

METHODOLOGY ARTICLE

Open Access



High throughput automatic muscle image segmentation using parallel framework

Lei Cui¹, Jun Feng^{1*} , Zizhao Zhang² and Lin Yang¹

Abstract

Background: Fast and accurate automatic segmentation of skeletal muscle cell image is crucial for the diagnosis of muscle related diseases, which extremely reduces the labor-intensive manual annotation. Recently, several methods have been presented for automatic muscle cell segmentation. However, most methods exhibit high model complexity and time cost, and they are not adaptive to large-scale images such as whole-slide scanned specimens.

Methods: In this paper, we propose a novel distributed computing approach, which adopts both data and model parallel, for fast muscle cell segmentation. With a master-worker parallelism manner, the image data in the master is distributed onto multiple workers based on the Spark cloud computing platform. On each worker node, we first detect cell contours using a structured random forest (SRF) contour detector with fast parallel prediction and generate region candidates using a superpixel technique. Next, we propose a novel hierarchical tree based region selection algorithm for cell segmentation based on the conditional random field (CRF) algorithm. We divide the region selection algorithm into multiple sub-problems, which can be further parallelized using multi-core programming.

Results: We test the performance of the proposed method on a large-scale haematoxylin and eosin (H&E) stained skeletal muscle image dataset. Compared with the standalone implementation, the proposed method achieves more than 10 times speed improvement on very large-scale muscle images containing hundreds to thousands of cells. Meanwhile, our proposed method produces high-quality segmentation results compared with several state-of-the-art methods.

Conclusions: This paper presents a parallel muscle image segmentation method with both data and model parallelism on multiple machines. The parallel strategy exhibits high compatibility to our muscle segmentation framework. The proposed method achieves high-throughput effective cell segmentation on large-scale muscle images.

Keywords: Muscle image segmentation, Cloud computing, Multi-core programming

Background

Skeletal muscle has been extensively recognized as the tissue related to many diseases such as heart failure and chronic obstructive pulmonary disease (COPD) [1, 2]. To accelerate the disease diagnosis at the cellular level and reduce the inter-observer variations, these exist increasing demands for accurate and efficient computer-aided muscle image analysis system [3]. Automatic muscle cell

segmentation is usually the first step for further image feature quantification. In recent years, several state-of-the-art algorithms have been reported for cell segmentation on skeletal muscle and various cancer images [4–10]. For example, unsupervised methods, such as the deformable model [4, 10, 11], Liu et al. [4] propose a deformable model-based segmentation algorithm, which uses color gradient for cell boundary seeking. Later a contour detection and region-based selection algorithm, which is able to deal with low quality skeletal muscle images, is presented in [12]. However, due to the high model complexity, these

*Correspondence: fengjun@nwu.edu.cn

¹Department of Information Science and Technology, Northwest University, Xi'an, China

Full list of author information is available at the end of the article



methods are not applicable to large-scale muscle images (e.g. 4000×4000).

Recently, there is an encouraging evidence that applying medical image analysis [13, 14] to high performance computing resources can significantly improve the running time of the algorithms. Meanwhile, analyzing the whole-slide images can provide much richer information, which is helpful to clinical diagnosis [15]. Therefore, there is an urgent need of efficient large-scale image analysis algorithms. High performance computing techniques emerge as one solution to tackle this challenge, and have attracted a great deal of research interests in medical image analysis [14, 16, 17]. In particular, we have successfully applied a cloud computing framework [13, 18, 19] to content-based sub-image retrieval on whole-slide tissue microarray images, and another application is reported in [14] for high throughput landmark based image registration. Although many high performance computing applications in medical image analysis have been presented in recent literatures, there exists very few reports focusing on cell segmentation.

In this paper, we first present an effective muscle cell segmentation framework, mainly consisting of three steps: 1) muscle cell contour detection using structured random forests (SRF); 2) region candidate generation using superpixel techniques, and 3) hierarchical tree based region selection. A similar framework is first presented in our previous conference version [12], and we extend this approach to a distributed computing framework in this paper. Figure 1 shows the time profile of each step of the framework running on a standalone machine. It indicates that the region selection dominates the running time (accounting for around 94%), this paper mostly focuses on accelerating this step with both data and model parallelism.

