal of Computer Vision*, vol. 86, no. 1, pp. 1–32, 2010.
- [22] R. Gonzalez and R. Woods, *Digital Image Processing*, Prentice-Hall, New York, NY, USA, 2nd edition, 2004.
- [23] L. Wang, G. Liu, S. Fu, L. Xu, K. Zhao, and C. Zhang, "Retinal image enhancement using robust inverse diffusion equation and self-similarity filtering," *PLoS One*, vol. 11, no. 7, article e0158480, 2016.
- [24] L. Rudin, *Shock Filters*, Tech. Rep., Defence Advanced Research Projects Agency (DARPA), Rockwell International Science Center, United States, Thousand Oaks, 1984.
- [25] S. Fu, C. Zhang, and X. Tai, "Image denoising and deblurring: non-convex regularization, inverse diffusion and shock filter," *Science China Information Sciences*, vol. 54, no. 6, pp. 1184–1198, 2011.
- [26] D. A. Jesus and D. R. Iskander, "Assessment of corneal properties based on statistical modeling of OCT speckle," *Biomedical Optics Express*, vol. 8, no. 1, pp. 162–176, 2017.
- [27] C.-A. Deledalle, L. Denis, and F. Tupin, "How to compare noisy patches? Patch similarity beyond Gaussian noise," *International Journal of Computer Vision*, vol. 99, no. 1, pp. 86–102, 2012.

Research Article

Medical Imaging Lesion Detection Based on Unified Gravitational Fuzzy Clustering

Jean Marie Vianney Kinani,¹ Alberto Jorge Rosales Silva,² Francisco Gallegos Funes,² Dante Mújica Vargas,³ Eduardo Ramos Díaz,⁴ and Alfonso Arellano⁵

¹Instituto Tecnológico Superior de Huichapan, Domicilio Conocido S/N, Col. El Saucillo, 42411 Huichapan, HGO, Mexico

²Instituto Politécnico Nacional de México, Avenida IPN s/n, Edificio Z, Acceso 3, 3er piso, SEPI-Electrónica, Col. Lindavista, 07738 Ciudad de México, Mexico

³Centro Nacional de Investigación y Desarrollo Tecnológico, Interior Internado Palmira S/N, Palmira, 62490 Cuernavaca, MOR, Mexico

⁴Universidad Autónoma de la Ciudad de México, Calle Prolongación San Isidro 151, Iztapalapa, San Lorenzo Tezonco, 09790 Ciudad de México, Mexico

⁵Instituto Nacional de Neurología y Neurocirugía, Av. Insurgentes Sur 3877, Col. La Farma, 14269 Ciudad de México, Mexico

Correspondence should be addressed to Alberto Jorge Rosales Silva; arosaless23@gmail.com

Received 5 April 2017; Accepted 31 July 2017; Published 12 October 2017

Academic Editor: Shujun Fu

Copyright © 2017 Jean Marie Vianney Kinani 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.

We develop a swift, robust, and practical tool for detecting brain lesions with minimal user intervention to assist clinicians and researchers in the diagnosis process, radiosurgery planning, and assessment of the patient's response to the therapy. We propose a unified gravitational fuzzy clustering-based segmentation algorithm, which integrates the Newtonian concept of gravity into fuzzy clustering. We first perform fuzzy rule-based image enhancement on our database which is comprised of T1/T2 weighted magnetic resonance (MR) and fluid-attenuated inversion recovery (FLAIR) images to facilitate a smoother segmentation. The scalar output obtained is fed into a gravitational fuzzy clustering algorithm, which separates healthy structures from the unhealthy. Finally, the lesion contour is automatically outlined through the initialization-free level set evolution method. An advantage of this lesion detection algorithm is its precision and its simultaneous use of features computed from the intensity properties of the MR scan in a cascading pattern, which makes the computation fast, robust, and self-contained. Furthermore, we validate our algorithm with large-scale experiments using clinical and synthetic brain lesion datasets. As a result, an 84%–93% overlap performance is obtained, with an emphasis on robustness with respect to different and heterogeneous types of lesion and a swift computation time.

1. Introduction

In broad terms, “brain lesion” can be defined as an abnormal damage or change in the brain tissue; this can be caused by injury, infection, exposure to certain chemicals, problems with the immune system, and many other factors. Due to their location, at the center of thought, physical function, and emotion, brain lesions are difficult to diagnose and treat. However, thanks to recent advances

in magnetic resonance imaging and computing, brain lesion diagnosis has made a great leap. The algorithm presented in this paper could make a huge impact in both diagnosing and monitoring processes. It can detect damaged or unhealthy regions on MRI scans and delineates them with high precision. This facilitates the decision making and the planning of surgical removal of the lesion (if necessary and possible). It also allows one to apply spatially localized radiotherapy, for example, Cyberknife and

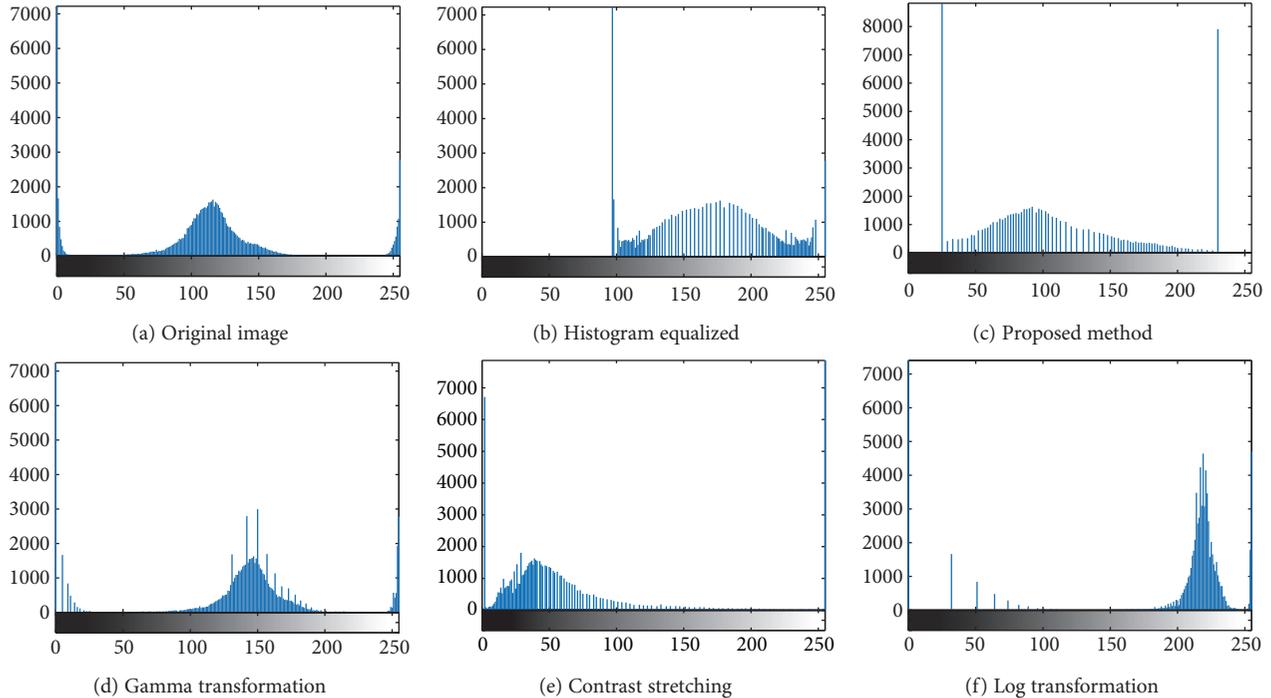


FIGURE 1: Corresponding histograms.

iMRT [1–3], which in current clinical practice is usually done manually on both contrast-enhanced T1-weighted or FLAIR images. Most of the medical imaging modalities give out images of gray scale intensities, including MRIs. These images are subject to noise, artifacts, and poor resolution and contrast due to instrument and reconstruction algorithm limitations or even patient movement. Consequently, auto-detection becomes so challenging that the algorithm advantages and disadvantages may vary depending on the properties of the image under examination. Therefore, due to the image deterioration factors mentioned, it is hard to develop a standard approach capable of working with all types of MR images [4]. As a result, trade-offs have always been present in computer-aided diagnosis systems. Yet, comparing our unified fuzzy-hard clustering-based system against classical approaches based on methods like classifier, region growing, neural networks, deformable models, and so forth, a big advantage of our approach is recognized especially when dealing with the deteriorating factors mentioned [5]. In Section 2, we present our framework for lesion detection. Starting with a brief background of fuzzy sets, we will show how this theory can be exploited to enhance the inhomogeneous MR images by adapting appropriate fuzzy rules [6]. Next, we will give an outline of the segmentation method used, which is based on a novel gravitational fuzzy clustering concept and level set evolution that defines the final location, shape, and size of the lesion. Experiments and evaluation studies that were carried out on both synthetic and expert-segmented data sets are presented in Section 3. We will then finish the paper with a discussion and a conclusion in Section 4.

2. Fuzzy and Gravitational Methods Proposed

2.1. The Proposed Fuzzy Set-Based MR Image Enhancement. Accurate diagnosis of brain lesions depends upon the quality of the MR scan; in particular, on the visibility of small, low-contrast objects within the brain image. Unfortunately, the contrast between these objects is often so low that the detection of some abnormalities becomes difficult, especially when dealing with dense tissues. To deal with this issue, contrast enhancement is normally carried out on original images before the detection process can take place.

A detailed investigation on image enhancement carried out by Bankman [7], González and Woods [8], Shih [9], and Russ [10] shows that classical methods like negative transformation, Log, Gamma, contrast stretching, or histogram-based transformations work effectively in enhancing ordinary images; however, when applied to MR images, they bring about tradeoffs between the enhancement and image detail preservation due to the loss of some basic characteristics in the original image histogram, as it is pronounced in Figure 1.

After an intensive study of MR image enhancement techniques, we came up with a much better method based on adapting fuzzy rules. This technique did correct the above-mentioned drawback; it has achieved this by enhancing the contrast of features of interest and improving the visibility of diagnostic details without creating artifacts or losing image details as a whole. To get a better understanding of this technique, we have to go back to the fuzzy set theory.

Normally, in set theory, we are used to the so called crisp sets whose membership can only be true or false in the traditional sense of bivalued Boolean logic. This classical

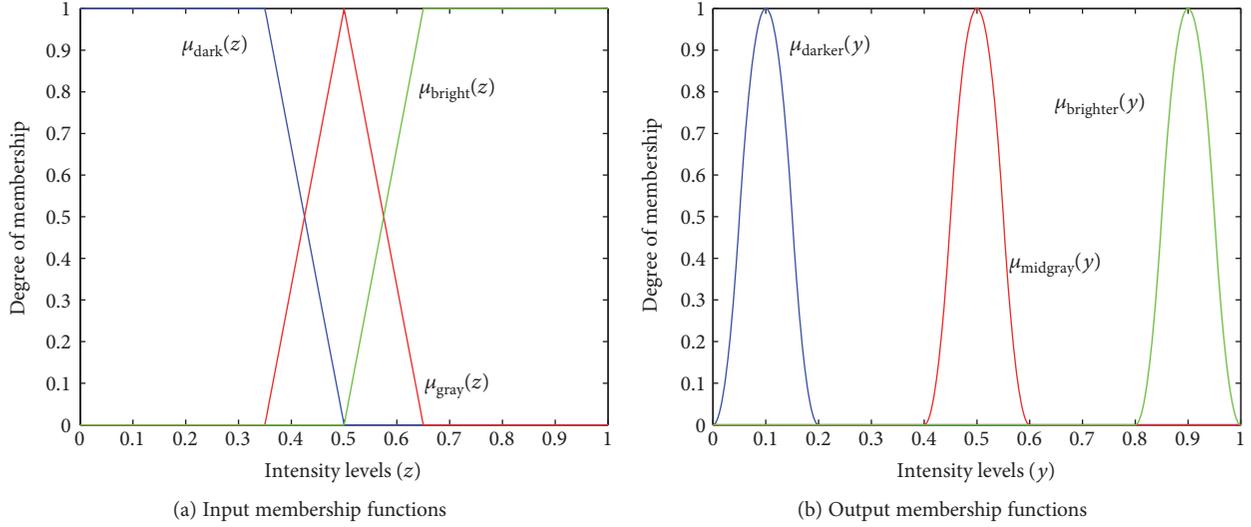


FIGURE 2: Membership functions.

set theory has limited use in practical applications due to its lack of flexibility [11]. Thus, Professor Zadeh proposed a more successful theory of fuzzy sets that introduced the idea of partial memberships described by membership functions [11]. Suppose Z is a set of elements, and each element is denoted by z , that is, $z \in Z$. A fuzzy subset A in Z is characterized by a membership function, $\mu_A(z)$, so

$$A = \{z, \mu_A(z) | z \in Z\}. \quad (1)$$

When the variables are continuous, A in this equation can have an infinite number of elements. When the values of z are discrete, the elements of A can be shown explicitly. The formalization of the problem into fuzzy rules consists in finding a way to increase the contrast of certain tissues in the brain, while leaving other tissues quasi-untouched to accentuate the difference between these tissues, so the following fuzzy rule is proposed:

Fuzzy Rule 1: *IF* a pixel is dark, *THEN* make it darker; *OR*,

Fuzzy Rule 2: *IF* a pixel is gray, *THEN* make it gray; *OR*,

Fuzzy Rule 3: *IF* a pixel is bright, *THEN* make it brighter.

(2)

The first antecedent of the fuzzy rule will seek to relate the fuzzy set dark to the set darker (the two sets are represented in blue in Figure 2), and the consequence is achieved using a fuzzy *AND* operation, implemented through a min operation [11], as shown in

$$\mu_1(z, y) = \min\{\mu_{\text{dark}}(z), \mu_{\text{darker}}(y)\}, \quad (3)$$

where z and y are scalar values representing the intensity levels of the pixels in the input and output fuzzy sets, respectively. z_0 denotes a specific intensity level in the interval near to the visible black color spectrum. The degree of membership of the dark set component in response to this input is a scalar value $\mu_{\text{dark}}(z_0)$. We find the output corresponding

to the first part of the fuzzy rule, and this specific input, by performing the *AND* operation between $\mu_{\text{dark}}(z_0)$ and the general result $\mu_1(z, y)$, evaluated also at z_0 , so therefore,

$$Q_1(y) = \min\{\mu_{\text{dark}}(z_0), \mu_1(z_0, y)\}, \quad (4)$$

where $Q_1(y)$ denotes the fuzzy output value due to the first part of the fuzzy rule and a specific input z_0 . Using the same line of reasoning, we obtain the fuzzy responses due to the other antecedents and consequences along with the input z_0 , which are as follows:

$$\begin{aligned} Q_2(y) &= \min\{\mu_{\text{gray}}(z_0), \mu_2(z_0, y)\}, \\ Q_3(y) &= \min\{\mu_{\text{bright}}(z_0), \mu_3(z_0, y)\}. \end{aligned} \quad (5)$$

These equations represent the result of the implication process. We should keep in mind that each of these responses is given in a fuzzy set, even though the input is a scalar value.

The application of aggregation method to the above fuzzy sets to obtain the overall response generated by the rule is carried out, and this is achieved through an *OR* operation as suggested by the proposed fuzzy rule base, that is,

$$Q(y) = \max_r \left\{ \min_s \{\mu_s(z_0), \mu_r(z_0, y)\} \right\}, \quad (6)$$

$r = \{1, 2, 3\}$ being the number of fuzzy outputs and $s = \{\text{dark}, \text{gray}, \text{bright}\}$. We can see that the overall response is the union of the three individual fuzzy sets. And this is the complete output corresponding to a specific input. But we are still dealing with a fuzzy set, so the last step is to obtain a crisp output y_0 . This is achieved through defuzzifying the final output fuzzy set Q obtained above; that is, obtaining a crisp,

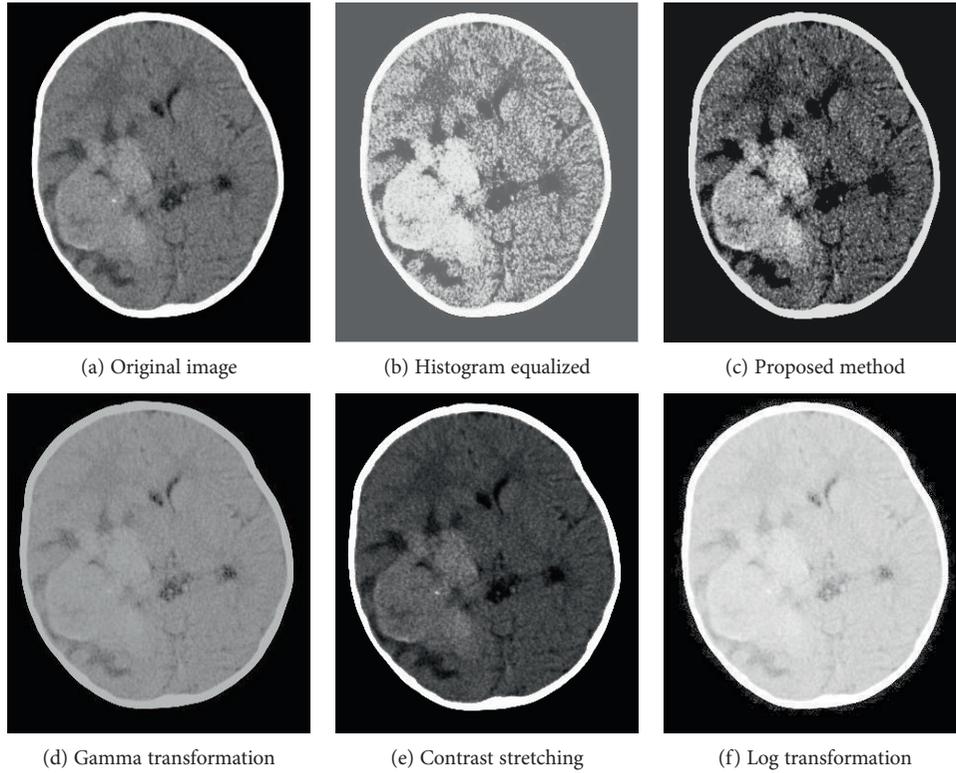


FIGURE 3: Comparison of different enhancement methods.

scalar output brought in through computing the center of gravity of Q , that is,

$$y_0 = \frac{\sum_{y=1}^p yQ(y)}{\sum_{y=1}^p Q(y)}, \quad (7)$$

y_0 being the crisp output, p being the number of all possible values that $Q(y)$ in (6) may have, and y being a scalar value representing the intensity levels of the pixels in output fuzzy sets.

In this way, we are able to achieve dynamic range expansion of the contrast in an efficient way using simple computation operations to conditionate the image to the future processes.

2.1.1. Experimental Results of the Contrast Enhancement. Our contrast enhancement experiment was carried out using different images from different MRIs, and the proposed fuzzy set-based technique proved to be the most effective and this can be proven by the histogram shapes obtained in Figure 1. By examining the outcome of our method and comparing it to that of other classical methods, a clear difference is noted.

Let us now examine our results. Figure 3(a) shows an image whose intensities span a narrow range of the gray scale, as the histogram in Figure 1(a) reveals. The next result is an image with low contrast; Figure 3(b) is the result of using the histogram equalization to increase image contrast. As shown by the histogram in Figure 1(b), the gray scale was a bit spread out, but shifted to the right, and it completely lost

the shape of the histogram of the original image. The result was an image with an overexposed appearance. We can see that it would be quite difficult to distinguish the healthy from the unhealthy tissues due to some gray details that are lost. As we can see in Figure 3(c), the result of the proposed method is an image having increased contrast and a rich tonality. The reason for this improvement can be explained by examining the histogram in Figure 1(c). Unlike the histograms produced by other techniques, this histogram has kept the same basic characteristics of the histogram of the original image [8]. And the spread of the gray scale occurred in all directions and this applies to all MR images tested. As for the gamma and Log transformations, we can see that a lot of details were lost, as proved by their histograms. Contrast stretching was a bit more productive, but its histogram lacked stretching and this resulted in a negligible enhancement.

2.2. Lesion Detection Process through Segmentation. Generally, the principal goal of segmentation is to partition an image into regions (also called classes or subsets) that are homogeneous with respect to one or more characteristics or features [11]. This is very important in medical imaging, since it allows feature extraction, image measurement, and display. Most importantly, it permits the classification of image pixels into anatomical or pathological regions such as lesion and tissue deformities, amongst others.

2.2.1. Unified Gravitational Fuzzy Clustering (UGFC). Some ground-breaking MR image segmentation approaches have been developed by the most prominent researchers and run on modern processors. These include the work by Prastawa

et al. [12] who, with their automatic, multimodal, atlas-based method, have reported 86.7% average overlap on a small dataset of three patients with an average 1.5 h processing time. In a recent study, Hamamci et al. [13] reported an 80%–90% overlap performance, with their method named “Tumor-cut,” which is based on the cellular automata (CA) algorithm. A serious drawback with this method, however, is that it requires diameter drawing initialization, which raises its computation time to about 16 minutes and prevents it from being fully automatic. Menze et al. [14] reported 60% average overlap on 25 glioma patients with their method which is based on the discriminative random decision forests framework. Gooya et al. [15] reported 74.5% average overlap on 15 glioma patients with about 6–14 hours of processing time, with their methods which are based on EM algorithm. Geremia et al. [16] adopted a discriminative random decision forest framework that gave good results in high-quality MRI with low noise level and high resolution. Another interesting work was by Liu et al. [17] who used the classical fuzzy clustering-based method and reported a 95.6% average overlap on a well-performing five-patient dataset of FLAIR images. The latter method needed intensive user interaction and correction at least 8.4 minutes per patient. Most recently, Shen et al. [18] extended the Fuzzy C-Means (FCM) approach by introducing an additional term describing the distance between the fuzzy membership and the prior tissue probability maps; they used a simulated image dataset, and an overlap varying from 34% to 79% was reported, depending on the signal reduction.

In this paper, we are proposing a novel segmentation method based on a combined hard and fuzzy clustering framework. This method adopts the Newtonian gravity concept from a clustering perspective in order to hard-cluster image pixels that otherwise could unnecessarily be assigned membership to clusters that they categorically do not belong to. However, this method fuzzy-clusters those pixels that are located in controversial areas in order to optimize the partial volume effect handling. The result is a well-defined region of interest (ROI) that is made up of unhealthy tissues on the MR image under consideration.

Generally, Newton’s gravity law can be formulated as follows:

$$F = l \frac{m_1 \cdot m_2}{d^2}. \quad (8)$$

F denotes the gravity force between object 1 with mass m_1 and object 2 whose mass is m_2 , d represents the distance between the two objects, and l is a coefficient that takes the place of Newton’s constant; we set $l = 2$ for computational convenience. To apply this law effectively, we make the following assumptions:

- (i) The quality of each pixel is 1.
- (ii) $m_i^{(t)}$ pixels have been clustered into cluster i at time t .
- (iii) Each pixel belonging to a cluster has the same potential, that is, equal preference.

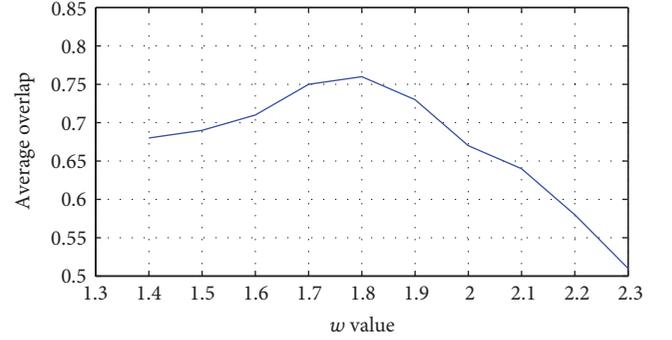


FIGURE 4: The effect of fuzzifier w on the segmentation performance.

All pixels in a cluster flow into an object whose mass is equal to the number of these pixels. Based on the above assumptions, the gravity force $F^{(t)}$ between k th pixel and i th cluster will be

$$F^{(t)} = l \frac{1 \times m_i^{(t)}}{d^2} = \frac{2m_i^t}{d^2} = \frac{2m_i^t}{(x_k - v_i)^2}. \quad (9)$$

x_k is the k th pixel and v_i is the i th cluster center.

Now, we will define the gravitational clustering objective function (J_{GC}) that should be minimized:

$$J_{GC} = \sum_{i=1}^c \sum_{x_k \in S_i} d^2[x_k, v_i], \quad (10)$$

where S_i denotes the set of pixels clustered to i th cluster and c stands for the total number of clusters in consideration. Now, as stated in [19], the standard Fuzzy C-Means objective function (J_{CM}) is defined as

$$J_{CM} = \sum_{i=1}^c \sum_{j=1}^N (\mu_{ij})^w (d_{ij})^2. \quad (11)$$

N is the number of pixels, and μ_{ij} is the membership of j th pixel in the i th cluster and is defined as

$$\mu_{ij}^{(t+1)} = \left[\sum_{k=1}^c \left(\frac{d_{ij}^{(t)}}{d_{kj}^{(t)}} \right)^{2/(w-1)} \right]^{-1}, \text{ with } \sum_{i=1}^c \mu_{ij} = 1, \quad (12)$$

where t is the iteration number, w being the fuzzifier, and $w \in (1.4, 2.6)$, as was recently proven by Ozkan and Turksen [20] in their study based on Taylor expansion analysis of the membership value calculation function. In our experiment, w was chosen to be 1.7 and 1.8 since these two numbers gave the best clustering. For our three datasets, the effect of this parameter on the segmentation performance in terms of Dice overlap measure is plotted in Figure 4. Now, we will define an integrated gravitational fuzzy clustering objective function (J_{GFC}):

$$J_{GFC} = J_{GC} \cdot J_{CM}. \quad (13)$$

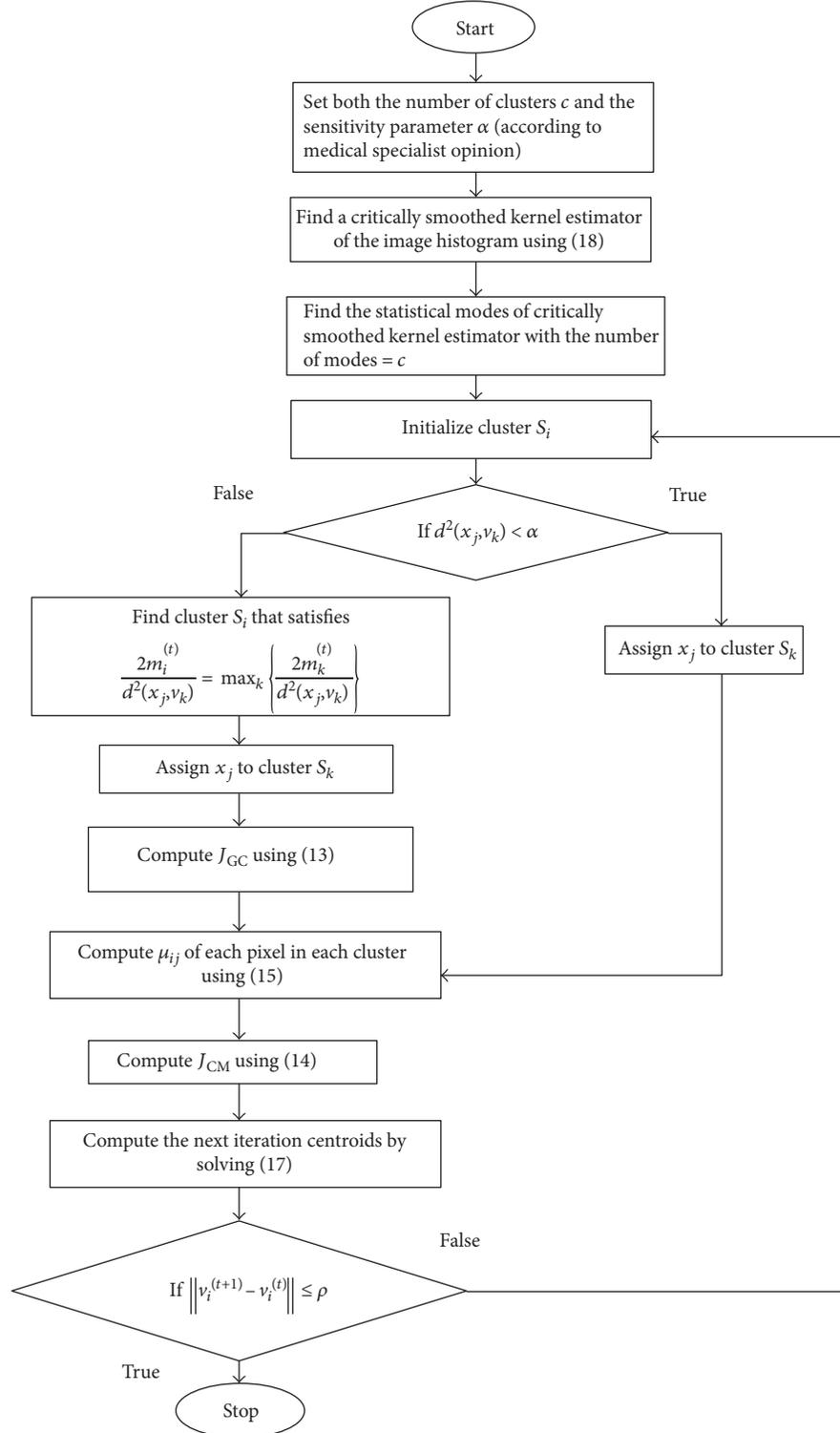


FIGURE 5: UGFC algorithm flow diagram.

Taking the first derivatives of J_{GFC} with respect to v_i and setting them to zero results in the following linear systems:

$$\frac{\partial J_{\text{GFC}}}{\partial v_i} = J_{\text{CM}} \cdot \frac{\partial J_{\text{GC}}}{\partial v_i} + J_{\text{GC}} \cdot \frac{\partial J_{\text{CM}}}{\partial v_i} = 0. \quad (14)$$

Based on these equations, the integrated gravitational fuzzy clustering algorithm is structured as indicated in Figure 5. ρ being the convergence criterion, in our experiments, it was set to 0.01. It should be noted that the kernel estimator of the image histogram which was used in defining

initial centroids, that is, $v^{(0)} = \{v_1^{(0)}, v_2^{(0)}, \dots, v_C^{(0)}\}^t$, at $t = 0$, is defined as

$$\hat{p}(x) = \frac{1}{Nh} \sum_{j=1}^N \psi\left(\frac{x - x_j}{h}\right), \quad (15)$$

where ψ is the kernel function—a Gaussian function with zero mean and variance equal to 1; N is the total number of pixels; x represents the intensity levels; and $h \in (0, 50)$ is the bandwidth. Moreover, the critical $h(h_{\text{crit}})$ is defined as the minimum value of h such that \hat{p} has c modes that correspond to the number of clusters.

2.2.2. Level Set Evolution on a Constructed Region of Interest. The contour definition of the constructed region of interest (ROI) which is based on the level set evolution without reinitialization constitutes an important part of our proposed method. This is because the clinical expert segmentation, particularly in neuroimaging, mainly outlines the edges of the ROI using manual contouring, for either surgery planning, radiotherapy, or treatment response analysis [21, 22]. A variational formulation was proposed which consists of an internal energy term that penalizes the deviation of the level set function from a signed distance function, along with an external energy term that drives the motion of the zero-level set toward the ROI boundaries set by the GFC algorithm. This approach had the following advantages over the classical level set formulation [23, 24]:

- (i) The contours represented by the level set function could break or merge naturally during the evolution, and the topological changes are thus automatically handled.
- (ii) The level set function always remains a function on a fixed grid, which permits efficient numerical schemes (a finite difference scheme in our case).
- (iii) Due to the internal energy, the level set function (LSF) is naturally and automatically kept as an approximate signed distance function during the evolution. Consequently, reinitialization is avoided.

Traditionally, if Ω_ϕ is a subset of the Euclidean space with a smooth boundary, then the signed distance function (SDF) of this subset is differentiable practically everywhere, and its gradient satisfies the Eikonal equation [25], that is,

$$|\nabla\phi| = 1. \quad (16)$$

So, any function ϕ satisfying this property is a SDF plus a constant. Now, we propose the following integral:

$$P(\phi) = \frac{1}{2} \int_{\Omega} (|\nabla\phi| - 1)^2 dx dy, \quad (17)$$

as a metric that defines how close ϕ is to a SDF in $\Omega \subset R^2$ and we call it “internal energy.” Having $P(\phi)$ on our disposal, we then propose the following variational formulation:

$$\varepsilon(\phi) = \beta P(\phi) + \varepsilon_m(\phi), \quad (18)$$

$\beta \in [0.04, 0.1]$ is a parameter controlling the effect of penalizing the deviation of ϕ from a SDF, and $\varepsilon_m(\phi)$ is the external energy that drives the motion of zero level curve of ϕ and depends upon the image data.

Now, we consider $\partial\varepsilon/\partial\phi$ as being the Gateaux derivative of ε [26], and the evolution equation,

$$\frac{\partial\phi}{\partial t} = -\frac{\partial\varepsilon}{\partial\phi}, \quad (19)$$

becomes the gradient flow that minimizes ε . Let I be the image generated by the GFC algorithm and g the edge indicator function that regularizes $\varepsilon(\phi)$ in order to stop level set evolution near the optimal solution [27]. The latter is defined as

$$g = \frac{1}{1 + |\nabla G_\sigma * I|^2}, \quad (20)$$

where G_σ is the Gaussian kernel with standard deviation σ . We then define external energy for $\phi(x, y)$ as

$$\varepsilon_{g,\lambda,\nu}(\phi) = \lambda L_g(\phi) + \nu A_g(\phi), \quad (21)$$

with

$$L_g(\phi) = \int_{\Omega} g\delta(\phi)|\nabla\phi| dx dy, \quad (22)$$

and

$$A_g(\phi) = \int_{\Omega} gH(-\phi) dx dy, \quad (23)$$

where $\lambda \in [2, 6]$ and $\nu \in [1, 3.5]$ are constants, δ is the univariate Dirac function, and H is the Heaviside function. So, the total energy becomes

$$\varepsilon(\phi) = \beta P(\phi) + \varepsilon_{g,\lambda,\nu}(\phi). \quad (24)$$

Now, to understand the geometrical meaning of $L_g(\phi)$, suppose that the zero level set of ϕ is represented by a differentiable parameterized curve $K(\tau)$ with $\tau \in [0, 1]$. Then, according to Vemuri and Chen [23], $L_g(\phi)$ computes the length of the zero level curve of ϕ in the conformal metric $ds = g(K(\tau))|K'(\tau)|d\tau$. Note that when function g is a constant one, $A_g(\phi)$ becomes the area of the region $\Omega_\phi = \{(x, y) | \phi(x, y) < 0\}$ [28]. ν of A_g should be positive if the initial contours are placed outside the ROI and negative when they are placed inside, to speed up the contraction or the expansion, respectively. By calculus of variations, the Gateaux derivative of $\varepsilon(\phi)$ can be written as [26]

$$\frac{\partial\varepsilon}{\partial\phi} = -\beta \left[\Delta\phi - \text{div} \left(\frac{\nabla\phi}{|\nabla\phi|} \right) \right] - \lambda\delta(\phi) \text{div} \left(g \frac{\nabla\phi}{|\nabla\phi|} \right) - \nu g\delta(\phi). \quad (25)$$

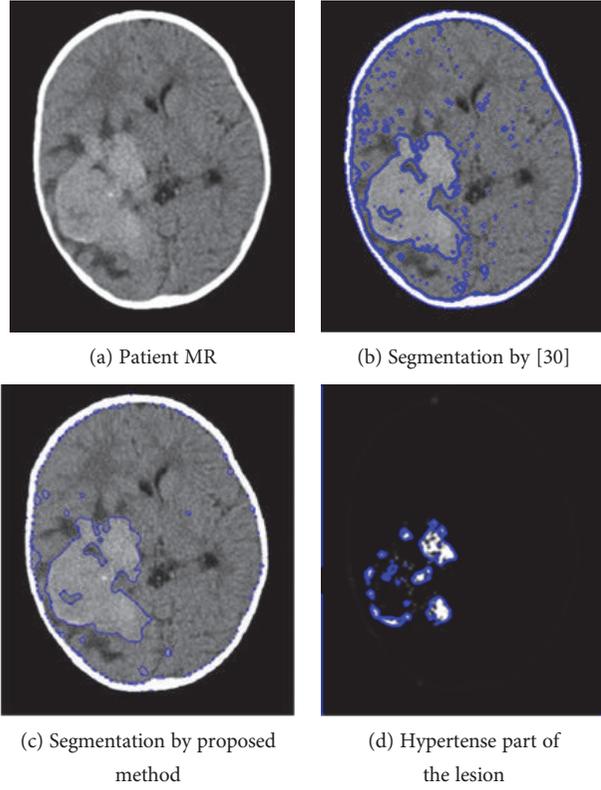


FIGURE 6: Lesion contour calculation.

Therefore, the function ϕ that minimizes this function satisfies the Euler-Lagrange equation:

$$\frac{\partial \varepsilon}{\partial \phi} = 0. \quad (26)$$

So, the gradient flow that minimizes ε is

$$\frac{\partial \phi}{\partial t} = -\beta \left[\Delta \phi - \operatorname{div} \left(\frac{\nabla \phi}{|\nabla \phi|} \right) \right] + \lambda \delta(\phi) \operatorname{div} \left(g \frac{\nabla \phi}{|\nabla \phi|} \right) - \nu g \delta(\phi), \quad (27)$$

which is the evolution equation of the level set function in the proposed algorithm. To explain the effect of the first term in the right-hand side of equation above, that is, $\beta P(\phi)$ which is the internal energy, we notice that the gradient flow:

$$\Delta \phi - \operatorname{div} \left(\frac{\nabla \phi}{|\nabla \phi|} \right) = \operatorname{div} \left[\left(1 - \frac{1}{|\nabla \phi|} \right) \nabla \phi \right], \quad (28)$$

has the factor $(1 - (1/|\nabla \phi|))$ as its diffusion rate. If $|\nabla \phi| > 1$, the diffusion rate is positive and the effect of this term is the usual diffusion, that is, making ϕ more even and therefore reduce the gradient $|\nabla \phi|$. If $|\nabla \phi| < 1$, then the term has effect of reverse diffusion and therefore increases the gradient [29].

3. Evaluation and Experimental Results

Quantitative and qualitative validation studies of the developed method were conducted over three different datasets. These sets comprise multimodal MR images (T1, T1Gd, T2,

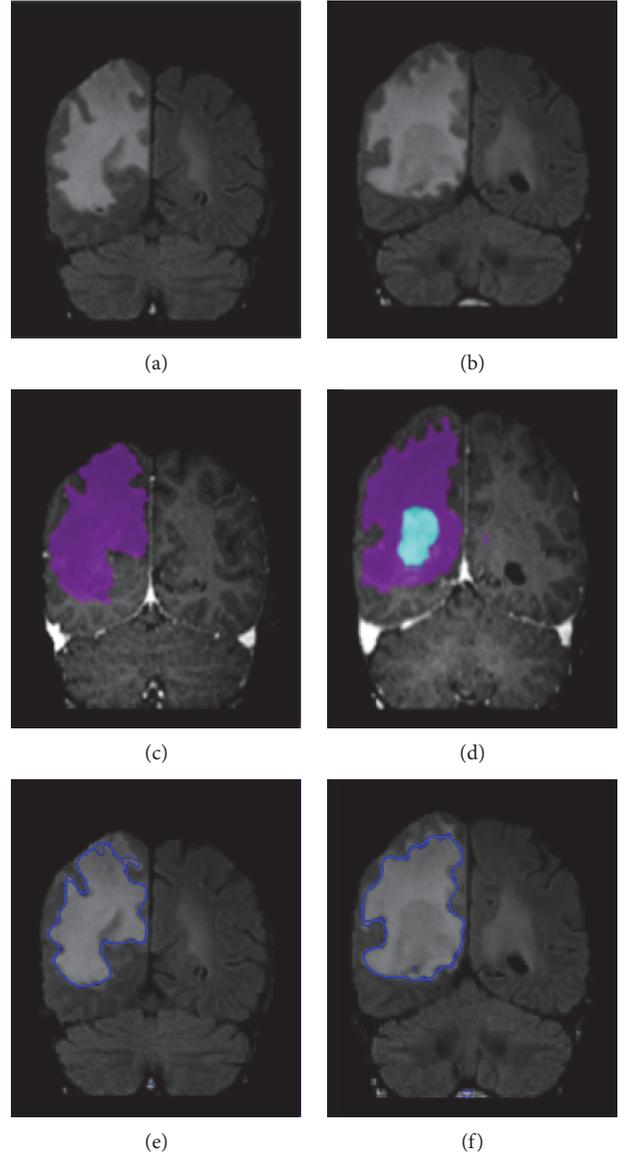


FIGURE 7: FLAIR images (a, d), T1-Gad ground truth (b, e), and T1-Gad proposed method (c, f). KIT001 (a, b, c) and KIT002 (d, e, f).

and FLAIR) of 80 low-grade and high-grade gliomas from synthetic and real patient cases, with 1 mm of isotropic resolution. These datasets were obtained from the following:

- (i) University of Utah database [30]: Synthetic brain tumor datasets were used in the first part of validation. This data simulates contrast-enhanced T1-weighted MR images with synthetically generated tumors. The tumor probability maps and levels of intensity nonuniformity (bias field) are also available. This dataset is included in the performance evaluations since the ground truth segmentation is readily available.
- (ii) Real MR images with ground truth from Kitware database (KIT) [31]: KIT offers both synthetic and

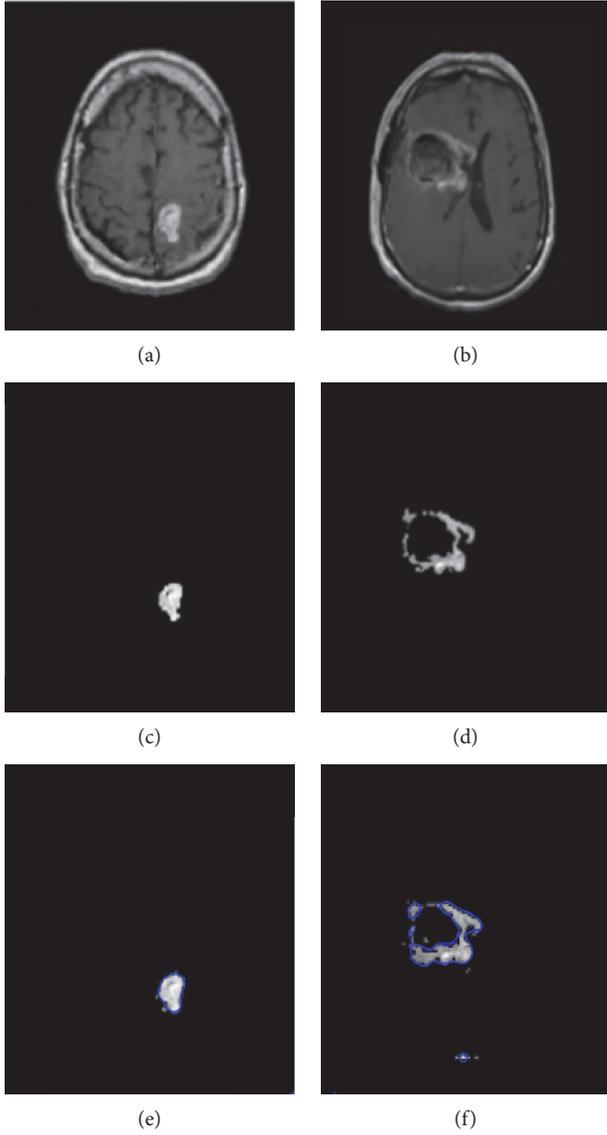


FIGURE 8: T1-image (a, b), ground truth (c, d), and proposed method (e, f). INNN1 (a, c, e) and INNN2 (b, d, f).

real brain images with manually guided expert segmentation results.

- (iii) Brain tumor datasets obtained from our INNN: A large dataset of brain tumor/lesion patients, who received treatment from the Instituto Nacional de Neurología y Neurocirugía (INNN), Mexico, was utilized in the second set of experiments. As a ground truth for our segmentation phase, we used the tumor contours outlined manually by a radio-oncology specialist in the same hospital. These images were both T1-weighted and FLAIR modalities, and they provided an effective way to suppress cerebrospinal fluid (CSF) to bring out periventricular hyperintense lesions. It is worth mentioning that a 1.5 T MRI scanner located at INNN was used to generate the images.

TABLE 1: Performance criteria of the proposed method applied on synthetic images from Utah database: Dice, Jaccard, sensitivity, and specificity score.

ID	Dice	Jaccard	Sensitivity	Specificity
Syn001	0.467	0.341	0.654	0.993
Syn002	0.324	0.284	0.525	0.991
Syn003	0.545	0.439	0.713	0.997
Syn004	0.442	0.322	0.611	0.991
Syn005	0.449	0.328	0.613	0.994
Syn006	0.621	0.564	0.784	0.998
Syn007	0.578	0.458	0.733	1
Syn008	0.233	0.169	0.499	0.989
Syn009	0.669	0.587	0.795	0.999
Syn010	0.611	0.536	0.743	0.997
Syn011	0.378	0.381	0.416	0.987
Syn012	0.558	0.431	0.701	0.992
Syn013	0.471	0.351	0.662	0.995
Syn014	0.342	0.299	0.546	0.994
Syn015	0.482	0.359	0.666	1
Mean	0.478	0.390	0.644	0.994
Standard deviation	0.122	0.114	0.109	0.004

To evaluate the segmentation quantitatively, we used the Jaccard coefficient (R) and the Dice overlap (D) [32], for similarity measurement and sensitivity (Se) and specificity (Sp) for success and error rate measurements [33]. These measures can be expressed as

$$R(T, S) = \frac{|T \cap S|}{|T \cup S|} = \frac{|TP|}{|TP| + |FP| + |FN|}, \quad (29)$$

$$D(T, S) = \frac{2|T \cap S|}{|T| + |S|} = \frac{2|TP|}{|TP| + |FP| + |FN|}, \quad (30)$$

$$Se = \frac{|TP|}{|T|} = \frac{|TP|}{|TP| + |FN|}, \quad (31)$$

$$Sp = \frac{|TN|}{|\bar{T}|} = \frac{|TN|}{|TN| + |FP|}, \quad (32)$$

where T = ground truth, S = pixels labelled by the algorithm, TP = true positive, FP = false positive, and FN = false negative. As was previously mentioned, experiments were conducted on both real and synthetic images.

Figure 6 shows the same challenging patient's MR image demonstrated in Figure 2, whose contrast is very poor and has very small intensity shift along the lesion edges, making the detection more complex. This can be witnessed in Figure 2(b), where after the enhancement, detection was carried out by a robust system described in [33] and ended up being misleading due to the intensity inhomogeneity present in the image. However, by applying the proposed method, the lesion was detected with greater accuracy, as seen in Figure 6(c). This improvement owes a lot to the hard-fuzzy

TABLE 2: Performance criteria of the proposed method applied on low-grade/high-grade images from Kitware: Dice, Jaccard, sensitivity, and specificity score.

ID	Dice	Jaccard	Sensitivity	Specificity
KIT001	0.742	0.611	0.793	0.999
KIT002	0.843	0.703	0.852	1
KIT003	0.798	0.731	0.884	1
KIT004	0.738	0.601	0.785	0.993
KIT005	0.913	0.852	0.908	1
KIT006	0.822	0.765	0.925	1
KIT007	0.877	0.797	0.932	1
KIT008	0.932	0.868	0.958	1
KIT009	0.668	0.526	0.790	0.994
KIT010	0.880	0.798	0.912	1
KIT011	0.929	0.862	0.947	1
KIT012	0.757	0.630	0.818	0.999
KIT013	0.862	0.756	0.924	1
KIT014	0.893	0.828	0.939	1
KIT015	0.971	0.890	0.975	1
KIT016	0.778	0.661	0.792	1
KIT017	0.832	0.758	0.922	1
KIT018	0.840	0.748	0.852	1
KIT019	0.952	0.898	0.968	1
KIT020	0.722	0.665	0.825	0.996
KIT021	0.852	0.795	0.925	1
KIT022	0.902	0.838	0.928	1
KIT023	0.834	0.769	0.923	1
KIT024	0.898	0.841	0.944	1
KIT025	0.618	0.478	0.750	0.992
Mean	0.834	0.747	0.887	0.999
Standard deviation	0.088	0.112	0.067	0.002

TABLE 3: Performance criteria of the proposed method applied on high-grade/low-grade tumor images from INNN: Dice, Jaccard, sensitivity, and specificity score.

ID	Dice	Jaccard	Sensitivity	Specificity
INNN001	0.942	0.851	0.943	1
INNN002	0.899	0.785	0.926	1
INNN003	0.907	0.839	0.938	1
INNN004	0.860	0.763	0.918	1
INNN005	0.968	0.882	0.978	1
INNN006	0.801	0.701	0.880	1
INNN007	0.795	0.699	0.834	1
INNN008	0.807	0.731	0.890	1
INNN009	0.861	0.725	0.914	1
INNN010	0.842	0.711	0.906	1
INNN011	0.828	0.701	0.886	1
INNN012	0.837	0.707	0.901	1
INNN013	0.871	0.765	0.933	1
INNN014	0.933	0.831	0.940	1
INNN015	0.911	0.876	0.939	1
INNN016	0.907	0.855	0.927	1
INNN017	0.732	0.658	0.811	1
INNN018	0.790	0.684	0.839	1
INNN019	0.801	0.713	0.846	1
INNN020	0.951	0.873	0.967	1
INNN021	0.511	0.440	0.816	0.901
INNN022	0.887	0.778	0.949	1
INNN023	0.819	0.718	0.850	1
INNN024	0.918	0.877	0.942	1
INNN025	0.851	0.728	0.944	1
Mean	0.850	0.756	0.905	0.996
Standard deviation	0.091	0.097	0.048	0.019

clustering, which reflects the gravity concept. Generally, in classical fuzzy clustering, all pixels in an image are assigned a membership to every structure, regardless of the magnitude of the corresponding membership. This makes it easier, as we discovered in our experiment, for some pixels to be misclassified especially when they fall into the intensity range of the structure. For this reason, when considering the presence of some random intensity-inhomogeneities that are spread across the image and physically linked not to the tissue problem but rather to the radiofrequency MR signal, one needs not only to consider the intensity level of the pixel but also its spatial location in order to guarantee a proper clustering. Spatial location of pixels is of great importance because it helps us carry out a gravitation-based clustering in areas where fuzzy clustering is not suitable. For instance, pixels at the center of WM/GM tissues could unnecessarily be assigned membership to other tissues that they do not in fact belong to; this is why it is preferable that they are hard-clustered. Applying the gravitation concept, these same pixels will attract all the other pixels located not far from this neighborhood as long as d^2 is small enough and $m_i^{(t)}$ is big enough

to outweigh the attraction by other clusters. Moreover, they will be joined together to form a stronger cluster as was explained in the algorithm. On the other hand, pixels located in the intersection zone of tissues, which have an intensity range that could make them fit in any of these tissues, would do better if fuzzy-clustered. This would eliminate every sort of uncertainty, take care of partial volume effect, and reach optimal state more promptly. The advantages of this gravitational fuzzy clustering phenomenon can be explicitly witnessed in Figure 6(c), where all the pixels wrongly assigned to the lesion in Figure 6(b) just because their intensity range matched that of the lesion were discounted in Figure 6(c) not because of their intensity but because of their spatial location.

The same scenario can be witnessed in Figure 7 where each of the two FLAIR images on the top row carries a lesion on the left frontal lobe, but there is also a seemingly detectable lesion on the right frontal lobe. However, the algorithm disregarded it, despite its convincing intensity range, which is in accordance with the ground truth in the second row.

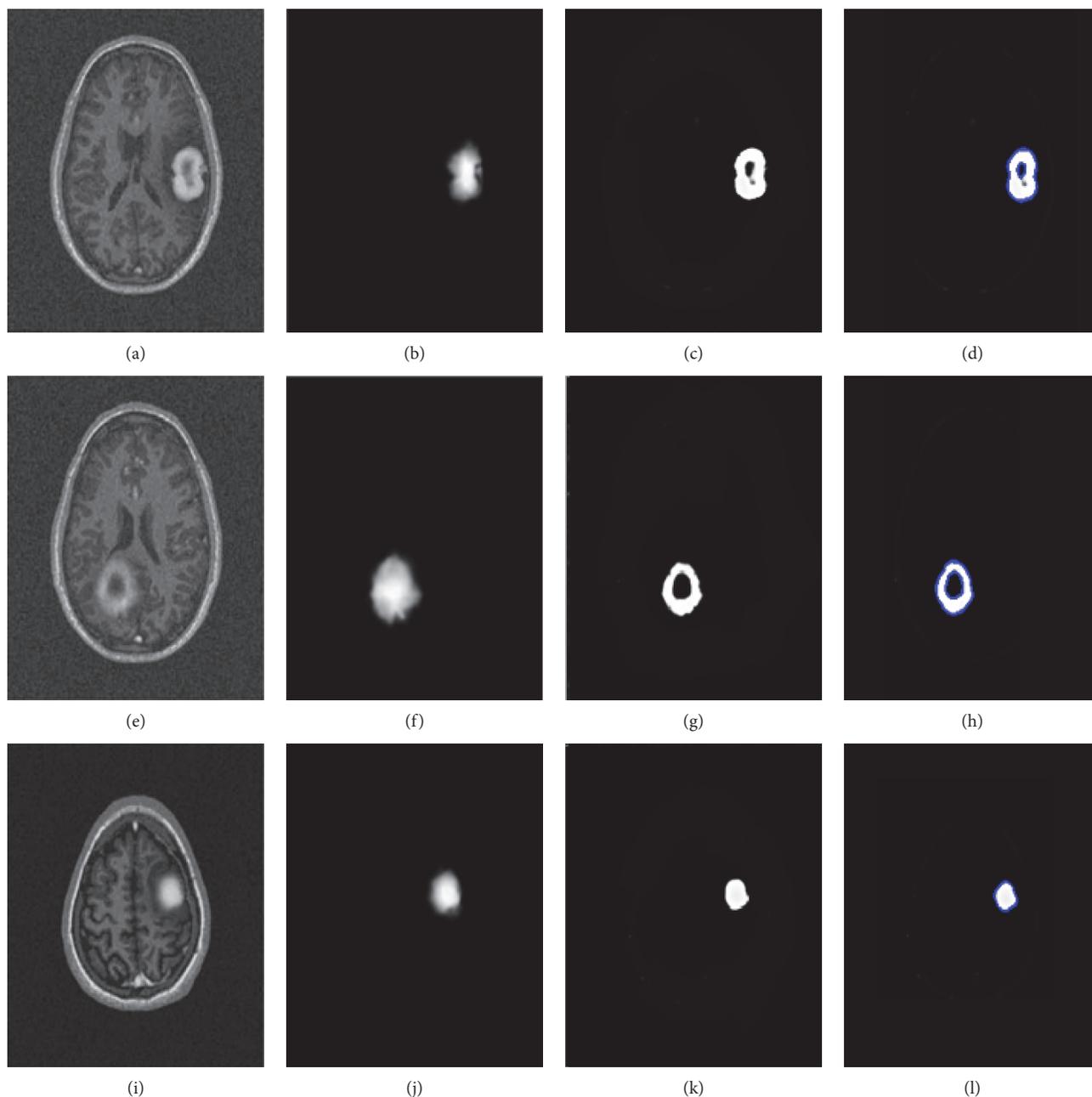


FIGURE 9: MR images of the synthetic dataset. syn001 (a, b, c, d), syn002 (e, f, g, h), and syn003 (i, j, k, l). Original image (a, e, i), ground truth (b, f, j), GFC-result (c, g, k), and final result (d, h, l).

Figure 8 portrays the segmentation results for the two high-grade gliomas, and we can see that the *UGFC* results in the third row perfectly match the ground truth in the second row. As mentioned before, evaluation studies were carried out using (29), (30), (31), and (32). Furthermore, performance measures, that is, Dice overlap, Jaccard, sensitivity, and specificity between the ground truth segmentation and the result of the algorithm are reported in Tables 1, 2, and 3 for both synthetic and clinical datasets. Good results on real MR images were obtained as can be seen in Tables 2 and 3, which demonstrate the performance on Kitware datasets and Instituto Nacional de Neurología y Neurocirugía

datasets. However, this was not the case on synthetic images. Figure 9 shows some synthetic MR images from the University of Utah database where the detection was successful; however, as demonstrated in Table 1, a poor overlap performance was obtained due to a misleading ground truth. Nevertheless, based on visual inspection, one can see that the detection was good enough. Figure 10 shows more challenging cases of patients with multiple sclerosis (INNN17, INNN21), where the intensity spectrum of the lesion almost matches that of the healthy tissues. Moreover, the lesions present a high degree of discontinuity which was a big challenge to this algorithm, thereby resulting in a very poor

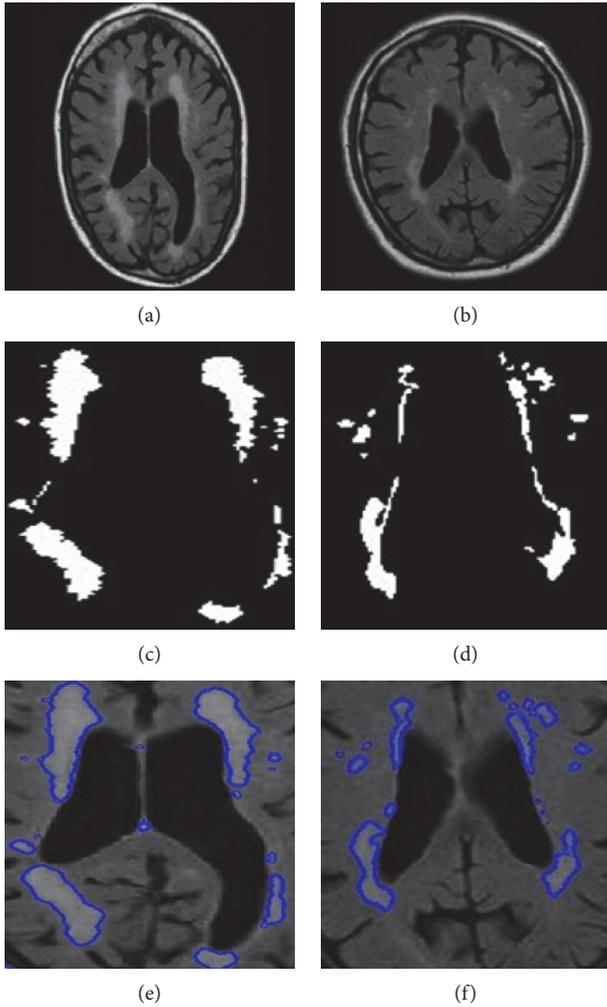


FIGURE 10: Multiple sclerosis cases. INNN17 (a, c, e) and INNN21 (b, d, f). Original image (a, b), ground truth (c, d), and proposed method (e, f).

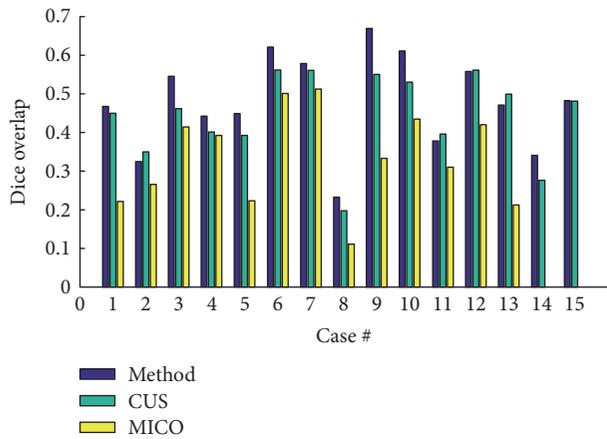


FIGURE 11: Comparison of the proposed method against CUS and MICO methods as carried out on the synthetic dataset.

performance as can be seen in Table 1 where INNN17 has a 0.732 overlap and INNN21 has a 0.511 performance. We also compared the proposed method against the Automated

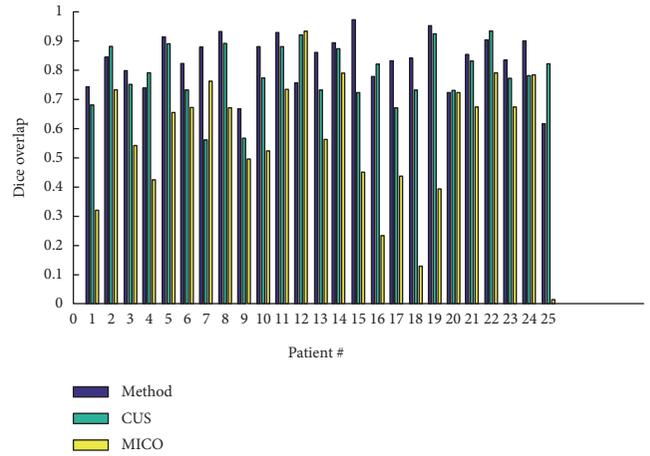


FIGURE 12: Comparison of the proposed method against CUS and MICO methods as carried out on the Kitware dataset.

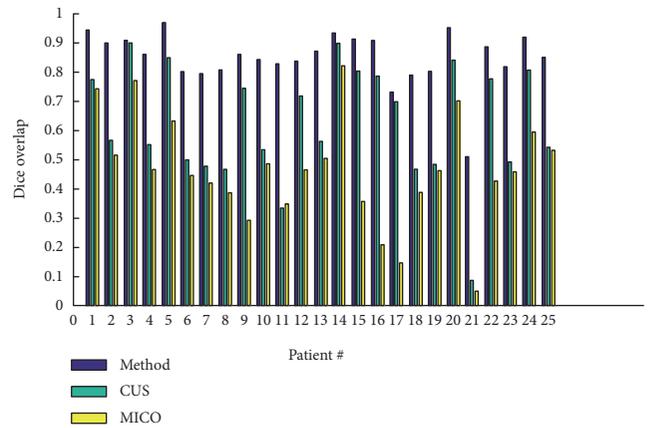


FIGURE 13: Comparison of the proposed method against CUS and MICO methods as carried out on INNN dataset.

Lesion Detection on MRI Scans Using Combined Unsupervised and Supervised Methods by Guo et al. [34] and Multiplicative Intrinsic Component Optimization (MICO) [35] method by Li et al., and their results are reported in Figures 11, 12, and 13.

Finally, Figure 14 shows two KIT images processed by the proposed method and the Multiplicative Intrinsic Component Optimization (MICO), and it can be seen that a visual inspection witnesses a much more precise detection by our method than the MICO method. We should emphasize that the intensity inhomogeneity presented in Figure 14 did confuse the MICO method, which went even beyond the lesion contour. Yet, our method got it right.

4. Discussion and Conclusions

We presented a brain lesion detection algorithm along with validation studies over a synthetic lesion dataset and two real datasets: one from Kitware Repository and another from a clinical database of tumors that underwent radiosurgery

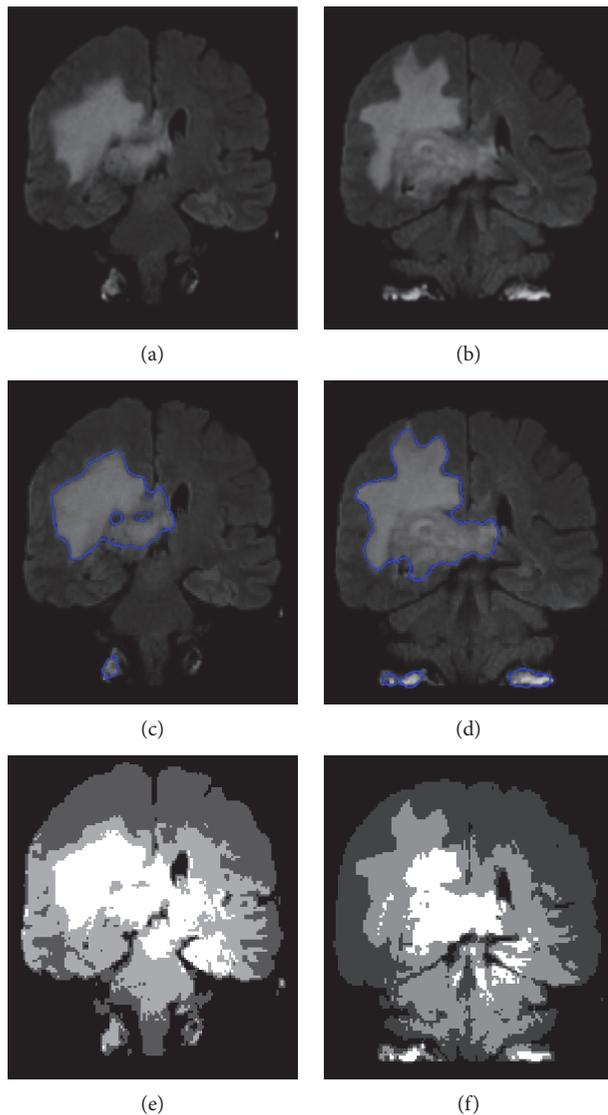


FIGURE 14: Comparison between proposed method (c, d) and Multiplicative Intrinsic Component Optimization (MICO) (e, f).

planning at Instituto Nacional de Neurología y Neurocirugía de México (INNN). The performance over these datasets of highly heterogeneous tissue content demonstrated an average overlap of 0.83 and 0.85. However, a poor performance of 0.48 was registered in synthetic MR images from the Utah database, mostly due to the poor ground truth presented. In all cases, the proposed algorithm provided superior quality segmentation when compared to the benchmark algorithms. On an average, the proposed automated algorithm takes about 8 seconds (measured using MATLAB R2007b on a 3.00 GHz, dual processor) in comparison to 21 and 12 seconds taken by the MICO and CUS respectively. Strengths of the proposed method include its automatic nature, its efficiency in terms of computation time, and its robustness with respect to different and heterogeneous lesion types. Another important aspect to consider is that the algorithm can work on different MR modalities.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Acknowledgments

The authors would like to thank the Instituto Politécnico Nacional de México for promoting this project, the Instituto Nacional de Neurología y Neurocirugía de México for providing medical guidance along with MR images, and CONA-CyT de México for funding this project.

References

- [1] J. R. Adler, S. D. Chang, M. J. Murphy, J. Doty, P. Geis, and S. L. Hancock, "The Cyberknife: a frameless robotic system for radiosurgery," *Stereotactic and Functional Neurosurgery*, vol. 69, no. 1–4, Part 2, pp. 124–128, 1997.
- [2] W. Kilby, J. R. Dooley, G. Kuduvalli, S. Sayeh, and C. R. Maurer Jr., "The CyberKnife® robotic radiosurgery system in 2010," *Technology in Cancer Research and Treatment*, vol. 9, 2010.
- [3] E. Laurence Court, P. Balter, and R. Mohan, *Principles of IMRT*, Springer, Japan, 2015.
- [4] R. N. Strickland, *Image-Processing Techniques for Tumor Detection*, CRC Press, 2002.
- [5] J. C. Bezdek, L. O. Hall, and L. P. Clarke, "Review of MR image segmentation techniques using pattern recognition," *Medical Physics*, vol. 20, pp. 1033–1048, 1993.
- [6] J. M. V. Kinani, A. J. Rosales-Silva, F. J. Gallegos-Funes, and A. Arellano, "Fuzzy C-means applied to MRI images for an automatic lesion detection using image enhancement and constrained clustering," in *2014 4th International Conference on Image Processing Theory, Tools and Applications (IPTA)*, pp. 1–7, Paris, 2014.
- [7] N. I. Bankman, *Handbook of Medical Image Processing and Analysis*, Academic Press, 2009.
- [8] R. C. González and R. E. Woods, *Digital image processing*, Addison-Wesley, 2008.
- [9] F. Y. Shih, *Image Processing and Pattern Recognition, Fundamentals and Techniques*, IEEE Press, 2010.
- [10] J. C. Russ, *The Image Processing Handbook*, CRC Press, Sixth edition, 2011.
- [11] L. A. Zadeh, "Fuzzy sets," *Information and Control*, vol. 8, no. 3, pp. 338–358, 1965.
- [12] E. Prastawa, S. H. Bullitt, and G. Gerig, "A brain tumor segmentation framework based on outlier detection," *Medical Image Analysis*, vol. 8, no. 3, pp. 275–283, 2008.
- [13] A. Hamamci, N. Kucuk, K. Karaman, K. Engin, and G. Unal, "Tumor-cut: segmentation of brain tumors on contrast enhanced MR images for radiosurgery applications," *IEEE Transactions on Medical Imaging*, vol. 31, no. 3, 2012.
- [14] B. Menze, K. V. Leemput, D. Lashkari, M. A. Weber, N. Ayache, and P. Golland, "A generative model for brain tumor segmentation in multimodal images," *Medical Image Computing and Computer-Assisted Intervention*, vol. 13, pp. 151–159, 2012.
- [15] A. Gooya, G. Biros, and C. Davatzikos, "Deformable registration of glioma images using EM algorithm and diffusion

- reaction modeling,” *IEEE Transactions on Medical Imaging*, vol. 30, no. 2, pp. 375–390, 2011.
- [16] E. Geremia, B. H. Menze, O. Clatz, E. Konukoglu, A. Criminisi, and N. Ayache, *Spatial Decision Forests for MS Lesion Segmentation in Multi-Channel MR Images*, Elsevier, 2010.
- [17] J. Liu, J. K. Udupa, D. Odhner, D. Hackney, and G. Moonis, “A system for brain tumor volume estimation via MR imaging and fuzzy connectedness,” *Computerized Medical Imaging and Graphics*, vol. 29, pp. 21–34, 2005.
- [18] S. Shen, A. J. Szameitat, and A. Sterr, “An improved lesion detection approach based on similarity measurement between fuzzy intensity segmentation and spatial probability maps,” *Magnetic Resonance Imaging*, vol. 28, 2009.
- [19] J. Bezdek, J. Keller, R. Krisnapuram, and N. R. Pal, *Fuzzy Models and Algorithms for Pattern Recognition and Image Processing*, Springer, 2005.
- [20] I. Ozkan and I. B. Turksen, *Upper and Lower Values for the Level of Fuzziness in FCM*, 2009.
- [21] P. Therasse, “Evaluation of response: new and standard criteria,” *Annals of Oncology*, vol. 13, pp. 127–129, 2002.
- [22] M. Chernov, M. Hayashi, J. Ganz, and K. Takakura, *Gamma Knife Neurosurgery in the Management of Intracranial Disorders*, Springer, 2013.
- [23] B. Vemuri and Y. Chen, “Joint image registration and segmentation,” in *Geometric Level Set Methods in Imaging, Vision, and Graphics*, pp. 251–269, Springer, 2003.
- [24] X. Han, C. Xu, and J. Prince, “A topology preserving level set method for geometric deformable models,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, pp. 755–768, 2003.
- [25] V. I. Arnold, *Geometrical Methods in the Theory of Ordinary Differential Equations*, Springer-Verlag, New York, 1983.
- [26] L. C. Evans, *Partial Differential Equations*, American Mathematical Society, 2010.
- [27] C. Li, C. Xu, C. Gui, and M. D. Fox, “Distance regularized level set evolution and its application to image segmentation,” *IEEE Transactions on Image Processing*, vol. 19, no. 12, 2010.
- [28] S. Osher and R. Fedkiw, *Level Set Methods and Dynamic Implicit Surfaces*, Springer-Verlag, New York, 2002.
- [29] X. Xie, “Active contouring based on gradient vector interaction and constrained level set diffusion,” *IEEE Transactions on Image Processing*, vol. 19, no. 1, pp. 154–164, 2010.
- [30] September 2013, <http://www.ucnia.org/softwaredata/5-tumor-data/10-simtumordb.html>.
- [31] <http://insight-journal.org/midas/gallery>.
- [32] L. R. Dice, “Measures of the amount of ecologic association between species,” *Ecology*, vol. 26, no. 3, pp. 297–302, 1945.
- [33] J. M. V. Kinani, A. J. S. Rosales, F. J. F. Gallegos, and A. Arellano, “Computer-aided diagnosis of brain tumors using image enhancement and fuzzy logic,” *DYNA*, vol. 81, no. 183, pp. 148–157, 2014.
- [34] D. Guo, J. Fridriksson, P. Fillmore et al., “Automated lesion detection on MRI scans using combined unsupervised and supervised methods,” *BMC Medical Imaging*, vol. 15, no. 1, p. 50, 2015.
- [35] C. Li, C. J. Gore, and C. Davatzikos, “Multiplicative intrinsic component optimization (MICO) for MRI bias field estimation and tissue segmentation,” *Magnetic Resonance Imaging*, vol. 32, no. 7, pp. 913–923, 2014.

Research Article

Semiautomatic Epicardial Fat Segmentation Based on Fuzzy c-Means Clustering and Geometric Ellipse Fitting

Vladimir Zlokolica,¹ Lidija Krstanović,¹ Lazar Velicki,² Branislav Popović,³ Marko Janev,⁴ Ratko Obradović,¹ Nebojsa M. Ralević,¹ Ljubomir Jovanov,⁵ and Danilo Babin⁵

¹Department of Fundamentals Sciences, Faculty of Technical Sciences, University of Novi Sad, Trg Dositeja Obradovića 6, 21000 Novi Sad, Serbia

²Faculty of Medicine, University of Novi Sad, Hajduk Veljkova 3, 21000 Novi Sad, Serbia

³Department of Power, Electronic and Telecommunication Engineering, Faculty of Technical Sciences, University of Novi Sad, Trg Dositeja Obradovića 6, 21000 Novi Sad, Serbia

⁴Mathematical Institute, Serbian Academy of Arts and Sciences, Kneza Mihaila 36, 11000 Beograd, Serbia

⁵Faculty of Engineering and Architecture, imec-IPI-UGent, Ghent University, Sint-Pietersnieuwstraat 41, 9000 Gent, Belgium

Correspondence should be addressed to Vladimir Zlokolica; vzlokolica@uns.ac.rs

Received 5 May 2017; Revised 6 August 2017; Accepted 20 August 2017; Published 20 September 2017

Academic Editor: Shujun Fu

Copyright © 2017 Vladimir Zlokolica 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 of particular heart parts plays an important role in recognition tasks, which is utilized for diagnosis and treatment. One particularly important application is segmentation of epicardial fat (surrounds the heart), which is shown by various studies to indicate risk level for developing various cardiovascular diseases as well as to predict progression of certain diseases. Quantification of epicardial fat from CT images requires advance image segmentation methods. The problem of the state-of-the-art methods for epicardial fat segmentation is their high dependency on user interaction, resulting in low reproducibility of studies and time-consuming analysis. We propose in this paper a novel semiautomatic approach for segmentation and quantification of epicardial fat from 3D CT images. Our method is a semisupervised slice-by-slice segmentation approach based on local adaptive morphology and fuzzy c-means clustering. Additionally, we use a geometric ellipse prior to filter out undesired parts of the target cluster. The validation of the proposed methodology shows good correspondence between the segmentation results and the manual segmentation performed by physicians.

1. Introduction

Epicardial fat presents a deposition of a visceral fat that surrounds the heart and is mainly located in the atrioventricular and interventricular groves surrounding the epicardial vessels [1]. There are studies that suggest that growing amount of epicardial fat can indicate the development of atherosclerotic diseases and correlate well with the parameters of the metabolic syndrome. Further on, some studies suggest that increased epicardial fat volume and its outspread distribution play an important part in the development of an unfavorable metabolic profile and potentiate cardiovascular risk [2]. Some of the conditions linked to increased epicardial

fat volume include the following: increased prevalence of coronary artery calcifications [3, 4], impaired diastolic filling [5], development of myocardial infarction [3] and atrial fibrillation [6], and increased carotid stiffness [7]. Finally, there is evidence that epicardial fat possesses characteristics of an endocrine tissue capable of promoting accelerated atherosclerosis [8, 9].

Due to the importance of the epicardial fat and its relation to various cardiovascular and metabolic diseases, epicardial fat detection, quantification, and 3D segmentation have become important topic in cardiac image analysis and cardiology in general. Different imaging technologies can be used for epicardial fat imaging, segmentation, and quantification

[10–12]. Specifically, CT has been found to be more accurate in the evaluation of fat tissues due to its higher spatial resolution when compared to ultrasound [13].

In Coppini et al. [13], a region growing strategy is proposed as preprocessing step, in order to extract the heart region from the rest of the target CT image. As an additional step, medical doctor was required to place up to 7 control points along the pericardium border, which are then used for spline interpolation in order to obtain a smooth closed pericardial contour. Finally, the epicardial fat is simply quantified by thresholding, since it is theoretically located within this generated contour. A hierarchical, multiclass, multifeature, fuzzy affinity-based computational framework for segmentation of abdominal adipose tissue was proposed in [14] and further extended in [15] in order to be applicable to segmentation of epicardial fat. Further on, the work of [13] additionally reduced the user intervention, although the expert was still required to scroll through slices and to place control points on the pericardium. An algorithm that uses the anatomy of the heart to detect the pericardium line and separate epicardial fat from pericardial fat is presented in [16]. In Dey et al. [17], the authors used an additional processing step for the identification of the pericardium by tracing lines going from the heart centroid to the pericardium layer and using splines to interpolate outer points. Shahzad et al. [1] proposed the first fully automated method for epicardial fat segmentation that is using multiatlas approach. In Zlokolica et al. [18], segmentation of the heart region from the rest of the image content was done by applying morphological opening by reconstruction operation followed by thresholding and subsequent morphological dilation operator. The size and the shape of the morphological operators are found experimentally for all image slices, while the threshold value is found by statistical analysis of the input image histogram and is in general different and adaptable for each image slice. That idea is further extended in [19] by imposing structural prior knowledge and subjective-objective correspondence, as an additional step for excluding the segmentation parts that were out of the predefined circle.

In this paper, we build on our previous work and propose a new method for epicardial fat segmentation which is based on a 2D segmentation approach that includes fuzzy *c*-means clustering and ellipse fitting. The fuzzy *c*-means clustering is done using a predefined number of clusters within a detected region of interest (ROI). The clustering is performed with reduced supervision, where the only input information from a physician are the coordinates of a single patch corresponding to the epicardial fat tissue. After the user marks a patch corresponding to the epicardial fat, the proposed fuzzy *c*-means clustering is applied to segment the correct target image cluster as the one with the minimal distance to the user-marked patch. We use the following facts as prior information in our method. Firstly, the epicardial fat tissue lies in the neighborhood of the ellipse placed at the center of the mass of the resulting cluster. This allows us to determine the epicardial fat pixels in segmentation process. Secondly, motivated by the fact that the epicardial fat is located on the surface of the ventricles (mainly anterolateral surface), we impose additional assumptions regarding the angle in

the parametrization of the ellipse. By doing so, we additionally filter parts of the resulting cluster not belonging to the epicardial fat tissue. We validate the proposed concept by comparing the resulting segmentation results to the manual segmentation provided by a cardiovascular surgeon. Additionally, we also demonstrate that the proposed approach outperforms the segmentation method of [18] and general image segmentation methods: snake active contour method [20] and region growing method [21].

The paper is organized as follows. In Section 2, we describe the proposed methodology for epicardial fat segmentation, where we first discuss input data set and introduce image preprocessing step in Section 2.2. After that, we explain the proposed segmentation approach in Section 2.3. Experimental results with discussion are provided in Section 3. Finally, the conclusions are drawn in Section 4.

2. Materials and Methods

In this section, we discuss data set used in our study and propose methodology for epicardial fat segmentation. The proposed method consists of two main steps: (i) image preprocessing and ROI extraction (Section 2.2) and (ii) our novel image segmentation approach (Section 2.3) used for final epicardial fat segmentation. Section 2.3 is organized in two subsections: (i) Section 2.3.1 explaining the proposed fuzzy *c*-means clustering step and (ii) Section 2.3.2 describing the cluster data filtering step by geometrical fitting scheme.

2.1. Data Set and Ethical Approval. The study included a total of ten patients with previously established coronary artery disease (stable and nonstable angina). All patients were submitted to standard contrast chest CT scan including CT coronarography performed on a Siemens SOMATOM Dual Source 256-slice scanner. In the proposed approach, we consider contrasted CT images with 12 bits pixel depth resolution, providing thus 4096 gray levels. The study was approved by the institutional review board, and all subjects signed the written informed consent. The study was conducted in accordance with the Declaration of Helsinki.

2.2. CT Image Enhancement and ROI Extraction. In this section, we build on our previous work [18] for enhancing input CT image slices and detection of heart (and epicardial fat) region of interest (ROI). The goal of the processing step is to remove the nonheart-related image content and retain image content only related to the whole heart region. An example of one slice of acquired 3D CT image of a heart, without previous contrast enhancement, is shown in Figure 1.

As can be seen from Figure 1, the input image slice is not of sufficient image quality to be used for manual or automatic image segmentation. In order to improve input image slice quality in terms of visual quality and the quality for feature extraction and segmentation, we propose to first preprocess the input image slices with image-adaptive tone mapping. The proposed tone mapping function is formed based on input image histogram based on which the transfer tone mapping function is optimized in terms of its parameters.

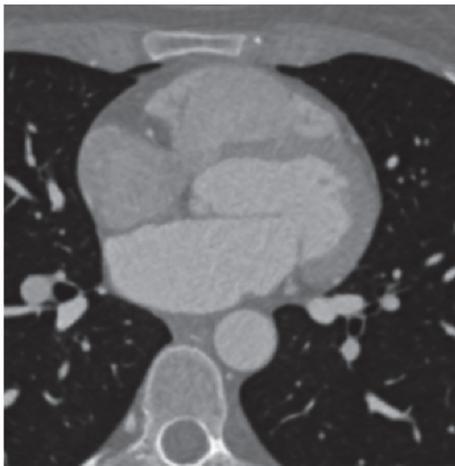


FIGURE 1: Input CT image slice of a heart, without additional contrast enhancement applied.

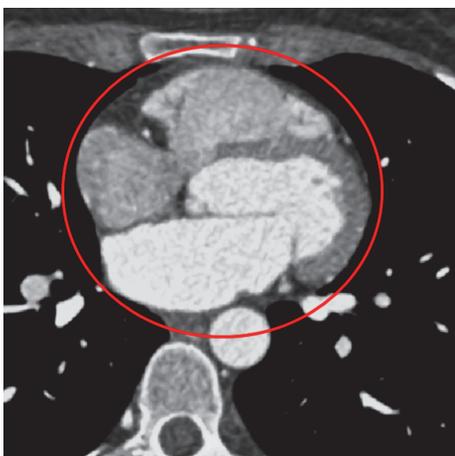


FIGURE 2: Output image after contrast enhancement with denotation of the region of interest that corresponds to the heart.

We aim at distinguishing between the whole heart region and surrounding (other) parts of the body. For this reason, we first detect ROI that includes the whole heart (denoted in Figure 2) within which we further perform 3D segmentation of epicardial fat. Prior to this, we apply contrast enhancement on input image to facilitate subsequent segmentation task.

In order to automatically extract the heart region from the rest of the image slice, we propose a morphology-based masking scheme. This includes morphological “opening by reconstruction” operation, followed by thresholding and subsequent morphological dilation operator. We use circular (disk-shaped) structuring element (SE) for morphological operators due to mainly circular shape of the heart (in 2D image slice), its chambers, and proximal vessels. We determine the optimal size of SEs based on the predicted size of the heart (input provided by cardiologists) and pixel spacing (contained in the DICOM header). In our case, the calculated SE size of the morphological operators is 40 pixels and 20 pixels for opening by reconstruction and dilation operator,

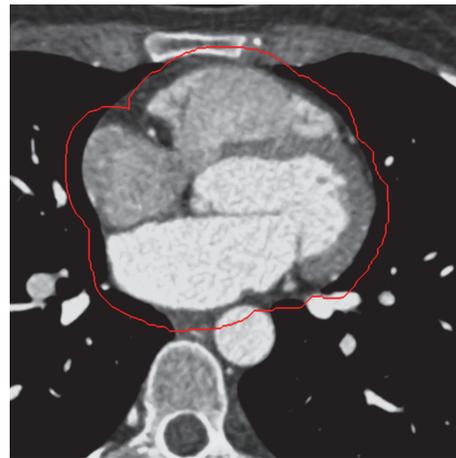


FIGURE 3: The result of automatic ROI extraction from the rest of the image content (rounded by red line), for a particular enhanced image slice.

respectively. The intermediate thresholding depends on each input image slice statistics and is therefore applied differently to each input image slice. The result of automatic ROI extraction is shown in Figure 3 for the image slice in Figure 2, showing the accurate delineation of the whole heart region.

Alternatively, the extracted ROI can be shown as a mask (binary) image X , where all pixels corresponding to nonheart region are assigned gray value 0, while the heart pixels are assigned gray value 1:

$$X(i, j) = \begin{cases} 1, & \text{if pixel } (i, j) \text{ belongs to the region of interest,} \\ 0, & \text{otherwise.} \end{cases} \quad (1)$$

2.3. Epicardial Fat Segmentation Scheme. In this section, we describe the slice-by-slice epicardial fat segmentation scheme featuring high level of adaptivity, accuracy, and automatization. In the proposed segmentation approach, we first perform image clustering on previously preprocessed input image within the extracted ROI as shown in Figure 3. This represents an initial segmentation corresponding to a single cluster, which is then further processed to only extract parts corresponding to the epicardial fat tissue. The main steps in the proposed scheme are fuzzy c -means clustering approach, ellipse fitting, and final subclustering/filtering of the ellipse-shaped cluster region corresponding to the epicardial fat only.

2.3.1. Fuzzy c -Means Clustering in the Region of Interest. The proposed initial segmentation of the preprocessed input image slice (within the ROI) is performed using fuzzy c -means clustering [22] with predefined number of clusters C . The feature vector used for clustering in the segmentation task is $u = [u_1 u_2]^T$, where u_1 represents the mean value of luminance of the particular $R \times R$ image pixel patch $U = \{U(i, j) | i, j = 1, \dots, R\}$, while u_2 represents the variance of the luminance in that patch:

$$\begin{aligned} u_1 &= \frac{1}{R^2} \sum_{i,j=1}^R U(i, j), \\ u_2 &= \frac{1}{R^2} \sum_{i,j=1}^R (U(i, j) - u_1)^2. \end{aligned} \quad (2)$$

Patches within the ROI are sampled and gathered for each image slice. Features are calculated so that we obtain a set $\mathcal{S} = \{u_k | k = 1, \dots, Q\}$ of features on which we perform clustering. Note that Q is the number of uniformly sampled patches from the image slice and that we have a unique correspondence between the index of features k in \mathcal{S} and centers (coordinates) of patches $p_k = [i^{(k)}, j^{(k)}]^T$, where $k = 1, \dots, Q$. The number of clusters C is predefined and is determined empirically depending of the particular CT scanning aperture. The clustering is performed without initial seed information and generates $\mathcal{G} = \{\mu_c(k) | k = 1, \dots, Q, c = 1, \dots, C\}$ fuzzy clusters according to their soft characteristic functions and satisfying $\mu_c(k) \in [0, 1]$, for $c = 1, \dots, C$ and $k = 1, \dots, Q$. Note that the closer $\mu_c(k)$ is to 1 indicates a higher correspondence of a particular center of the patch p_k the cluster c .

In order to obtain reference image patch, user is required to inspect single representative CT slice (from the target patient data) and place a seed point. The seed point represents the image coordinates of the center (i^*, j^*) of the reference patch U^* . The proposed unsupervised clustering approach determines the resulting cluster, that is, selects the cluster that corresponds to the epicardial fat. We obtain that by calculating the distance between the reference patch U^* and all C fuzzy clusters contained in \mathcal{G} . For the fixed fuzzy cluster with index c , we determine the distance as follows:

$$d(c, U^*) = \|u^* - \hat{u}_c\|_2, \quad (3)$$

where \hat{u}_c is the mean of the feature vectors from \mathcal{S} , pondered by corresponding soft characteristic function $\mu_c(k)$, that is,

$$\hat{u}_c = \frac{1}{Q} \sum_{k=1}^Q \mu_c(k) u_k. \quad (4)$$

We obtain the resulting cluster, as the cluster with minimal distance to the reference patch U^* , that is, the cluster with index c^{res} given by

$$c^{\text{res}} = \arg \min_c d(c, U^*). \quad (5)$$

Note that each pixel with spatial coordinate (i, j) , $i = 1, \dots, M$ and $j = 1, \dots, N$ (image slices are of $M \times N$ size), can be covered by multiple patches (unless patch sampling is disjoint). Since we need the resulting fuzzy cluster on a discrete set of pixels, the resulting pixel-wise fuzzy cluster c^{res} is obtained using the following equation:

$$\mu_{c^{\text{res}}}(i, j) = \frac{1}{|H_{i,j}|} \sum_{h \in H_{i,j}} \mu_{c^{\text{res}}}(k(h)), \quad (6)$$

where $H_{i,j}$ is the set of all patches that cover pixel (i, j) and $k(h)$ is the index of h th patch within the fuzzy cluster set in

\mathcal{G} . The resulting fuzzy cluster with index c^{res} , recalculated using (6) and defined on the set of particular image slice pixels, is the output cluster from the clustering step. The resulting cluster contains pixels corresponding to epicardial fat tissue, and its cluster data is used in further processing.

In Figure 4(a), we show the resulting fuzzy cluster for the input image slice shown in Figure 3. The resulting cluster contains epicardial fat tissue with other outlier data that we need to remove further.

2.3.2. Cluster Data Postfiltering Using Geometric Ellipse Fitting. After completing the fuzzy c -mean-based initial segmentation (as explained in Section 2.3.1), we perform postfiltering of the extracted cluster data using a geometric ellipse fitting. This step is used for filtering out undesired parts of the target cluster data.

As shown in Figure 4(b), the upper parts of the cluster data can be well fitted with an ellipse denoted with yellow dashed lines. As confirmed by physicians, the epicardial fat distribution corresponds to the fitted ellipse in Figure 4. Hence, our assumption is that significant parts of the resulting cluster, which the medical doctors would classify as epicardial fat tissue, are mainly located around the fitted ellipse. In this assumption, we denote *Geometric Ellipse Prior* (GEP). The ellipse, denoted as *Referent Ellipse* (RE), is placed at the center of the cluster and passing through the bordering parts, as depicted in Figure 4(b). As we expect the epicardial fat tissue to be placed around the ventricles, mainly anterolateral surface, we incorporate the mentioned assumption in the additional geometric prior, also related to the GEP. In addition to the geometric ellipse fitting, we apply constrain on the angle that parametrizes RE so that parts of the resulting cluster within the predefined arc of the RE are selected as desirable and the other parts are filtered out as undesirable. The latter is shown in Figure 4(c) with red colored arc.