We propose a parallel approach using cloud computing techniques which is able to handle very large-scale muscle images. A master-worker parallelism manner is exploited to distribute image data onto multiple worker

nodes of a cloud cluster. On each worker node in the cluster, we propose a hierarchical tree based region selection leveraging on the conditional random field (CRF) algorithm. Its optimization process is divided into multiple sub-problems, which can be solved using multi-core programming techniques. Our proposed method achieves more than 10 times speed improvement on very large-scale muscle images containing hundreds to thousands of cells. Meanwhile, our proposed method produces superior segmentation results compared with several state-of-the-art muscle image segmentation methods on our H&E skeletal muscle image dataset.

The rest of the paper is constructed as follows: we start by introducing our muscle image segmentation method and analyze its characteristics for parallelism; then we present the parallel approach to accelerate the overall segmentation efficiency; next, the “[Experimental results](#)” section evaluates the speed and accuracy of our proposed muscle image segmentation method; the last section concludes this paper.

Contour detection and region candidate generation

We present the proposed cell segmentation method in this section. Effective contour detection is the first step of most region-based image segmentation methods [20–22]. We start by introducing a structured random forest (SRF) based method for fast and accurate muscle contour detection, SRF is selected because its: 1) fast prediction ability for high-dimensional data, 2) robustness to label noise [23], and 3) good support to arbitrary size of outputs. Next, a superpixel algorithm is used to generate region candidates. Finally, we present a hierarchical tree based method to select the optimal candidate regions based on CRF, Fig. 2 shows the entire process.

Contour detection

Random forest (RF) classifier is an ensemble learning technique which combines t decision trees to form a forest

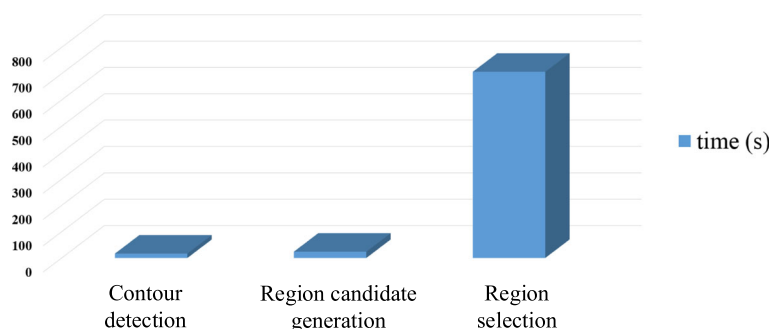
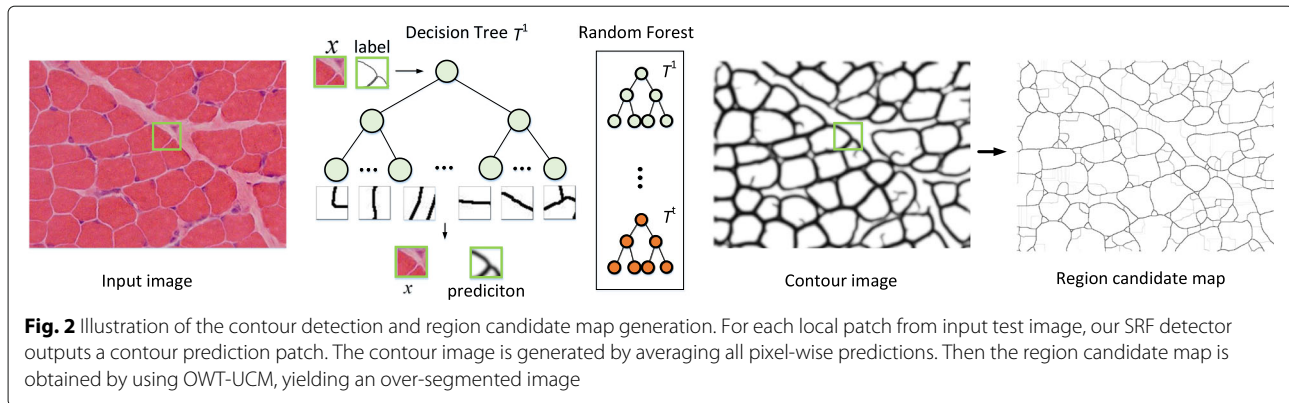