Computation of the GEP and corresponding RE for the particular image slice is done in the following three steps:

- (1) Estimation of the RE center (x_c, y_c)
- (2) Estimation of the angle φ_0 between the major axes of the RE and the x referent axes of image slice
- (3) Estimation of the RE parameters a and b corresponding to two characteristic axes that define the shape of the ellipse.

The RE is parameterized and represented in polar coordinates and given in the following form:

$$\begin{aligned} x(\varphi) &= x_c + a \cos(\varphi + \varphi_0), \\ y(\varphi) &= y_c + b \sin(\varphi + \varphi_0), \\ \varphi &\in [0, 2\pi), \end{aligned} \quad (7)$$

where $(x_c, y_c) \in \mathbb{R}^2$ is the center of an ellipse, $a, b > 0$ are parameters corresponding to characteristic axes of an ellipse, and $\varphi \in [0, 2\pi)$ is an angular parameter (see Figure 5). For estimating (x_c, y_c) , we use *the center of gravity* approach, where the estimates (x_c, y_c) are determined as

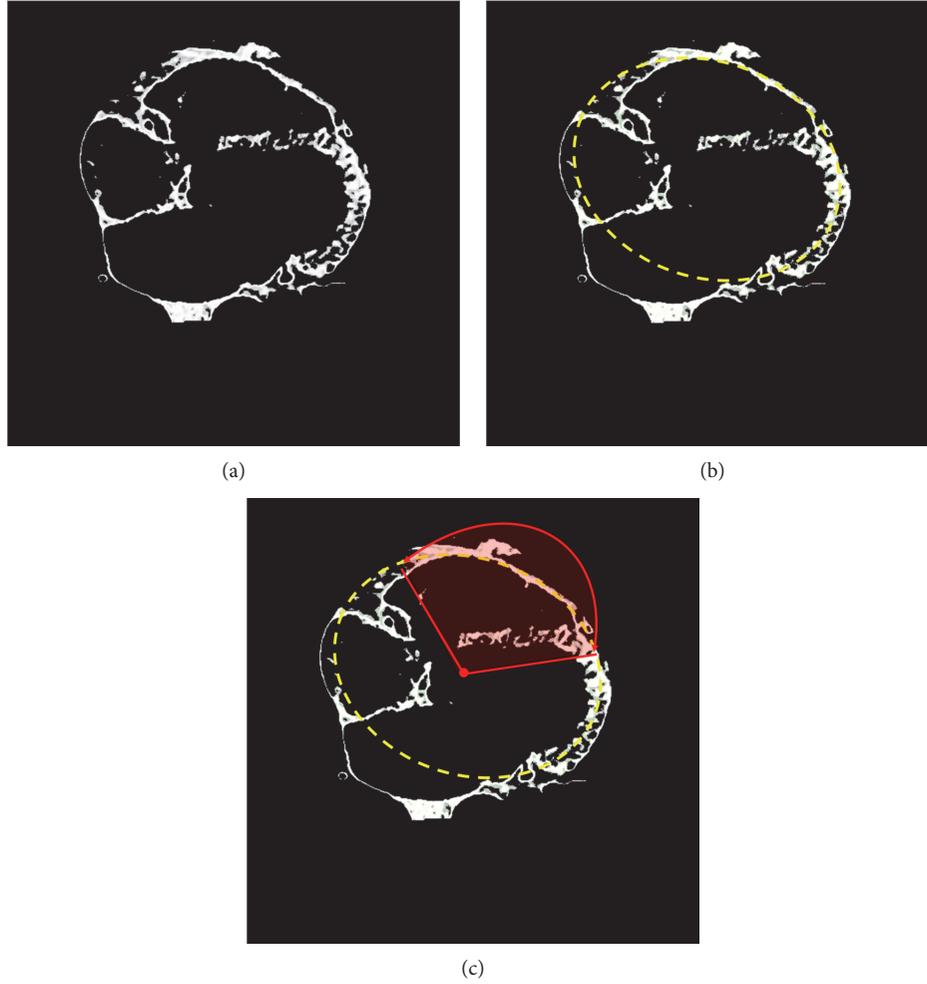


FIGURE 4: The resulting fuzzy cluster that is considered to contain epicardial fat tissue, for the image slice example shown in Figure 3. White color-marked regions show clustered data (a), yellow color-marked pixels indicate ellipsoidal shape fitting approach (b), and red color-marked lines denote the angular bandwidth constrain (c).

$$\begin{aligned}\hat{x}_c &= \frac{M_{\mu_{eres}I}(1, 0)}{M_{\mu_{eres}I}(0, 0)}, \\ \hat{y}_c &= \frac{M_{\mu_{eres}I}(0, 1)}{M_{\mu_{eres}I}(0, 0)},\end{aligned}\quad (8)$$

with moments $M_{\mu_{eres}I}(p, q)$, $p, q \in \mathbb{N}$, defined by

$$M_{\mu_{eres}I}(p, q) = \sum_{i=1}^M \sum_{j=1}^N i^p j^q X(i, j) \mu_{eres}(i, j), \quad (9)$$

where i and j are discrete spatial coordinates corresponding to x - and y -axes of the image slice, that is, referent coordinate system in which the RE is supposed be positioned. Further on, in (8), X stands for the pixel luminance that is weighted by the characteristic function of the resulting fuzzy cluster μ_{eres} . In such manner, we do not only take into account pixel luminance value but also the membership degree of the particular pixel to the resulting fuzzy cluster (if we consider pixels to be statistical observations with mass density function μ_{eres} , $M(1, 0) = E_i(\cdot)$, and $M(0, 1) = E_j(\cdot)$, the previous

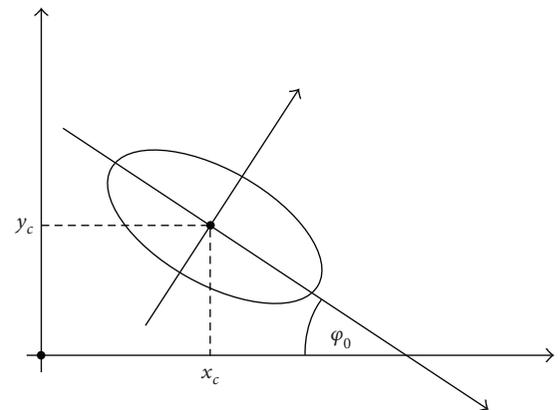


FIGURE 5: Figure depicting ellipse which parameters need to be estimated.

expression can be considered as mathematical expectation for given discrete spatial coordinates i and j , resp.).

Next, we need to estimate the angle φ_0 between the major axes of the RE and the x referent axes of the image

slice (see Figure 5), which defines the main direction of the “soft point cloud” in the resulting cluster as well as orientation of the RE. Therefore, we define the vector of observations \mathcal{P} as $\{p_i = \mu_{c^{\text{res}}}(i, j)[i, j]^T | (i, j) \text{ belongs to ROI}\}$ and estimate the covariance matrix $C_{P,P} = E((P - E(P))(P - E(P))^T)$, as

$$\begin{aligned} \widehat{C}_{P,P} &= \frac{1}{|\mathcal{P}|} \sum_{i=1}^{|\mathcal{P}|} \left((p_i - \widehat{P})(p_i - \widehat{P})^T \right), \\ \widehat{P} &= \sum_{i=1}^{|\mathcal{P}|} p_i, \end{aligned} \quad (10)$$

where we use $\mu_{c^{\text{res}}}$ in order to incorporate additional information regarding the soft characteristic function for the resulting cluster c^{res} (5). For the purpose of determining φ_0 , we apply principle component analyses (PCA): we decompose the positive definite matrix $\widehat{C}_{P,P}^\varepsilon = \widehat{C}_{P,P} + \varepsilon I$ which is regularized by diagonal loading of $\widehat{C}_{P,P}$ with small positive $\varepsilon > 0$. As a result, we obtain the following: $\widehat{C}_{P,P}^\varepsilon = V\Lambda V^T$. Here, $\Lambda = \text{diag}(\lambda_1, \lambda_2)$ and $V = [v_1 | v_2]$, where $\lambda_i > 0$ and $v_i \in \mathbb{R}^2$, $i = 1, 2$ are eigenvalues and eigenvectors of the matrix $\widehat{C}_{P,P}^\varepsilon$, respectively. Additionally, V is unitary matrix, where $v_1 \perp v_2$, and $\|v_i\| = 1$, $i = 1, 2$.

The main direction of the “soft point cloud,” representing the resulting cluster, is actually the direction of eigenvector v^* corresponding to the largest eigenvalue, namely, $v^* = v_{\widehat{i}}$, such that $\widehat{i} = \arg \max_i \{\lambda_i\}$. Finally, we obtain the estimate of φ_0 as

$$\widehat{\varphi}_0 = \arccos\left(\mathbb{P}\text{roj}_{e_x}(v^*)\right), \quad (11)$$

where $\mathbb{P}\text{roj}_{e_x}(v^*) = (v^*)^T e_x$ is the projection of v^* to the unit x -axis vector e_x . We approximate the projection by simplified affine transformation applied to the data points belonging to the ROI:

$$\begin{bmatrix} \tilde{x} \\ \tilde{y} \end{bmatrix} = R(-\widehat{\varphi}_0) \begin{bmatrix} x \\ y \end{bmatrix} - \begin{bmatrix} \widehat{x}_c \\ \widehat{y}_c \end{bmatrix}, \quad (12)$$

with

$$R(-\widehat{\varphi}_0) = \begin{bmatrix} \cos \widehat{\varphi}_0 & \sin \widehat{\varphi}_0 \\ -\sin \widehat{\varphi}_0 & \cos \widehat{\varphi}_0 \end{bmatrix} \quad (13)$$

being rotation matrix for angle $-\widehat{\varphi}_0$. Such projection matrix is applied on each (x, y) position in the image slice. Further on, we also apply (12) to the original RE defined by (7), which will then be centered in $(0, 0)$ with x -axis aligned with the principle axes ($\widehat{\varphi}_0 = 0$). The new transformed RE represented by the following parametric equation:

$$\begin{aligned} x_t(\varphi) &= a \cos \varphi, \\ y_t(\varphi) &= b \sin \varphi, \\ \varphi &\in [0, 2\pi). \end{aligned} \quad (14)$$

Finally, in order to estimate the $a, b > 0$ parameters, which define the two characteristic axes of the RE, we aim at solving the following constrained optimization problem:

$$\begin{aligned} \{\widehat{a}, \widehat{b}\} &= \arg \min_{a,b} \sum_{i=1}^M \sum_{j=1}^N X(i, j) \mu_c^{\text{res}}(i, j) \\ &\cdot \left[\left(x_t(\varphi_{i,j}) - \tilde{x}(i, j) \right)^2 + \left(y_t(\varphi_{i,j}) - \tilde{y}(i, j) \right)^2 \right] \\ &\text{s.t. } 0 < B_l < a, b < B_u, \end{aligned} \quad (15)$$

where $x_t(\varphi)$ and $y_t(\varphi)$ are given in parametric form by (14), so that (x_t, y_t) belongs to the transformed RE (whose parameters a and b we actually estimate) and \tilde{x} and \tilde{y} correspond to transformed coordinates (by (12)) of the image slice data samples. Additionally, $\varphi_{i,j} = \arctan(\tilde{y}(i, j)/\tilde{x}(i, j))$, $i = 1, \dots, M$, $j = 1, \dots, N$, while terms $B_l, B_u > 0$ bound the admissible set parameters from above, in order to make search space smaller. Since the optimization problem (15) is not convex, we solved it by finding the local minimum using standard gradient procedures.

After we obtain the optimal parameters for the RE, we aim at filtering out the nonepicardial fat parts of the resulting cluster. According to the GEP assumption, these are the segments of the resulting fuzzy cluster that are out of the ε neighborhood of the RE. Parameter $\varepsilon > 0$ has been found empirically and has been fixed for all slices and all patients in all experiments. In addition to the obtained RE and ε neighborhood ring, we introduce two arcs to constrain the part of the RE within which the resulting cluster will be marked as desirable (to be retained), while the rest will be undesirable (to be removed), as shown in Figure 6. The angle between the two arcs is defined as $\varphi_C = \varphi_2 - \varphi_1$, $\varphi_1 < \varphi_2$, where φ_1 and φ_2 represent the first and second arc, starting from x -axis, respectively.

In order to determine the φ_1 and φ_2 , that is, the desirable area, we first perform the defuzzification of the resulting cluster. We compute defuzzified crisp cluster $\mathcal{M} = (\mu_{c^{\text{res}}})_\alpha$ by α -cutting the resulting fuzzy cluster $\mu_{c^{\text{res}}}$ with predefined $\alpha \in (0, 1)$. (The parameter α has been found empirically, and it has been fixed for all slices and all patients in all experiments.) As a result, we have $(\mu_{c^{\text{res}}})_\alpha = \{(i, j) | \mu_{c^{\text{res}}}(i, j) \geq \alpha\}$ based on which we obtain the desirable parts \mathcal{M}^* of the resulting fuzzy cluster $\mu_{c^{\text{res}}}$, that is, crisp cluster \mathcal{M} , in which pixels satisfy the following:

$$\mathcal{M}^* = \left\{ (i, j) \in \mathcal{M} | \varphi_{i,j} \in [\varphi_1, \varphi_2], d((i, j), \text{RE}) \leq \varepsilon \right\}, \quad (16)$$

where $\varphi_{i,j} = \arctan(\tilde{y}(i, j)/\tilde{x}(i, j))$ represents the angular polar coordinate corresponding to particular pixel $(i, j) \in \mathcal{M}$ and $d((i, j), \text{RE})$ represents the distance between the pixel $(i, j) \in \mathcal{M}$ and the RE:

$$d((i, j), \text{RE}) = \left| \tilde{\rho}_{i,j} - \rho_{\text{RE}}(\varphi_{i,j}) \right|, \quad (17)$$

where

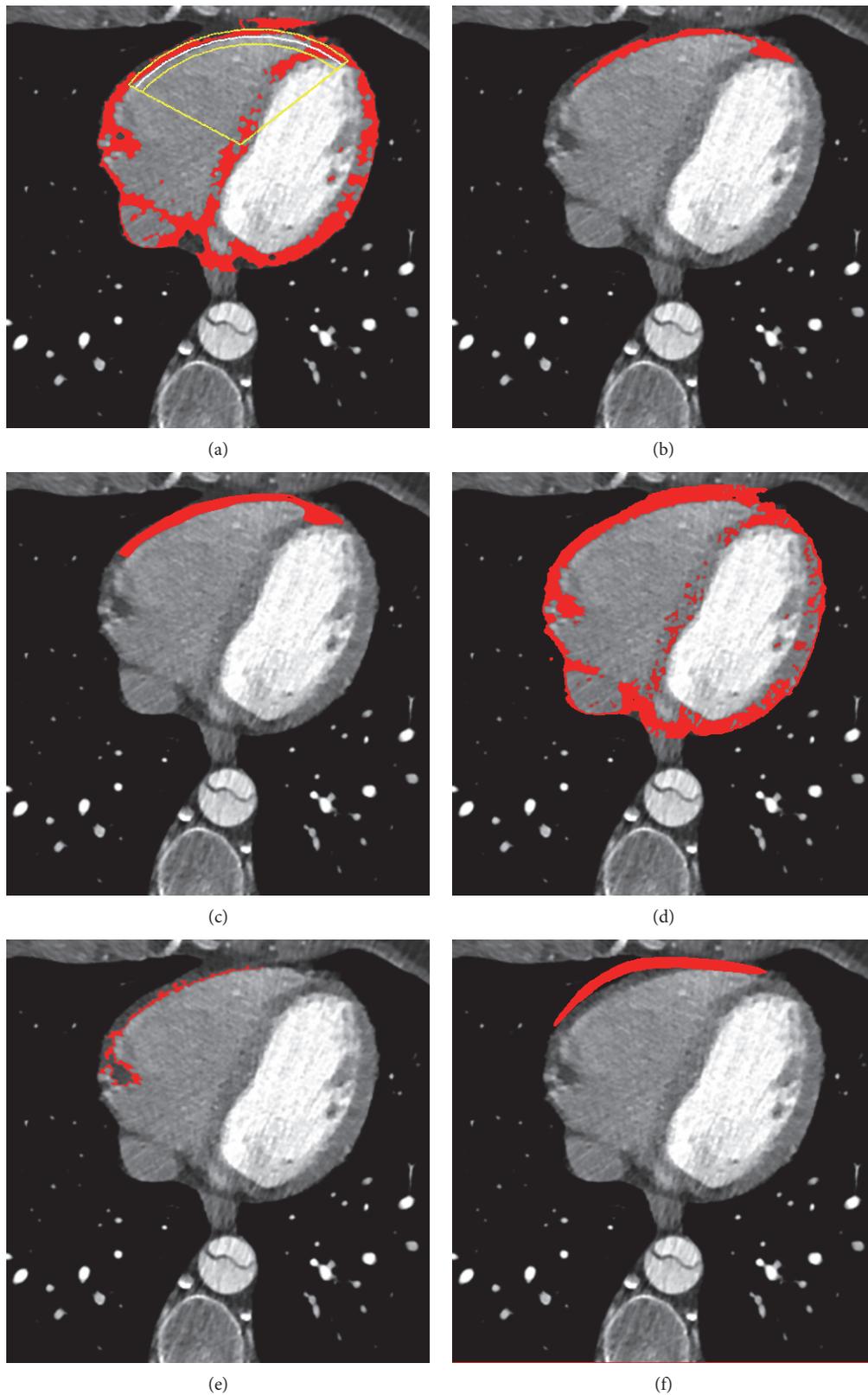


FIGURE 6: Results for one particular CT image slice by the proposed segmentation method in comparison to the manual segmentation obtained by medical doctors and the three reference methods (snake, region growing, and LAMT): (a) α -cut of the resulting cluster; (b) the proposed segmentation approach; (c) manual segmentation performed by medical doctors; (d) LAMT [18]; (e) region growing [21]; (f) snake [20].

$$\tilde{\rho}_{i,j} = \sqrt{\tilde{x}^2(i,j)\cos^2\varphi_{i,j} + \tilde{y}^2(i,j)\sin^2\varphi_{i,j}} \quad (18)$$

represents the radius polar coordinate corresponding to $(i, j) \in \mathcal{M}$ and

$$\rho_{RE}(\varphi_{i,j}) = \sqrt{\hat{a}^2\cos^2\varphi_{i,j} + \hat{b}^2\sin^2\varphi_{i,j}} \quad (19)$$

represents the radius polar coordinate corresponding to point on the RE, with the angular polar coordinate equal to $\varphi_{i,j}$. (\hat{a} and \hat{b} are defined in (15).)

The \mathcal{M}^* actually represents the final 2D segmentation of epicardial fat for particular CT image slice. In our experiments, the angles φ_1 and φ_2 have been determined empirically for particular type of patients and CT imaging parameters. For this purpose, we have used assistance of physicians to set *the control angles* on training set of patients. The control angles are represented by φ_1^l and φ_2^l , $l=1, \dots, L$, which are manually set for reduced number of equidistant control image slices. The actual φ_1 and φ_2 for each CT slice are then determined as a linear interpolation of the existing neighboring control angles. The final φ_1 and φ_2 (for all CT slices) are then used in test phase for segmentation of the data corresponding to same type of patients and CT imaging parameters. As the final postprocessing step, we perform the morphological image closure operation (with the circle-shaped structuring element) of the region \mathcal{M}^* , thus yielding the final result of epicardial fat segmentation on a particular CT image slice.

3. Results and Discussion

We present in this section the results of the proposed semiautomatic epicardial fat segmentation in terms of objective segmentation measure and clinical evaluation. The manual segmentation performed by physicians is used as ground truth data for computation of objective 3D segmentation measures. The “leave one out” cross-validation principle was used in training and testing phase. We compare the segmentation results of the proposed approach to the results obtained by our previous method for epicardial fat segmentation [18] (which we call local adaptive morphology-thresholding (LAMT), snake active contour method [20], and region growing method [21]). (The implementation of the snake and region growing algorithms has been used from open-source Matlab code freely available on <http://www.mathworks.com>.)

We seek the optimal values for the following parameters: $ms1$, $ms2$, α , and ε . Parameter (diameter in the number of pixels) $ms1$ is the size of the disk-shaped structural element for the morphological operation “opening by reconstruction,” used in the preprocessing phase of extracting the heart region (region of interest) from the rest of the image slice. Parameter $ms2$ is the size of the disk-shaped structural element for the morphological operation “dilation,” which follows the “opening by reconstruction” operation, both performed on contrast-enhanced input image, as explained in Section 2.2.

The parameter $\alpha \in (0, 1)$ is used for obtaining the α -cut in the defuzzification phase, as the part of obtaining the final segmentation result for the particular image slice, which is described in Section 2.3.2. Parameter ε , given in the number of pixels, defines the neighborhood of geometric ellipse used in postfiltering phase, described in Section 2.3.2. In order to obtain optimal values for the mentioned parameters, we form the following discrete combination set: $ms1 \in \{30, 35, 40, 45, 50, 55, 60\}$, $ms2 \in \{5, 10, 15, 20, 25, 30\}$, $\alpha \in \{0.4, 0.5, 0.6, 0.7, 0.8\}$, and $\varepsilon \in \{5, 10, 15, 20, 25, 30\}$. On the previously described combination set (passing all possible values), and along the lines of “leave one out” cross-validation training/testing procedure, we evaluate the cost function. The cost function, for this purpose, is defined as the average of the Dice similarity coefficients between the processed and the manually segmented (by the medical expert) ground truth image slices belonging to the training set of patient. As we use “leave one out” cross-validation procedure, in each round of the procedure, we randomly chose, by uniform distribution, one patient for testing and the rest to be in the training set. The described cost is then evaluated on all CT image slices belonging to the training set. We obtained “the best” values of those parameters as $(ms1, ms2, \alpha, \varepsilon) = (40, 20, 0.7, 20)$.

The number L of equidistant control CT image slices for obtaining the control bordering angles φ_1 and φ_2 is chosen to be large enough, but still small enough as well, not to be too demanding and time consuming for the medical expert. The medical experts had to label those L slices for each patient, and as a result, the bordering control angles φ_1 and φ_2 were obtained, along with the previously mentioned $(ms1, ms2, \alpha, \varepsilon)$ within the “leave one out” cross-validation training/testing procedure. Namely, in each round of the cross-validation procedure, for each $l=1, \dots, L$, we obtain the resulting bordering control angles $\varphi_1^{(l)}$ and $\varphi_2^{(l)}$, by simple averaging over all patients chosen to be in the training set in that particular round. Then, the testing is performed, for that particular round. The final control angles are chosen to be those obtaining the lowest cost (the average of the Dice similarity coefficients between the processed and the manually segmented slices) of all rounds. As stated in Section 2.3.2, for all other patients and CT slices not belonging to the resulting bordering control slices, the bordering angles φ_1 and φ_2 are determined by interpolation of the existing neighboring control angles. Finally, the lower and upper bounds B_l and B_u , used in (15), are obtained empirically and set to $B_l = 200$ and $B_u = 450$ in all our experiments.

In Figures 6(a), 7(a), and 8(a), we show, for three different examples of the input CT slices, the α -cut of the resulting fuzzy cluster along with the ε , φ_1 , and φ_2 neighborhood of RE corresponding to the desirable area, with applied postprocessing step of morphological operation. In Figures 6(a), 7(a), and 8(a), red color-marked region denotes resulting fuzzy cluster that is subsequently fitted by the proposed RE. Additionally, the proposed α -cut RE is shown in the same figures with yellow and white lines, where white line corresponds to the α -cut part of the estimated RE and yellow lines indicate ε neighborhood and angular bandwidth of the selected part.

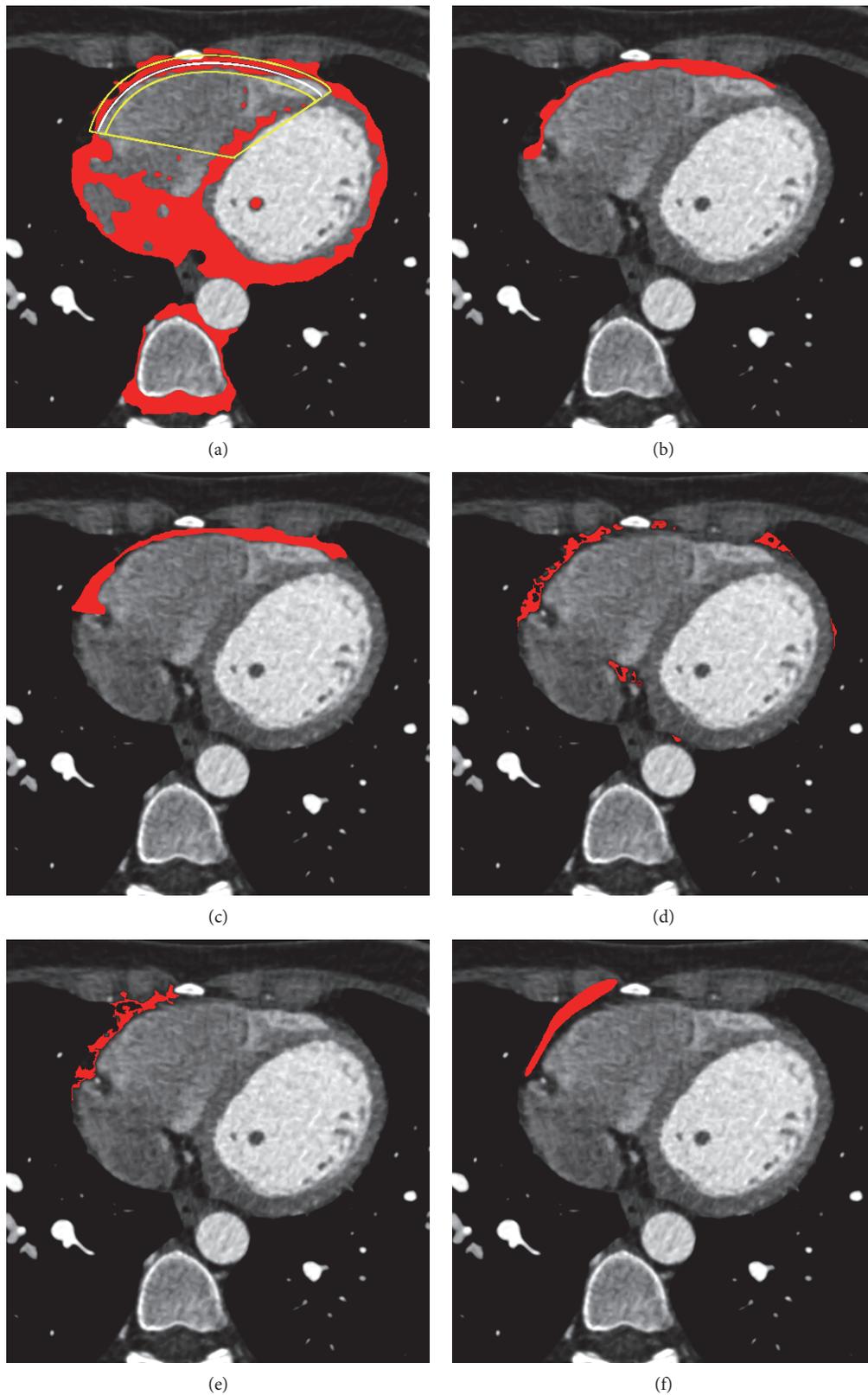


FIGURE 7: Results for one particular CT image slice by the proposed segmentation method in comparison to the manual segmentation obtained by medical doctors and the three reference methods (snake, region growing, and LAMT): (a) α -cut of the resulting cluster; (b) the proposed segmentation approach; (c) manual segmentation performed by medical doctors; (d) LAMT [18]; (e) region growing [21]; (f) snake [20].

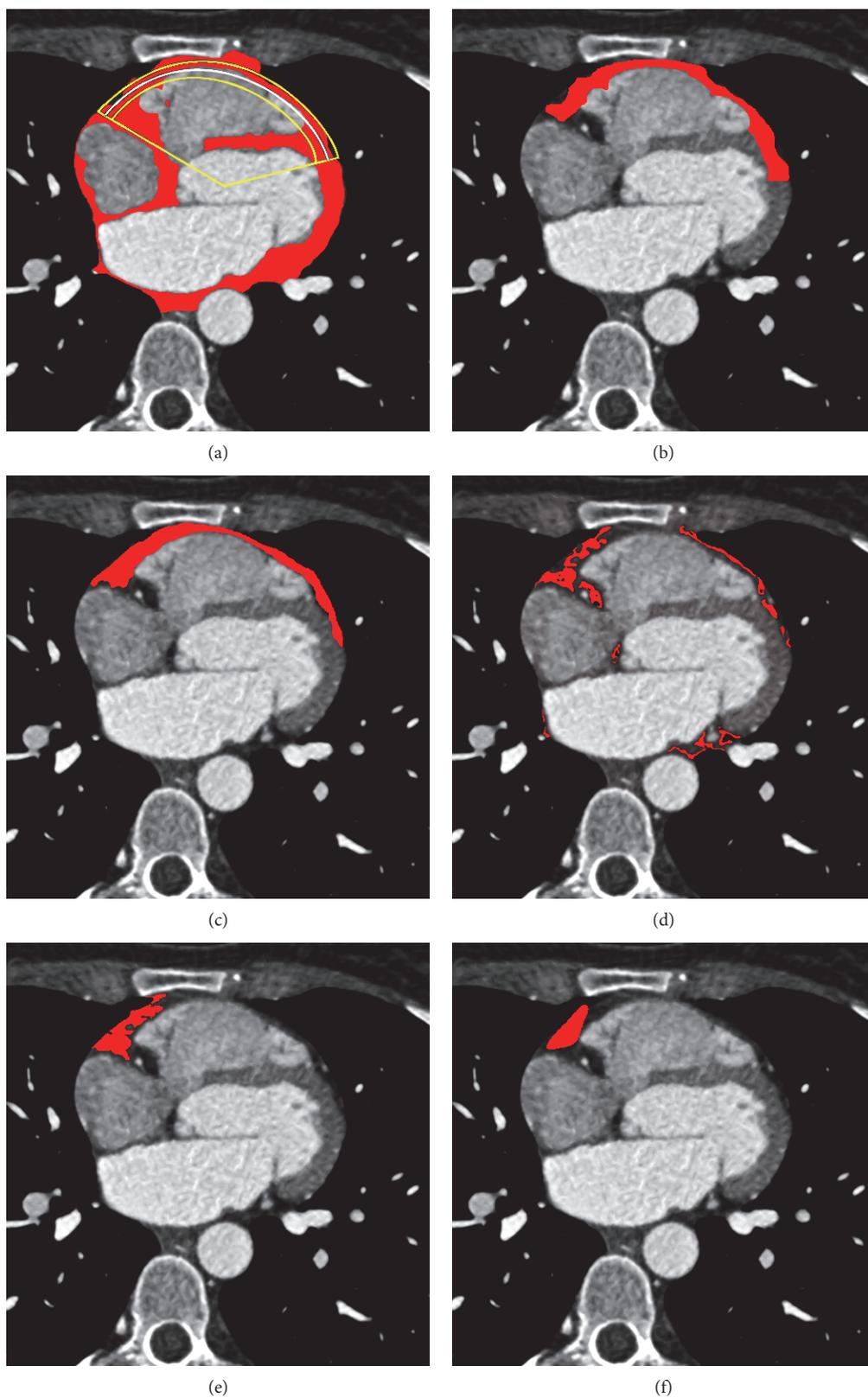


FIGURE 8: Results for one particular CT image slice by the proposed segmentation method in comparison to the manual segmentation obtained by medical doctors and the three reference methods (snake, region growing, and LAMT): (a) α -cut of the resulting cluster; (b) the proposed segmentation approach; (c) manual segmentation performed by medical doctors; (d) LAMT [18]; (e) region growing [21]; (f) snake [20].

TABLE 1: Comparison of methods using Dice similarity coefficient with manually segmented ground truth images (approved by a cardiovascular surgeon). Our proposed method outperforms the LAMT method, the classical region growing, and the snake methods.

Patient	1	2	3	4	5	6	7	8
Proposed method	0.76	0.65	0.6	0.7	0.73	0.54	0.73	0.82
LAMT	0.3	0.32	0.38	0.24	0.5	0.2	0.3	0.07
Region growing	0.39	0.15	0.34	0.34	0.28	0.24	0.39	0.17
Snake	0.3	0.15	0.16	0.32	0.36	0.29	0.38	0.3

TABLE 2: Volume measurement error between our proposed method and manually segmented ground truth images.

Patient	1	2	3	4	5	6	7	8
Normalized volume (voxels)	1.1	0.73	0.43	0.68	0.96	0.95	1.38	0.91
Volume relative error (%)	1.5	15	14.5	29.6	33.8	8.7	14.6	9.5
Coronary artery disease (vessels)	3	2	3	2	3	2	3	3
Hypertension ejection	No	Yes	No	Yes	No	Yes	No	Yes
Fraction (vessels)	60	37	26	63	42	55	56	60

The segmentation results for a single CT slices are further shown in Figures 6, 7, and 8, where the proposed segmentation approach is compared to manual segmentation (done by medical doctors) and the reference three baseline methods. Red color-marked regions in the figures denote resulting segmented epicardial fat regions, obtained by applying particular method. It can be observed that the proposed method outperforms the three baseline methods ([18, 20, 21]), in terms of lower number of erroneously detected regions and higher degree of connectivity of the detected region, relatively to the reference (manually segmented ground truth).

In order to compare the segmentation results of our method to the ground truth, we use Dice similarity coefficient [23], which is the set similarity measure defined for two sets A and B (in our case, set of voxels belonging to segmented images and the ground truth image) as $D(A, B) = (2|A \cap B|)/(|A| + |B|)$, where $|\cdot|$ of voxels within the particular set.

In Table 1, we show the results for the computed Dice coefficients, by which we measure the difference between automatic and manually obtained segmentation (approved by the medical expert) for the whole 3D image/volume, for the proposed and three reference segmentation methods, namely, snake active contour method [20], region growing method [21], and the LAMT method [18]. It can be observed that the results of our proposed method yield high Dice coefficient values and that are significantly higher than the Dice coefficients obtained for the reference methods that we compared to. We also note that the proposed method is less time consuming for the medical expert in comparison to the reference methods. Namely, snake active contour method requires initial contour to be delivered for every CT image slice; region growing method demands coordinates of at least one patch to be delivered for each CT slice; and LAMT method requires coordinates of four patches to be provided.

In order to draw clinical conclusions from the measured epicardial fat volumes, we introduce *Normalized Volume* (NV) of epicardial fat as a ratio of the epicardial fat volume

and weight of the patient. The calculated NV values are given in Table 2 together with relative error of epicardial fat measured by comparing the segmented (by the proposed method) and manually annotated volume. The results of relative error calculation show acceptable error rates (less or equal to 15%) with two outliers.

Our experiments indicate higher NV values for 3-vessel coronary artery disease than in case of 2-vessel coronary artery disease with average values 0.96 and 0.75, respectively. Also, patients with hypertension have exhibited slightly lower NV values than patients with normal blood pressure, with average values 0.82 and 0.96, respectively. Out of all test subjects, 3 patients had reduced ejection fraction (below 42%). However, the reduced ejection fraction does not correlate well enough with calculated NV fat volume.

4. Conclusion

Segmentation and quantification of epicardial fat in cardiac CT images play an important role in assisting physicians in diagnosing cardiovascular diseases and progressing their outcomes. The higher level of automation of this process helps to reduce interobserver variability and shorten execution times. In this work, a novel epicardial fat segmentation scheme is proposed. The method is based on local adaptive morphology and fuzzy c-means clustering, where the only user input is the coordinates of the particular image patch belonging to the target epicardial fat area. Additionally, geometric ellipse prior is used in order to filter out the undesired parts of the target cluster. Experimental results presented demonstrate the good correspondence between the proposed method and the manual ground truth segmentation.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Acknowledgments

This paper has been partially supported by the Provincial Secretariat for Science and Technological Development of the Autonomous Province of Vojvodina (Serbia) (Grant no. 114-451-873/2015-01).

References

- [1] R. Shahzad, D. Bos, C. Metz et al., "Automatic quantification of epicardial fat volume on non-enhanced cardiac CT scans using a multi-atlas segmentation approach," *Medical Physics*, vol. 40, no. 9, article 091910, 2013.
- [2] A. Rosito, J. M. Massaro, U. Hoffman et al., "Pericardial fat, visceral abdominal fat, cardiovascular disease risk factors, and cardiovascular calcification in a community-based sample: the Framingham Heart Study," *Circulation*, vol. 117, no. 5, pp. 605–613, 2008.
- [3] A. A. Mahabadi, M. H. Berg, N. Lehmann et al., "Association of epicardial fat with cardiovascular risk factors and incident myocardial infarction in the general population: the Heinz Nixdorf Recall study," *Journal of the American College of Cardiology*, vol. 61, no. 13, pp. 1388–1395, 2013.
- [4] P. M. Gorter, A. M. de Vos, Y. Graaf et al., "Relation of epicardial and pericoronary fat to coronary atherosclerosis and coronary artery calcium in patients undergoing coronary angiography," *The American Journal of Cardiology*, vol. 102, no. 4, pp. 380–385, 2008.
- [5] S. Dabbah, H. Komarov, A. Marmor, and N. Assy, "Epicardial fat, rather than pericardial fat, is independently associated with diastolic filling in subjects without apparent heart disease," *Nutrition Metabolism and Cardiovascular Diseases*, vol. 24, no. 8, pp. 877–882, 2014.
- [6] U. Hoffmann, C. L. Schlett, M. Ferencik et al., "Association of pericardial fat and coronary high-risk lesions as determined by cardiac CT," *Atherosclerosis*, vol. 222, no. 1, pp. 129–134, 2012.
- [7] T. E. Brinkley, F.-C. Hsu, J. J. Carr et al., "Pericardial fat is associated with carotid stiffness in the multi-ethnic study of atherosclerosis," *Nutrition Metabolism and Cardiovascular Diseases*, vol. 21, no. 5, pp. 332–338, 2011.
- [8] A. Yerramasua, D. Dey, S. Venuraju et al., "Increased volume of epicardial fat is an independent risk factor for accelerated progression of sub-clinical coronary atherosclerosis," *Atherosclerosis*, vol. 220, no. 1, pp. 223–230, 2012.
- [9] T. Choi, N. Ahmadi, S. Sourayanezhad, I. Zeb, and M. J. Budoff, "Relation of vascular stiffness with epicardial and pericardial adipose tissues, and coronary atherosclerosis," *Atherosclerosis*, vol. 229, no. 1, pp. 118–123, 2013.
- [10] R. Sicari, A. M. Sironi, R. Petz et al., "Pericardial rather than epicardial fat is a cardiometabolic risk marker: an MRI vs Echo study," *Journal of the American Society of Echocardiography*, vol. 24, no. 10, pp. 1156–1162, 2011.
- [11] G. Iacobellis, M. C. Ribaud, F. Assael et al., "Echocardiographic epicardial adipose tissue is related to anthropometric and clinical parameters of metabolic syndrome: a new indicator of cardiovascular risk," *Journal of Clinical Endocrinology and Metabolism*, vol. 88, no. 11, pp. 5163–5168, 2003.
- [12] T. S. Polonsky, R. L. McClelland, N. W. Jorgensen et al., "Coronary artery calcium score and risk classification for coronary heart disease prediction," *Journal of the American Medical Association*, vol. 303, no. 16, pp. 1610–1616, 2010.
- [13] G. Coppini, R. Favilla, P. Marraccini, D. Moroni, and G. Pieri, "Quantification of epicardial fat by cardiac CT imaging," *The Open Medical Informatics Journal*, vol. 4, pp. 126–135, 2010.
- [14] A. Pednekar, "Automatic segmentation of abdominal fat from CT data," in *Application of Computer Vision*, pp. 308–315, Breckenridge, CO, USA, January 2005, IEEE.
- [15] I. Kakadiaris, A. Bandekar, K. Gul et al., "Automated pericardial fat quantification in CT data," *Proceedings of Journal of the American College of Cardiology*, vol. 47, no. 4, p. 264, 2006.
- [16] J. G. Barbosa, B. Figueiredo, N. Bittencourt, and J. M. R. S. Tavares, "Towards automatic quantification of the epicardial fat in non-contrasted CT images," *Computer Methods in Biomechanics and Biomedical Engineering*, vol. 14, no. 10, pp. 905–914, 2011.
- [17] D. Dey, Y. Suzuki, S. Suzuki et al., "Automated quantitation of pericardial fat from noncontrast CT," *Investigative Radiology*, vol. 43, no. 2, pp. 145–153, 2008.
- [18] V. Zlokolica, L. Velicki, M. Janev et al., "Epicardial fat registration by local adaptive morphology-thresholding based 2D segmentation," in *2014 56th International Symposium ELMAR (ELMAR)*, pp. 1–4, Zadar, Croatia, September 2014.
- [19] V. Zlokolica, L. Velicki, B. Banjac et al., "Epicardial fat segmentation optimization based on structural prior knowledge and subjective-objective correspondence," in *2015 IEEE 15th International Conference on Bioinformatics and Bioengineering (BIBE)*, pp. 1–6, Belgrade, Serbia, November 2015.
- [20] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: active contour models," *International Journal of Computer Vision*, vol. 1, no. 4, pp. 321–331, 1988.
- [21] R. C. Gonzalez and R. E. Woods, *Digital Image Processing 2nd Edition*, Prentice Hall, New Jersey, 2002.
- [22] G. J. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic: Theory and Applications*, UK, London, 1995.
- [23] L. R. Dice, "Measures of the amount of ecologic association between species," *Ecology*, vol. 26, no. 3, pp. 297–302, 1945.

Research Article

Frame-Based CT Image Reconstruction via the Balanced Approach

Weifeng Zhou^{1,2} and Hua Xiang³

¹College of Mathematics and Physics, Qingdao Science and Technology University, Qingdao, Shandong 266071, China

²School of Computer Science and Technology, Shandong University, Jinan, Shandong 250101, China

³College of Cooperation, Qingdao Agricultural University, Qingdao, Shandong 266109, China

Correspondence should be addressed to Weifeng Zhou; wfzhouqd@163.com

Received 22 April 2017; Accepted 15 August 2017; Published 17 September 2017

Academic Editor: Chengpo Mu

Copyright © 2017 Weifeng Zhou and Hua Xiang. 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.

Frame-based regularization method as one kind of sparsity representation method has been developed in recent years and has been proved to be an efficient method for CT image reconstruction. However, most of the developed CT image reconstruction methods are analysis-based frame methods. This paper proposes a novel frame-based balanced hybrid model with two sparse regularization terms for CT image reconstruction. We generalize the fast alternating direction method to solve the proposed model so that every subproblem can be easily solved. The numerical experiments suggest that the proposed hybrid balanced-based wavelet regularization scheme is efficient in terms of reducing the defined reconstruction root mean squared error and improving the signal to noise ratio in CT image reconstruction.

1. Introduction

X-ray computed tomography (CT) image reconstruction is an indispensable tool for diagnosing diseases and research requirements. However, X-ray radiation is harmful, and high-dose X-ray radiation may induce genetic mutation, cell canceration, and so on [1]. Therefore, more and more attentions are paid on the low-dose X-ray CT image reconstruction. Since X-ray CT imaging quality depends on the X-ray dose, reducing the X-ray dose will result to poor image reconstruction quality. Consequently, how to decrease the X-ray dose not to affect the diagnosis is a hot topic in recent years. Mathematically, CT image reconstruction often can be formulated as a linear inverse problem. For the detected measurements data b , the objective is to find the targeted image u from the following equation:

$$b = Au + \varepsilon, \quad (1)$$

where A is the discrete radon linear transform operator and ε denotes the noise with variance δ .

One strategy for X-ray radiation-dose reduction is to reduce the projection data, but this few-view method will result in insufficient data. As a result of the undersampling of this strategy and the system errors, the above-mentioned problem (1) is usually ill-posed from the mathematical point of view. Therefore, traditional filtered back projection (FBP) method [2] cannot yield desirable imaging quality. Sparse regularization methods are developed in recent years to overcome the ill-posedness of these problems, and moreover, these methods can acquire higher quality images in few-view circumstance. The sparse representation methods assume that images are sparse in some transformed domains. Discrete gradient [3], that is, the so-called total variation (TV), is such a sparse transform domain to solve the few-view image reconstruction problems. Although the TV-based regularization sparse domain is useful in reducing radiation dose [4, 5], there are also some shortcomings, for example, the staircase effect [6], and the power of it is still limited [7]. Therefore, many improved methods are proposed, such as PWLS-TGV proposed by Niu et al. [8], TVS-POCS proposed by Liu et al. [9], and the method proposed by Ritschl et al.

[10]. Besides, some other sparse transform domains are developed, such as the gamma regularization-based method [11], some nonlocal domains [12, 13], different kinds of wavelet frame domains [14–16], and some dictionary learning sparse methods [17, 18].

This paper mainly considers the sparse representation by wavelet tight frame. Wavelet tight frame can ensure the given signal be perfectly represented as a linear combination of the sparse wavelet coefficients which is also called the perfect reconstruction property [19]. Due to the flexibility of decomposition and reconstruction and the well performance, wavelet tight frame-based methods have been widely used in almost every branch of image processing [14, 20–22]. In recent years, most of the developed wavelet tight frame methods for CT image reconstruction are analysis-based wavelet frame methods. This paper proposes a novel balanced-based method. The remainder of this paper is organized as follows. In Section 2, we present some necessary preliminaries and theories about the wavelet tight frame. Section 3 presents the proposed balanced model for CT image reconstruction and develops the efficient computational algorithm for solving the proposed strategy. Numerical simulations to demonstrate the improvement of our proposed method in terms of RMSE and PSNR are given in Section 4. In the end, we make the conclusion in Section 5.