Fig. 1 The time profile for each step of the proposed entire segmentation algorithm running on a standalone machine with a 6000×6000 image. The hierarchical tree based region selection step dominates the running time, around 94% of the total time cost



$\mathcal{F} = \{T_j\}_{j=1}^t$ [24]. Each tree T_j is trained independently and the final classification is determined by applying a majority voting to all the outputs of trees.

However, conventional RF can not capture the inherent contour structures inside local image patches so that it is difficult to obtain satisfactory contour detection performance [25]. In order to capture rich structures of contours during the SRF training, we propose to deploy SRF [26], a variation of RF, to detect the muscle cell contours. SRF is trained with a set of training data $D = \{(x, y) \in \mathcal{X} \times \mathcal{Y}\}$, where $\mathcal{X} = \mathbb{R}^{(d \cdot d) \times c}$ is the feature space of a $d \times d$ image patch, so that each pixel in the image patch is featured by a c -dimensional vector. The structured label $y \in \mathcal{Y} \in \mathbb{Z}^{d \cdot d}$ corresponding to x is a patch cropped from the ground truth image, which is a binary image having the value of 1 in contour pixels and 0 otherwise.

To enable the training of SRF with structured labels, in node i where training data D_i falls, we adopt a mapping function proposed by [26] to map structured labels into a discrete space for each $x \in D_i$, which intrinsically consider the contour structure information. Then a split function $h(x, \theta) = \mathbf{1}[x(k) < \tau]$ splits and propagates the data $D_i \subset X \times Y$ to the left L (when $h = 0$) or right R ($h = 1$) subtree of node i , which is the same as the node splitting procedure of RF. The τ and k are determined by maximizing the standard information gain criterion C_i at node i [24]:

$$C_i = H(D_i) - \sum_{o \in \{L, R\}} \frac{|D_i^o|}{|D_i|} H(D_i^o), \quad (1)$$

where $H(D_i)$ is the Gini impurity measure, $H(D_i) = \sum_l c_l(1 - c_l)$. c_l denotes the proportion of data in D_i with label l . After the data in D_i is propagated to the child nodes, the above steps are performed recursively until leaf nodes are reached (i.e., the stopping criteria is satisfied [24]). The most representative structural label y (close to mean) in each node is stored as its structured prediction [27].

In practice, following [25], we utilize three color channels computed using the CIE-LAB color space. Two gradient magnitude channels are computed with varying amounts of blur (we use Gaussian blurs with $\sigma = 0$ and $\sigma = 1.5$). Additionally, eight orientation channels in two image scales to represent the features of image patches. Such that in total $c = 13$ channels in \mathcal{X} are extracted by using optimized code from [28] available online¹. To prevent overfitting when training SRF, each tree randomly selects a subset of training samples and features for training. In the testing stage (see Fig. 2), since the prediction of each tree for each pixel is independent, we can parallelize this stage using a multi-thread technique [26].

Region candidate generation

Based on the contour image detected by our SRF contour detector, region candidates can be generated using super-pixel techniques, which is able to group similar pixels in terms of color, edge strength (referring to our detected contour image), and spatial cues.