2. Preliminaries

For convenience, we first present some basic definitions and some results of the wavelet tight frame used in the proposed model. More details can be seen in [20, 21].

In the 1D discrete circumstance, a set of vectors $\{x_i \in \mathbb{R}^M\}_{i=1}^N$ ($N \geq M$) is called a wavelet tight frame if for each $b \in \mathbb{R}^M$,

$$\sum_{i=1}^N |\langle x_i, b \rangle|^2 = \|b\|_2^2, \quad (2)$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product. The corresponding analysis operator denoted as W is written as

$$W = [x_1, x_2, \dots, x_N]^T. \quad (3)$$

Then, $Wb = \{\langle b, x_i \rangle\}_{i=1}^N$ is called the wavelet coefficients. Another operator W^T which is usually called the synthesis operator is the synthesis of the wavelet coefficients, that is, if c denotes the wavelet coefficients, then

$$W^T c = \sum_{i=1}^N c(i)x_i. \quad (4)$$

Then

$$W^T W = I_M, \quad (5)$$

which can be derived by the identity (2). Here, $I_M : \mathbb{R}^M \rightarrow \mathbb{R}^M$ is the identity operator. This property is often called the “perfect reconstruction property,” which can reduce the calculation amount in some applications.

Choose: parameters $\gamma > 0, \mu > 0, \nu > 0, \text{tol}$;
Input: iteration initial values x^0, f^0, z^0 ;
for $k = 1, 2, 3, \dots, n$, or $\|u^{k+1} - u^k\|_2 / \|u^k\|_2 \leq \text{tol}$
 $u^{k+1} = (A^T A + \gamma + \mu)^{-1} [A^T b + \gamma(W^T x^k - b)];$
 $f^{k+1} = A u^{k+1} - b + f^k;$
 $x^{k+1} = \text{Shrink}_{\nu/\gamma} W(u^{k+1} + b);$
 $z^{k+1} = z^k + u^{k+1} - W^T x^{k+1};$
end
Output: μ

ALGORITHM 1: Balanced-based wavelet framelet approach for CT image reconstruction.

Based on the Unitary Extension Principle (UEP) condition [20], the wavelet tight frame often can be generated by some filters $\{a_i\}_{i=0}^{r-1}$ that satisfy

$$\sum_{i=0}^{r-1} \sum_{n \in \mathbb{Z}} a_i(l+n)a_i(n) = \delta_l, \quad \forall l \in \mathbb{Z}. \quad (6)$$

Here, $\delta_l = 1$ when $l = 0$; otherwise it is zero. Piecewise linear B-spline framelet is such a wavelet tight frame whose associated filters are

$$\begin{aligned} a_0 &= \frac{1}{4} [1, 2, 1], \\ a_1 &= \frac{\sqrt{2}}{4} [1, 0, -1], \\ a_2 &= \frac{1}{4} [-1, 2, -1]. \end{aligned} \quad (7)$$

The wavelet-based method has been used in almost every image processing branch [22–24]. Two-dimensional wavelet tight frame filters can be obtained by the tensor product of the corresponding one-dimensional filters.

Given the observed data b ; the sparse regularization methods for image processing based on the wavelet tight frame can be summarized as

$$x^* = \arg \min_x \frac{1}{2} \|A W^T x - b\|_2^2 + \frac{\kappa}{2} \|(I - W W^T)x\|_2^2 + \lambda \|x\|_p, \quad (8)$$

where A denotes a linear transform operator, which is a discrete radon transform in CT image reconstruction, a Fourier transform in MR image reconstruction, and a convolution operator in image deblurring. $\|\cdot\|_p$ denotes the p -norm so as to finally obtain a sparse solution. $\|\cdot\|_1$ as the approximation of $\|\cdot\|_0$ is often used to realize the sparse regularization. Compared with $\|x\|_0$, $\|x\|_1$ is convex so that the well-posed property can be guaranteed. Then, the target image $u^* = W^T x^*$. Here, the middle term in (8), that is, $(\kappa/2) \|(I - W W^T)x\|_2^2$, is used to balance the distance between the target image $W^T x$ and the coefficient x . Then, in terms of different values of κ , three approaches are distinguished, that is, the analysis approach ($\kappa = \infty$), the balanced approach ($0 < \kappa < \infty$), and the synthesis approach ($\kappa = 0$).

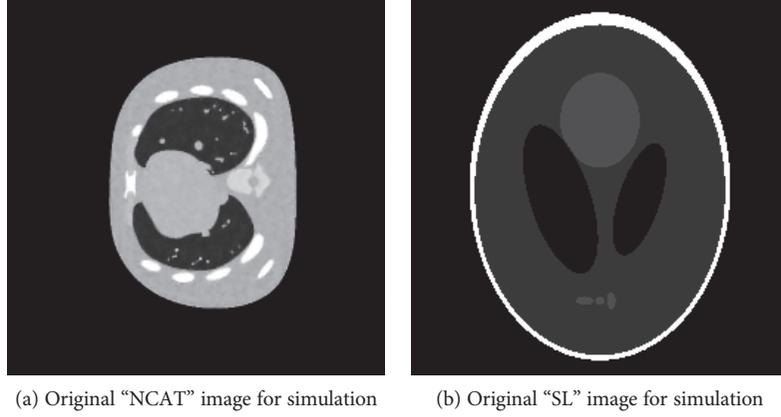


FIGURE 1: Original images used for simulation.

Obviously, the three approaches are the same when W is orthogonal. Generally speaking, it is difficult to make a conclusion that which approach among the three approaches described in (8) is better. Every approach has its own favorite image sets [25].

For the balanced-based approach, we have the following result [19].

Lemma 1. *Let W and W^T , respectively, denote the analysis operator and the synthesis operator of a wavelet tight frame; then $\|(I - WW^T)x\|_2^2 + \|W^T x - b\|_2^2 = \|x - Wb\|_2^2$ holds.*

3. The Proposed Model and Method

In recent years, most of the developed methods for CT image reconstruction are analysis-based frame methods [16, 21, 26], and many state-of-the-art methods such as Split Bregman method [27], alternative direction [28], and augmented Lagrangian method [29] are implemented to solve these problems. The researches of image reconstruction modeled based on balanced method [25] and synthesis method [30] are relatively few. Since every approach has its own favorite image sets, it is difficult to make a conclusion which one is better [25].

This paper proposes a novel-constrained balanced-based model CT image reconstruction as follows:

$$\begin{aligned} \min_x \frac{\kappa}{2} \|(I - WW^T)x\|_2^2 + \frac{\alpha}{2} \|W^T x\|_2^2 + \beta \|x\|_1, \\ \text{s.t. } \|AW^T x - b\|_2^2 \leq \delta^2, \end{aligned} \quad (9)$$

where A denotes a radon transform operator, b is the obtained data from the scanner, and W is the wavelet tight frame. Then, the reconstructed image $u^* = W^T x^*$, where x^* denotes the solution of (9). Compared with the model in (8), the added term $(\alpha/2)\|W^T x\|_2^2$ is used to regularize the solution further and avoid the Gibbs defects bring from the wavelet tight framelet. In general, the parameter α is chosen to be smaller than the parameter β . Actually, bigger α will result in the overall smoothness of the results' image. In recent years, some efficient algorithms are developed for the

balanced-based models, for example, proximal linearized alternating direction method in which the linearization of quadratic term of the augmented function for the model with one regularization term was used [31].

Next, let us investigate the corresponding flexible iteration algorithm for our proposed balanced-based model. We also can use the fast alternating direction method for solving our proposed balanced-based model. The corresponding convergence analysis of the alternating direction method can resort to the [32, 33]. By introducing $W^T x = u$, the constrained minimization problem (9) can be changed into the following unconstrained one:

$$\begin{aligned} \min_{u,x,z,f} \frac{\iota}{2} \|Au - b + f\|_2^2 + \frac{\kappa}{2} \|(I - WW^T)x\|_2^2 \\ + \frac{\delta}{2} \|u - W^T x + z\|_2^2 + \frac{\alpha}{2} \|u\|_2^2 + \lambda \|x\|_1. \end{aligned} \quad (10)$$

Actually, (10) is equivalent to the following:

$$\begin{aligned} \min_{u,x,z,f} \frac{1}{2} \|Au - b + f\|_2^2 + \frac{\tau}{2} \|(I - WW^T)x\|_2^2 \\ + \frac{\gamma}{2} \|u - W^T x + z\|_2^2 + \frac{\mu}{2} \|u\|_2^2 + \nu \|x\|_1, \end{aligned} \quad (11)$$

where τ , γ , and μ are parameters. Here, we omit the relationship between these parameters and the parameters in (10).

Here, we choose $\tau = \gamma$ for the convenience of calculation. Then, based on Lemma 1, (11) is equivalent to the following:

$$\min_{u,x,z,f} \frac{1}{2} \|Au - b + f\|_2^2 + \frac{\gamma}{2} \|x - W(u + z)\|_2^2 + \frac{\mu}{2} \|u\|_2^2 + \nu \|x\|_1. \quad (12)$$

Then by the alternative direction method, the minimization problem (12) can be decomposed into the following four subproblems:

$$u^{k+1} = \arg \min_u \frac{1}{2} \|Au - b + f^k\|_2^2 + \frac{\gamma}{2} \|u - W^T x^k + z^k\|_2^2 + \frac{\mu}{2} \|u\|_2^2, \quad (13)$$

$$f^{k+1} = Au^{k+1} - b + f^k, \quad (14)$$

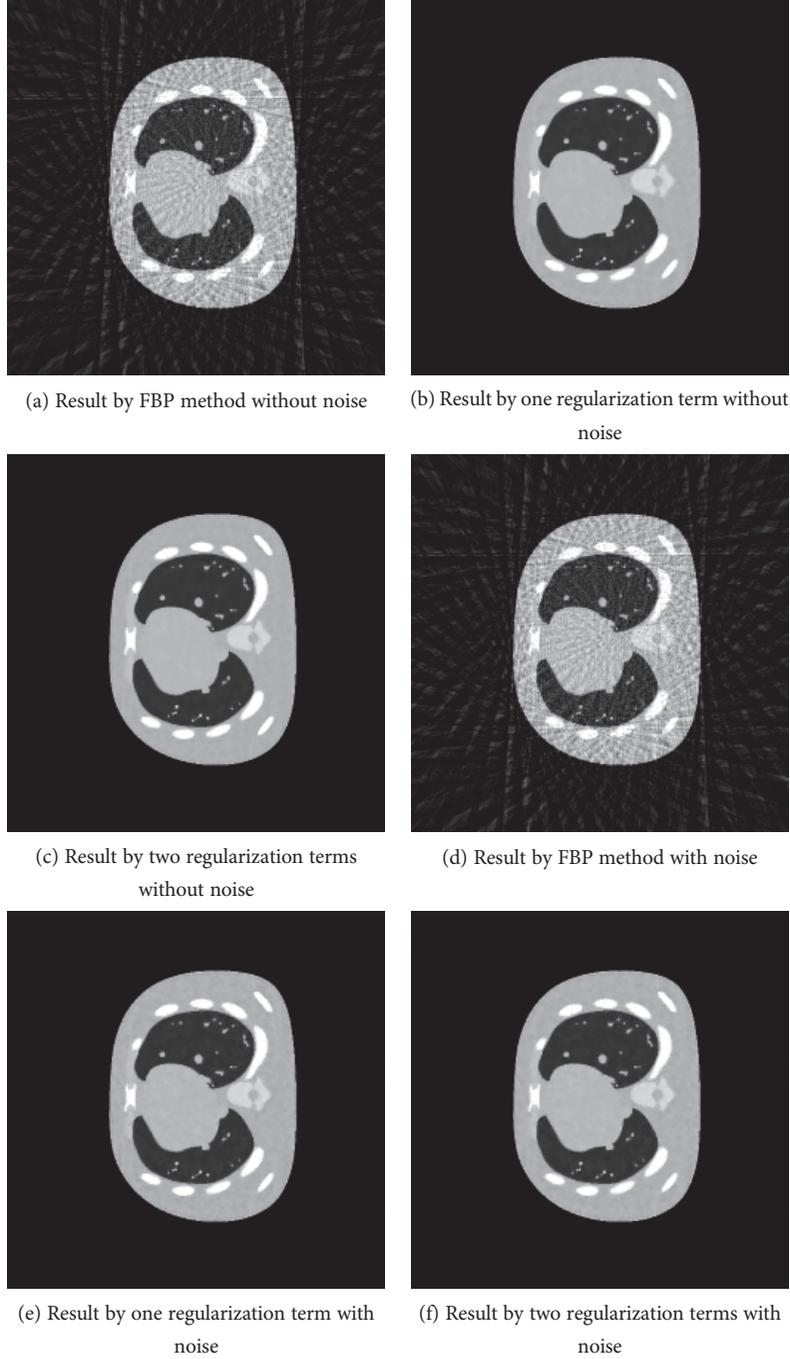


FIGURE 2: Recovered results by FBP method and models with two regularization terms and one regularization term based on our proposed balanced approach with 50 projection views.

$$x^{k+1} = \arg \min_x \frac{\gamma}{2} \|u^{k+1} - W^T x + z^k\|_2^2 + \nu \|x\|_1, \quad (15)$$

$$z^{k+1} = z^k + u^{k+1} - W^T x^{k+1}. \quad (16)$$

As for the first subproblem, by the KKT condition, $u^{k+1} = (A^T A + \gamma + \mu)^{-1} [A^T b + \gamma(W^T x^k - b)]$ can be easily solved by the well-known conjugate gradient (CG) method. On one hand, the added term $(\mu/2)\|u\|_2^2$ in (12) can regularize the solution further; on the other hand, we can see from

the solution of the u -subproblem (13) that parameter μ can further overcome the ill-conditioning of the operator A in CT scanning. Then, we summarize our proposed algorithm based on balanced wavelet tight frame approach for CT image reconstruction as in Algorithm 1.

4. Numerical Simulations

In this section, two numerical studies are presented to illustrate the well performance of our proposed scheme.

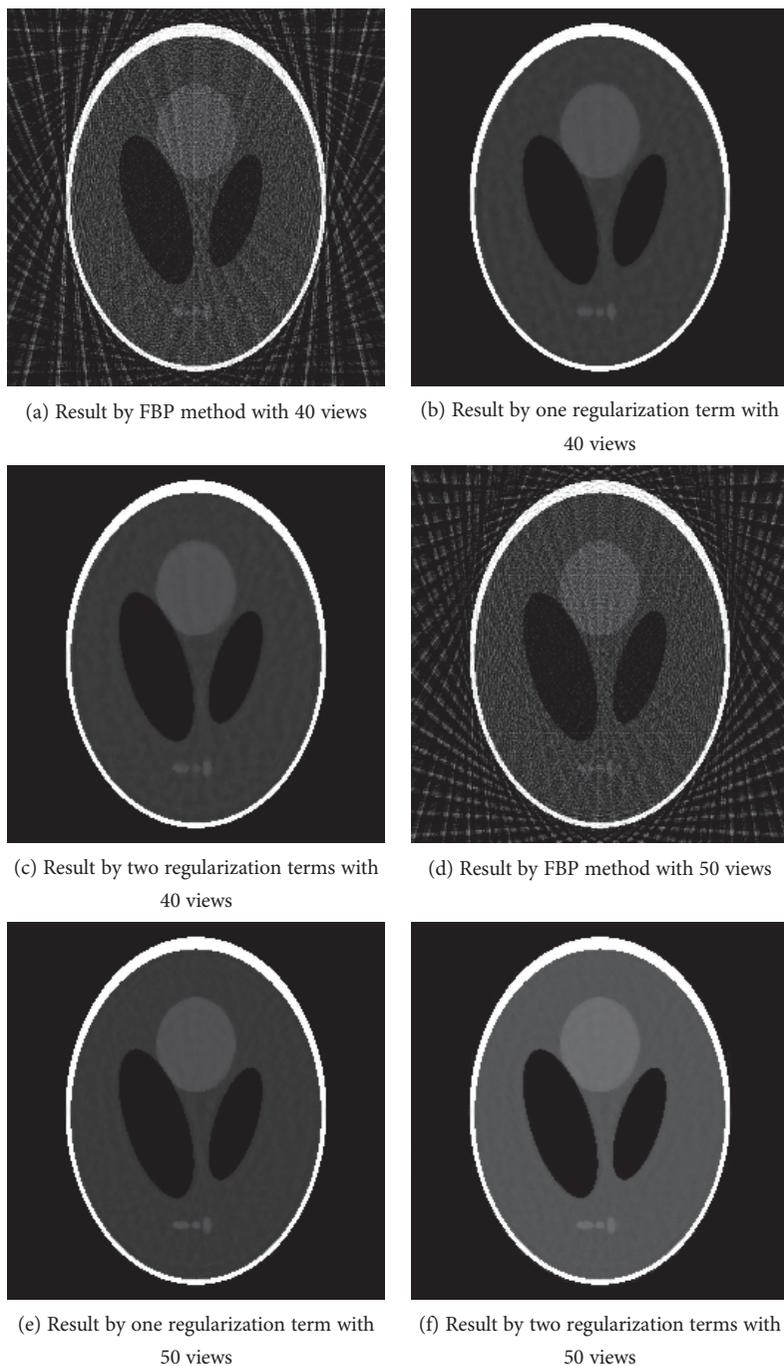


FIGURE 3: Recovered results by FBP method and models with two regularization terms and one regularization term based on our proposed balanced approach under different projection views.

We compare our proposed hybrid regularization scheme ($\mu \neq 0$) in (12) with the traditional FBP algorithm and the one with only one regularization term ($\mu = 0$) by our proposed balanced-based frame algorithm. In the following two simulations, we adapt the piecewise linear B-spline tight frame as the transform operator. For evaluating the quality of the reconstructed images, the following root mean squared error (RMSE) and the peak value signal to noise ratio (PSNR) are used:

$$\begin{aligned} \text{RMSE} &= \frac{\|u - f\|_2}{\|f\|_2}, \\ \text{PSNR} &= 20 * \log_{10} \frac{255}{(1/MN) * \|u - f\|_2}. \end{aligned} \quad (17)$$

Here, u and f denote the reconstructed image and the original image, respectively, and M and N denote the size of the image. Generally speaking, smaller RMSE means better reconstructed quality and higher PSNR means more closer to

the original images. Some other appraisal criteria can also be used, such as the SSIM [34]. The iteration is stopped when $\|u^{k+1} - u^k\|_2 / \|u^k\|_2 \leq 10^{-3}$. In these experiments, we use GPU (graphics processing units) to accelerate the computation of the A and A^T for more fast reconstruction [35]. In the following simulations, the projection views were equally distributed over 360° and detector offset was not to be considered. All the experiments are performed by Matlab 2009 on the PC with 64-bit operating system and 2.90 GHz processor. In addition, after a great deal of our simulations, we found that the parameter μ should be much smaller than the parameter ν , since larger μ will result in the overall smoothness of the reconstruction image. The parameter γ can be set as 5 or 6 when the noise and the projection views are not changed greatly. The stronger the noise, the greater the parameter ν should be used. Next, we will present two concrete examples.

Example 4.1. We use “NCAT” shown in Figure 1(a) to evaluate the proposed model and algorithm. In recent years, “NCAT” has been widely used to evaluate medical imaging technology. The reconstructions RMSE and PSNR by FBP method (Figure 2(a)) with 50 projection views are 0.2701 and 69.0921, respectively. Figures 2(b) and 2(c) are, respectively, the recovered results with 50 projection views by models with one regularization ($\mu = 0$) and two regularizations ($\mu \neq 0$) in (12) based on our proposed balanced approach with algorithm parameters $\gamma = 6$, $\nu = 0.2$, and $\mu = 0.01$. These parameters are optimized. 23 iterations are carried out with about 3.6 s. The values of RMSE and PSNR are 0.0271 and 89.0138, respectively, for two regularization term balanced algorithm and 0.0305 and 87.9635, respectively, for the corresponding one regularization term balanced algorithm. We can see that in terms of RMSE, 12.5% error was reduced with the added regularization terms $(\mu/2)\|u\|_2^2$. The reconstruction quality is more pleasing for the balanced algorithm with two regularization term than the corresponding balanced algorithm with one regularization term.

Due to the device defect or the low-exposing dose, noise is often inevitable. So, we add 1% noise to the projection data of “NCAT.” The corresponding reconstructions RMSE and PSNR by FBP method (Figure 2(d)) are 0.2751 and 68.8931, respectively. The corresponding reconstructions RMSE and PSNR are 0.0338 and 87.1018, respectively, for two regularization term balanced-based algorithm (Figure 2(f)) and 0.0362 and 86.2801, respectively, for the one regularization term balanced-based algorithm (Figure 2(e)). Consequently, in the environments both with noise and without noise, the balanced algorithm with two regularization terms can obtain more better reconstruction quality in terms of RMSE and PSNR.

Example 4.2. Shepp-Logan image “SL” as our second test image can be seen in Figure 1(b). We evaluate algorithms under different views. The corresponding reconstructions RMSE and PSNR by FBP method with 40 projection views

(Figure 3(a)) are 0.5906 and 64.8555, respectively. The corresponding reconstructions RMSE and PSNR by FBP method with 50 projection views (Figure 3(d)) are 0.5726 and 65.1283, respectively. Obviously, the streak artifacts are introduced by FBP algorithm with few-view projections. Figures 3(c) and 3(d) are, respectively, the recovered results with 40 projection views by models with one regularization ($\mu = 0$) in (12) and two regularizations ($\mu \neq 0$) in (12) based on our proposed balanced approach with algorithm parameters $\gamma = 5$, $\nu = 0.2$, and $\mu = 0.01$. Every parameter has been optimized. The values of RMSE are 0.0780 and 0.0759. The values of PSNR are 82.1841 and 82.6393. The reconstruction results with 50 projection views are displayed in Figures 3(e) and 3(f). The values of RMSE are 0.0581 and 0.0557. The values of PSNR are 84.2616 and 85.9924. The RMSE and PSNR of our proposed hybrid regularization method are improved under different views which means more pleasing reconstruction results. Both the two kinds of the balanced-based image reconstruction schemes yield more desiring results than the FBP method.

5. Conclusion

In this paper, we propose a balanced-based wavelet CT reconstruction model with two regularization term. We investigate the fast algorithm for the proposed hybrid model based on the alternative direction method. Simulation results evidently demonstrate the superiority of our proposed scheme in reducing the values of RMSE and promoting the PSNR. This method is very flexible and can also be easily generalized to some other image processing problems.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Authors' Contributions

Weifeng Zhou and Hua Xiang are co-first authors with equal contribution.

Acknowledgments

This work was supported by the Natural Science Foundation of China (NSFC) (11526118 and 51605238), Natural Science Foundation of Shandong Province (ZR2015PF001), Science and Technology Planning Project of Shandong Province (2015GGX101020), the Foundation of China Postdoctoral (2015M570594), and the Talent Start-up Fund of Qingdao Science and Technology University (010022670).

References

- [1] D. J. Brenner and E. J. Hall, “Computed tomography—an increasing source of radiation exposure,” *New England Journal of Medicine*, vol. 357, pp. 2277–2284, 2007.
- [2] I. Elbakri and J. Fessler, “Statistical image reconstruction for polyenergetic X-ray computed tomography,” *IEEE transactions on medical imaging*, vol. 21, no. 2, pp. 89–99, 2002.

- [3] L. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D*, vol. 60, no. 1–4, pp. 259–268, 1992.
- [4] H. Yu and G. Wang, "Compressed sensing based interior tomography," *Physics in Medicine and Biology*, vol. 54, no. 9, pp. 2791–2805, 2009.
- [5] H. Yu, J. S. Yang, M. Jiang, and G. Wang, "Supplemental analysis on compressed sensing based interior tomography," *Physics in Medicine and Biology*, vol. 54, no. 18, pp. 425–432, 2009.
- [6] J. Tang, B. Nett, and G. Chen, "Performance comparison between total variation (TV)-based compressed sensing and statistical iterative reconstruction algorithms," *Physics in Medicine and Biology*, vol. 54, no. 19, pp. 5781–5804, 2009.
- [7] Y. Liu, J. Ma, Y. Fan, and Z. Liang, "Adaptive-weighted total variation minimization for sparse data toward low-dose X-ray computed tomography image reconstruction," *Physics in Medicine and Biology*, vol. 57, no. 23, pp. 7923–7956, 2012.
- [8] S. Niu, Y. Gao, Z. Bian et al., "Sparse-view x-ray CT reconstruction via total generalized variation regularization," *Physics in medicine and biology*, vol. 59, no. 12, pp. 2997–3017, 2014.
- [9] Y. Liu, Z. Liang, J. Ma et al., "Total variation-stokes strategy for sparse-view X-ray CT image reconstruction," *IEEE transactions on medical imaging*, vol. 33, no. 3, pp. 749–763, 2014.
- [10] L. Ritschl, F. Bergner, C. Fleischmann, and M. Kachelrie, "Improved total variation-based CT image reconstruction applied to clinical data," *Physics in Medicine and Biology*, vol. 56, no. 6, pp. 1545–1561, 2011.
- [11] J. Zhang, Y. Hu, J. Yang, Y. Chen, J. L. Coatrieux, and L. Luo, "Sparse-view X-ray CT reconstruction with gamma regularization," *Neurocomputing*, vol. 230, pp. 251–269, 2017.
- [12] Y. Wang, F. Shu, W. Li, and C. Zhang, "An adaptive nonlocal filtering for low-dose CT in both image and projection domains," *Journal of computational design and engineering*, vol. 2, no. 2, pp. 113–118, 2015.
- [13] X. Jia, Y. Lou, B. Dong, Z. Tian, and S. Jiang, "4D computed tomography reconstruction from few-projection data via temporal non-local regularization," *Medical Image Computing and Computer-Assisted Intervention—MICCAI*, vol. 13, Part 1, pp. 143–150, 2010.
- [14] B. Zhao, H. Gao, H. Ding, and S. Molloy, "Tight-frame based iterative image reconstruction for spectral breast CT," *Medical Physics*, vol. 40, no. 3, article 031905, 2013.
- [15] H. Gao, R. Li, Y. Lin, and L. Xing, "4D cone beam CT via spatiotemporal tensor framelet," *Medical Physics*, vol. 39, no. 11, pp. 6943–6946, 2012.
- [16] W. Zhou, J. Cai, and H. Gao, "Adaptive tight frame based medical image reconstruction: a proof-of-concept study for computed tomography," *Inverse Problems*, vol. 29, no. 12, article 125006, 2013.
- [17] Q. Xu, H. Yu, X. Mou, L. Zhang, J. Hsieh, and G. Wang, "Low-dose X-ray CT reconstruction via dictionary learning," *IEEE Transactions on Medical Imaging*, vol. 31, no. 9, pp. 1682–1697, 2012.
- [18] Y. Lu, J. Zhao, and G. Wang, "Few-view image reconstruction with dual dictionaries," *Physics in Medicine and Biology*, vol. 57, no. 1, pp. 173–189, 2012.
- [19] J.-F. Cai, S. Huang, H. Ji, Z. Shen, and G. Ye, "Data-driven tight frame construction and image denoising," *Applied and Computational Harmonic Analysis*, vol. 37, no. 1, pp. 89–105, 2014.
- [20] I. Daubechies, B. Han, A. Ron, and Z. Shen, "Framelets: MRA-based constructions of wavelet frames," *Applied and Computational Harmonic Analysis*, vol. 14, pp. 1–6, 2003.
- [21] Y. Wang, G. Wang, S. Mao et al., "A spectral interior CT by a framelet-based reconstruction algorithm," *Journal of X-ray Science and Technology*, vol. 24, no. 6, pp. 771–785, 2016.
- [22] C. Bao, B. Dong, L. Hou, Z. Shen, X. Zhang, and X. Zhang, "Image restoration by minimizing zero norm of wavelet frame coefficients," *Inverse problems*, vol. 32, no. 1, pp. 1–28, 2016.
- [23] X. Zhang, W. Zhou, X. Zhang, and H. Gao, "Forward-backward splitting method for quantitative photoacoustic tomography," *Inverse problems*, vol. 30, article 125012, pp. 1–19, 2014.
- [24] Y. Liu, J.-F. Cai, Z. Zhan et al., "Balanced sparse model for tight frames in compressed sensing magnetic resonance imaging," *PLoS One*, vol. 10, no. 4, article e0119584, 2015.
- [25] Z. Shen, K. Toh, and S. Yun, "An accelerated proximal gradient algorithm for frame-based image restoration via the balanced approach," *SIAM Journal on Imaging Sciences*, vol. 4, no. 2, pp. 573–596, 2011.
- [26] H. Gao, H. Yu, S. Osher, and G. Wang, "Multi-energy CT based on a prior rank, intensity and sparsity model (PRISM)," *Inverse problems*, vol. 27, no. 11, article 115012, 2011.
- [27] J. F. Cai, S. Osher, and Z. Shen, "Split Bregman methods and frame based image restoration," *Multiscale Modeling and Simulation*, vol. 8, no. 2, pp. 337–369, 2010.
- [28] M. Figueiredo and J. Bioucas-Dias, "Restoration of poissonian image using alternating direction optimization," *IEEE transactions on Image Processing*, vol. 19, no. 12, pp. 3133–3145, 2010.
- [29] E. Esser, "Applications of Lagrangian-based alternating direction methods and connections to split Bregman method," *Computational Applied Mathematics Tech. Rep*, University of California, Los Angeles, 2009.
- [30] N. Pustelnik, C. Chau, and J.-C. Pesquet, "Parallel algorithm and hybrid regularization for image restoration using hybrid regularization," *IEEE transactions on Image Processing*, vol. 20, no. 9, pp. 2450–2462, 2011.
- [31] T. Jeong, H. Woo, and S. Yun, "Frame-based poisson image restoration using a proximal linearized alternating direction method," *Inverse problems*, vol. 29, no. 7, article 075007, 2013.
- [32] C. Wu and X. Tai, *Augmented Lagrangian Method, Dual Methods, and Split Bregman Iteration for ROF, Vectorial TV, and High Order Models*, UCLA, CAM Report, 2009.
- [33] J. Eckstein and D. Bertsekas, "On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators," *Mathematical Programming*, vol. 55, no. 1–3, pp. 293–318, 1992.
- [34] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.
- [35] H. Gao, "Fast parallel algorithms for the x-ray transform and its adjoint," *Medical Physics*, vol. 39, no. 11, pp. 7110–7120, 2012.

Research Article

Automatic Radiographic Position Recognition from Image Frequency and Intensity

Ning-ning Ren,^{1,2} An-ran Ma,^{1,2} Li-bo Han,^{1,2} Yong Sun,¹ Yan Shao,^{1,3} and Jian-feng Qiu¹

¹College of Radiology, Taishan Medical University, Taian, Shandong, China

²College of Information and Engineering, Taishan Medical University, Taian, Shandong, China

³Department of Radiology, Affiliated Hospital of Taishan Medical University, Taian, Shandong, China

Correspondence should be addressed to Jian-feng Qiu; jfqu100@gmail.com

Received 15 February 2017; Accepted 17 July 2017; Published 17 September 2017

Academic Editor: Shujun Fu

Copyright © 2017 Ning-ning Ren 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.

Purpose. With the development of digital X-ray imaging and processing methods, the categorization and analysis of massive digital radiographic images need to be automatically finished. What is crucial in this processing is the automatic retrieval and recognition of radiographic position. To address these concerns, we developed an automatic method to identify a patient's position and body region using only frequency curve classification and gray matching. *Methods.* Our new method is combined with frequency analysis and gray image matching. The radiographic position was determined from frequency similarity and amplitude classification. The body region recognition was performed by image matching in the whole-body phantom image with prior knowledge of templates. The whole-body phantom image was stitched by radiological images of different parts. *Results.* The proposed method can automatically retrieve and recognize the radiographic position and body region using frequency and intensity information. It replaces 2D image retrieval with 1D frequency curve classification, with higher speed and accuracy up to 93.78%. *Conclusion.* The proposed method is able to outperform the digital X-ray image's position recognition with a limited time cost and a simple algorithm. The frequency information of radiography can make image classification quicker and more accurate.

1. Introduction

Digital X-ray imaging technique has generated massive amounts of clinical image data in radiology departments every day. These data need to be classified, retrieved, and analyzed in Picture Archiving and Communication Systems (PACS) or Radiology Information Systems (RIS). The urgent requirements to process these massive image data demand an automated and computationally efficient approach [1, 2]. Among these approaches, image classification, radiographic position identification, and artificial intelligence analysis are the most widely used ones. In this sense, the retrieval of images and the learning of radiographic position are the most fundamental parts.

Traditional medical image retrieval is semimanual, which obtains clinical information from manually retrieved image annotations and databases. The disadvantage of this approach involves human errors and operator variations, which is labor intensive and results in lower accuracy [1].

Automated methods using image retrieval technique are based on image features such as color [2, 3], texture [4], and shape [5]. Wang et al. proposed a dynamic interpolation method to achieve stereo microscopic measurements, but the scheme required a large quantity of matching elements [6]. Histograms were also widely used for image retrieval but had its relevant disadvantages [7, 8]. Other image retrieval techniques such as wavelet transform (WT) [9], Fourier transform (FT) [10], local binary pattern (LBP) [11], and Tamura texture features [12] can recognize an image type through library searching and image classification. However, position information cannot be automatically determined with these algorithms. Besides, these methods lack recognition on imaging organ tracking as researched by Jiao et al. [13].

Pattern recognition can automatically process and analyze digital images as mentioned by Paparo et al. [14, 15]. Feature selection method reported by Silva et al. [16] and Hussain [17] was used in traditional learning algorithms

such as support vector machine (SVM) and k-means for image retrieval but needs large datasets for training. Medical expert systems as discussed elsewhere [18, 19] used mixed algorithm to extract target area. Multilayer perceptron neural networks (MLPNN) can identify tissues and diseases as discussed in other places [20–22]; however, the process is complex and the processing time is too long for clinical use. Recently, the well-known deep learning algorithm has also been introduced to medical image processing and achieved equivalent results compared with professional expertise [23–25], but the data quantity and accuracy have remained a debate [26].

Therefore, in this paper, a method that combines frequency curve classification with gray scale matching for image retrieval and matching is proposed. It uses a whole-body phantom image as the template mask for anatomical and radiographic location marking, with shorter time cost and higher accuracy.

2. Materials and Methods

2.1. Image Preprocessing. Raw digital radiographic image data typically has large dynamic range and gray level features. Therefore, we use linear histogram stretching and a median filter for noise reduction. The respective equations are

$$F_H(x, y) = \left(\frac{65535}{B - A} \right) \cdot [f(x, y) - A], \quad (1)$$

where

$$\begin{aligned} A &= \min[f(x, y)], \\ B &= \max[f(x, y)], \end{aligned} \quad (2)$$

and

$$F_w(x, y) = \text{med}\{f(x - R, y - l), (R, l) \in w\}. \quad (3)$$

w is 5.

2.2. The Phantom X-Ray Image Masks. X-ray imaging phantoms are physical analogs of human body shapes and tissues as studied by Dewerd and Kissick [27]. Plastic and nylon are used to simulate the outline of the human body, bones, and primary tissues for whole-body radiography. We took X-ray imaging of the brain, cervical spine, chest, lumbar spine, pelvis, and limbs of a whole-body phantom (Whole Body Phantom PBU-50, Kyoto Kagaku, Japan) by using Digital Radiology DR (Wan dong HF50, Beijing, China). Each of the images was processed by adjusting the histogram, filtering, performing rigid translations, and scaling [28] and then fitted into a whole-body radiographic image. We also performed contrast-limited adaptive histogram equalization (CLAHE) for handling the variation in X-ray exposures.

For recognition of the radiographic positions after completing the input image matching, we performed the anatomical definition to a phantom template; the matrix of images is 2000×800 , and the height of the corresponding body is 165 cm without gender. For the information of



FIGURE 1: The whole-body phantom's X-ray mask and the examples of partial anatomical definition.

the image, diagnostician can use different ranges to define different organs, such as head size ranging from [260, 1] to [540, 285] and lung size ranging from [250, 130] to [560, 365], as shown in Figure 1. For the frontal image, there are seven radiographic positions and six radiographic target organs. The phantom template defines the target template for subsequent matching based on automatic identification and X-ray photography posture.

2.3. Classification Based on Image Frequency. Radiographic images have special frequency and amplitude characteristics, which are position dependent. These characteristics of the frequency curve can be used for classifying the type of image (for a given radiological position) and extract the texture of the organ.

2.4. The Characteristics of X-Ray Image Frequency. We use the fast Fourier transform (FFT) of the organ images to obtain the frequency spectrum as follows:

$$F(u, v) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) \cdot e^{-j2\pi(ux/M+vy/N)}, \quad (4)$$

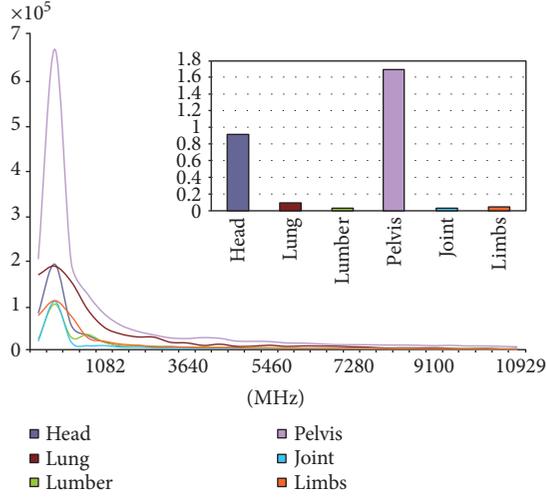


FIGURE 2: Frequency curves and the AUCs for various anatomical regions.

where M and N are the image resolution and u and v are coordinates in the frequency domain. From the frequency image and 2D curve, we find that the effective anatomical contours concentrate on the minimum 2% of the frequency curve. In Figure 2, the frequency curve at each position is the average of 10 images of the same radiological position in the same coordinate system, and the curve features shown are significantly different among the positions. Partly, the differences of some positions such as the lungs and limbs are not reflected in the frequency curve; thus, we offer the areas under the curve (AUCs), whose values of the lungs and limbs have obvious differences. Combined with frequency curve and AUCs, differences of positions can be obviously shown. The radiological positions are the head, lungs, lumbar (spine), pelvis (abdominal), joint (knee), and limbs.

In X-ray images, organs or tissue has a characteristic frequency response, even in different samples and different radiological positions. For example, the chest imaging using appropriate exposure parameters shows lung texture details and the lung signal captured in certain frequency bands. As shown in Figure 3, which shows the average frequency curve of 10 lung X-ray images, there is a peak in the low-frequency range, which corresponds to a lung texture detail (extracted using a Butterworth filter). For comparison, a similar peak in the averaged knee curve corresponds to bone trabeculae as plotted in Figure 4.

2.5. Classification Based on Image Frequency. The frequency curves for six radiographic positions were used as the standard library for comparison with arbitrary input images, and the similarity between input image and standard library was determined by the mean variance of the vector frequency curve. The input image is $f(w)$; the corresponding amplitude-arranged vector is $[A1, A2, \dots, An]$; the six frequency curves, $F(w)$, are used as a standard for comparison with arbitrary input images in the library and have

amplitudes of $[B1, B2, \dots, Bn]$. The mean-variance similarity between the input image and the reference organ image is

$$a = \text{sqrt} \left(\frac{\sum_{i=1}^n ((f(w) - F(w)) \cdot \wedge 2)}{\text{length}(f(w))} \right), \quad (5)$$

$$a = \text{sqrt} \left(\frac{\text{sum}((f(w) - F(w)) \cdot \wedge 2)}{\text{length}(f(w))} \right).$$

The cosine of the angle θ between the two images can be described as follows:

$$\cos \theta = \frac{\sum_{i=1}^n (f(w) \cdot F(w))}{\sqrt{\sum_{i=1}^n (f(w))^2} \cdot \sqrt{\sum_{i=1}^n (F(w))^2}} = \frac{f(w) \cdot F(w)}{|f(w)| \cdot |F(w)|}. \quad (6)$$

The smaller mean-variance is and the closer the cosine value is to 1 (indicates an angle closer to zero), the greater the similarity is. Matching 6 curves yields 6 mean-variance values, and then bubble sort is performed to determine the two mean-variances with the highest absolute value. The absolute values of the top two are less than 0.02, comparing the cosine similarity between the wave curves of the source image and organs which corresponds to the top two mean-variances. The organ which is the closest to mean-variance is considered the same as the organ of the source image. Six organs had mean-variances with standard frequency curves, and the reciprocal of that for all organs is plotted in Figure 5, as histograms. Higher reciprocal of mean-variance signifies greater similarity.

2.6. Image Matching Based on Matrix Multiplication and Correlation Coefficient. After the vector calculations based on image frequency have been performed, we determine the types of the images that are the most similar to the standard organ curve according to the shape of their curves and mean variances. The input image will be matched against the whole-body phantom mask so that the organ field is defined. This step involves matrix multiplication and the correlation coefficient.

In (7), (8), and (9), the input image after preprocessing is $F_{\text{input}}(x, y)$ and the 2% part of the frequency curve is $f(w)$. The image has been finished by classification based on image frequency, and the phantom image is denoted as $p(m, n)$. (In (7), (8), and (9), $p(m, n)$ represents image patches whose frequency is not within the minimum 2% range. The range of $p(m, n)$ represents the image from top to bottom.) By finding the maximum values of M and R , region T can be found, which is the intersection of M and R shown in the phantom image and also the target recognized region.

$$M = \frac{F_{\text{input}}(x, y) \cdot p(m, n)}{\sqrt{F_{\text{input}}(x, y)^2}} = \frac{\sum_m \sum_n F_{\text{input}}(x, y) \cdot p(m, n)}{\sqrt{\sum_m \sum_n (F_{\text{input}}(x, y))^2}}, \quad (7)$$

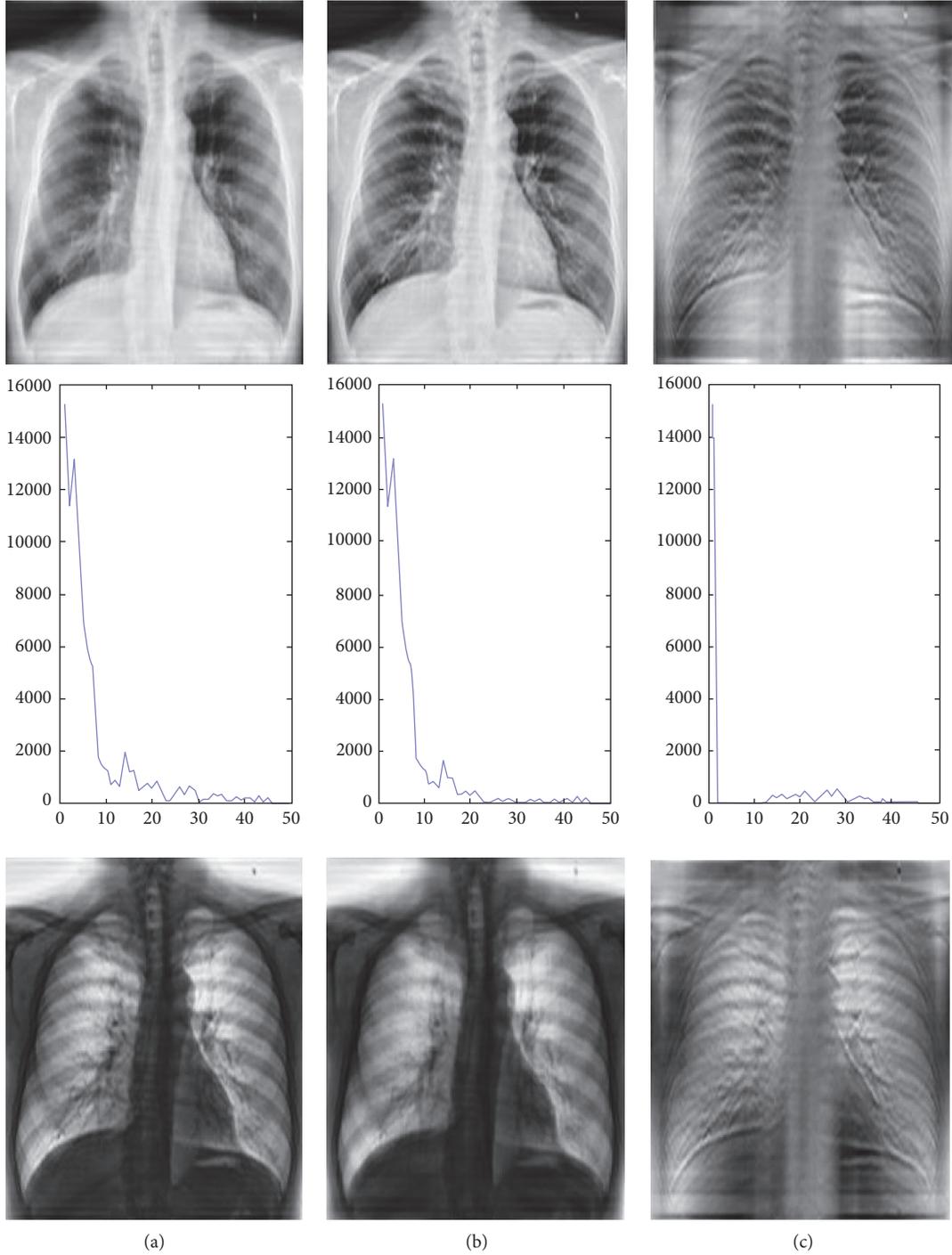


FIGURE 3: (a) From top to bottom: the chest X-ray image, the image frequency curve, and the chest X-ray image with inversed gray scale. (b) From top to bottom: chest X-ray image by Butterworth filtering, the image frequency curve, and the chest X-ray image with inversed gray scale. (c) From top to bottom: lung texture image reconstructed by the filtered frequency information, the frequency curve, and the lung texture image with inversed gray scale.

$$R = \frac{\sum_m \sum_n F_{\text{input}}((x, y) - \overline{\sum_m \sum_n F_{\text{input}}(x, y)}) \cdot (p(m, n) - \overline{\sum_m \sum_n p(m, n)})}{\sqrt{\sum_m \sum_n (F_{\text{input}}(x, y) - \overline{\sum_m \sum_n F_{\text{input}}(x, y)})^2 \cdot (p(m, n) - \overline{\sum_m \sum_n p(m, n)})^2}}, \quad (8)$$

$$T = M|R. \quad (9)$$

The maximum values of M and R have been solved, respectively, by the matrix multiplication and correlation coefficient, between the input image and phantom image. T is a region corresponding to the phantom area and is indicated by a bright box. To improve the processing speed, the matrix of the input and phantom images is reduced (maintaining image proportions).