In this paper we use the well-known oriented watershed transform and ultra-metric contour map (OWT-UCM) [29] algorithm to obtain our region candidate maps for three main reasons: 1) it is very efficient to handle large-scale images; 2) regions in a map are well nested at different thresholds; 3) it guarantees that the boundaries of each region are closed and single-pixel wide. These characteristics can facilitate the parallelism of the subsequent proposed hierarchical tree based region selection algorithm. OWT-UCM takes a contour image as input and outputs an over-segmented region candidate map [30], which is illustrated in Fig. 2. The next step is to select those regions using our proposed hierarchical tree-based region selection algorithm.

Hierarchical tree-based region selection

Given the over-segmented region candidate map, our region selection algorithm aims to select region candidates as final segmentation by merging or discarding the segments in the region candidate maps.

First we build a hierarchical tree structure using the region candidate map. In our hierarchical tree structure, the leaf nodes represent the initial regions of the region candidate map. The regions are pair-wised merged using a simple hierarchical clustering algorithm to construct the tree. The root node of the tree corresponds to the whole image region. Each node in the tree is a candidate region. The tree structure satisfies the “non-overlapping” criteria, a common way for hierarchical tree based segmentation methods [12, 31, 32].

Suppose there are N base candidate regions, the total number of nodes in the tree would be $2N - 1$. We denote $R = \{R_1, R_2, \dots, R_{2N-1}\}$ as the region candidate map consisting of a set of region candidates R_i . Our goal is to select nodes in the tree as our final muscle cell segments. We show that this can be achieved by the condition random field (CRF) algorithm [33].

CRF has been widely used in image segmentation. It is a probabilistic graphical model aiming at maximizing a posterior given a defined energy function. In our method, the energy function is defined as

$$E(R) = \sum_{i=1}^{2N+1} U_i(R_i) + \sum_{(i,j) \in \hat{R}} V_i(R_i, R_j), \quad (2)$$

where \hat{R} is the subset of R contains all adjacent regions (i.e., any leaf nodes of a common father node) in leaves of the hierarchical tree. $U_i(R_i)$ is the unary term for region R_i , which is a score to evaluate the probability of R_i covering a complete cell segment. We adopt our previously developed method [12] to evaluate U_i by training a cell scoring classifier, which is able to assign a probability value to determine whether a segment is a good region candidate. In brief, a set of features based on multiple cues are proposed to represent the candidate regions and a standard RF classifier is trained to classify the cell regions. $V_i(R_i, R_j)$ is the pair-wise term to evaluate the dissimilarity between two regions R_i and R_j . We define $V_i(R_i, R_j)$ as

$$V_i(R_i, R_j) = \mu e^{-B(R_i, R_j)} \times L(R_i, R_j), \quad (3)$$

where $B(R_i, R_j)$ is the boundary strength and $L(R_i, R_j)$ is the boundary length. μ is a constant to trade-off the contribution of the two terms. These two terms can be calculated based on the single-pixel wide and closed region candidate maps generated by OWT-UCM [29].

The inference procedure is to minimize the energy function E so as to assign a label (1 means this region is a complete cell segment and 0, otherwise) to each region in the node and, at the same time, satisfy the “non-overlapping” criteria, i.e., any subtree can only has one label. We deploy the pylon model, a hierarchical CRF model, to minimize E [34]. However, the tree will become very big as the number of initial segments inside increases.

In the next section, we propose a strategy to divide the inference procedure into several sub-problems which can be parallelized using multi-core programming.