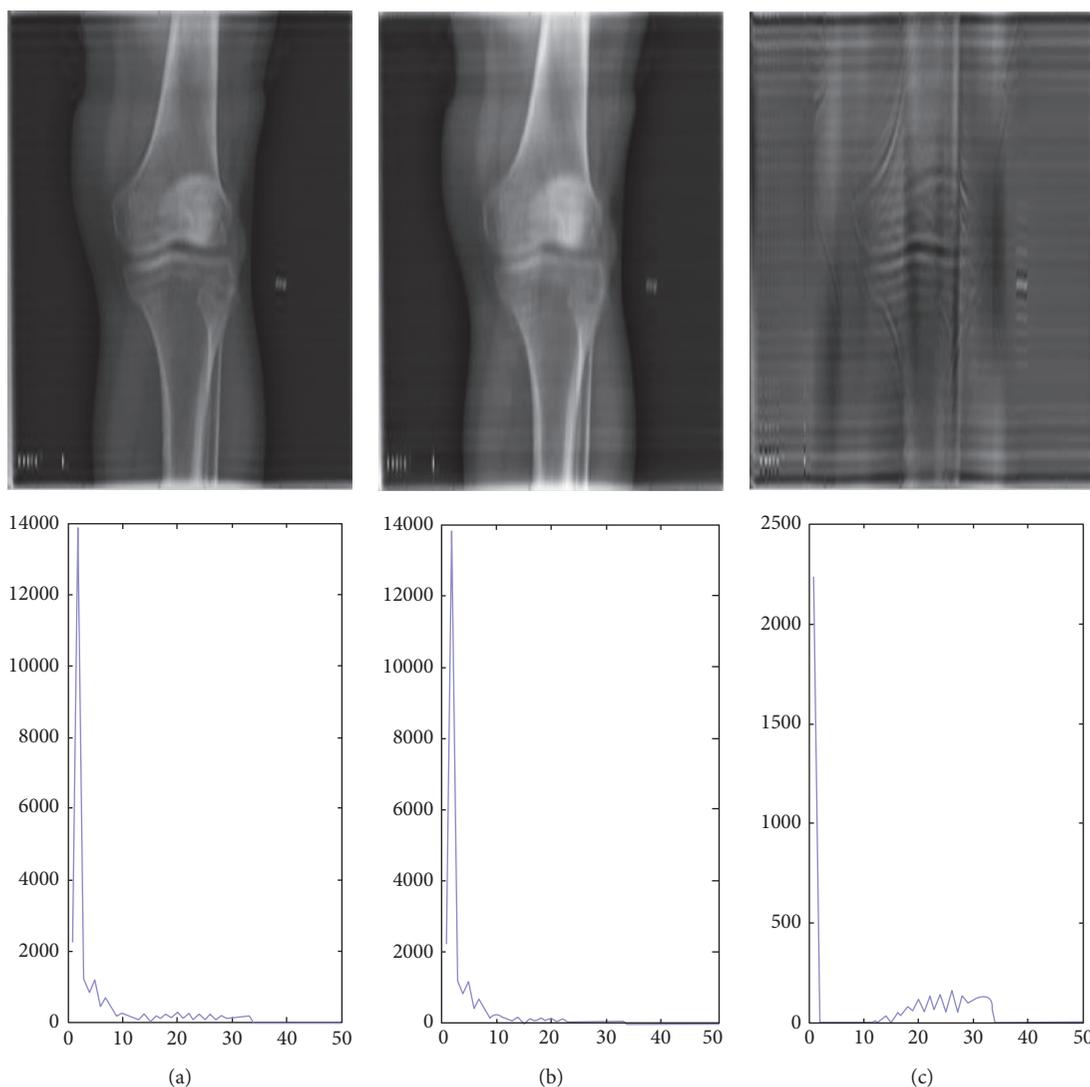


FIGURE 4: (a) From top to bottom: the knee X-ray image and the knee image frequency curve. (b) From top to bottom: the knee X-ray image by Butterworth filtering and the image frequency curve by filtering. (c) From top to bottom: the trabeculae texture image reconstructed by the filtered frequency information and the frequency curve.

2.7. Implementation of the Overall Algorithm. For any input image being preprocessed, the 2D Fourier transform will be taken and the lowest 2% frequency curve of the image is obtained. Compared with 6 predefined curve types and the input image type (radiographic position), the curve is classified by calculating the curve similarity and the mean variance. Next, the image is matched in the phantom image by finding the maximum value of matrix similarity. The final matching region, which corresponds to a priori knowledge of the patient's anatomical field, is shown in the phantom as the result. The workflow is shown in Figure 6.

3. Results and Discussion

217 clinical radiological images were randomly collected in this study, from the Radiology Department of Taishan Medical University. The radiological position and body

region in all images have been automatically recognized by our method. The results were verified by the clinical physicians of the Radiology Department. For comparison, the input images were also processed by dot matrix matching, correlation matching, and histogram retrieval algorithms. The accuracy rates and the processing times are shown in Table 1. The accuracy between the proposed method and any other methods has a statistically significant difference ($p < 0.005$).

The results have shown that the proposed algorithm has the highest accuracy and robustness for all images (6 position types); the average organ recognition accuracy was 93.78% and the average judgment time was 0.2903 s.

The proposed method is better than other benchmark methods; moreover, the method can obtain the radiographic position's description from the anatomical knowledge in the phantom image and reduce the processing time and recognition accuracy. What is more, compared with some effective

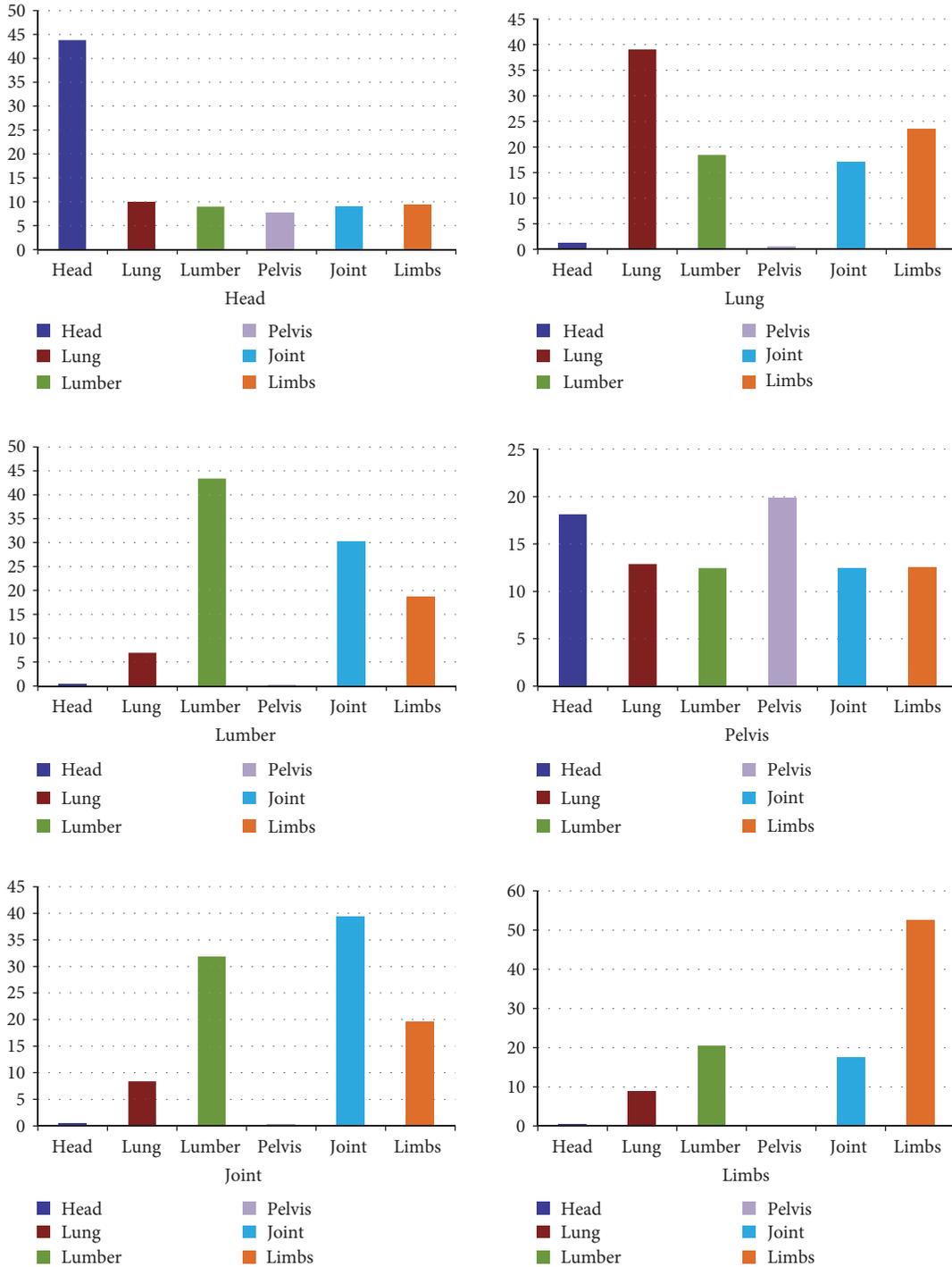


FIGURE 5: The reciprocal of mean-variance between 6 organs and the standard frequency curve.

approach such as the large margin local estimate (LMLE) [15] and deep learning network [24], the LMLE method only achieved less than 90% accuracy with 10% data as the training set. Although the convolution neural network in [24] achieved more than 90% accuracy in most image data, the approach needs 7000+ image slices and a most recently equipped computer (i7 3.4GHz, 16 GB RAM) for neural network training, while our method only needs simple matrix multiplication and correlation coefficient which can be

calculated on a multicore computer with less time and more than 90% of the accuracy.

The sample results of the radiographic position recognition are shown in Figure 7, by matching rectangular areas and annotated text. This integrated method can accurately mark the photograph site on the phantom images. We can get the information of photography range and photography sites according to early anatomical definition in phantom-pixel area. For different images with the same

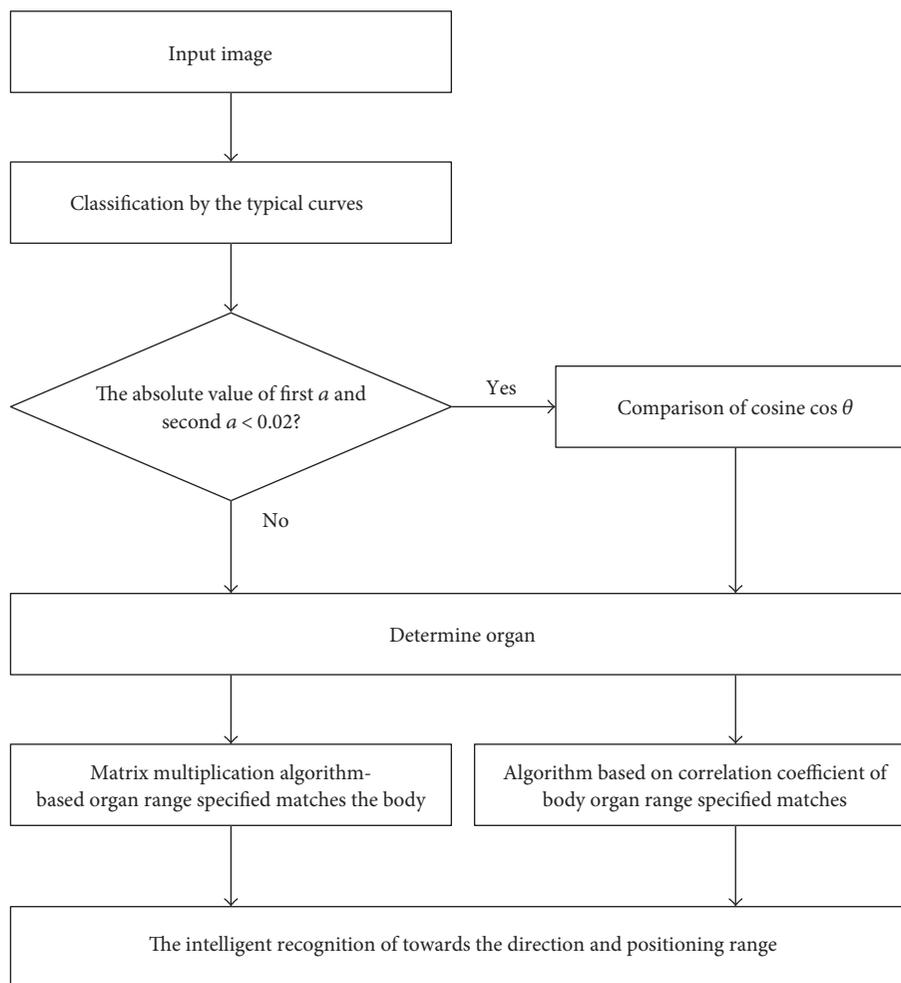


FIGURE 6: Workflow.

TABLE 1: The accuracy rate of four different radiographic position matching methods.

Radiographic positions	Dot matrix matching algorithm (%)	Correlation matching algorithm (%)	Histogram retrieval algorithm (%)	Current algorithm (%)	Average time of current algorithm (s)
Head	83.3	100.0	50.0	100.0	0.2808
Lungs	47.4	71.9	45.6	100.0	0.2918
Lumbar	45.6	66.7	40.5	100.0	0.2934
Pelvis	35.3	41.2	41.2	66.7	0.2919
Joint	90.9	100.0	27.3	100.0	0.2903
Limbs	75.8	56.8	56.6	96.0	0.2936
Average	63.1	72.7	43.5	93.7	0.2903

position type, the image matching can show the regional differences in the whole-body phantom image. For example, in Figure 7, three different cervical spine images have been identified and shown in different cover areas.

The human body model was represented by a phantom template X-ray image. The phantom was developed to mimic the human body X-ray attenuation parameters. The radiography of the phantom was closely approximated to the real human, even though the model structure was only simplified to the macroscopic shape of the organs. For example, the lung phantom made of plastic can simulate the lung contour and

segments but did not include the pulmonary veins and nodules. In the X-ray image of the phantom, the macroscopic profile of the lung is authentic for the imaging modality. The majority of conventional radiography sites are matched accurately by using this phantom image approach. For the detection of the contours of the lungs and the heart, the independent frequency or gray information is not sufficient.

The histogram and gray intensity are widely used for image similarity detection. Histogram matching has the advantage of being fast and no limitation by image size. However, it cannot determine the position and scope information.

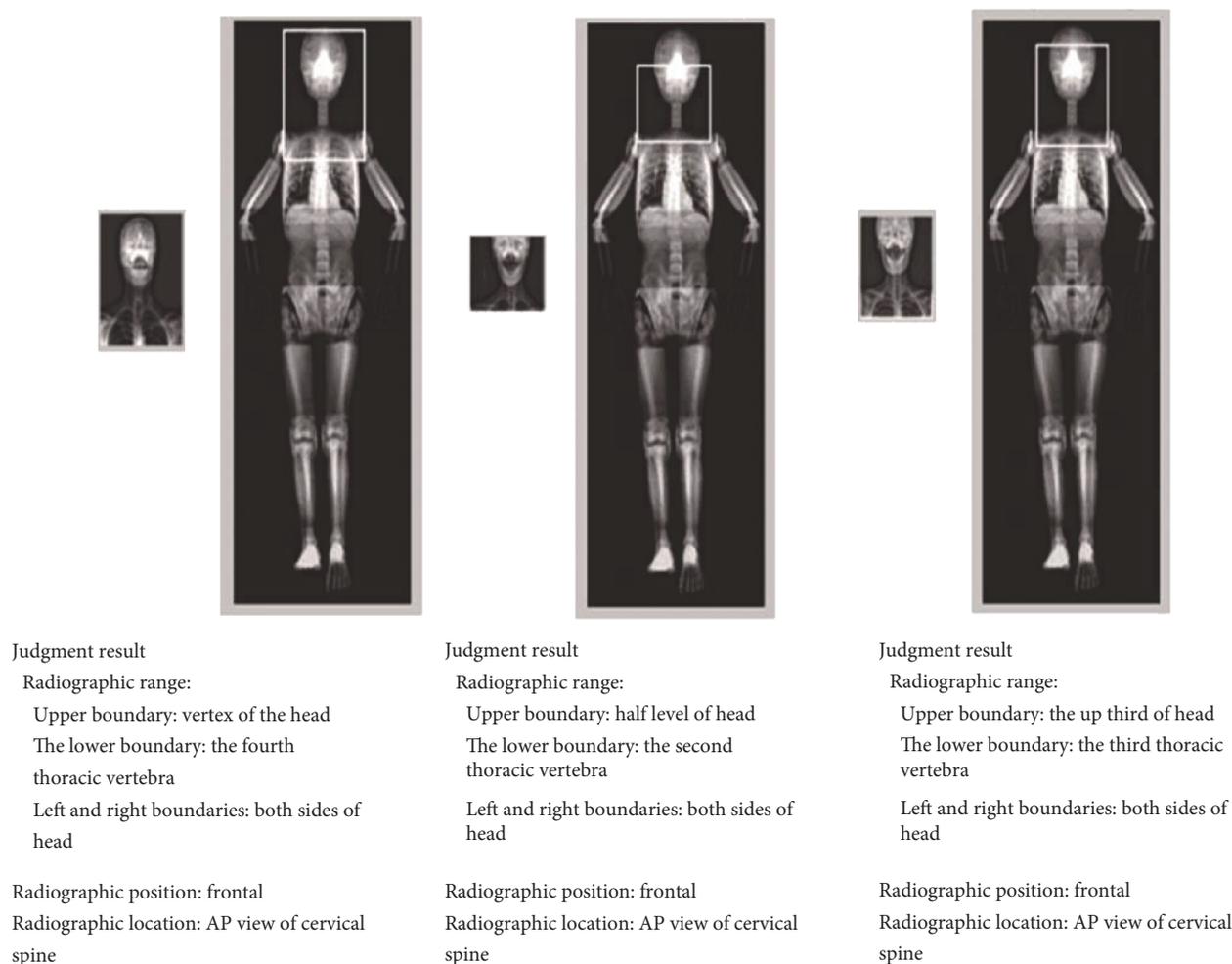


FIGURE 7: The automatic recognition results for three cervical spine images.

The method presented in this paper obtains robust frequency characteristic curves from X-ray information. The templates of different anatomical features have distinct frequencies and amplitudes. Comparison of input images and template only needs to take 2% effective frequency characteristics.

We extract a 1D curve from a 2D image, which accelerates and simplifies the image-matching algorithm. For 5.5 GB image data consisting of 217 images, the total processing time was 414.6 s.

Although our method was performed well for all of the test images, the algorithm has some limitations. The major obstacle is the poor result for nonstandard radiography; the matched result will be in the wrong position in the phantom image. For these cases, in a subsequent study, we plan to develop more standard phantom models, such as for babies, animals, and separate male and female bodies, in order to obtain more appropriate phantom images.

4. Conclusions

In this paper, we proposed a method for the automatic recognition of a radiographic position and body field, based on frequency curve classification and gray information of

digital radiographic images. Compared with image analysis methods based on complex pattern recognition algorithm, the proposed method can extract more information about the patient's position. The frequency classification in this work has good sensitivity and robustness to reduce the errors, which is caused by variations in the lighting environment (image exposure, detector sensitivity). This method is a fast 1D classification for 2D images and can be used for automatic feature extraction and be applied to big data calculations.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Acknowledgments

The authors thank the funding support by the China National Key Research and Development Program (2016YFC0103400) and the Natural Science Foundation of Taishan Medical University (no. GCC003). The authors also thank Weizhao Lu for his help with the English writing in the paper.

References

- [1] K. H. Hwang, H. Lee, and D. Choi, "Medical image retrieval: past and present," *Healthcare Informatics Research*, vol. 18, no. 1, pp. 3–9, 2012.
- [2] K. E. van de Sande, T. Gevers, and C. G. Snoek, "Evaluating color descriptors for object and scene recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 9, pp. 1582–1596, 2010.
- [3] J. Kajimura, R. Ito, N. R. Manley, and L. P. Hale, "Optimization of single- and dual-color immunofluorescence protocols for formalin-fixed, paraffin-embedded archival tissues," *Journal of Histochemistry & Cytochemistry*, vol. 64, no. 2, pp. 112–124, 2016.
- [4] X. Zhang, X. Gao, B. J. Liu, K. Ma, W. Yan, and H. Fujita, "Effective staging of fibrosis by the selected texture features of liver: which one is better, CT or MR imaging?," *Computerized Medical Imaging & Graphics the Official Journal of the Computerized Medical Imaging Society*, vol. 46, Part 2, pp. 227–236, 2015.
- [5] A. K. Jain and A. Vailaya, "Image retrieval using color and shape," *Pattern Recognition*, vol. 29, no. 8, pp. 1233–1244, 1996.
- [6] Y. Wang, G. Jiang, M. Yu, S. Fan, and J. Deng, "A study of stereo microscope measurements based on interpolated feature matching," *Biomedical Materials and Engineering*, vol. 26, Supplement 1, pp. S1473–S1481, 2015.
- [7] G. Pass and R. Zabih, "Histogram refinement for content-based image retrieval," in *IEEE Workshop on Applications of Computer Vision*. IEEE Computer Society, p. 96, 1996.
- [8] D. Shen, "Image registration by local histogram matching," *Pattern Recognition*, vol. 40, no. 4, pp. 1161–1172, 2007.
- [9] M. Kokare, P. K. Biswas, and B. N. Chatterji, "Texture image retrieval using new rotated complex wavelet filters," *IEEE Transactions on Systems Man & Cybernetics Part B Cybernetics A Publication of the IEEE Systems Man & Cybernetics Society*, vol. 35, no. 6, pp. 1168–1178, 2005.
- [10] D. Zhang and G. Lu, "Shape-based image retrieval using generic Fourier descriptor," *Signal Processing Image Communication*, vol. 17, no. 10, pp. 825–848, 2002.
- [11] J. Sun, S. Zhu, and X. Wu, "Image retrieval based on an improved CS-LBP descriptor," in *2010 2nd IEEE International Conference on Information Management and Engineering*. IEEE, pp. 115–117, 2010.
- [12] M. X. Xin, Z. X. Shi, G. B. Cui, and H. B. Lu, "Algorithm improvement of Tamura texture features in content-based medical image retrieval," *Chinese Medical Equipment Journal*, 2010.
- [13] X. Jiao, D. R. Einstein, V. Dyedov, and J. P. Carson, "Automatic identification and truncation of boundary outlets in complex imaging-derived biomedical geometries," *Medical & Biological Engineering & Computing*, vol. 47, no. 9, pp. 989–999, 2009.
- [14] F. Paparo, A. Piccardo, L. Bacigalupo et al., "Multimodality fusion imaging in abdominal and pelvic malignancies: current applications and future perspectives," *Abdominal Imaging*, vol. 40, no. 7, pp. 2723–2737, 2015.
- [15] Y. Song, W. Cail, H. Huang, Y. Zhou, D. D. Feng, and M. Chen, "Large margin aggregation of local estimates for medical image classification," *Medical Image Computing and Computer-Assisted Intervention*, vol. 17, Part 2, pp. 196–203, 2014.
- [16] S. F. D. Silva, M. X. Ribeiro, J. D. E. S. Batista Neto, C. Traina Jr., and A. J. M. Traina, "Improving the ranking quality of medical image retrieval using a genetic feature selection method," *Decision Support Systems*, vol. 51, no. 4, pp. 810–820, 2011.
- [17] S. A. Hussain, "A novel feature selection mechanism for medical image retrieval system," *International Journal of Advances in Engineering & Technology*, vol. 6, pp. 1283–1298, 2013.
- [18] L. Rourke, V. Willenbockel, L. Cruickshank, and J. Tanaka, "The neural correlates of medical expertise," *Journal of Vision*, vol. 15, no. 12, p. 1131, 2015.
- [19] Y. Shen and W. Zhu, "Medical image processing using a machine vision-based approach," *International Journal of Signal Processing Image Processing & Pattern Recognition*, vol. 6, 2013.
- [20] C. Y. Chang, "A contextual-based Hopfield neural network for medical image edge detection," in *2004 IEEE International Conference on Multimedia and Expo (ICME) (IEEE Cat. No. 04TH8763)*, vol. 2, pp. 1011–1014, 2004.
- [21] J. C. Fu, C. C. Chen, J. W. Chai, S. T. Wong, and I. C. Li, "Image segmentation by EM-based adaptive pulse coupled neural networks in brain magnetic resonance imaging," *Computerized Medical Imaging and Graphics*, vol. 34, no. 4, p. 308, 2010.
- [22] T. Kondo and J. Ueno, "Medical image diagnosis of lung cancer by a revised GMDH-type neural network using various kinds of neuron," *Artificial Life and Robotics*, vol. 16, no. 3, pp. 301–306, 2011.
- [23] A. Esteva, B. Kuprel, R. A. Novoa et al., "Dermatologist-level classification of skin cancer with deep neural networks," *Nature*, vol. 542, no. 7639, p. 115, 2017.
- [24] Z. Yan, Y. Zhan, Z. Peng et al., "Body part recognition using multi-stage deep learning," *Information Processing in Medical Imaging*, vol. 24, pp. 449–461, 2015.
- [25] H. C. Shin, H. R. Roth, M. Gao et al., "Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning," *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, p. 1285, 2016.
- [26] J. Cho, K. Lee, E. Shin et al., "How much data is needed to train a medical image deep learning system to achieve necessary high accuracy," *Computer Science*, 2016.
- [27] L. A. Dewerd and M. Kissick, *The Phantoms of Medical and Health Physics*, Springer New York, Heidelberg, Dordrecht, London, 2014.
- [28] <http://www.mathworks.com/matlabcentral/fileexchange/authors/24896>.

Research Article

The Edge Detectors Suitable for Retinal OCT Image Segmentation

Su Luo,¹ Jing Yang,² Qian Gao,¹ Sheng Zhou,² and Chang'an A. Zhan¹

¹Guangdong Provincial Key Laboratory of Medical Image Processing, School of Biomedical Engineering, Southern Medical University, Guangzhou 510515, China

²State Key Laboratory of Ophthalmology, Zhongshan Ophthalmic Center, Sun Yat-sen University, 54S Xianlie Road, Guangzhou 510060, China

Correspondence should be addressed to Chang'an A. Zhan; changan.zhan@gmail.com

Received 4 May 2017; Accepted 19 July 2017; Published 17 August 2017

Academic Editor: Shujun Fu

Copyright © 2017 Su Luo 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.

Retinal layer thickness measurement offers important information for reliable diagnosis of retinal diseases and for the evaluation of disease development and medical treatment responses. This task critically depends on the accurate edge detection of the retinal layers in OCT images. Here, we intended to search for the most suitable edge detectors for the retinal OCT image segmentation task. The three most promising edge detection algorithms were identified in the related literature: Canny edge detector, the two-pass method, and the EdgeFlow technique. The quantitative evaluation results show that the two-pass method outperforms consistently the Canny detector and the EdgeFlow technique in delineating the retinal layer boundaries in the OCT images. In addition, the mean localization deviation metrics show that the two-pass method caused the smallest edge shifting problem. These findings suggest that the two-pass method is the best among the three algorithms for detecting retinal layer boundaries. The overall better performance of Canny and two-pass methods over EdgeFlow technique implies that the OCT images contain more intensity gradient information than texture changes along the retinal layer boundaries. The results will guide our future efforts in the quantitative analysis of retinal OCT images for the effective use of OCT technologies in the field of ophthalmology.

1. Background

Optical coherence tomography (OCT) is the optical equivalent of ultrasonography, with the capability of capturing the depth-resolved cross-sectional images of biological tissues in vivo at near-histologic resolution [1]. Due to its noninvasiveness and high resolution, in combination with the characteristics of the eye and retinal anatomy, OCT has a rapid development of clinical applications in ophthalmology in recent years.

Quantitative analysis of retinal OCT image has been critical for reliable and efficient diagnosis of diseases such as glaucoma, age-related macular degeneration, and macular edema caused by diabetic retinopathy and for the evaluation of development of diseases, medical treatment responses, drug effectiveness, visual functions, and so forth [2–4]. Among others [5–7], automatic and semiautomated measurement of retinal layer thickness is considered as a class of key

quantitative analysis. Numerous research efforts have been devoted to this topic [8–11], and these efforts have significantly promoted the clinical understanding of ocular diseases and improved the OCT technologies and their applications.

Retinal layer thickness measurement relies on accurate OCT image segmentation. For many automated segmentation algorithms, edge detection is an essential foundation [12–19], notwithstanding some methods resort to other features of the images [18]. Literature shows that diverse types of edge detection algorithms can be employed as a key step in image segmentation. Table 1 summarizes the commonly used algorithms for retinal OCT image segmentation, including the Canny edge detector [12, 13], two-pass edge detection method [14, 15], local mean gradient-based edge tracking [16], peak detection method, Gaussian smoothing in combination with the Sobel kernel method [17, 18], and EdgeFlow technique [19]. Based on the nature of the information used in their algorithms, we can classify these different

TABLE 1: Summary of edge detection techniques involved in retinal OCT image segmentation studies.

Principles	Edge detection algorithms	Applications in OCT image analysis	References
Intensity-based	Canny edge detector	To obtain global gradient information	Yang et al. [12]
		To extract global boundary	Shijian et al. [13]
	Two-pass edge detection algorithm	To extract global boundary	Bagci et al. [14]
	Edge-tracking algorithm	To detect ILM	Bagci et al. [15]
	Peak detection method	To detect the easily detected boundary ILM and OS/RPE	Rossant et al. [16]
Gaussian smoothing + Sobel kernel	Cha and Han [17]		
Texture-based	EdgeFlow technique	To detect global boundary of retinal OCT image	Lang et al. [18]
			Niu et al. [19]

edge detection techniques into two categories, namely, the intensity-based and texture-based methods. The former category utilizes the intensity gradient in the images, whereas the latter tracks the texture changes rather than the intensity gradient.

Given the importance of edge detection for retinal image segmentation and the diversity of edge detectors, a natural question is which method gives the best edge detection outcome for current application. To our best knowledge, there is no study so far systematically evaluating the performance of the edge detectors applied for OCT image analysis. General performance evaluation of edge detection has long been interesting many researchers, but prior studies have not reached a unanimous conclusion yet because of the complexity of the problem. According to Heath et al., the challenge lies in that the edge detection performance depends not only on the algorithms themselves but also on the images applied to, the parameters used in the specific case, and the evaluation metrics [20].

Our goal in this paper is to evaluate what type of edge detectors best suit for retinal OCT image segmentation with given equivalent parameters, when measured using performance metrics that are meaningful for retinal layer thickness quantification. To this end, we first collect the commonly used edge detection methods in the literature on OCT image analysis and choose the most representative ones for comparison. We then research the edge detection performance evaluation literature to select the most relevant performance metrics that are meaningful for OCT image segmentation and adapt them when necessary. Using these metrics, we examine which of the selected edge detectors gives the best edge detection outcome when they are applied to the OCT images that we have collected from healthy subjects.

The remaining parts of the article are organized as follows. In the next section, we describe our research materials and methods, including retinal OCT image data collection, screening edge detectors for comparison by reviewing and analyzing the edge detection techniques used in prior studies on the retinal OCT image segmentation, and determining the most relevant performance evaluation metrics. In the third section, we present the comparison of the representative advanced approaches to retinal OCT image edge detection against different performance evaluation metrics. Finally, we discuss the findings and research opportunities.

2. Materials and Methods

We conducted this study in accordance with the Tenets of the World Medical Association’s Declaration of Helsinki [21]. Ethical approvals were obtained from the Ethical Review Board of Southern Medical University, the Ethical Review Board of Sun Yat-sen University, and the Research Ethical Committee of Zhongshan Ophthalmic Center. After an introduction about the purpose of the study and explanation of the process and risks, the voluntary participants signed the informed consent for this data collection.

2.1. Image Data Choice and Data Collection. In this article, our goal is to evaluate the performance of edge detection algorithms for retinal OCT image segmentation. Because of their image dependence [20], edge detectors perform well for other types of images but may not give as good results when applied to retinal OCT image. Therefore, all interested algorithms are tested on retinal OCT image in this study.

We collected the image data from 11 healthy volunteers (age ranges from 21 to 29, 7 males) using Topcon 3D OCT-2000 (Topcon Corporation, Tokyo, Japan) at Zhongshan Ophthalmic Center, a tertiary specialized hospital affiliated to Sun Yat-sen University, Guangzhou, China. Using 7 line 6.0 scanning mode to scan the macular area with a resolution of 1024 A-scans, we obtained the raw retinal OCT images. Due to the limitation of the software (software version: 8.20.003.04, Topcon Corporation, Tokyo, Japan), the raw images saved into bmp files have a size of 759×550 pixels. As the best available from our OCT equipment, this resolution is above average within those of OCT images reported in the most recent literature [3, 6, 7, 17–19]. Figure 1 shows a typical OCT image in our dataset. In this image, the 6 layer boundaries, namely, ILM, NFL-GCL, IPL-INL, OPL-ONL, ONL-IPS, and RPE-choroid are readily observable.

2.2. Edge Detection Algorithms. Three criteria for our choice of edge detectors for evaluation were

- (1) to include the algorithms that have been most commonly used in the OCT image segmentation literature,
- (2) to give preference to the ones representing the state of the art in edge detection, which were usually used to detect more than 3 retinal layer boundaries,

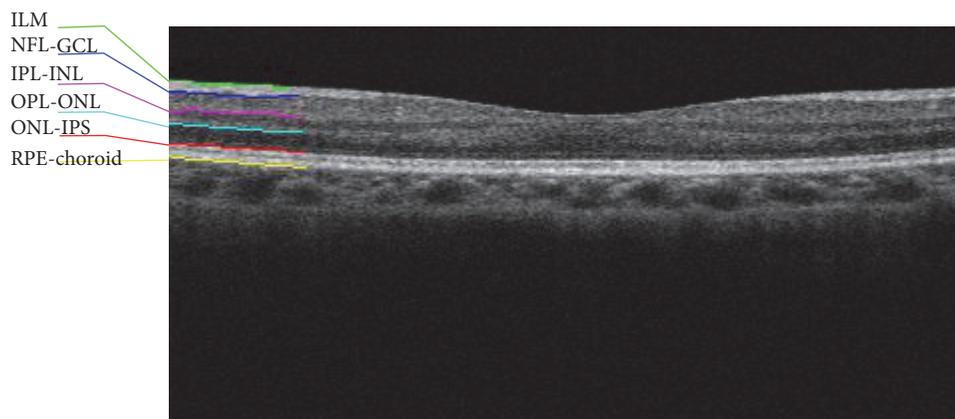


FIGURE 1: The retinal layers in a representative cross-sectional SD-OCT image. The region from the top layer ILM to the bottom layer RPE is of interest in this study and most clinical applications. Six layer boundaries are marked: ILM = the inner limiting membrane; NFL-GCL = the boundary between retinal nerve fiber layer and ganglion cell layer; IPL-INL = the boundary between inner plexiform layer and inner nuclear layer; OPL-ONL = the boundary between outer plexiform layer and outer nuclear layer; ONL-IPS = the boundary between outer nuclear layer and inner photoreceptor segment; RPE-choroid = the boundary between retinal pigment epithelium and choroid.

- (3) to include a diverse mix of algorithms utilizing different image feature information.

Using these criteria, three algorithms reviewed in Table 1 were chosen for further analysis. They are the Canny edge detector [21], two-pass algorithm [14, 15], and EdgeFlow method [22]. The former two are based on the image intensity gradient, whereas the last one is based on the image texture changes. We refer to the first two as the intensity-based edge detection methods and the third as the texture-based edge detection method. The principles of the three algorithms will be outlined below. The rest three edge detection methods, all intensity-based, were excluded from further analysis for different reasons. Among them, the edge-tracking algorithm based on the maximization of the local mean gradient was used only to detect ILM [16]; the peak detection method [17] and the Sobel kernel in combination of Gaussian smoothing [18] only detect the easily detectable boundaries ILM and OS/RPE.

2.2.1. Canny Edge Detector. The Canny edge detection algorithm [23] is now generally regarded as the “standard” for edge detection in the field of digital image processing ([24], Chapter 10). The Canny edge detector works in a multistep process to detect a wide range of edges in images. At first, the image is smoothed using a linear Gaussian filter. Then, a 2D first derivative operator is utilized on the smoothed image to compute the derivatives in both the vertical and the horizontal orientations. The gradient magnitude is calculated as the root sum of squares of the derivatives in two orthogonal directions and the gradient phase as the arctangent of their ratio. Candidate edge pixels are identified as the pixels that survive after a thinning process called non-maximal suppression. In this process, the edge strength of each candidate edge pixel is set to zero if its gradient magnitude is not larger than the gradient magnitude of the two adjacent pixels in the gradient direction, and the pixel whose gradient magnitude is the local maximum is preserved. At

last, hysteresis thresholding is used to eliminate weak edge points and track the possible edge pixels. In this step, double-threshold T1 and T2 with $T2 > T1$ are applied; all candidate edge pixels below the lower threshold T1 are set to zero, and all pixels above the lower threshold T1 can be connected to any pixels above the higher threshold T2 through a chain of edge pixels which are labeled as edge pixels. The hysteresis helps in ensuring that the noisy edges are not broken into multiple edge fragments. In the Canny edge detection algorithm, three parameters are incorporated, which play a decisive role for detecting the result. One is the width of Gaussian filter (i.e., standard deviation of the Gaussian, σ). An increase in the width of Gaussian filter reduces the detector’s sensitivity to noise, but blurs the image and results in loss of finer edge details. The other two are the lower threshold (T1) and the higher threshold (T2), respectively. The higher threshold should be set reasonably high and the lower threshold quite low for good detection results, because if it is too high, the lower threshold causes edge fragments and if too low, the higher threshold increases false alarms and undesirable edge fragments in the edge detection output.

2.2.2. Two-Pass Edge Detection Algorithm. Two-pass edge detection algorithm is designed exclusively for detection of retinal layer in OCT images by Bagci et al. [14, 15]. The feature of edges in retinal OCT image, extending along the horizontal direction with a gentle up and downslope, was taken into account in the algorithm. The edge detection kernel $L(x, y)$ is based on the first derivative of Gaussian in the vertical direction:

$$L(x, y) = -p \frac{x}{\pi\sigma^2} e^{-((x^2+y^2)/2\sigma^2)}. \quad (1)$$

The parameter p determines the polarity of edges and takes values either 1 or -1 . The edge detection kernel is applied twice with alternating values of p . On the first pass, the boundaries between each pair of adjacent bright and dark

regions, with bright on the top, such as NFL-GCL, IPL-INL, and OPL-ONL, are extracted with $p = 1$. On the second pass, boundaries between each pair of adjacent bright and dark regions, with dark on the anterior, such as ILM, INL-OPL, and ONL-IPS, were detected with $p = -1$. The peak values are marked as edges, using nonmaximal suppression and hysteresis thresholding. Satisfactory results can be obtained by adjusting the value of σ .

2.2.3. EdgeFlow Technique. The EdgeFlow technique is a novel boundary detection scheme proposed by Ma and Manjunath [22]. The technique for boundary detection based on EdgeFlow utilizes a predictive coding model to characterize the direction of change in color (intensity of grey image) and texture at each image location at a given scale and constructs an EdgeFlow vector. By propagating the EdgeFlow vectors, the boundaries can be detected at image locations which encounter two opposite directions of flow in the stable state. Differing from intensity-based detection methods that focus on finding the local gradient maximum, EdgeFlow technique computes the directions of edge energy according to intensity or texture in an image and associated probabilities. The edge energy and corresponding probabilities obtained from different image attributes are pooled together to form a single edge field for boundary detection:

$$\begin{aligned} E(s, \theta) &= \sum_{a \in A} E_a(s, \theta) \cdot \omega(a), \quad \sum_{a \in A} \omega(a) = 1, \\ P(s, \theta) &= \sum_{a \in A} P_a(s, \theta) \cdot \omega(a), \end{aligned} \quad (2)$$

where $E_a(s, \theta)$ and $P_a(s, \theta)$ represent the energy and probability of the EdgeFlow computed from image attribute a , $a \in \{\text{intensity/color, texture}\}$; $\omega(a)$ is the weighting coefficient associated with image attribute a . The edge flow direction is estimated as follows:

$$\Theta(s) = \arg \max_{\theta} \left\{ \sum_{\theta \leq \theta' < \theta + \pi} P(s, \theta') \right\}. \quad (3)$$

The EdgeFlow vector is then defined as

$$\vec{F}(s) = \sum_{\Theta(s) \leq \theta < \Theta(s) + \pi} E(s, \theta) \cdot \exp(j\theta), \quad (4)$$

where $\vec{F}(s)$ is a complex number with its magnitude representing the resulting edge energy and phase representing the flow direction. After the EdgeFlow vector of an image is computed, boundary detection can be performed by propagating the EdgeFlow vector and identifying the locations where two opposite flow directions encounter each other. The scheme facilitates integration of intensity and texture into a single framework for boundary detection.

2.3. Performance Evaluation Using Ground Truth. According to Heath et al. [20], edge detection performance evaluation can be classified into theoretical and empirical approaches. The former uses pure mathematical analysis without the

algorithms ever being applied to an image. It has major limitation for not being able to deal with the complexity of modern edge detection algorithms. The latter can be further classified into (1) evaluation using ground truth and (2) evaluation without ground truth. Our goal in this study is to examine how different edge detectors give the best results for OCT retinal layer segmentation. Ultimately, we hope to identify the most reliable and efficient edge detectors to help doctors automate the measurement of retinal layer thickness in order to make quantitatively informed medical decisions. Our human vision systems are the most complex and efficient machine for image analysis, including edge detection. Therefore, for our intended application, the most appropriate evaluation approach should be evaluation using ground truth, which measures the difference between the algorithm-detected edges and the human-detected edges.

Due to the importance of edge detector performance evaluation using ground truth, researchers have developed numerous metrics. These metrics can be largely classified into three categories. The first category, which we refer to as the edge presence accuracy metrics (EPAM), focuses on to which extent the detected edges coincide with the ground truth without considering location shift. EPAM include mainly four metrics [25–28], namely, true positive rate, false positive rate, false negative rate, and total edge detection accuracy. The first three, respectively, measure the ratio of true edge pixels, falsely detected edge pixels, and missed edge pixels to the number of total edge pixels in the ground truth, and the fourth is the ratio between the total true edge pixels and true nonedge pixels and the total number of pixels in the region of interest. The second category, which we refer to as the edge location accuracy metrics (ELAM), focuses on the extent of edge shifts [29] introduced by the edge detection algorithms as compared to the ground truth. Metrics in this category include Hausdorff's distance [30], which measures the similarity between two images, and mean localization deviation (MLD). The third category takes into account both location accuracy and edge presence. Typical metrics include figure of merit (FOM) [31] and its expanded version (expanded FOM) [32] and multifeature quality measurement [25].

2.3.1. Criteria for Performance Evaluation Metrics. OCT retinal layer segmentation aims to automate retinal layer thickness measurement in order to free ophthalmologists from laborious manual tracing of the layer boundaries. The ideal layer edge detector would give the same thickness measures to those from ground truth specified by human observers. However, even experts could not arrive at the same segmentation for a given retinal OCT image [33]. This is because manual segmentation is subject to human subjectiveness. The ground truth used for the evaluation is not really the ultimate truth. Thus, it is important to note that the traditional edge presence accuracy metrics, the probabilities of true positive, false positive (spurious edges), and missing edge, cannot offer the complete evaluation of edge detector performance. We propose that the performance metrics need to meet the criteria as follows:

- (1) To measure the edge presence accuracy by calculating the rates of true positive, false positive, true negative, and false negative (missing)
- (2) To measure edge location accuracy by calculating the signed and unsigned edge shift distance
- (3) To allow edge shift when calculating edge presence accuracy
- (4) To examine the computational costs.

2.3.2. Evaluation Metrics for this Study. Based on our analysis of existing performance metrics in the literature and the metric criteria discussed previously, we choose the figure of merit (FOM, Pratt) [31], true positive rate (TPR), false positive rate (FPR), accuracy (ACC), and mean localization deviation (MLD) [34] as the basis to develop our procedure for comprehensive evaluation of the chosen edge detectors. In the paragraphs to follow, we outline the principles of these metrics.