Parallel muscle image segmentation

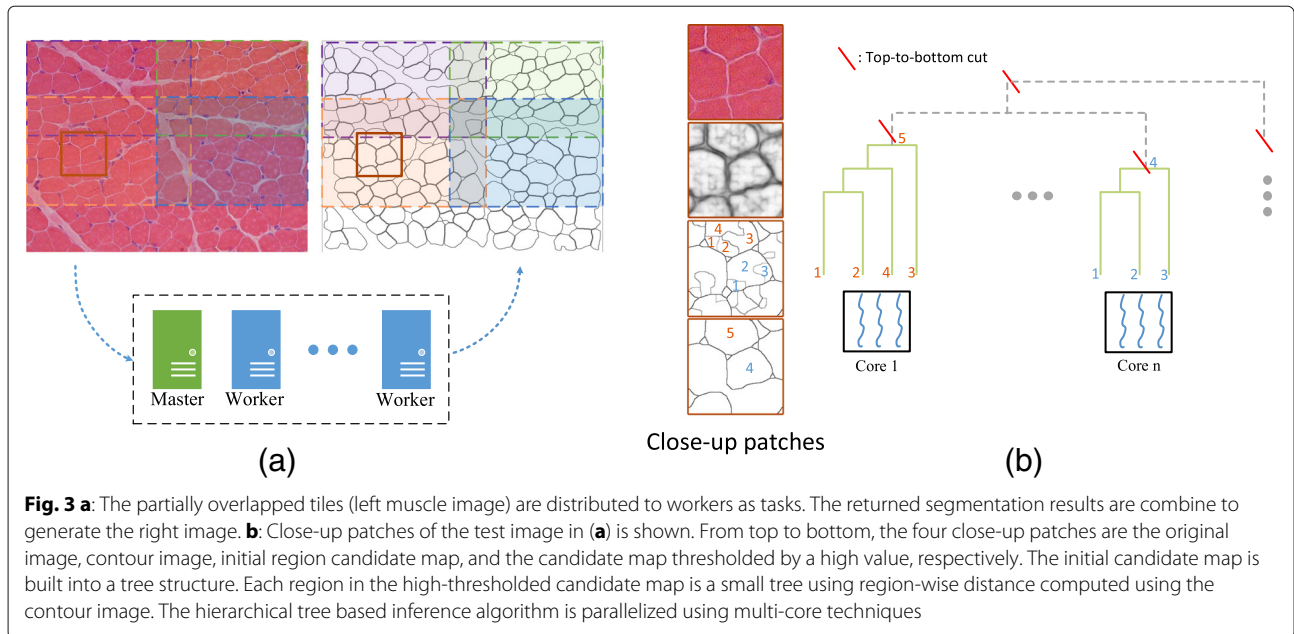
In this section, we present the proposed data distributed and model parallelized approach for muscle cell segmentation. We first introduce the data distribution procedure, which assigns the data (i.e., non-overlapped image tiles) to multiple workers using a master-worker parallelism manner. Then we introduce the method to parallelize the proposed hierarchical tree based region selection method using multi-core programming. Figure 3 illustrates the two steps.

Data distribution using spark

Due to the extremely high resolution of muscle images, the running time cost on a standalone machine is computational expensive. Since the segmentation of different image regions is independent with each other, we propose to divide the image into multiple partially-overlapped tiles and distribute them onto multiple worker nodes for concurrent processing.

To this end, we implement this parallel strategy in a master-worker manner with the Spark cloud computing platform [35]. In comparison with other distributed computing frameworks, Spark has the following advantages: 1) it has a flexible cluster management mechanism such that a parallel system can be easily built and run on local clusters; 2) it uses an Resilient Distributed Datasets (RDDs) technique [36] to perform in-memory computations, which is suitable for applications requires large storage space; 3) it exhibits strong compatibility, supporting multiple standard programming languages.

Our parallel muscle image segmentation algorithm consists of three steps: 1) data distribution: the test image I is divided into w tiles, $\mathbb{I}_1, \dots, \mathbb{I}_w$, and the master dynamically maps \mathbb{I}_w to all worker nodes using a user-defined map function; 2) segmentation: on each worker node, the proposed cell segmentation algorithm will be executed on multi-cores to perform contour detection, region candidate generation, and region selection; 3) data collection: the segmentation results returned by each worker node are collected to form the final segmentation. To avoid the loss of cell segments crossing the stitching positions of different tiles, we simply pad the tiles to make neighborhood tiles partially overlapped (the padding size is empirically set to 300×300). In order to reduce the overhead of data transfer between master-worker and alleviate extra cost of combing results returned from workers, we only require workers to return masked binary images which will be concatenated as the final segmentation results as shown in Fig. 3a.