(1) *Pratt's Figure of Merit.* FOM [31] is a classical metric utilized by numerous researchers for evaluating the performance of edge detection algorithms [25, 26, 32, 35, 36]. The definition of the FOM is given by

$$\text{FOM} = \frac{1}{\max(N_I, N_A)} \sum_{i=1}^{N_A} \frac{1}{1 + \alpha d^2(i)}, \quad (5)$$

where N_I and N_A represent the number of ideal and actual detected edge pixels, $d(i)$ denotes the distance between the i th detected edge pixel and its correct position, and α is the scaling constant (normally set at 1/9) that is applied to provide a relative penalty between smeared edges and isolated, but offset, edges.

(2) *Edge Presence Accuracy.* The criteria on which the FOM of Pratt is based include missed valid edges, localization errors, and false alarms. Different configurations of detected edges may yield equal FOM value [35]. In order to decompose the sources of difference, Yin et al. [27] developed three metrics (TPR, FPR, and ACC) that are defined as follows.

True positive rate (TPR):

$$\text{TPR} = \frac{\text{TP}}{\text{TP} + \text{FN}} = \frac{\text{TP}}{N_I}. \quad (6)$$

False positive rate (FPR):

$$\text{FPR} = \frac{\text{FP}}{\text{FP} + \text{TN}} = \frac{\text{FP}}{N - N_I}. \quad (7)$$

Accuracy (ACC):

$$\text{ACC} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}} = \frac{\text{TP} + \text{TN}}{N}. \quad (8)$$

In these equations, TP (true positive) and TN (true negative) represent the numbers of correctly detected edge pixels and nonedge pixels. FP (false positive) is the number of pixels

not belonging to edge but recognized as one by the algorithm, and FN (false negative) is the number of pixels belonging to edge but failed to be recognized by the algorithm. N is the total number of pixels within the ROI of the image, and N_I is the number of ideal edge pixels.

Given the large number ($N - N_I$) of nonboundary pixels in the images, FPR calculated in the form of (7) is close to zero, making the metrics insensitive to the change of edge detection algorithms. We redefined it as

$$\text{FPR} = \frac{\text{FP}}{\text{FP} + \text{TP}} = \frac{\text{FP}}{N_A}, \quad (9)$$

where N_A denotes the number of pixels of the actually detected edges.

(3) *Edge Location Accuracy Metrics.* It is known that some image processing procedures cause the shift of detected edges ([29], Chapter 3, p. 56). In order to characterize the extent to which the results from edge detection algorithms deviate from the ground truth, we introduce the location accuracy metrics, the mean localization deviation (MLD) in the context of OCT image analysis:

$$\text{MLD} = \frac{1}{N_I} \sum_{N_I} \left(\frac{1}{N_b} \sum_{N_b} d(i) \right), \quad (10)$$

where N_b is the number of edge pixels in searching neighborhood of a ground truth edge pixel, N_I the number of pixels in the ideal edge, and $d(i)$ the Euclidean distance of the current edge pixel in ground truth and edge pixels in searching neighborhood. For the retinal OCT images, we limited the searching neighborhood to be within 3 pixels of the true edge along each A-scan.

(4) *Adjusted TPR, FPR, and ACC.* Some procedures of image processing can introduce edge shift ([29], Chapter 3, p. 56–74). As a result, the detected edge may not match the position of actual edge. As the goal of retinal OCT image segmentation is to extract the contours of retinal layer boundaries and measure the thicknesses of different retinal layers, small and constant shifts do not have effective impact when the layer thicknesses are of the only interest. Therefore, edge pixels in the neighborhood detected by the algorithms may be accepted into true positive edge pixels when calculating the edge presence metrics. In this case, FOM, the true positive rate, and false positive rate and accuracy measures need also to be adjusted. We define these adjusted metrics as FOM_{ADJ} , TPR_{ADJ} , FPR_{ADJ} , and ACC_{ADJ} :

$$\begin{aligned} \text{FOM}_{\text{ADJ}} &= \frac{1}{\max(N_I, N_A)} \sum_{i=1}^{N_A} \frac{1}{1 + \alpha d_{\text{ADJ}}^2(i)}, \\ \text{TPR}_{\text{ADJ}} &= \frac{\text{TP}_{\text{ADJ}}}{N_I}, \\ \text{FPR}_{\text{ADJ}} &= \frac{\text{FP}_{\text{ADJ}}}{N_I}, \\ \text{ACC}_{\text{ADJ}} &= \frac{\text{TP}_{\text{ADJ}} + \text{TN}_{\text{ADJ}}}{N}, \end{aligned} \quad (11)$$

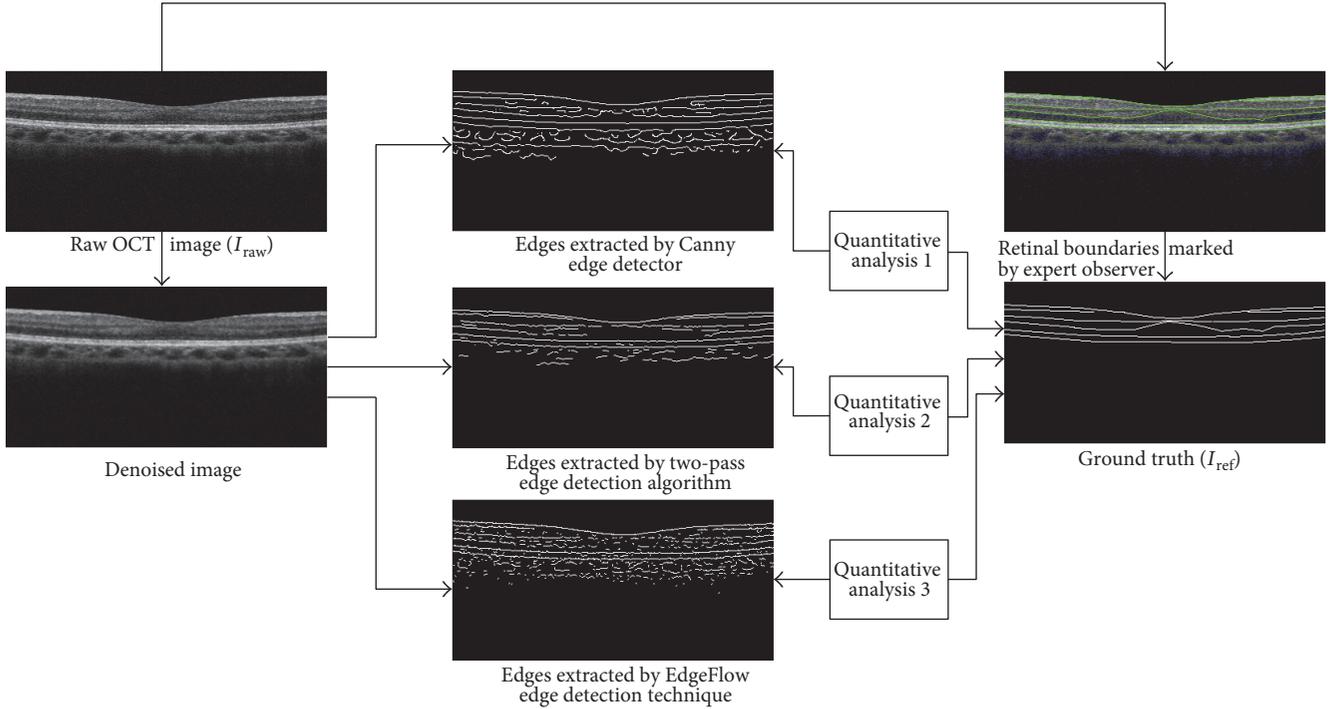


FIGURE 2: Procedures of performance evaluation of edge detectors. The ground truth of OCT retinal layer boundaries is labeled by an expert observer. Raw OCT images are firstly denoised and then applied with three automatic edge detection algorithms (i.e., Canny, two-pass, and EdgeFlow) to obtain the algorithm-detected boundaries, which are compared with the ground truth.

where N_I and N_A represent the number of ideal and actual detected edge pixels, $d_{\text{ADJ}}(i)$ is the distance between the i th detected edge pixel and its correct position, TP_{ADJ} is the number of edge pixels detected by an algorithm that are considered as edges within the neighborhood of the ground truth, FP_{ADJ} is the number of false positive pixels after the neighborhood searching and edge pixel adjustment, and TN_{ADJ} equals TN as it is not affected by the neighborhood adjustment. These adjusted metrics allowing the edge shift can better reflect the amount of detected edge points.

2.3.3. Evaluation Procedure. The major steps for performance measurement include the preparation of ground truth, the preprocessing of OCT images, the application of edge detectors with appropriate parameters to obtain the near optimal outcome for each detector, and using the performance metrics to evaluate the goodness of edge detectors against the ground truth. Figure 2 summarizes the flow of performance evaluation.

(1) Ground Truth Preparation. We asked an expert observer to manually delineate the edges for representative retinal OCT images to form a base dataset of ground truth, as noted by I_{ref} in Figure 2. Because the ILM and RPE are the outer boundaries of the retinal structure and they are strong edges that can usually be reliably detected, we define the images between ILM and RPE (included) as the region of interest (ROI). Only those edges within the ROI are extracted for comparison with the ground truth.

(2) Image Preprocessing. Before applying the computer algorithm for each edge detector, we conducted necessary image preprocessing. Due to constructive or destructive interference of the light waves from the object, spectral domain retinal OCT images suffer from the inherit speckle noise [37], which decreases the quality of image and causes unreliable retinal layer segmentation. In order to improve the quality of edge detection, preprocessing becomes a necessary step. We first converted the raw OCT image bmp files into gray-scale images and cropped the images to the region of interest (ROI, 200 by 400 pixels) in this study. The literatures have suggested the use of filters like mean, median, and Gaussian [38–40] for noise removal. We choose median filtering to remove the speckle noise. The original retinal OCT image and the denoised image are shown in Figure 3.

(3) Edge Detection. We randomly chose 8 images from our database of raw OCT retinal images and apply the three edge detectors. As discussed earlier, the edge detection outcomes may be influenced not only by the algorithm itself but also by the input parameters [20]. We varied the parameters systematically to obtain the optimal possible edge outcomes for each of the edge detection algorithms.

(4) Performance Evaluation. In edge detection performance evaluation step, we compared the edge detection outcomes from the three computer algorithms against the human manually traced retinal layer boundaries. We applied the metrics that were broadly used in the literature and relevant to our specific research context and purpose. We also applied the

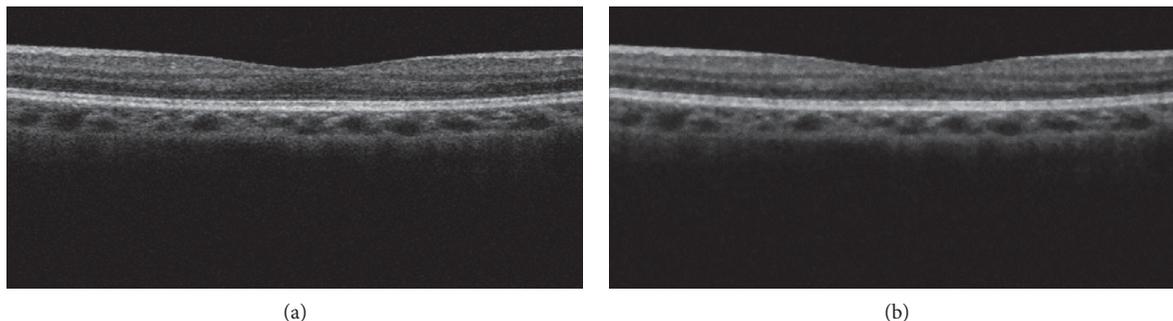


FIGURE 3: Median filtering of the retinal OCT image. (a) Raw retinal OCT image and (b) the denoised image.

adjusted metrics developed and discussed in the previous section. Finally, we examined the evaluation outcomes in terms of both differences and relationships.

3. Experiments and Results

We implemented all the data analysis in MATLAB R2012a (The Mathworks Inc., MA, USA) on a personal computer running Windows 7 operating system with an 3.60 GHz Intel® Core™ i7 CPU and 4 GB of memory. The raw image data was acquired and stored into bmp files. Data in the intermediate analysis steps were computed and stored in the double precision data format in order to minimize digitization errors. Three edge detection algorithms were carefully coded and double checked for the correctness.

3.1. Edge Detection

3.1.1. Input Parameters. Input parameters can significantly influence the resulting edge quality for given edge detection algorithm [20]. In selecting values for these parameters, we are interested in finding the set that provides good edge detection accuracy, that is, the boundaries coinciding the six retinal layers as shown in Figure 1 with high signal-to-noise ratio. In terms of performance evaluation metrics, good parameters give the high values of FOM, TPR, ACC, and their adjusted forms and lower values of FPR and MLD. Through applying the three detection methods with multiple sets of parameters on OCT retinal images ($n = 21$), the edge detection results were obtained and compared. By observing the outcomes, we chose the parameter set for each algorithm as follows. In Canny edge detection, the best result can be obtained by setting the width of the Gaussian filter $\sigma = 3$, the lower threshold $T1 = 0.005$, and the upper threshold $T2 = 0.1$; for two-pass edge detection algorithm, the width of Gaussian $\sigma = 3$, consistent with Bagci et al. [14], the lower threshold $T1 = 0.005$, and the upper threshold $T2 = 0.15$, similar to those for Canny edge detector; for the EdgeFlow algorithm, we followed Ma and Manjunath [22] and chose the equal weighting coefficients for intensity and texture, that is, $\omega(\text{intensity}) = \omega(\text{texture}) = 0.5$.

3.1.2. Edge Detection Results. Figure 4(a) shows an original retinal OCT image. Figures 4(b), 4(c), and 4(d) show the edge detection outcomes from the Canny edge detector, two-pass edge detection technique, and EdgeFlow algorithm,

respectively. From these edge detection results, six retinal layer boundaries of our interest are readily identifiable, although with some noises caused by false positives and boundary breakages caused by the false negatives.

For our purpose in this study, we were mainly interested in how the three edge detectors performed in detecting the 6 retinal layer boundaries, which were also the key information in the literatures for retinal layer thickness measurement [8–11]. We defined the region of interest (ROI) to be the area between the ILM and RPE that are the most outer boundaries of the retinal structure. Figure 5(a) is the OCT retinal image with overlaid ground truth edges marked by an expert observer. Figures 5(b), 5(c), and 5(d) show the edges within ROI detected by the three algorithms, which will be the basis of performance evaluation in the next section. Visually, the result from the Canny edge detector in Figure 5(b) shows a well-defined six boundaries, although with some breakages and noises. The result from the two-pass method shown in Figure 5(c) gives more than 6 layers in some locations, but in general, the six layer boundaries of interest are very clear with less breakages compared to those in Figure 5(b). The result from EdgeFlow algorithm depicted in Figure 5(d) shows more breakages and more noises, although all 6 layer boundaries are still recognizable.

3.2. Performance Evaluation. To quantify the performance of the three edge detectors, we use three sets of measurements discussed in Materials and Methods. The first set of metrics include FOM, TPR, FPR, and ACC, which have been broadly used in the literature [20, 25–28, 30–32] when evaluating edge detectors on images other than the OCT retinal images. In order to calculate TPR, FPR, and ACC, we use the ground truth as a template to screen the coincided edge pixels from the three edge detection algorithms. Figure 6 shows the edge points overlapping with the manually traced edges (ground truth).

The calculated performance metrics are summarized in Table 2. Based on the mean values of FOM for the three edge detectors, it seems that the best performer is the two-pass method. Table 3 summarizes the results of the statistical analysis. Two-sample t -tests confirmed the impression that two-pass method significantly outperforms both the Canny edge detector ($p = 1.19e - 6$), and EdgeFlow ($p = 5.28e - 4$) in terms of FOM. In addition, the EdgeFlow algorithm outperforms the Canny edge detector ($p = 4.65e - 3$). The

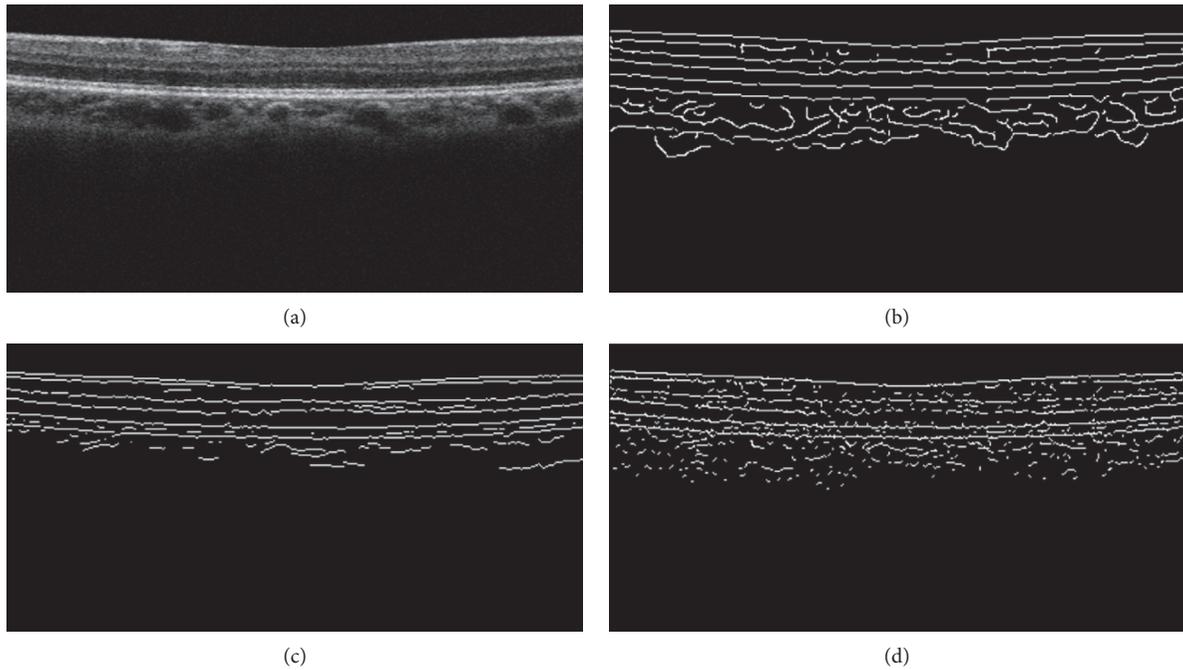


FIGURE 4: Detection results from three edge detection methods. (a) Original retinal OCT image, (b) Canny detector, (c) two-pass edge detection algorithm, and (d) EdgeFlow edge detection technique.

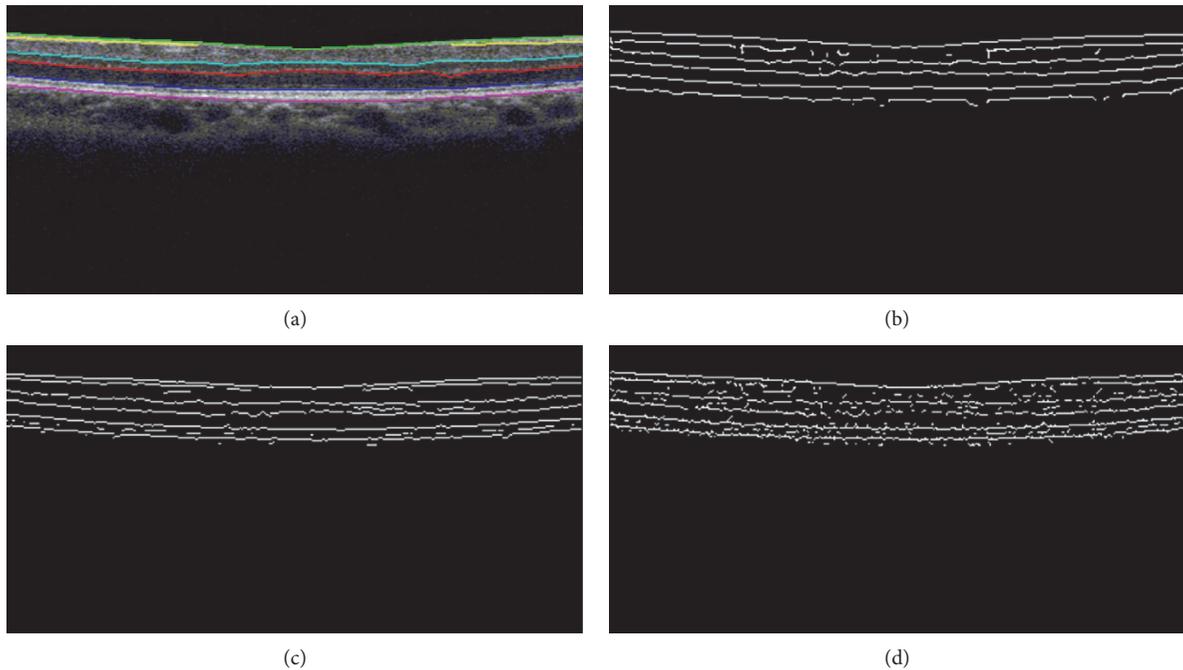


FIGURE 5: Edge extraction in ROI, from human observer and three algorithms. (a) Retinal layer boundaries marked by a human observer, (b) edges within ROI for the Canny edge detector, (c) edges within ROI for two-pass edge detection algorithm, and (d) edge within ROI for EdgeFlow technique.

TPR for two-pass is also significantly higher than that for Canny ($p = 4.29e - 4$) and that for the EdgeFlow method ($p = 0.0058$). The same pattern occurs when measured with ACC. ACC for the two-pass method is significantly higher than that for the Canny ($p = 3.15e - 3$) and

EdgeFlow method ($p = 5.61e - 4$). On the other hand, FPR for two-pass method is significantly lower than that for Canny ($p = 0.00024$) and EdgeFlow ($p = 2.73e - 04$). All the metrics suggest that two-pass method is the best among the three. However, Canny and EdgeFlow methods

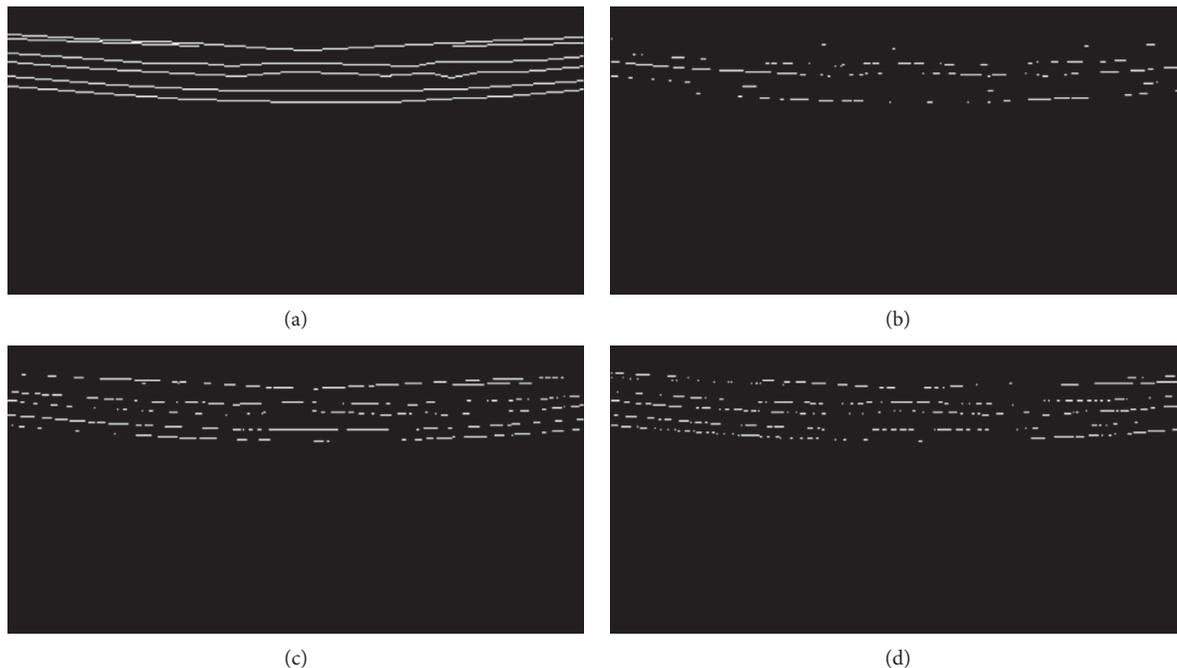


FIGURE 6: Edge pixels from three edge detectors overlapping with the ground truth. (a) The ground truth edges marked by human observer. (b) The true positive points obtained by Canny edge detector. (c) The true positive points obtained by two-pass edge detection algorithm. (d) The true positive points obtained by EdgeFlow technique.

TABLE 2: Performance evaluation using metrics FOM, TPR, FPR, and ACC.

Edge detectors	FOM	TPR	FPR	ACC
Canny edge detector	0.36 ± 0.0865	0.22 ± 0.0816	0.75 ± 0.0881	0.98 ± 0.0026
Two-pass edge detection algorithm	0.67 ± 0.0726	0.41 ± 0.0968	0.57 ± 0.1182	0.98 ± 0.0033
EdgeFlow technique	0.50 ± 0.0942	0.27 ± 0.0875	0.79 ± 0.094	0.97 ± 0.0027

TABLE 3: Two-sample *t*-test on comparison of the edge detection performance.

Samples	FOM	TPR	FPR	ACC
Canny and two-pass	$H_0: m(c) > m(t)$ $p = 1.19e - 06$	$H_0: m(c) > m(t)$ $p = 4.29e - 04$	$H_0: m(c) < m(t)$ $p = 0.0024$	$H_0: m(c) > m(t)$ $p = 3.15e - 03$
Canny and EdgeFlow	$H_0: m(c) > m(e)$ $p = 4.65e - 03$	$H_0: m(c) = m(e)$ $p = 0.2045$	$H_0: m(c) = m(e)$ $p = 0.3727$	$H_0: m(c) = m(e)$ $p = 0.3304$
Two-pass and EdgeFlow	$H_0: m(t) < m(e)$ $p = 5.2769e - 04$	$H_0: m(t) < m(e)$ $p = 0.0058$	$H_0: m(t) > m(e)$ $p = 2.73e - 04$	$H_0: m(t) < m(e)$ $p = 5.6127e - 04$

are not significantly different when measured using TPR ($p = 0.2045$), FPR ($p = 0.3727$), or ACC ($p = 0.3304$).

However, when comparing Figures 5(b), 5(c), and 5(d) and Figures 6(b), 6(c), and 6(d) one by one, the Canny edge detector results in much less edge pixels on locations of the ground truth, although its outcome seems good as well in Figure 5(b). This suggests that the Canny edge detector introduced edge shifts and resulted in lower performance score in Tables 2 and 3. To examine the possibility, we calculated the mean localization deviation (MLD) for the three edge detectors. The results in Table 4 indeed show the largest

TABLE 4: Mean localization deviation (MLD).

Canny edge detector	Two-pass edge detection	EdgeFlow technique
1.27 ± 0.3340	1.02 ± 0.2484	1.05 ± 0.1831

mean value of MLD for the Canny edge detector. Two-sample *t*-tests (Table 5) confirmed that the MLD value for the Canny edge detector is significantly higher than that for the two-pass edge detection algorithm ($p = 0.04$).

TABLE 5: Two-sample t -test on the adjusted performance metrics and MLD.

Samples	FOM_{ADJ}	TPR_{ADJ}	FPR_{ADJ}	ACC_{ADJ}	MLD
Canny and two-pass	$H_0: m(c) > m(t)$ $p = 4.41e - 11$	$H_0: m(c) = m(t)$ $p = 0.3933$	$H_0: m(c) = m(t)$ $p = 0.4613$	$H_0: m(c) = m(t)$ $p = 0.4175$	$H_0: m(c) < m(t)$ $p = 0.04003$
Canny and EdgeFlow	$H_0: m(c) > m(e)$ $p = 8.13e - 07$	$H_0: m(c) < m(e)$ $p = 0.01945$	$H_0: m(c) = m(e)$ $p = 0.1471$	$H_0: m(c) = m(e)$ $p = 0.1635$	$H_0: m(c) < m(e)$ $p = 0.0581$
Two-pass and EdgeFlow	$H_0: m(t) < m(e)$ $p = 2.6132e - 06$	$H_0: m(t) < m(e)$ $p = 8.1322e - 05$	$H_0: m(t) > m(e)$ $p = 2.41e - 04$	$H_0: m(t) < m(e)$ $p = 3.2125e - 05$	$H_0: m(t) = m(e)$ $p = 0.1074$

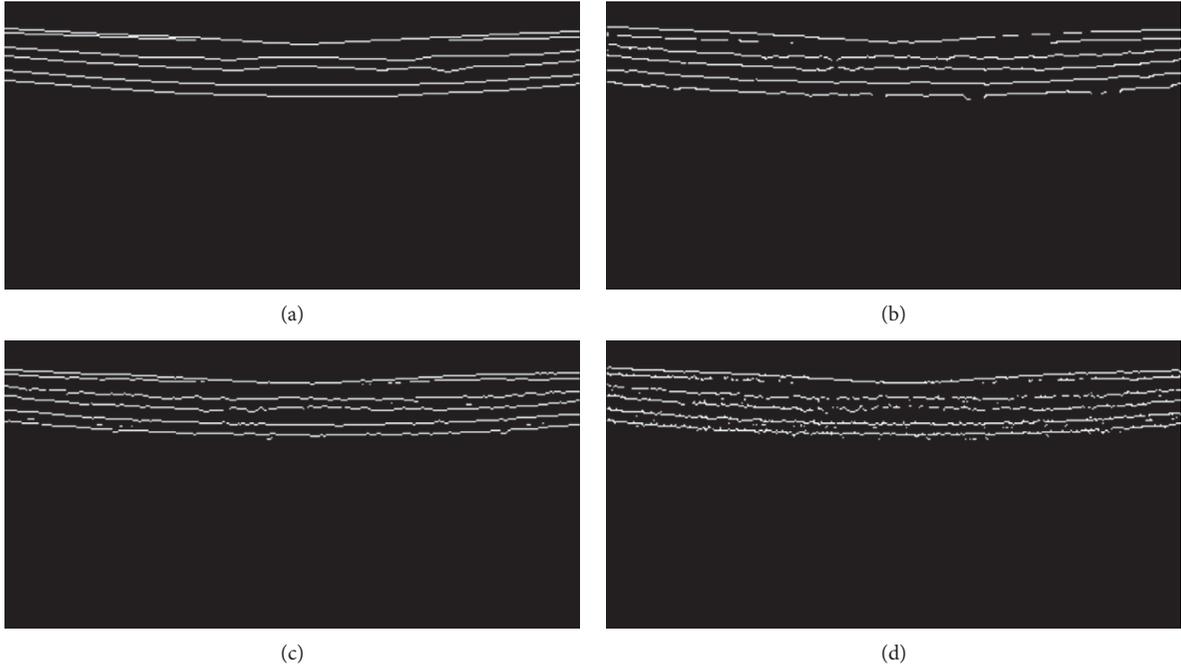


FIGURE 7: Edges for the adjusted performance metrics. (a) The ground truth edges, (b) the edge pixels for Canny edge detector, (c) the edge pixels for the two-pass edge detection algorithm, and (d) the edge pixels for the EdgeFlow technique.

and almost significantly higher than that for the EdgeFlow method ($p = 0.058$).

For OCT retinal image layer thickness measurement, if the edge shift is within a small range and to the same direction, the outcome may not be significantly influenced. Therefore, we examine how the edge detectors perform when measured with a second set of metrics; the adjusted measures include FOM_{ADJ} , TPR_{ADJ} , FPR_{ADJ} , and ACC_{ADJ} , which were developed in Materials and Methods. Figure 7(a) is the ground truth edges; Figures 7(b), 7(c), and 7(d), respectively, show detected edge pixels within 2 pixels searching neighborhood corresponding to its ground truth edge for the three edge detectors. Visually, the outcome for Canny edge detector is much improved when compared to that in Figure 6(b); the edges for EdgeFlow method are much noisy. Quantitative metrics for the edge detectors based on Figure 7 are summarized in Table 6.

The results for FOM, TPR, ACC, and FPR after the adjustment (Table 6) are all better than those before the adjustment (Table 2). This finding is reasonable in that the

false alarm reduces, and FOM, TPR, and ACC increase, when more detected pixels are considered as correct edge. Note that in both tables, the higher values for FOM, TPR, and ACC mean the better performance, whereas for FPR, the lower the value, the better the performance.

Detailed statistical analysis on the performance of three edge detectors measured with the adjusted metrics is summarized in Table 5. For FOM_{ADJ} measure, the two-pass is again significantly better than both Canny edge detector ($p = 4.41e - 11$) and the EdgeFlow method ($p = 2.61e - 6$); EdgeFlow is also significantly better than Canny ($p = 8.13e - 7$). When measured with TPR_{ADJ} , FPR_{ADJ} , and ACC_{ADJ} , Canny does not differ significantly from the two-pass method ($p = 0.3933$, $p = 0.4613$, and $p = 0.4175$, resp.); a similar pattern (except for TPR_{ADJ}) occurs for Canny and the EdgeFlow method ($p = 0.01945$, $p = 0.1471$, and $p = 0.1635$, resp.). However, these measures show that the two-pass method is significantly better than the EdgeFlow technique ($p = 8.1322e - 05$, $p = 2.41e - 04$, and $p = 3.2125e - 05$, resp.).

TABLE 6: FOM_{ADJ} , TPR_{ADJ} , FPR_{ADJ} , and ACC_{ADJ} after redefining the detected edges.

Edge detectors	FOM_{ADJ}	TPR_{ADJ}	FPR_{ADJ}	ACC_{ADJ}
Canny edge detector	0.51 ± 0.026	0.88 ± 0.0841	0.12 ± 0.0848	0.99 ± 0.0024
Two-pass edge detection algorithm	0.81 ± 0.0416	0.91 ± 0.0277	0.093 ± 0.0355	0.99 ± 0.0009
Edge flow technique	0.66 ± 0.0409	0.85 ± 0.0125	0.16 ± 0.014	0.99 ± 0.0004

Finally, we noticed in our experiments that the EdgeFlow technique took much longer time for each edge detection task. When the time was measured for processing a sample of OCT retinal images (200 by 400 pixels in the ROI), the average computational time is 2.77 ± 1.24 seconds, 3.85 ± 0.18 seconds, and 467.66 ± 1.33 seconds, respectively, for two-pass, Canny, and EdgeFlow methods, confirming that the EdgeFlow approach took a significantly larger amount of time than the other two algorithms.

4. Discussion and Conclusions

In this study, we intended to search for the edge detectors that best suit for the OCT retinal image segmentation task. With the analysis of literature and our experiment, we have identified the most promising candidate algorithms, namely, Canny edge detector, the two-pass method, and the EdgeFlow technique. Using the performance evaluation metrics (FOM, TPR, FPR, and ACC) and their adjusted versions (FOM_{ADJ} , TPR_{ADJ} , FPR_{ADJ} , and ACC_{ADJ}), we examined the three methods applied to the realistic OCT retinal images. Our results show that the two-pass method consistently outperforms the other two. In addition, the MLD metrics shows that the two-pass method caused smaller edge shifting problem. Although the computational cost for the two-pass method is slightly higher than the Canny edge detector, it is over 100 times lower than that for the EdgeFlow technique. Based on the above analysis and findings, we conclude that the two-pass method is among the three the best approach to edge detection for the OCT retinal layer image segmentation task. Furthermore, the outperformance of two-pass method measured by the original and adjusted metrics and the advantage of Canny edge detector over EdgeFlow technique in terms of FOM_{ADJ} and TPR_{ADJ} and MLD lead to another conclusion that the intensity-based edge detectors outperform the texture-based edge detector for OCT retinal image analysis.

The findings in the study suggest that it is critical to use the most appropriate algorithms to detect the retinal layer boundaries in the OCT images in order to automate the quantitative analysis of retinal OCT images. Combined with the findings in the literature that EdgeFlow method significantly outperformed Canny algorithm in texture segregation tasks [22], this study offers support to the idea that the performance of edge detectors is image property dependent [20] as both Canny and two-pass methods surpass EdgeFlow in the current application. In line with this thought and findings, it is necessary argue that the best performer for normal retinal OCT images also work best for pathological retinal images. Additionally, the intensity-gradient based methods (two-pass and Canny algorithms) outperforming texture-

based method (EdgeFlow) might suggest that the OCT images contain more intensity gradient changes than texture changes along the longitudinal direction. The relative weight of intensity and texture information in OCT retinal image warrants further study in the future.

With the development of OCT technologies and their applications in the field of ophthalmology, more and more data is readily available. Extracting meaningful information from the ever-increasing volume of clinical data reliably and efficiently forms the basis for modern medical decision making and research. Reliable and efficient OCT retinal image segmentation will contribute to the development of this trend. Future research efforts would need to overcome several limitations in this study. First, the input parameters used in our experiments were selected over a relatively small sample space and the decisions on the “optimal” parameters were subject to human subjectiveness. Although it is almost impossible to identify the absolutely optimal input parameters for each edge detector [20], the choice of optimal input parameters may be improved by conducting a large number of experiments and averaging opinions from more expert viewers. The second limitation in our study is the use of a single expert observer to define the ground truth. Individual subjectiveness may be reduced by averaging across multiple decisions for the ground truth. Moreover, our data were all collected from voluntary healthy subjects. If the edge detectors perform differently for different types of images, it is necessary to examine how they perform on pathological retinal images in future studies.

Conflicts of Interest

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

Authors’ Contributions

Su Luo and Jing Yang contributed equally to this work.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (no. 61271154 and no. 61172033). The authors would like to thank the peer reviewers for their insightful and constructive suggestions.

References

- [1] D. Huang, E. A. Swanson, C. P. Lin et al., “Optical coherence tomography,” *Science*, vol. 254, no. 5035, pp. 1178–1181, 1991.
- [2] P. Hrynchak and T. Simpson, “Optical coherence tomography: an introduction to the technique and its use,” *Optometry &*

- Vision Science Official Publication of the American Academy of Optometry*, vol. 77, no. 7, pp. 347–356, 2000.
- [3] W. Geitzenauer, C. K. Hitzengerger, and U. M. Schmidt-Erfurth, “Retinal optical coherence tomography: past, present and future perspectives,” *British Journal of Ophthalmology*, vol. 95, no. 2, pp. 171–177, 2011.
 - [4] M. Sonka and M. D. Abramoff, “Quantitative analysis of retinal OCT,” *Medical Image Analysis*, vol. 33, pp. 165–169, 2016.
 - [5] G. Quellec, K. Lee, M. Dolejsi, M. K. Garvin, M. D. Abramoff, and M. Sonka, “Three-dimensional analysis of retinal layer texture: identification of fluid-filled regions in SD-OCT of the macula,” *IEEE Transactions on Medical Imaging*, vol. 29, no. 6, pp. 1321–1330, 2010.
 - [6] H. Bogunovic, M. Sonka, Y. H. Kwon, P. Kemp, M. D. Abramoff, and X. Wu, “Multi-surface and multi-field co-segmentation of 3-D retinal optical coherence tomography,” *IEEE Transactions on Medical Imaging*, vol. 33, no. 12, pp. 2242–2253, 2014.
 - [7] X. Chen, P. Hou, and C. Jin, “Quantitative analysis of retinal layer optical intensities on three-dimensional optical coherence tomography,” *Investigative Ophthalmology & Visual Science*, vol. 54, no. 10, pp. 6846–6851, 2013.
 - [8] D. Koozekanani, K. Boyer, and C. Roberts, “Retinal thickness measurements in optical coherence tomography using a Markov boundary model,” *IEEE Transactions on Medical Imaging*, vol. 20, no. 9, pp. 900–916, 2001.
 - [9] A. Lang, A. Carass, and P. A. Calabresi, “An adaptive grid for graph-based segmentation in retinal OCT,” *Proceedings of SPIE - the International Society for Optical Engineering*, vol. 9034, no. 10, pp. 1314–1323, 2014.
 - [10] R. J. Zawadzki, A. R. Fuller, and D. F. Wiley, “Adaptation of a support vector machine algorithm for segmentation and visualization of retinal structures in volumetric optical coherence tomography data sets,” *Journal of Biomedical Optics*, vol. 12, no. 4, article 041206, 2007.
 - [11] F. D. Cabrera, H. M. Salinas, and C. A. Puliafito, “Automated detection of retinal layer structures on optical coherence tomography images,” *Optics Express*, vol. 13, no. 25, pp. 10200–10216, 2005.
 - [12] Q. Yang, C. A. Reisman, and Z. Wang, “Automated layer segmentation of macular OCT images using dual-scale gradient information,” *Optics Express*, vol. 18, no. 20, pp. 21293–21307, 2010.
 - [13] S. Lu, C. Y. Cheung, J. Liu, J. H. Lim, C. K. Leung, and T. Y. Wong, “Automated layer segmentation of optical coherence tomography images,” *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 10, pp. 2605–2608, 2010.
 - [14] A. M. Bagci, R. Ansari, and M. Shahidi, “A method for detection of retinal layers by optical coherence tomography image segmentation,” in *2007 IEEE/NIH Life Science Systems and Applications Workshop*, pp. 144–147, Bethesda, MD, USA, 2007.
 - [15] A. M. Bagci, M. Shahidi, R. Ansari, M. Blair, N. P. Blair, and R. Zelkha, “Thickness profiles of retinal layers by optical coherence tomography image segmentation,” *American Journal of Ophthalmology*, vol. 146, no. 5, pp. 679–687, 2008.
 - [16] F. Rossant, I. Ghorbel, and I. Bloch, “Automated segmentation of retinal layers in OCT imaging and derived ophthalmic measures,” in *2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pp. 1370–1373, Boston, MA, USA, 2009.
 - [17] Y. M. Cha and J. H. Han, “High-accuracy retinal layer segmentation for optical coherence tomography using tracking kernels based on Gaussian mixture model,” *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 20, no. 2, pp. 1–10, 2014.
 - [18] A. Lang, A. Carass, and M. Hauser, “Retinal layer segmentation of macular OCT images using boundary classification,” *Biomedical Optics Express*, vol. 4, no. 7, pp. 1133–1152, 2013.
 - [19] S. Niu, Q. Chen, and S. L. De, “Automated retinal layers segmentation in SD-OCT images using dual-gradient and spatial correlation smoothness constraint,” *Computers in Biology & Medicine*, vol. 54, no. C, pp. 116–128, 2014.
 - [20] M. D. Heath, S. Sarkar, and T. A. Sanocki, “A robust visual method for assessing the relative performance of edge-detection algorithms,” *IEEE Transactions on Pattern Analysis & Machine Intelligence*, vol. 19, no. 12, pp. 1338–1359, 1997.
 - [21] A. Nicogossian, O. Kloiber, and B. Stabile, “The Revised World Medical Association’s Declaration of Helsinki 2013: enhancing the protection of human research subjects and empowering ethics review committees,” *World Medical & Health Policy*, vol. 6, no. 1, pp. 1–3, 2014.
 - [22] W. Y. Ma and B. S. Manjunath, “EdgeFlow: a technique for boundary detection and image segmentation,” *IEEE Transactions on Image Processing*, vol. 9, no. 8, pp. 1375–1388, 2000.
 - [23] J. Canny, “A computational approach to edge detection,” *IEEE Transactions on Pattern Analysis & Machine Intelligence*, vol. 8, no. 6, pp. 679–698, 1986.
 - [24] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, Publishing House of Electronics Industry, 2nd edition, 2002.
 - [25] R. Román-Roldán, J. F. Gómez-Lopera, and C. Atae-Allah, “A measure of quality for evaluating methods of segmentation and edge detection,” *Pattern Recognition*, vol. 34, no. 5, pp. 969–980, 2001.
 - [26] I. E. Abdou and W. Pratt, “Quantitative design and evaluation of enhancement/thresholding edge detectors,” *Proceedings of the IEEE*, vol. 67, no. 5, pp. 753–763, 1979.
 - [27] X. Yin, B. W. Ng, and J. He, “Accurate image analysis of the retina using hessian matrix and binarisation of thresholded entropy with application of texture mapping,” *PLoS One*, vol. 9, no. 4, article e95943, 2014.
 - [28] E. Nadernejad, S. Sharifzadeh, and H. Hassanpour, “Edge detection techniques: evaluations and comparisons,” *Applied Mathematics Sciences*, vol. 2, no. 31, pp. 1507–1520, 2008.
 - [29] E. R. Davies, *Computer and Machine Vision: Theory, Algorithms, Practicalities*, China Machine Press, 4th edition, 2013.
 - [30] M. Beauchemin, “On the Hausdorff distance used for the evaluation of segmentation results,” *Canadian Journal of Remote Sensing*, vol. 24, no. 1, pp. 3–8, 1998.
 - [31] W. K. Pratt, *Digital Image Processing*, Wiley-Inter-science, New York, NY, USA, 1978.
 - [32] K. C. Strasters and J. J. Gerbrands, “Three-dimensional image segmentation using a split, merge and group approach,” *Pattern Recognition Letters*, vol. 12, no. 5, pp. 307–325, 1991.
 - [33] S. J. Chiu, X. T. Li, and P. Nicholas, “Automatic segmentation of seven retinal layers in SDOCT images congruent with expert manual segmentation,” *Optics Express*, vol. 18, no. 18, pp. 19413–19428, 2010.
 - [34] Z. Xinyan and Z. Ranyang, “An adaptive edge-detection algorithm based on Canny and its performance evaluation,”

- Computer Technology and Development*, vol. 25, no. 11, pp. 32–37, 2015.
- [35] F. van der Heyden, “Evaluation of edge detection algorithms,” in *Third International Conference on Image Processing and its Applications, 1989*, pp. 618–622, Warwick, UK, 1989.
- [36] S. Chabrier, H. Laurent, and B. Emile, “A comparative study of supervised evaluation criteria for image segmentation,” in *2004 12th European Signal Processing Conference*, pp. 1143–1146, Vienna, Austria, 2004.
- [37] E. Götzinger, M. Pircher, B. Baumann et al., “Speckle noise reduction in high speed polarization sensitive spectral domain optical coherence tomography,” *Optics Express*, vol. 19, no. 15, pp. 14568–14585, 2011.
- [38] Y. Huang, R. P. Danis, and J. W. Pak, “Development of a semi-automatic segmentation method for retinal OCT images tested in patients with diabetic macular edema,” *PLoS One*, vol. 8, no. 12, article e82922, 2013.
- [39] A. M. Abhishek, T. T. J. M. Berendschot, and S. V. Rao, “Segmentation and analysis of retinal layers (ILM & RPE) in optical coherence tomography images with edema,” in *2014 IEEE Conference on Biomedical Engineering and Sciences (IECBES)*, pp. 204–209, Kuala Lumpur, Malaysia, 2014.
- [40] G. R. Wilkins, O. M. Houghton, and A. L. Oldenburg, “Automated segmentation of intraretinal cystoid fluid in optical coherence tomography,” *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 4, pp. 1109–1114, 2012.

Research Article

Deep Learning- and Transfer Learning-Based Super Resolution Reconstruction from Single Medical Image

YiNan Zhang^{1,2} and MingQiang An³

¹School of Computer Science and Technology, Beijing Institute of Technology, Beijing 100081, China

²College of Computer Science and Information Engineering, Tianjin University of Science and Technology, Tianjin 300222, China

³College of Science, Tianjin University of Science and Technology, Tianjin 300222, China

Correspondence should be addressed to YiNan Zhang; zh.y.n@163.com

Received 23 January 2017; Revised 5 April 2017; Accepted 10 May 2017; Published 6 July 2017

Academic Editor: Shujun Fu

Copyright © 2017 YiNan Zhang and MingQiang An. 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.

Medical images play an important role in medical diagnosis and research. In this paper, a transfer learning- and deep learning-based super resolution reconstruction method is introduced. The proposed method contains one bicubic interpolation template layer and two convolutional layers. The bicubic interpolation template layer is prefixed by mathematics deduction, and two convolutional layers learn from training samples. For saving training medical images, a SIFT feature-based transfer learning method is proposed. Not only can medical images be used to train the proposed method, but also other types of images can be added into training dataset selectively. In empirical experiments, results of eight distinctive medical images show improvement of image quality and time reduction. Further, the proposed method also produces slightly sharper edges than other deep learning approaches in less time and it is projected that the hybrid architecture of prefixed template layer and unfixed hidden layers has potentials in other applications.

1. Introduction

Medical imaging [1] is an important tool to determine the presence of many diseases and analysis of experimental results. Enlarging medical images [2] can provide medical experts with more details for elevating diagnosis accuracy [3] in pathology research [4]. Therefore, the medical image enhancement becomes to be a hotspot. Further, enlarged medical images may substantially help computer-aided automatic detection [5]. For example, majority of single-detector spiral computer tomography (CT) [6] scanners and magnetic resonance imaging (MRI) [7] produce medical images as effective noninvasive examinations. Because of technical restrictions, such types of medical images are obtained in relatively low resolution (LR) and not suitable for further analyzing. Therefore, super resolution reconstruction (SRR) [8] methods can be used for solving this type of problems.

For medical images, many SRR methods were proposed. Those methods fuse some LR images from the same scene to one high-resolution (HR) image. Many corresponding

machine learning [9] methods were proved effectively. Wang et al. [10] proposed a sparse representation-based SRR method in 2014. Rousseau et al. [11] proposed a SRR approach from multiple low-resolution images. In 2011, Liyakathunisa [12] presented a SRR method which uses progressively DCT- and zonal filter-based denoising. Those researches were proved effective and efficient; however, those methods are limited by two conditions: (1) the input images must be a set of LR images and (2) the training dataset must be big enough which contains LR and HR medical images in pairs.

Those two conditions raise the following problems in reality: (1) a set of low-resolution medical images cannot be obtained for potential reasons, such as costs to patients, and (2) to machine learning approach, a big training dataset of LR and HR medical images is a giant cost in commercial project. For solving those problems, three conventional methods have been widely used. Nearest neighbor interpolation, bilinear interpolation, and cubic convolution interpolation are widely used [13, 14]. However, those three reconstruction approaches produce

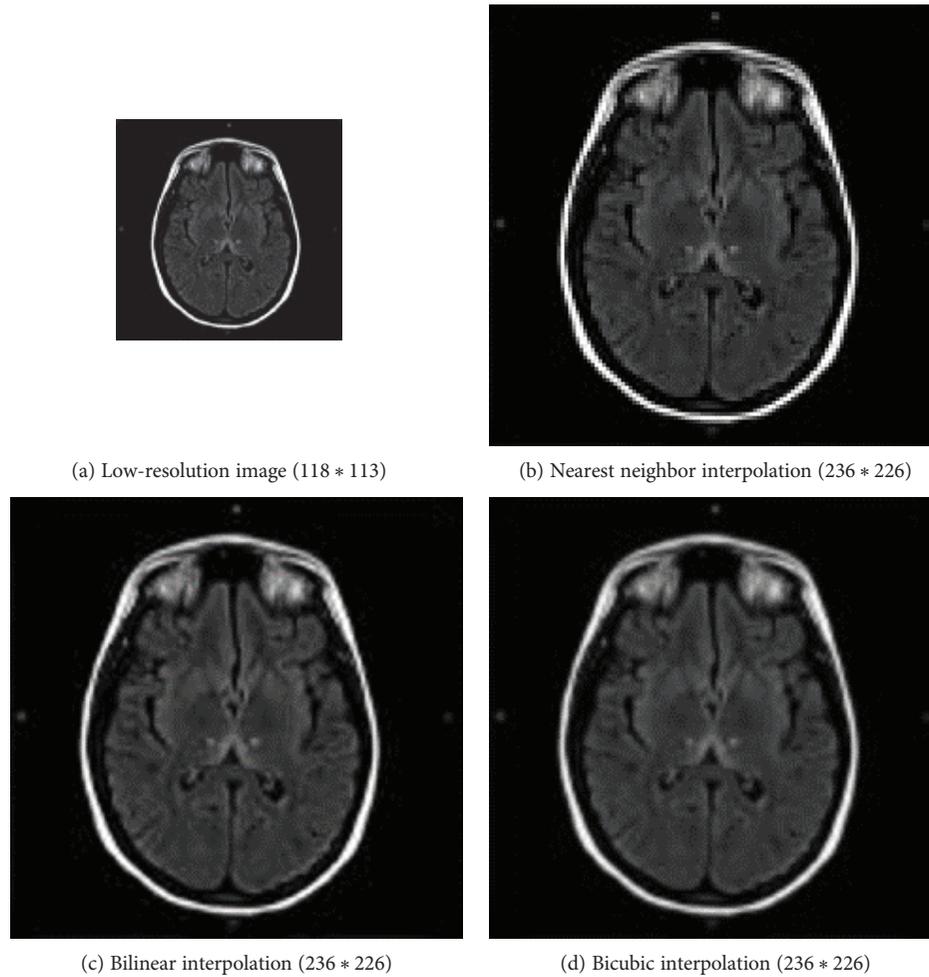


FIGURE 1: Three conventional SRR methods (MRI: brain).

high-resolution image with blur edges or image aliasing [15]. In medical field, a better SRR method is needed in urgent. Combining deep learning with transfer learning to reconstruct a HR image from one single LR image becomes to be a feasible and practical solution.

In this paper, we adapt deep learning [16] and transfer learning [17] to achieve SRR for medical images. First, a deep conventional neural network (DCNN) is designed for fulfilling SRR. Second, SIFT feature-based transfer learning is used to enlarge the training dataset of the DCNN by using the public image dataset. Finally, the trained DCNN can reconstruct a HR medical image from a given LR medical image. The major contributions of this paper are the following:

- (1) Using deep learning to achieve better SRR result than conventional methods
- (2) Using transfer learning to enlarge training dataset for DCNN
- (3) Transfer learning can reduce costs of preparing medical images in reality
- (4) SIFT feature-based transfer learning and DCNN can offer sharper edge

- (5) Proposing a hybrid DCNN structure which contains a prefixed template layer.

The rest of this paper is organized as follows: Section 2 shows SRR in medical research and related works, Section 3 presents the proposed method, and Section 4 shows experiments and results. In Section 5, a conclusion is drawn.

2. Backgrounds and Related Works

There are three conventional super resolution reconstruction methods: nearest neighbor (NN) interpolation [18], bilinear interpolation [19], and bicubic interpolation [20]. An example of SRR for medical image (MRI: brain) is shown in Figure 1.

Figure 1(a) is a LR. Figures 1(b), 1(c), and 1(d) are results of those three conventional SRR methods. Advantages and disadvantages of those conventional SRR methods can be summarized in Table 1.

Plenge et al. [21] proposed a SRR method by using cross-scale self-similarity in multislice MRI. Tao et al. [22] showed an SSR method which is SRR of late gadolinium-enhanced MRI from multiple views in 2014. And Zhao et al. [23]

TABLE 1: Advantages and disadvantages of three conventional SRR methods.

Interpolation method	Advantages	Disadvantages
Nearest neighbor	Easy to implement Very fast	Problem of image aliasing
Bilinear	Antialiasing Considering with 4 nearest pixels	Blur edges
Bicubic	Antialiasing Considering with 16 nearest pixels	Slightly blur edges Relatively slow

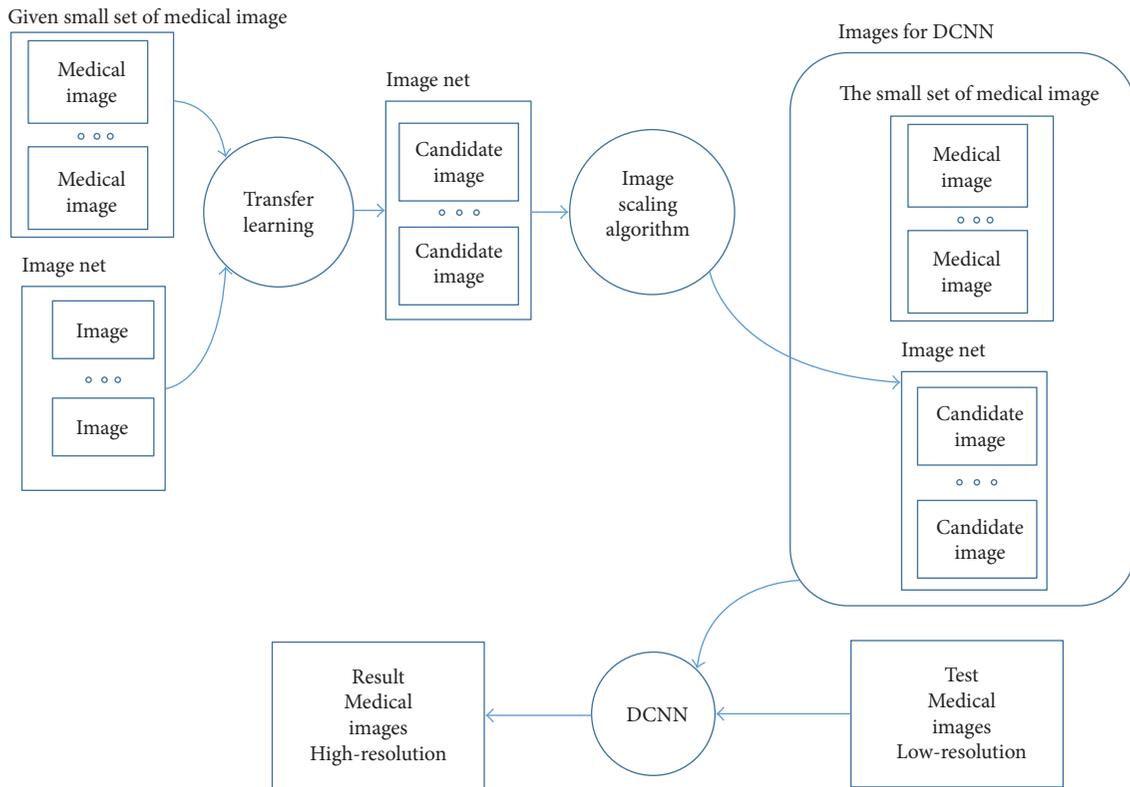


FIGURE 2: Overview: deep learning- and transfer learning-based SRR for single medical image.

proposed a multiframe SRR algorithm based on diffusion tensor regularization term in 2014.

Many effective SRR methods are proposed, but most of them are based on a group of LR medical images. Furthermore, nearest neighbor interpolation, bilinear interpolation, and bicubic interpolation can achieve SRR result for single medical image task now, but a better SRR method is needed in medical research and clinical diagnosis.

3. The Proposed Method

The proposed method includes three distinctive parts of algorithms/techniques: (1) SIFT feature-based transfer learning, (2) image scaling-down algorithm, and (3) deep learning (deep convolutional neural network (DCNN)). Image scaling-down algorithm is a conventional algorithm. Using SIFT feature-based transfer learning and a hybrid DCNN

structure is the major contribution in this paper. As Figure 2 shows, a small set of medical image samples is offered first. After the processing of SIFT feature-based transfer learning and image scaling-down algorithm, a big set of training dataset is prepared to train deep neural network. Finally, one LR image can be reconstructed to one HR image. Details are shown in Sections 3.1.1–3.1.5. Because this work is an improvement of SRCNN which was proposed by Dong et al. [24], Sections 3.1.2–3.1.5 have likely expressions as in reference [24].

3.1. Hybrid Deep Convolutional Neural Network Architectures. We improve the SRCNN [24] structure and present a hybrid structure of DCNN in this section. The proposed DCNN contains three layers; it has one prefixed bicubic interpretation layer by mathematics deduction and two convolutional layers which learn from given training

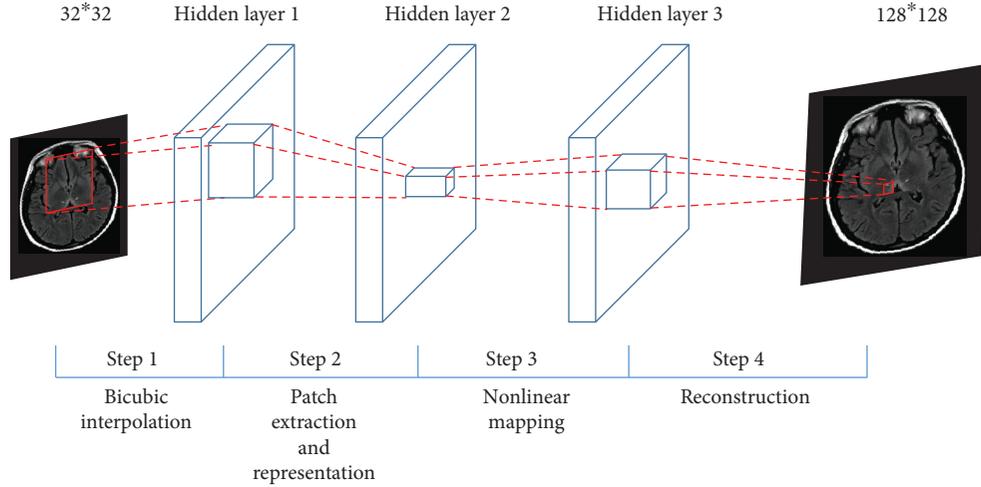


FIGURE 3: Deep convolutional neural network for medical image SRR.

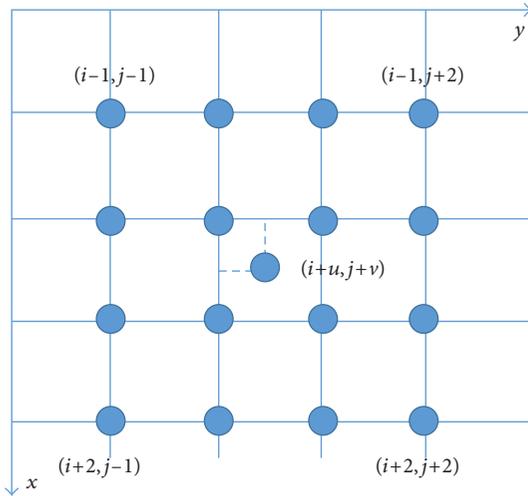


FIGURE 4: Bicubic interpretation.

images. As Figure 3 shows, the LR image is the input and the HR image is the output. Three hidden layers are designed for different tasks. The hidden layer 1 is a bicubic interpolation template layer which is used to fulfill bicubic interpolation. The hidden layer 2 is designed for patch extraction and representation, and the hidden layer 3 is designed for nonlinear mapping. Finally, a reconstruction result is produced at the end of DCNN. The size of hidden layer 1 is $f_1 * f_1 * n_1$, the size of hidden layer 2 is $f_2 * f_2 * n_2$, and the size of hidden layer 3 is $f_3 * f_3 * n_3$.

Steps of Figure 3 are as follows:

Step 1. This step fulfills the fast-bicubic interpolation; details of building this bicubic interpretation template layer are shown in Section 3.1.1.

Step 2. In this step, a convolutional layer achieves patch extraction and representation; details are shown in Section 3.1.2.

TABLE 2: Bicubic interpretation template parameters.

Template	Conditions	Discretized parameters
T_1	$v \in [0, 1/4], u \in [0, 1/4]$	Let $v = 1/8$ and $u = 1/8$
T_2	$v \in [0, 1/4], u \in [1/4, 2/4]$	Let $v = 1/8$ and $u = 3/8$
T_3	$v \in [0, 1/4], u \in [2/4, 3/4]$	Let $v = 1/8$ and $u = 5/8$
T_4	$v \in [0, 1/4], u \in [3/4, 1]$	Let $v = 1/8$ and $u = 7/8$
T_5	$v \in [1/4, 2/4], u \in [0, 1/4]$	Let $v = 3/8$ and $u = 1/8$
T_6	$v \in [1/4, 2/4], u \in [1/4, 2/4]$	Let $v = 3/8$ and $u = 3/8$
T_7	$v \in [1/4, 2/4], u \in [2/4, 3/4]$	Let $v = 3/8$ and $u = 5/8$
T_8	$v \in [1/4, 2/4], u \in [3/4, 1]$	Let $v = 3/8$ and $u = 7/8$
T_9	$v \in [2/4, 3/4], u \in [0, 1/4]$	Let $v = 5/8$ and $u = 1/8$
T_{10}	$v \in [2/4, 3/4], u \in [1/4, 2/4]$	Let $v = 5/8$ and $u = 3/8$
T_{11}	$v \in [2/4, 3/4], u \in [2/4, 3/4]$	Let $v = 5/8$ and $u = 5/8$
T_{12}	$v \in [2/4, 3/4], u \in [3/4, 1]$	Let $v = 5/8$ and $u = 7/8$
T_{13}	$v \in [3/4, 1], u \in [0, 1/4]$	Let $v = 7/8$ and $u = 1/8$
T_{14}	$v \in [3/4, 1], u \in [1/4, 2/4]$	Let $v = 7/8$ and $u = 3/8$
T_{15}	$v \in [3/4, 1], u \in [2/4, 3/4]$	Let $v = 7/8$ and $u = 5/8$
T_{16}	$v \in [3/4, 1], u \in [3/4, 1]$	Let $v = 7/8$ and $u = 7/8$

Step 3. This step maps high-dimensional vectors into another high-dimensional space; details are shown in Section 3.1.3.

Step 4. This reconstruction step produces the final HR image; details are shown in Section 3.1.4.

3.1.1. Step 1: Prefixed Bicubic Interpolation Layer. Bicubic interpolation methods use extra points to fit the sampling functions, a critical problem is time cost. As Figure 4 shows, bicubic interpretation method uses 16 adjacent pixels as inputs and cubic polynomial approximation function of the best interpretation function $\sin(\pi\omega)/(\pi\omega)$ in theory.

TABLE 3: Bicubic interpretation template solutions.

$T_1 = \frac{1}{2^{18}} \begin{bmatrix} 2401 & -24353 & -3479 & 343 \\ -24353 & 247009 & 35287 & -3479 \\ -3479 & 35287 & 5041 & -497 \\ 343 & -3479 & -497 & 49 \end{bmatrix}$	$T_2 = \frac{1}{2^{18}} \begin{bmatrix} 3675 & -19355 & -11613 & 2205 \\ -37275 & 196315 & 117789 & -22365 \\ -5325 & 28045 & 16827 & -3195 \\ 525 & -2765 & -1659 & 315 \end{bmatrix}$
$T_3 = \frac{1}{2^{18}} \begin{bmatrix} 2205 & -11613 & -19355 & 3675 \\ -22365 & 117789 & 196315 & -37275 \\ -3195 & 16827 & 28045 & -5325 \\ 315 & -1659 & -2765 & 525 \end{bmatrix}$	$T_4 = \frac{1}{2^{18}} \begin{bmatrix} 343 & -3479 & -24353 & 2401 \\ -3479 & 35287 & 247009 & -24353 \\ -497 & 5041 & 35287 & -3479 \\ 49 & -497 & -3479 & 343 \end{bmatrix}$
$T_5 = \frac{1}{2^{18}} \begin{bmatrix} 3675 & -37275 & -5325 & 525 \\ -19355 & 196315 & 28045 & -2765 \\ -11613 & 117789 & 16827 & -1659 \\ 2205 & -22365 & -3195 & 315 \end{bmatrix}$	$T_6 = \frac{1}{2^{18}} \begin{bmatrix} 5625 & -29625 & -17775 & 3375 \\ -29625 & 156025 & 93615 & -17775 \\ -17775 & 93615 & 56169 & -10665 \\ 3375 & -17775 & -10665 & 2025 \end{bmatrix}$
$T_7 = \frac{1}{2^{18}} \begin{bmatrix} 3375 & -17775 & -29625 & 5625 \\ -17775 & 93615 & 156025 & -29625 \\ -10665 & 56169 & 93615 & -17775 \\ 2205 & -10665 & -17775 & 3375 \end{bmatrix}$	$T_8 = \frac{1}{2^{18}} \begin{bmatrix} 525 & -5325 & -37275 & 3675 \\ -2765 & 28045 & 196315 & -19355 \\ -1659 & 16827 & 117789 & -11613 \\ 315 & -3195 & -22365 & 2205 \end{bmatrix}$
$T_9 = \frac{1}{2^{18}} \begin{bmatrix} 2205 & -22365 & -3195 & 315 \\ -11613 & 117789 & 16827 & -1659 \\ -19355 & 196315 & 28045 & -2765 \\ 3675 & -37275 & -5325 & 525 \end{bmatrix}$	$T_{10} = \frac{1}{2^{18}} \begin{bmatrix} 3375 & -17775 & -10665 & 2025 \\ -17775 & 93615 & 56169 & -10665 \\ -29625 & 156025 & 93615 & -17775 \\ 5625 & -29625 & -17775 & 3375 \end{bmatrix}$
$T_{11} = \frac{1}{2^{18}} \begin{bmatrix} 2205 & -10665 & -17775 & 3375 \\ -10665 & 56169 & 93615 & -17775 \\ -17775 & 93615 & 156025 & -29625 \\ 3375 & -17775 & -29625 & 5625 \end{bmatrix}$	$T_{12} = \frac{1}{2^{18}} \begin{bmatrix} 315 & -3195 & -22365 & 2205 \\ -1659 & 16827 & 117789 & -11613 \\ -2765 & 28045 & 196315 & -19355 \\ 525 & -5325 & -37275 & 3675 \end{bmatrix}$
$T_{13} = \frac{1}{2^{18}} \begin{bmatrix} 343 & -3479 & -497 & 49 \\ -3479 & 35287 & 5041 & -497 \\ -24353 & 247009 & 35287 & -3479 \\ 2401 & -24353 & -3479 & 343 \end{bmatrix}$	$T_{14} = \frac{1}{2^{18}} \begin{bmatrix} 525 & -2765 & -1659 & 315 \\ -5325 & 28045 & 16827 & -3195 \\ -37275 & 196315 & 117789 & -22365 \\ 3675 & -19355 & -11613 & 2205 \end{bmatrix}$
$T_{15} = \frac{1}{2^{18}} \begin{bmatrix} 315 & -1659 & -2765 & 525 \\ -3195 & 16827 & 28045 & -5325 \\ -22365 & 117789 & 196315 & -37275 \\ 2205 & -11613 & -19355 & 3675 \end{bmatrix}$	$T_{16} = \frac{1}{2^{18}} \begin{bmatrix} 49 & -497 & -3479 & 343 \\ -497 & 5041 & 35287 & -3479 \\ -3479 & 35287 & 247009 & -24353 \\ 343 & -3479 & -24353 & 2401 \end{bmatrix}$

$S(\omega)$ can be expressed as

$$S(\omega) = \begin{cases} 1 - 2|\omega|^2 + |\omega|^3, & 0 \leq |\omega| < 1 \\ 4 - 8|\omega| + 5|\omega|^2 - |\omega|^3, & 1 \leq |\omega| < 2 \\ 0, & |\omega| \geq 2. \end{cases} \quad (1)$$

The pixel $(i + u, j + v)$ can be calculated by

$$f(i + u, j + v) = [A] \times [B] \times [C], \quad (2)$$

where

$$[A] = [S(1 + v) S(v) S(1 - v) S(2 - v)], \quad (3)$$

$$[B] = \begin{bmatrix} f_{i-1,j-1} & f_{i-1,j} & f_{i-1,j+1} & f_{i-1,j+2} \\ f_{i,j-1} & f_{i,j} & f_{i,j+1} & f_{i,j+2} \\ f_{i+1,j-1} & f_{i+1,j} & f_{i+1,j+1} & f_{i+1,j+2} \\ f_{i+2,j-1} & f_{i+2,j} & f_{i+2,j+1} & f_{i+2,j+2} \end{bmatrix}, \quad (4)$$

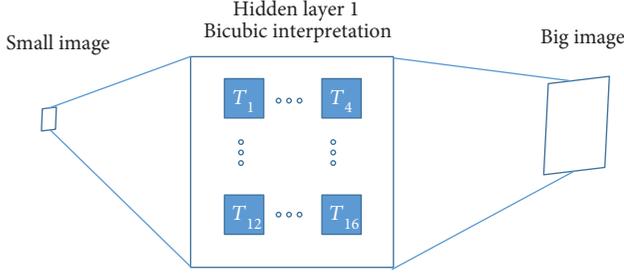


FIGURE 5: Hidden layer 1: fast bicubic interpretation.

$$[C] = \begin{bmatrix} S(1+u) \\ S(u) \\ S(1-u) \\ S(2-u) \end{bmatrix}. \quad (5)$$

Then, the bicubic interpretation templates can be deduced as follows:

$$f(i+u, j+v) = [A] \times [B] \times [C] = [S(1+v) S(v) S(1-v) S(2-v)] \cdot \begin{bmatrix} f_{i-1,j-1} & f_{i-1,j} & f_{i-1,j+1} & f_{i-1,j+2} \\ f_{i,j-1} & f_{i,j} & f_{i,j+1} & f_{i,j+2} \\ f_{i+1,j-1} & f_{i+1,j} & f_{i+1,j+1} & f_{i+1,j+2} \\ f_{i+2,j-1} & f_{i+2,j} & f_{i+2,j+1} & f_{i+2,j+2} \end{bmatrix} \begin{bmatrix} S(1+u) \\ S(u) \\ S(1-u) \\ S(2-u) \end{bmatrix} \quad (6)$$

$$= f_{i-1,j-1}S(1+u)S(1+v) + f_{i-1,j}S(u)S(1+v) + f_{i-1,j+1}S(1-u)S(1+v) + f_{i-1,j+2}S(2-u)S(1+v) \\ + f_{i,j-1}S(1+u)S(v) + f_{i,j}S(u)S(v) + f_{i,j+1}S(1-u)S(v) + f_{i,j+2}S(2-u)S(v) + f_{i+1,j-1}S(1+u)S(1-v) \\ + f_{i+1,j}S(u)S(1-v) + f_{i+1,j+1}S(1-u)S(1-v) + f_{i+1,j+2}S(2-u)S(1-v) + f_{i+2,j-1}S(1+u)S(2-v) \\ + f_{i+2,j}S(u)S(2-v) + f_{i+2,j+1}S(1-u)S(2-v) + f_{i+2,j+2}S(2-u)S(2-v) = [B]^* T,$$

where $*$ is the convolutional operation and

$$T = \begin{bmatrix} S(1+u)S(1+v) & S(u)S(1+v) & S(1-u)S(1+v) & S(2-u)S(1+v) \\ S(1+u)S(v) & S(u)S(v) & S(1-u)S(v) & S(2-u)S(v) \\ S(1+u)S(1-v) & S(u)S(1-v) & S(1-u)S(1-v) & S(2-u)S(1-v) \\ S(1+u)S(2-v) & S(u)S(2-v) & S(1-u)S(2-v) & S(2-u)S(2-v) \end{bmatrix}. \quad (7)$$

The u and v ($u \in [0, 1], v \in [0, 1]$) can be discretized as follows.

If $u \in [0, 1/4], v \in [0, 1/4]$, u and v can be set as its mean values; therefore, $u = 1/8$ and $v = 1/8$. From (1), (2), (3), (4), (5), (6), and (7), it can be deduced that

$$T_1 = \frac{1}{2^{18}} \begin{bmatrix} 2401 & -24353 & -3479 & 343 \\ -24353 & 247009 & 35287 & -3479 \\ -3479 & 35287 & 5041 & -497 \\ 343 & -3479 & -497 & 49 \end{bmatrix}. \quad (8)$$

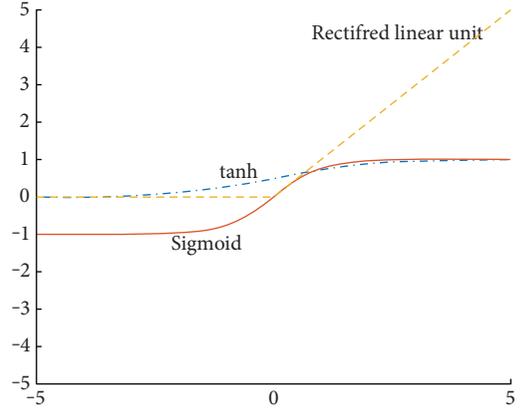


FIGURE 6: Plot of ReLU and classic activation function.

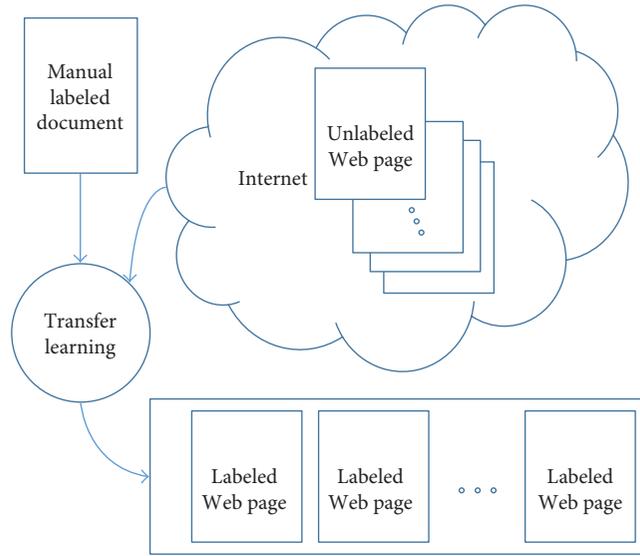


FIGURE 7: Instance-based transfer learning example: Web document classification.

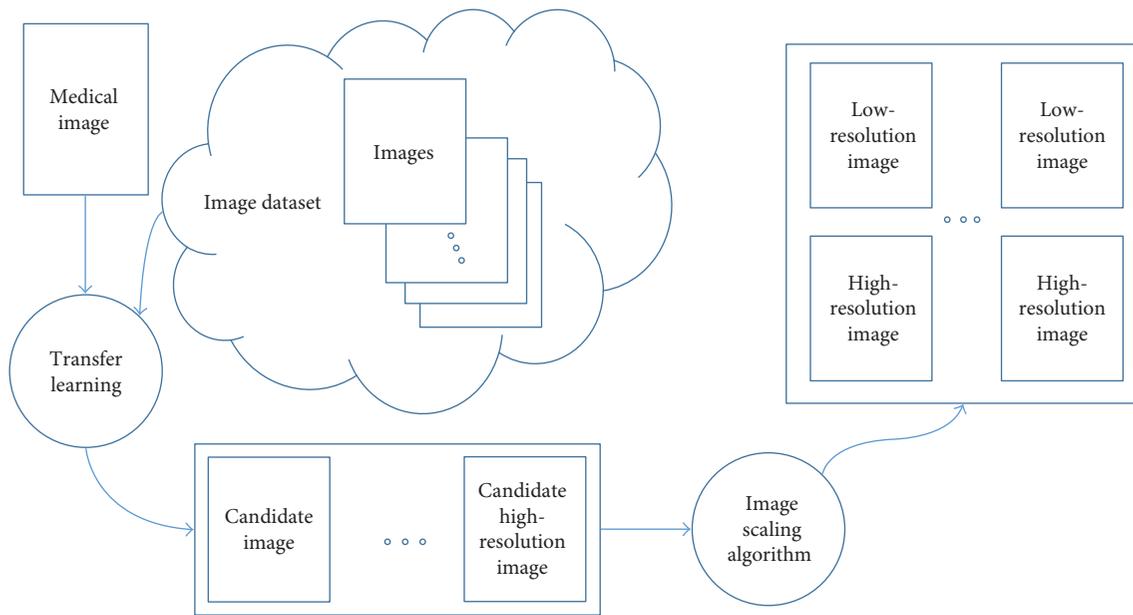


FIGURE 8: Transfer learning: SRR for medical images.

Therefore, $f(i + u, j + v) = [B][T_1]$. Similarly, it can deduce 16 templates in total. Table 2 shows all 16 conditions and corresponding discretized parameters.

Table 3 are the solutions of those 16 templates in Table 2. The hidden layer 1 can be built by those 16 templates, and those precalculated templates are weights of neurons. In the DCNN training step, this layer is constant and the weights of hidden layers 2-3 need to be updated in the following steps.

Therefore, 16 templates in hidden layer 1 can fulfill the bicubic interpretation method as Figure 5 shows. Further,

between the input and output images, there is only a multiple layer neural networks rather than a combination of bicubic interpretation and neural networks [24].

3.1.2. Step 2: Patch Extraction and Representation. This step extracts patch from the results of fast bicubic interpolation layer and maps patch into high-dimensional space as vectors. The dimensionality of these vectors equals a set of feature mapping.

Hidden layer 2 contains a group of rectified linear unit (ReLU, $\max(0, x)$) [25]. Classic activation function and the

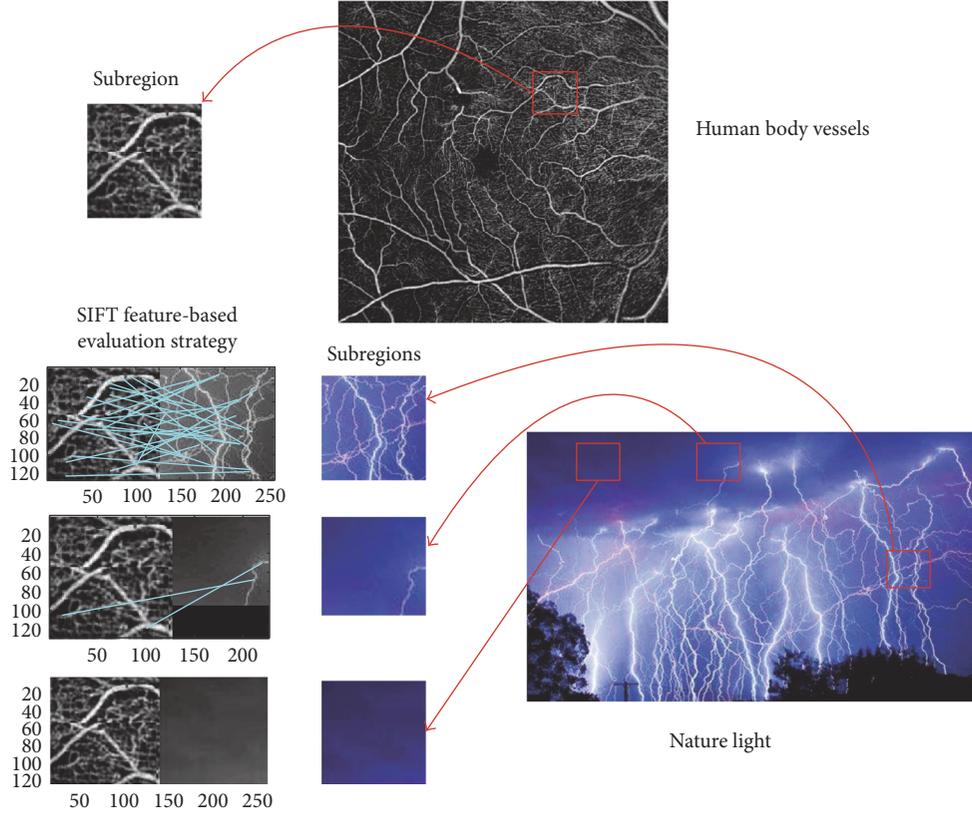


FIGURE 9: SIFT feature-based transfer learning example (vessels and light).

ReLU are plotted in Figure 6. To simplify, this layer can be expressed like an equation of a single ReLU unit. Therefore, F_2 can be defined as follows:

$$F_2(Y) = \max(0, W_2 Y + B_2), \quad (9)$$

where W_2 is the weight of neurons of hidden layer 2, B_2 is biases, W_2 is of a size $f_2 * f_2 * n_2$, f_2 is the spatial size of a filter, n_2 is the number of filters, and B_2 is an n_2 -dimensional vector. Equation (9) implies that a group of $f_2 * f_2 * n_2$ ReLU neurons can be denoted as a single ReLU for simplicity.

3.1.3. Step 3: Nonlinear Mapping. This step fulfills nonlinear mapping. The vectors are mapped in another high-dimensional space for patch extraction and representation layer. This mapping is for representing another set of features.

Like layer 2, the operation of the hidden layer 3 is

$$F_3(Y) = \max(0, W_3 Y + B_3). \quad (10)$$

Here, W_3 is the weight and it is of a size $n_1 * 1 * n_3$ and B_3 is n_3 -dimensional.

3.1.4. Step 4: Reconstruction. This step generates the result of HR image by aggregating the above patch representations.

$$F_4(Y) = W_4 F_3(Y) + B_4. \quad (11)$$

Here, W_4 is of a size $n_2 * f_3 * f_3$ and B_4 is an n_4 -dimensional vector. This step does not involve ReLU activation function.

3.1.5. Learning Procedure of the Proposed DCNN. In the learning procedure, the mapping function F requires the estimation of a set of parameters $\theta = \{W_i, B_i\}$, $i = 2, 3, 4$. W_i and B_i are weights and biases of neurons that should be obtained as above steps mentioned, and W_1 and B_1 of hidden layer 1 are precalculated in Section 3.1.2. For minimizing the loss between ground-truth HR image and the reconstructed image, it can use mean squared error (MSE) to indicate the loss function:

$$L(\theta) = \frac{1}{n} \sum_{i=1}^n \|F(Y_i; \theta) - X_i\|^2, \quad (12)$$

where n is the number of training samples. The loss can be minimized by using stochastic gradient descent with back propagation algorithm.

3.2. SIFT Image Feature-Based Evaluation Strategy for Transfer Learning. Many machine learning approaches hypothesize that the training and test dataset are drawn from the same feature space and they are in the same distribution. Once the distribution changes, lots of modules have to be rebuilt. In real-world applications, to collect sufficient training data is expensive or impossible. Therefore, it would be nice if reducing the effort of collecting the training data and the transfer learning would be desirable.

One typical example is Web document classification; this is an instance-based transfer learning example. Once a document in the area of Web document classification is offered

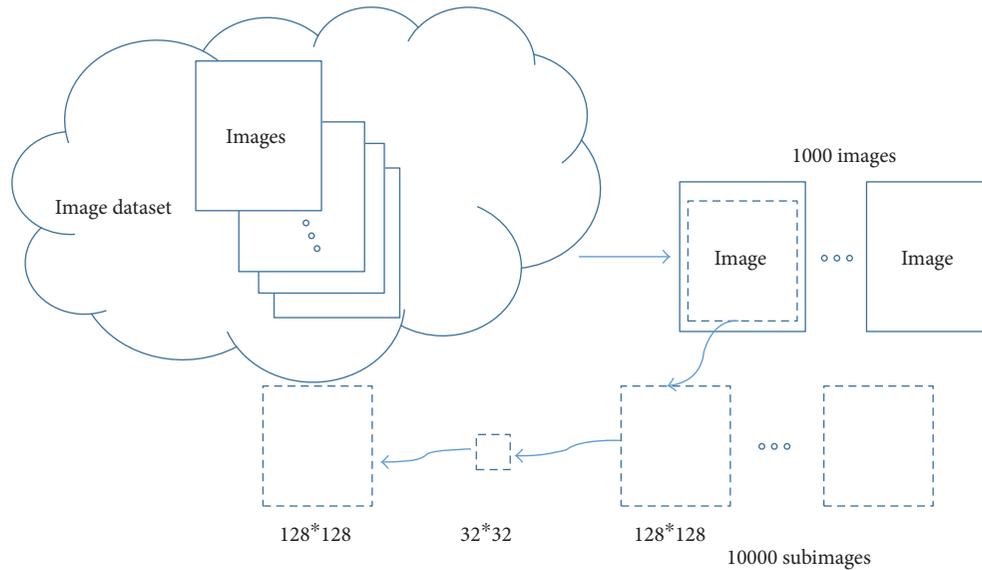


FIGURE 10: Preparation of training images.

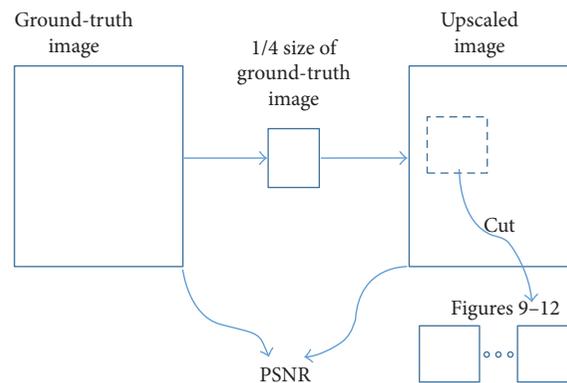


FIGURE 11: How to get PSNR and comparative figures.

with manual labeling, it would be helpful if the classification knowledge could be transferred into new Web pages with that manual labeled document. As Figure 7 shows, transfer learning can help classifier to obtain bigger training dataset and which is similar to the manual labeled samples.

Inspired by the transfer learning as mentioned above, SRR for medical images can also employ transfer learning methodology. As Figure 7 shows, a big size of images can be obtained by using transfer learning, and then the DCNN can learn from sufficient medical images. However, input image and output image are in pairs and only one HR image can be easily obtained from public image dataset. With reversed thinking of SRR, we adapt image scaling-down algorithm to produce corresponding LR image; therefore, pair images are ready for DCNN to learn. As the assumption mentioned above “the training and test dataset are drawn from the same feature space and the same distribution,” we propose a SIFT image feature-based transfer learning as Figure 8 presents.

For obtaining sharper outlines, we use scale-invariant feature transform (SIFT) feature as the base which provides the distinctiveness, the robustness, and the generality. SIFT feature descriptor can capture structural properties robustly, and its points dominantly distribute among regions even color and texture change.