With above data level parallelism, we can speed up the segmentation algorithm with no more than K times (because of data communication overhead) with K worker nodes in the cluster. To further speed up our segmentation algorithm, we parallelize the proposed hierarchical tree based region selection algorithm.

Hierarchical inference in parallel

The proposed hierarchical tree based inference method is mainly composed of: 1) building a tree structure using the region candidate map, 2) extracting feature representation for each R_i in the tree node, 3) computing $U_i(R_i)$ for each R_i , and 4) minimizing the energy function E . Based on our experiments, we observe that steps 2 and 3 dominate the time cost when number of nodes in the tree

grow to a large size. This is usually owing to two reasons. First, there are a large number of cells in an muscle image. Second, the low muscle image quality causes contour image having many false positive detections, which make the generated region candidate map contain numerous initial over-segments. However, we can still use the intensity of the contour image to evaluate the probability of real cell contours. We cut the tree from top-to-bottom by the region-wise distance computed from the detected contour image. We regard two adjacent regions whose common contour intensity above a certain threshold as two separate cells, and thus this two regions are not necessary to be clustered to a single subtree, so as their ancestor nodes. Figure 3b illustrates the idea. Therefore, the tree is separated into several subtrees and the energy

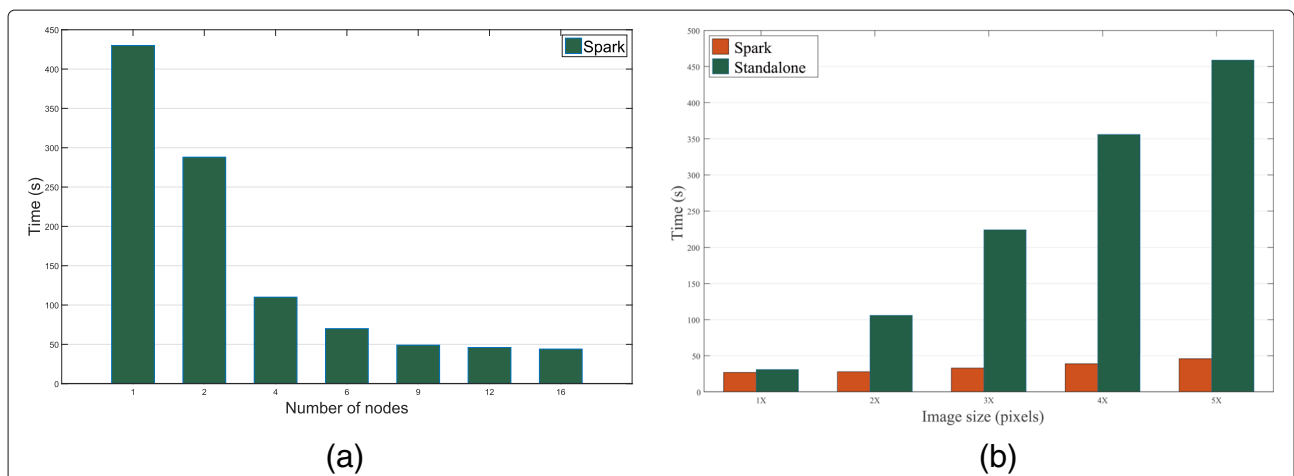


Fig. 4 a: The running time cost using different number of nodes on Spark. **b:** The comparison of time cost between the proposed parallel method and the standalone version. The x-axis is the image size (1x = 1000 * 800)