To a given image $I(x,y)$, the Gaussian scale-space $L(x,y,D)$ can be defined as conventional of scale variant Gaussian function $G(x,y,D)$ and $I(x,y)$, and they can be formulated as follows:

$$L(x,y,\sigma) = G(x,y,\sigma) \times I(x,y), \quad (13)$$

$$G(x,y,\sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2}. \quad (14)$$

In scale space, each pixel is compared to its surrounding 8 adjacent points and 18 neighboring points which are corresponding positions of two images adjacent scale of

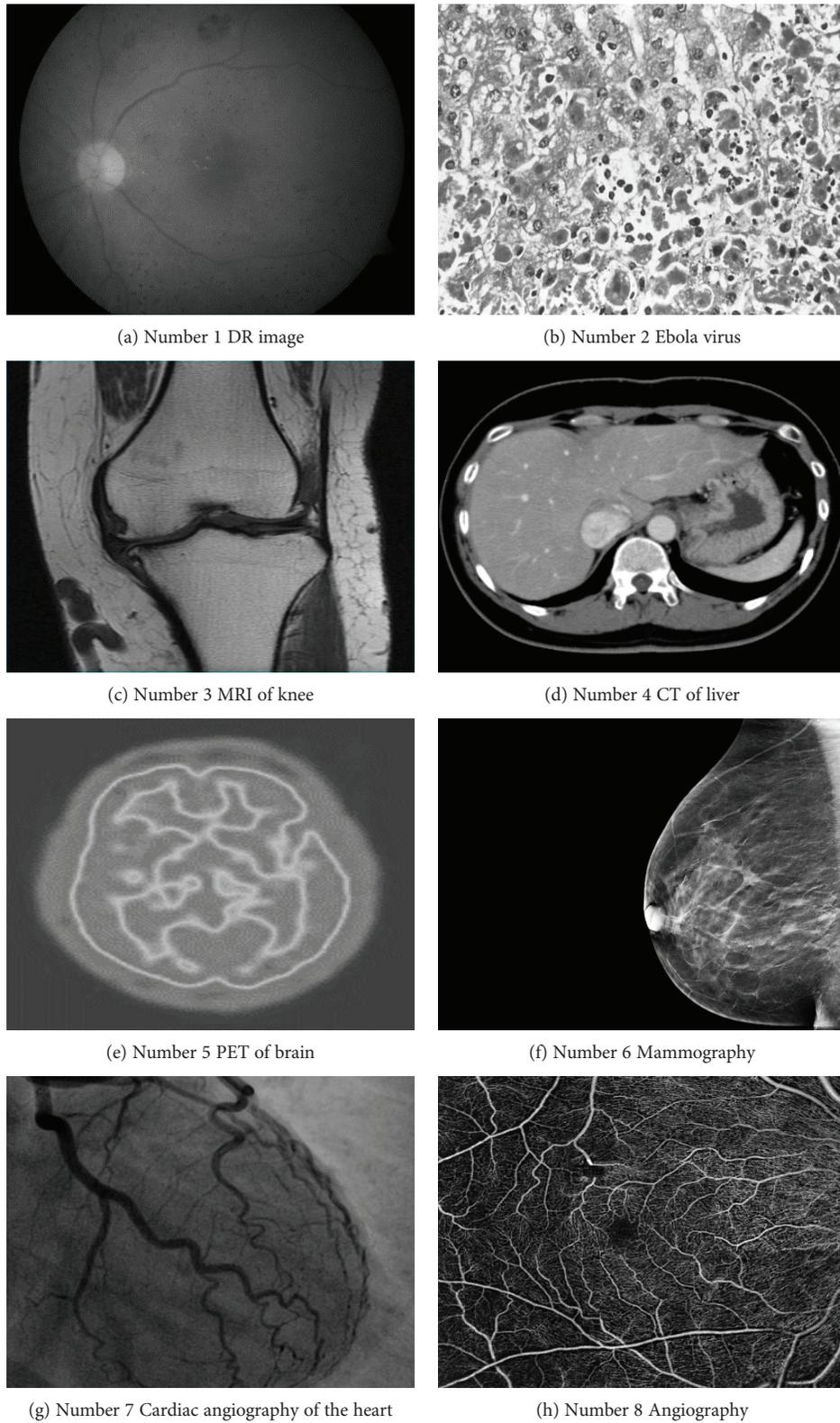


FIGURE 12: Ground-truth medical image.

up and down in pyramid. If the pixel value is different to any of the 26 points, then this pixel is the candidate feature point.

It can be sampled in the neighborhood window and centered at the candidate feature points. Then, histogram is used to count the gradient direction of neighborhood pixels. The

TABLE 4: Public medical images for comparison.

Number	File name	Description	Provider	Download weblink
(1)	image008.png	Size: 1500 * 1152 This image is from a public database of diabetic retinopathy detection	Lappeenranta University of Technology (LUT, in Finland)	http://www.it.lut.fi/project/imageret/diaretdb1/resources/images/ddb1_fundusimages/image008.png
(2)	7058_lores.jpg	Size: 700 * 466 This micrograph of human liver tissue infected with the Ebola virus.	Centers for Disease Control and Prevention (CDC, in U.S.)	https://phil.cdc.gov/PHIL_Images/20050318/21149f4ebdee43fcae990f40021450a/7058_lores.jpg
(3)	MRIofKnee.jpg	Size: 693 * 779 MRI image of the knee	National Institute of Health (NIH, in U.S.)	https://www.nibib.nih.gov/sites/default/files/MRI%20of%20Knee.jpg
(4)	CaseKS11-CT-liverSOL-3.JPG	Size: 1114 * 905 CT image of the liver	Department of Nuclear Medicine Biotracer Medicine, Kanazawa university (ac.jp, in Japan)	http://nucmed.w3.kanazawa-u.ac.jp/NMC/CaseKS11/CaseKS11-CT-liverSOL-3.JPG
(5)	300px-PET-image.jpg	Size: 300 * 339 PET (positron emission tomography) image of the brain	http://wikimedia.org	https://upload.wikimedia.org/wikipedia/commons/thumb/c/c6/PET-image.jpg/300px-PET-image.jpg
(6)	ht_141204_senoclaire_3d_mammography_800x600.jpg	Size: 800 * 600 Mammography	http://medscape.com	http://img.medscape.com/news/2014/ht_141204_senoclaire_3d_mammography_800x600.jpg
(7)	heart-385x330.jpg	Size: 385 * 330 Cardiac angiography of the heart	http://ideviate.org	http://ideviate.org/wp-content/uploads/2012/05/heart-385x330.jpg
(8)	panoramic-cnvs-octa.jpg	Size: 1360 * 1346 Angiography	http://www.oct-optovue.com	http://www.oct-optovue.com/angioOCT-ARMD/oct-angiography-armd7_files/panoramic-cnvs-octa.jpg

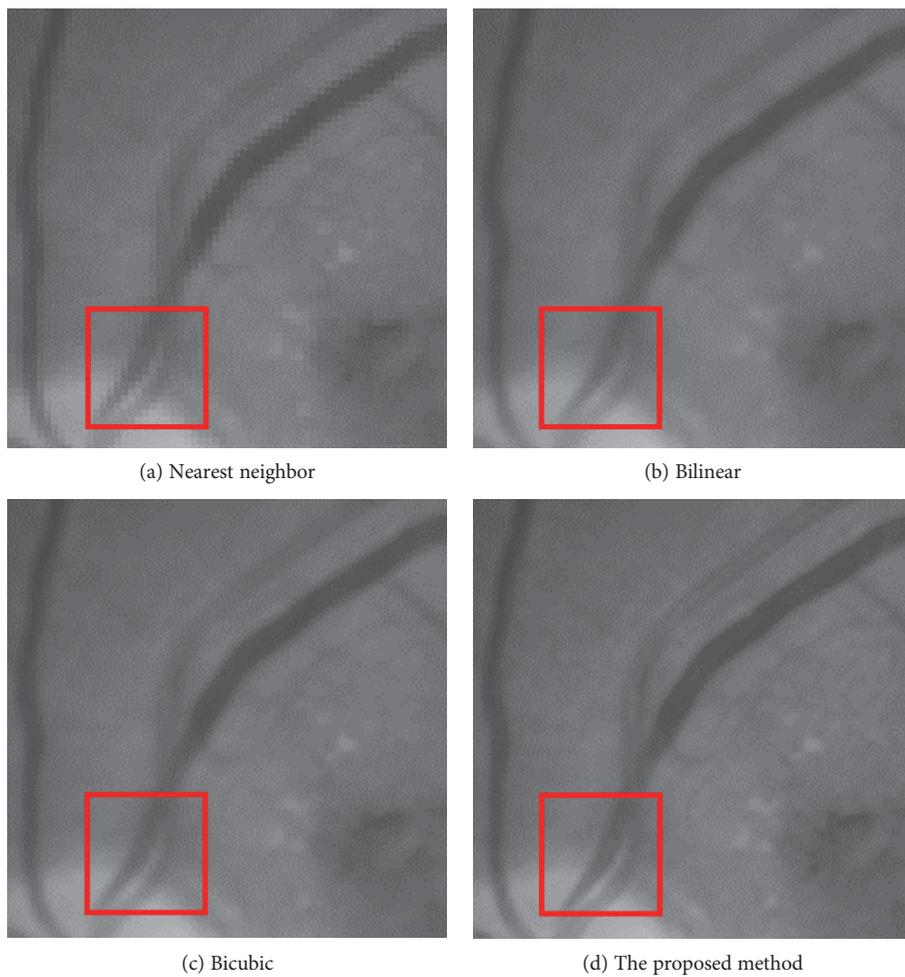


FIGURE 13: Diabetic retinopathy image.

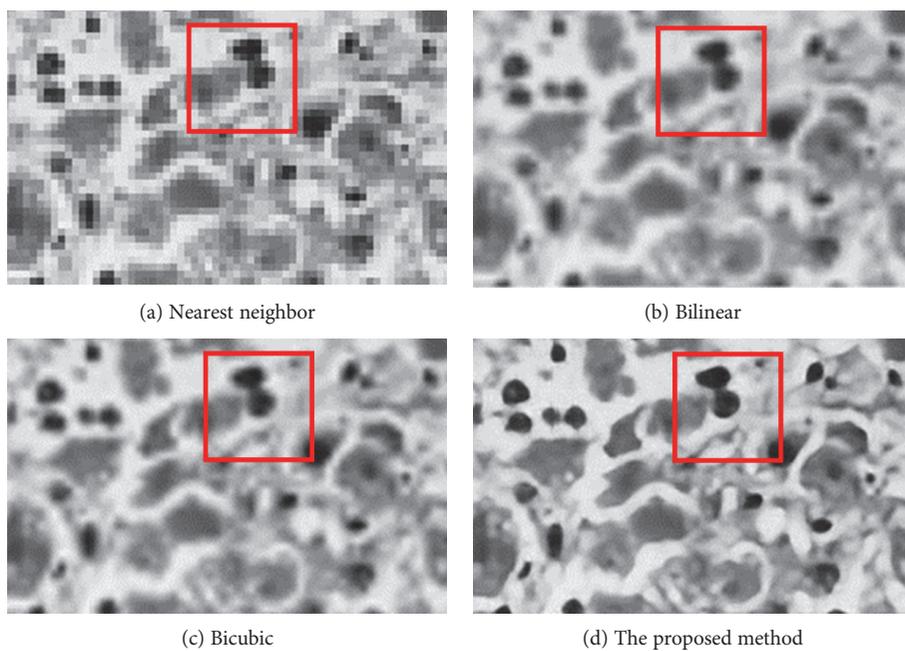


FIGURE 14: Ebola virus.

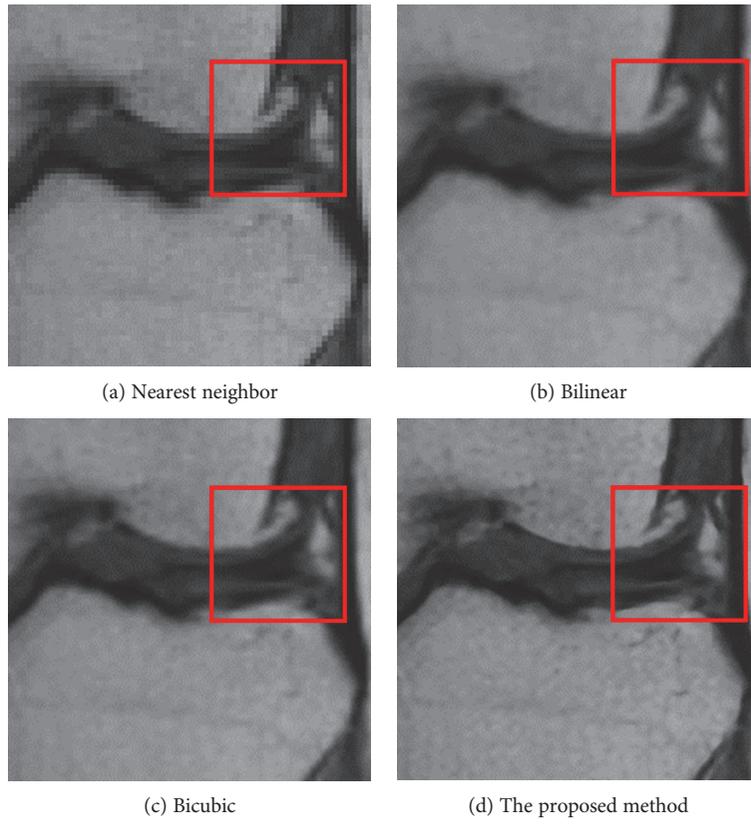


FIGURE 15: MRI knee.

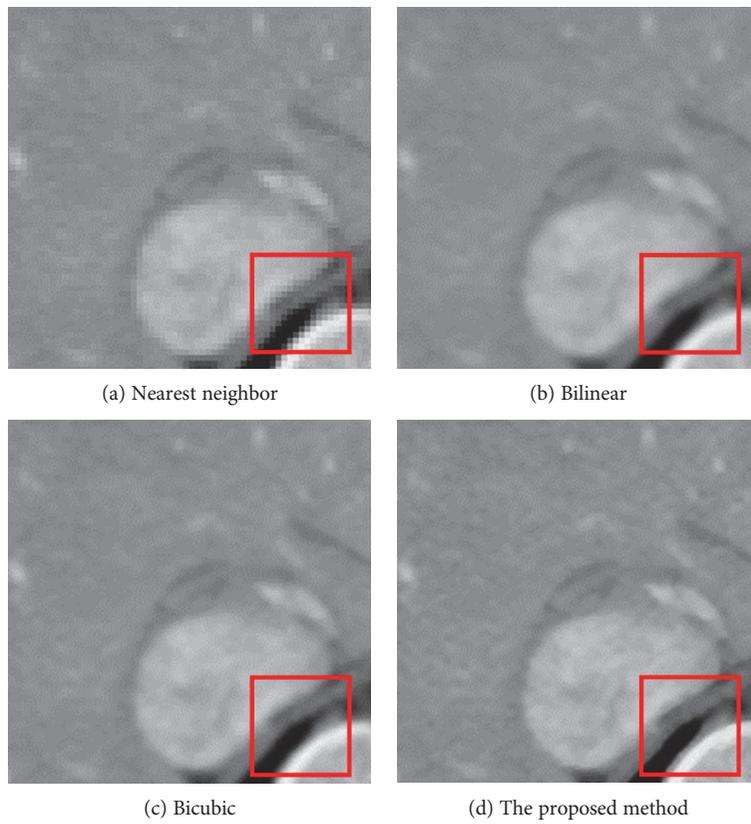


FIGURE 16: CT liver.

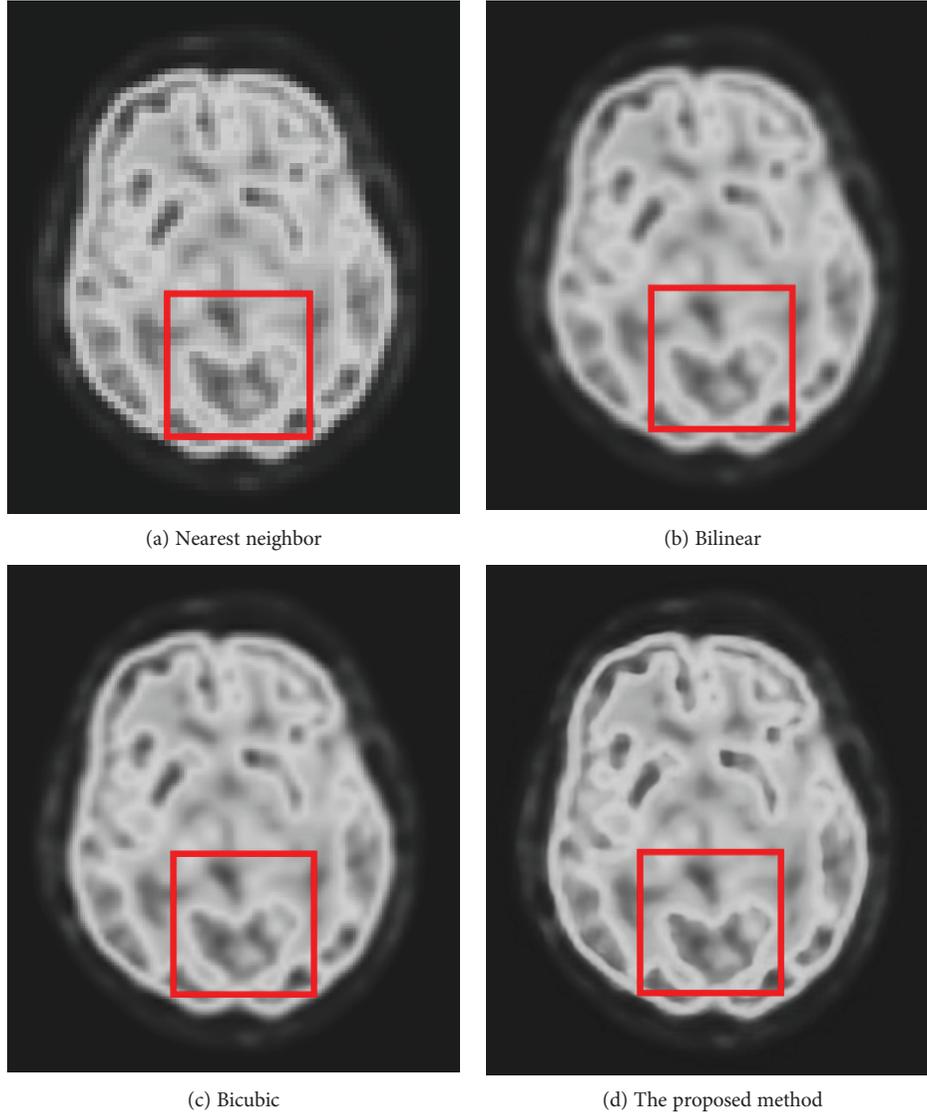


FIGURE 17: PET brain.

range of gradient histogram is from 0 to 360 degrees, and each column represents a direction in histogram.

Rotate the axis to the direction of the feature point which can ensure rotation invariance. Then, select an $8 * 8$ window centered at the feature point to calculate the gradient histogram of 8 directions in each small $4 * 4$ square. The accumulated values of each gradient direction are stored; therefore, a SIFT feature vector of $4 * 4 * 8$ equal to 128-dimensional vector is constituted.

We traverse all SIFT features in the given training image set. For each subregion of candidate images, we calculate the Euclidean distances to all the SIFT features, then sum up all of the distance, and define a mean of the sum as the distance among training images. It can be calculated as

$$D(S, T) = \frac{1}{m} \sum_{i=1}^m \left(\min_{1 \leq j \leq n} \sqrt{\sum_{k=1}^t (s_{ik} - t_{jk})^2} \right), \quad (15)$$

where m is the number of SIFT features of image S , $t = 128$, n is the number of SIFT features of the training images, s_{ik} is the k_{th} element of the i_{th} feature vector of S , and t_{jk} is the k_{th} element of j_{th} feature vector of T .

$D(S, T)$ denotes an average distance regarding all SIFT features among training images, and it is suitable for the case of detecting feature points from various images when applying instance-based transfer learning. In transfer learning set, images can be evaluated as follows:

$$\eta = \frac{N_m}{N_t}, \quad (16)$$

where N_m and N_t indicate the number of matched SIFT feature pairs. Finally, the evaluation criterion is $\eta \geq D(S, T)$.

Figure 9 is an example for illustrating the proposed SIFT feature-based transfer learning. There are a medical image of angiography and a nature image of light. We cut three distinctive subregions of the image "light" and

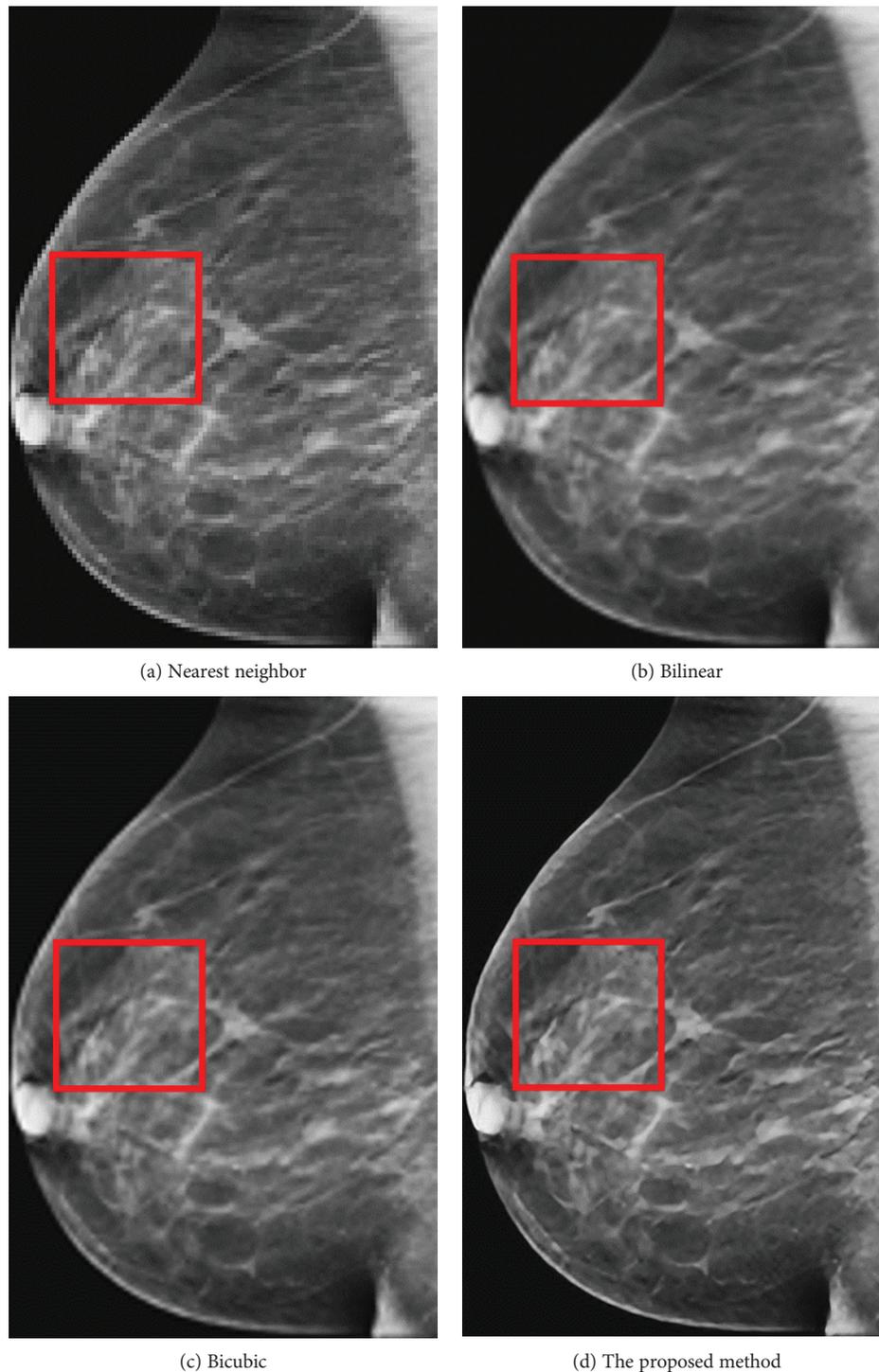


FIGURE 18: Mammography.

cut a subregion of vessels. We use SIFT feature to find matched points, and it can be found that the lights and vessels have lots of connections and those connections are key points in similar. If there is only one light, the connections reduce to 2 lines. However, if we compare the vessels and cloud part, there is no connected feature points. The proposed SIFT feature-based transfer learning intends to find a best match subregion and adds it into

training set. Therefore, SIFT features can help to enlarge the training set for the DCNN.

4. Experiments and Results

4.1. Computation Environment and Training Set. Involved tools contain CUDA, Python language, and Open CV. The method selects candidate images from Image-Net dataset

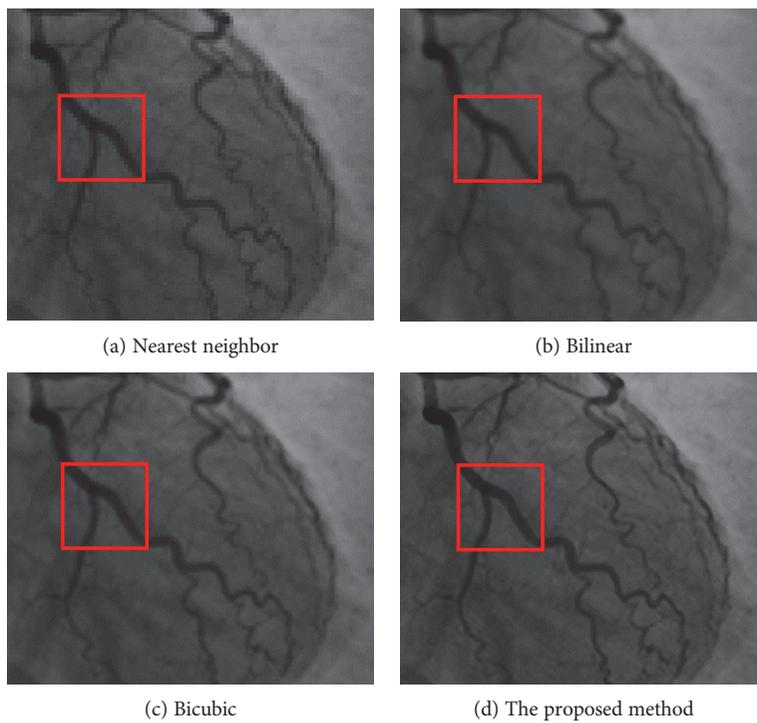


FIGURE 19: Cardiac angiography of the heart.

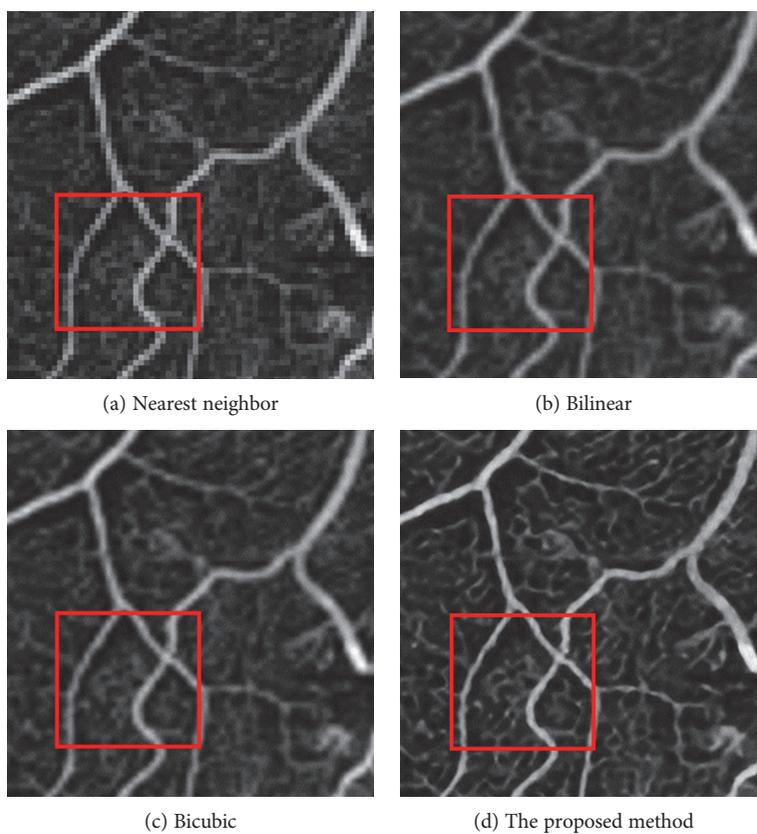


FIGURE 20: Angiography.

TABLE 5: The results of PSNR (dB).

Number and image name	Bicubic	The proposed method	Bilinear	NN	SRCNN
Number 1 DR	46.35	46.74	45.77	44.35	46.79
Number 2 Ebola	25.61	27.29	24.31	22.79	27.25
Number 3 MRI knee	36.67	37.74	35.33	33.04	37.77
Number 4 CT liver	39.27	39.81	36.96	32.49	27.74
Number 5 PET brain	37.71	37.21	34.61	30.11	37.26
Number 6 Mammography	30.64	31.65	30.13	29.49	31.60
Number 7 Cardiac angiography	36.28	37.23	35.56	34.47	37.19
Number 8 Angiography	25.96	27.27	24.98	24.13	27.21

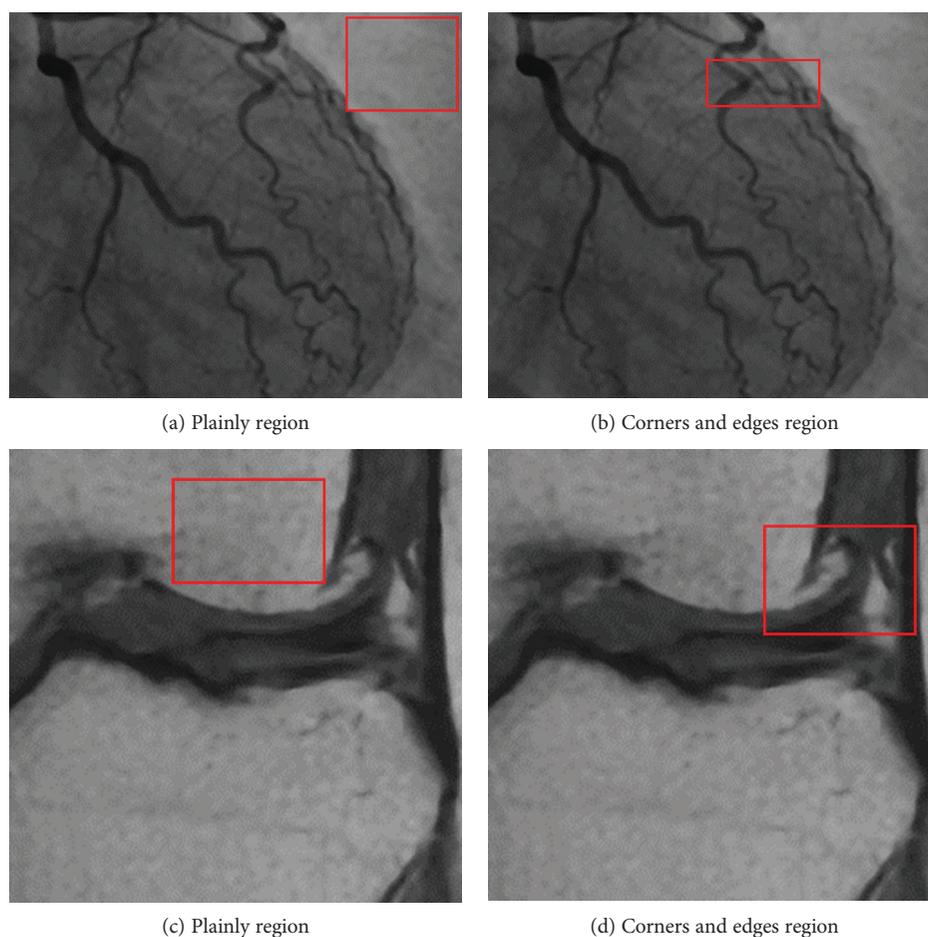


FIGURE 21: Comparison of different regions (PSNR).

[26]. For comparison, the SRCNN [24] has been trained but without SIFT feature-based transfer learning. The training set for SRCNN is from the public medical image database [27].

As Figure 10 shows, a group of 1000 images are collected for the next step. Each training image is cut by the size of 128×128 randomly and repeatedly. Finally, a set of 10000 images is done. The ground truth image which size is 128×128 is scaled to 32×32 as low-resolution image. Therefore, the 32×32 images are used as low-resolution input image, and the ground truth images are used for training as the output of the deep convolutional neural network.

4.2. Quantitative Evaluation. We use peak signal-to-noise ratio (PSNR) for quantitatively image restoration quality. PSNR can be calculated by

$$\text{PSNR} = 10 \times \log_{10} \left(\frac{(2^n - 1)^2}{\text{MSE}} \right), \quad (17)$$

where MSE is the mean squared error

For a fair comparison with conventional methods and SRCNN [24], we use medical images which were downloaded

TABLE 6: Comparison of edge and noncorner region in PSNR index (dB).

Image	Region	The proposed method	SRCNN
Number 4 MRI knee	Figure 21(d) Corners and edges region	34.57	34.12
	Figure 21(c) Plainly region	37.51	37.65
Number 7 Cardiac angiography	Figure 21(b) Corners and edges region	36.22	36.13
	Figure 21(a) Plainly region	42.79	42.85

from research organization/authorities. Those four images can be downloaded freely by anyone. Brief information and the download weblink is listed in Table 2. We first scale those four images down to 1/4 size; therefore, PSNR can be calculated with results of SRR methods, as Figure 11 illustrates.

Figure 12 is the ground-truth images as Table 4 listed, and the resolution is the same as Table 2 presented. Figures 13, 14, 15, 16, 17, 18, 19, and 20 show the super resolution results of four different methods by upscaling factor 4. For intuitive comparison, each upscaled image is cut by a rectangle zone. As those images shown, the proposed method produces much shaper edges than other methods.

4.3. Experiment Results and Discussion. As an overview of Table 5, the proposed method gives the highest PSNR in 5 experiments. It is also can be seen from Figures 13–20 intuitively. Nearest neighbor interpolation gives the lowest PSNR for four images. Bilinear interpolation is slightly better than nearest neighbor interpolation. The bicubic interpolation is the best conventional method for four images. Completed with SRCNN, the proposed method achieves 0.04–0.07 higher in PSNR index. But, numbers 1, 3, and 4 experiments show that SRCNN still gives better PSNR than the proposed DCNN and the PSNR is between 0.03 and 0.05.

In Figure 13, experimental results are presented. The image is about diabetic retinopathy research. Obviously, image aliasing exists in Figure 13(a) which is the result of nearest neighbor method. In Figure 13(b), the bilinear method gives more smooth edge of microvascular vessel (microblood vessel), but comparing Figure 13(b) with Figure 9(c), the bicubic method gives a higher contrast method. In Figure 13(d), the proposed method gives a sharper image and it should be noticed that the microvascular vessel and background had higher contrast than the other three methods.

In Figure 14, experimental results are about Ebola virus in microscope. Nearest neighbor method still gives an image with saw-tooth; this result cannot satisfy the needs of further analysis in medical research. Bicubic method shows better result than bilinear method, but the proposed method produces the best image among comparative methods.

SRR result of MRI image (knee) is shown in Figure 15. The proposed method shows a clearer meniscus and bone edges than the other methods. Bicubic method gives better results than nearest neighbor and bilinear method.

Figure 16 is the result of SRR for CT image (liver). Like Figures 13 and 14, the proposed method gives the best result. However, visual contrast among Figures 16(a), 16(b), 16(c), and 16(d) is not so obvious as Figures 14 and 15 show, the rectangle area of Figure 16(d) still presents a clear edge.

TABLE 7: Running costs of bicubic interpretation and hidden layer 1 (integer and floating-point arithmetic).

Computational operation	Hidden layer 1	Bicubic interpretation
Integer addition	15 times	0 times
Integer division	1 time	0 times
Integer multiplications	16 times	0 times
Floating-point additions	4.6 times	41 times
Floating-point multiplications	0 times	28 times

Figures 17–20 are results of PET brain, mammography, cardiac angiography, and angiography. Those medical images are processed by four methods, and subregions are cut for comparison. It can be found that the proposed DCNN can produce enlarged images with sharper edges in visual contract.

Figure 21 shows comparison between the proposed method and SRCNN, and Table 6 presents the PSNR index of Figure 21. We choose two types of subregions which contain sharp edges or plainly texture, and we call them as “corners and edges region” and “plainly region.” By comparing the PSNR of those two types of regions, interesting details can be found. The SRCNN has bigger PSNR than the proposed DCNN in Figures 21(a) and 21(c); those two subregions are compared in plainly regions. It can be seen that the proposed DCNN is slightly higher than SRCNN in Figures 17(b) and 17(d) which are corner regions. As the previous example of Figure 9 illustrated in Section 3.2, the proposed DCNN learns from more subimages from the image database with the SIFT feature-based selection strategy.

$$\begin{aligned}
f(i+u, j+v) = & f_{i-1, j-1} S(1+u) S(1+v) + f_{i-1, j} S(u) S(1+v) \\
& + f_{i-1, j+1} S(1-u) S(1+v) + f_{i-1, j+2} S(2-u) S(1+v) \\
& + f_{i, j-1} S(1+u) S(v) + f_{i, j} S(u) S(v) \\
& + f_{i, j+1} S(1-u) S(v) + f_{i, j+2} S(2-u) S(v) \\
& + f_{i+1, j-1} S(1+u) S(1-v) + f_{i+1, j} S(u) S(1-v) \\
& + f_{i+1, j+1} S(1-u) S(1-v) + f_{i+1, j+2} S(2-u) S(1-v) \\
& + f_{i+2, j-1} S(1+u) S(2-v) + f_{i+2, j} S(u) S(2-v) \\
& + f_{i+2, j+1} S(1-u) S(2-v) + f_{i+2, j+2} S(2-u) S(2-v).
\end{aligned} \tag{18}$$

Another comparison is running time. The proposed convolutional hidden layer 1 saves approximately half time costs than classic bicubic interpretation; therefore, it can save

TABLE 8: Overall comparisons of running time (in milliseconds).

Image	Methods				
	NN	Bicubic	Bilinear	SRCNN	The proposed method
Number 1 DR	8.075	14.790	14.655	63579.328	63565.847
Number 2 Ebola	8.000	10.518	11.613	11244.818	11234.735
Number 3 MRI knee	7.003	12.180	12.029	19231.110	19217.921
Number 4 CT liver	7.362	11.270	12.214	36315.960	36304.727
Number 5 PET brain	8.993	9.067	12.954	1986.092	1977.153
Number 6 Mammography	6.762	12.769	11.296	17342.371	17330.239
Number 7 Cardiac angiography	6.123	8.940	9.965	3336.666	3328.433
Number 8 Angiography	8.929	13.580	13.626	66651.142	66638.937

much time. As (18) shows, we can deduce the bicubic interpretation which has 28 floating-point multiplications and 41 floating-point additions. The proposed bicubic interpretation hidden layer needs 16 integer multiplications, 15 integer additions, 1 integer division, and 4.6 floating-point additions. Computational needs are shown in Table 7. Those values in Table 7 are derived from (1), (7), and (18) and Table 2. Therefore, the proposed DCNN is faster than SRCNN not only in practice but also in theory.

Table 8 shows the time costs of the eight distinctive medical images on a PC (Intel XEON E3-1230V3, 16G RAM). All images are preloaded in RAM, and then we processed those medical images by five methods and recorded the time costs. As Table 8 shows in milliseconds, NN is the fastest method and the SRCNN is slowest method. However, conventional methods such as NN, bicubic, and bilinear interpretation methods cannot meet the needs of medical image enlargement. Comparing SRCNN and the proposed method, the proposed method is slightly faster than SRCNN. It should be noticed that the training procedure processes lots of images; therefore, to save various time slices is meaningful for compute-intensive methods, such as deep learning.

The SRCNN is a novel super resolution reconstruction method, and we tried to improve three parts:

- (1) Prefixed template layer: A prefixed template layer saves costs by using mathematic deduction. On the other hand, training a convolutional layer with given training samples to fulfill bicubic interpretation is feasible. However, to train a convolutional layer requires various training images in pair. Therefore, we suggest to prefix the bicubic interpretation layer by mathematics deduction. Moreover, the proposed fixed templates may help other researchers and engineers to use them in real application easily, and they can deduce and verify those templates by their own.
- (2) Hybrid DCNN structure: Most researches focus on training deep neural networks (NN), and the whole NN is composed of unfixed parameters. The structure of our method combines fixed and unfixed parameters. Maybe, the combination of fixed and unfixed NN structure has undiscovered potentials in other applications.

- (3) Reducing costs and enhanced edge: The prefixed template layer can save more time than SRCNN, and the proposed SIFT feature-based transfer learning method guarantees the proposed DCNN can produce enlarged medical images with sharper edges.

To conclude, nearest neighbor method may be suitable for some occasions, but it definitely cannot satisfy SRR needs of medical images, such as microscope, CT, MRI, mammography, cardiac angiography, and angiography. Compared with the other conventional methods, bicubic gives better results than bilinear result; however, bicubic method still yields to the proposed method. The SRCNN is an effective and efficient DCNN architecture, but it lacks a faster convolutional interpretation layer. The bicubic interpretation hidden layer in the proposed method ensures a faster running speed than SRCNN, and the SIFT feature-based transfer learning provides sharper edges and corners region than SRCNN by selectively choosing training samples. Moreover, the bicubic interpretation hidden layer can provide an enlarged image which has continuous first and second derivative. This novel bicubic interpretation hidden layer has the potential to solve other image enhancement problems.

5. Conclusion

In this paper, a deep learning- and transfer learning-based super resolution reconstruction method has been presented. The proposed method aims to reconstruct a high-resolution image from one single low-resolution image. We propose a fast bicubic interpretation layer and SIFT feature-based transfer learning to speed up DCNN and to obtain sharper outlines; therefore, the proposed method can avoid collecting a great number of various medical images. Empirical experiments show that the proposed method can achieve better performance than other conventional methods. We suggest that this enhancement method is meaningful for clinical diagnosis, medical research, and automatic image analysis.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] E. J. Hall, "Radiobiology for the radiologist," *Radiation Research*, vol. 116, no. 1, 1988.
- [2] J. Mifflin, "Visual archives in perspective: enlarging on historical medical photographs," *The American Archivist*, vol. 70, no. 1, pp. 32–69, 2007.
- [3] P. C. Cosman, R. M. Gray, and R. A. Olshen, "Evaluating quality of compressed medical images: SNR, subjective rating, and diagnostic accuracy," *Proceedings of the IEEE*, vol. 82, no. 6, pp. 919–932, 1994.
- [4] K. Kayser, J. Gortler, T. Goldmann, E. Vollmer, P. Hufnagel, and G. Kayser, "Image standards in tissue-based diagnosis (diagnostic surgical pathology)," *Diagnostic Pathology*, vol. 3, no. 1, pp. 17–17, 2008.
- [5] B. S. Ramakrishna, W. Liu, G. Saiprasad et al., "An automatic computer-aided detection system for meniscal tears on magnetic resonance images," *IEEE Transactions on Medical Imaging*, vol. 28, no. 8, pp. 1308–1316, 2009.
- [6] D. J. Brenner and E. J. Hall, "Computed tomography—an increasing source of radiation exposure," *The New England Journal of Medicine*, vol. 357, no. 22, pp. 2277–2284, 2009.
- [7] W. D. Foltz and D. Jaffray, "Principles of magnetic resonance imaging," *Radiation Research*, vol. 177, no. 4, pp. 331–348, 2012.
- [8] M. K. Ng and N. K. Bose, "Mathematical analysis of super-resolution methodology," *IEEE Signal Processing Magazine*, vol. 20, no. 3, pp. 62–74, 2003.
- [9] T. G. Dietterich, "Machine learning," *ACM Computing Surveys*, 1996.
- [10] Y. Wang, J. Qiao, J. Li, P. Fu, S. C. Chu, and J. F. Roddick, "Sparse representation-based MRI super-resolution reconstruction," *Measurement*, vol. 47, pp. 946–953, 2014.
- [11] F. Rousseau, K. Kim, C. Studholme, M. Koob, and J. L. Dietemann, "On super-resolution for fetal brain MRI," *Medical Image Computing and Computer-Assisted Intervention*, vol. 13, Part 2, pp. 355–362, 2010.
- [12] K. C. N. Liyakathunisa, "A novel super resolution reconstruction of low resolution images progressively using DCT and zonal filter based denoising," *International Journal of Computer Science and Information Technology*, vol. 3, no. 1, pp. 54–72, 2011.
- [13] N. Jiang and L. Wang, "Quantum image scaling using nearest neighbor interpolation," *Quantum Information Processing*, vol. 14, no. 5, pp. 1559–1571, 2015.
- [14] R. G. Keys, "Cubic convolution interpolation for digital image processing," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 29, no. 6, pp. 1153–1160, 1981.
- [15] A. V. Tonisson, A. Dorrell, N. Mehrseresht, and R. Zillman, "Image scaling with aliasing cancellation," *Proceedings of SPIE*, 2011.
- [16] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Neural Information Processing Systems*, pp. 1097–1105, 2012.
- [17] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 10, pp. 1345–1359, 2010.
- [18] W. X. Wei, T. Quan, Y. Wang, H. J. Wang, and P. Z. Li, "Application of three-dimensional interpolation methods in contaminated site evaluation," *Advanced Materials Research*, vol. 878, pp. 782–790, 2014.
- [19] Y. Liu, H. Wu, and S. Yu, "Adaptive DE-based reversible steganographic technique using bilinear interpolation and simplified location map," *Multimedia Tools and Applications*, vol. 52, no. 2-3, pp. 263–276, 2011.
- [20] T. Kim, J. S. Yang, S. Shin, and J. Lee, "Statistical torsion angle potential energy functions for protein structure modeling: a bicubic interpolation approach," *Proteins*, vol. 81, no. 7, pp. 1156–1165, 2013.
- [21] E. Plenge, D. H. Poot, W. J. Niessen, and E. Meijering, "Super-resolution reconstruction using cross-scale self-similarity in multi-slice MRI[C]," *Medical Image Computing and Computer Assisted Intervention*, vol. 16, Part 3, pp. 123–130, 2013.
- [22] Q. Tao, O. Dzyubachyk, D. H. Poot et al., "Super-resolution reconstruction of late gadolinium enhanced MRI from multiple views," *Journal of Cardiovascular Magnetic Resonance*, vol. 16, no. 1, 2014.
- [23] X. D. Zhao, Z. F. Zhou, J. Z. Cao et al., "Multi-frame super-resolution reconstruction algorithm based on diffusion tensor regularization term," *Applied Mechanics and Materials*, vol. 543, pp. 2828–2832, 2014.
- [24] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 2, pp. 295–307, 2016.
- [25] G. E. Dahl, T. N. Sainath, and G. E. Hinton, "Improving deep neural networks for LVCSR using rectified linear units and dropout," in *International Conference on Acoustics, Speech, and Signal Processing*, pp. 8609–8613, Vancouver, BC, May 2013.
- [26] <http://www.image-net.org/>.
- [27] <https://grand-challenge.org/>.