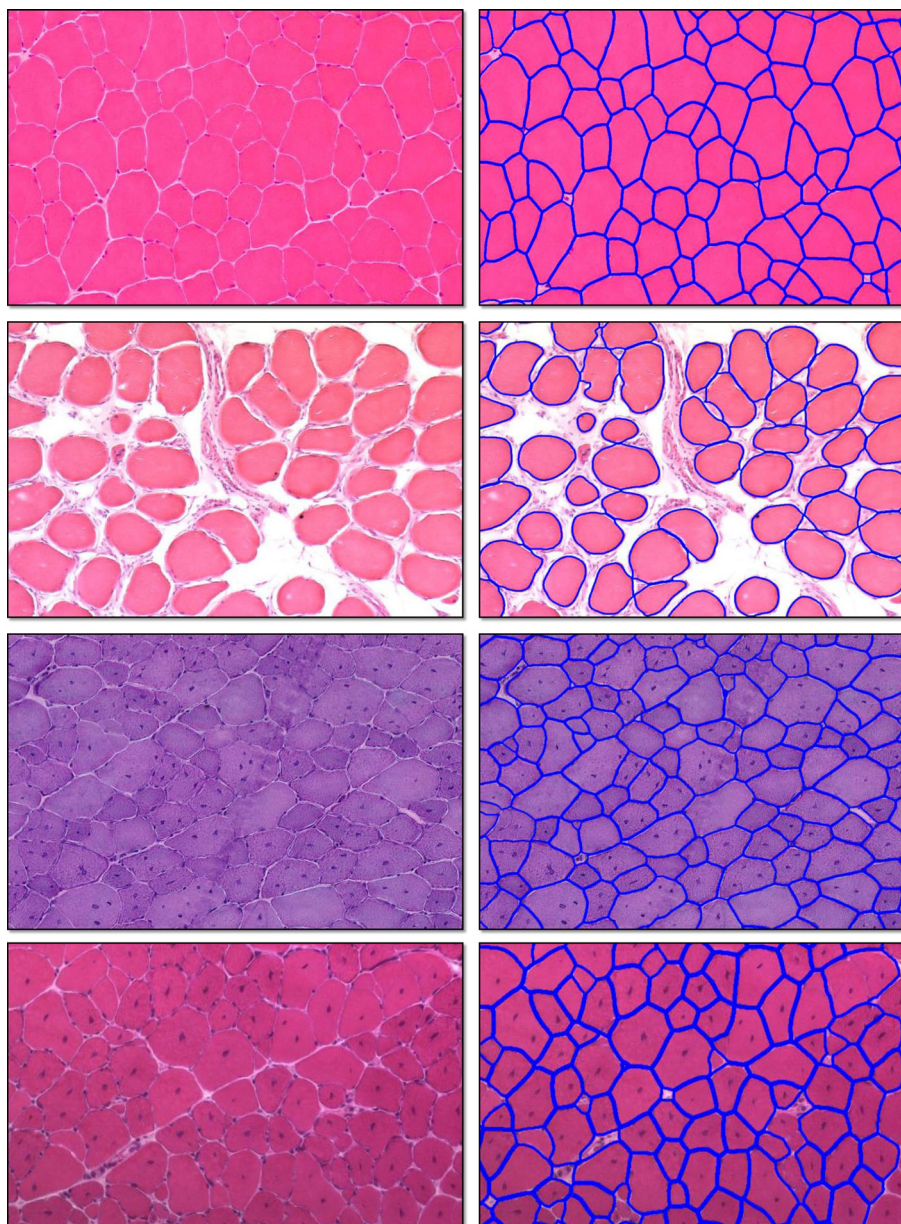


Fig. 5 Segmentation results on four sample H&E stained skeletal muscle image patches. The left column is the original images and the right column is the corresponding overlaid segmentation results. The blue lines are the contours of segmented cells overlaid on the original images for better visualization

minimization process (step 4) between subtrees is independent. We parallelize the inference algorithm using a multi-core programming technique on all worker nodes.

Experimental results

In this section, we demonstrate the efficiency of our proposed parallel approach compared with the standalone mode for large-scale muscle image segmentation. We also evaluate the segmentation accuracy compared

with other methods on a H&E stained skeletal muscle image dataset, which are captured by the whole-slide digital scanner from the cooperative institution Muscle Miner and the segmentation ground truth is annotated by several experts.

Data preparation

The images are cropped from a set of whole-slide scanned skeletal muscle images. We evaluate the efficiency of

the proposed method using a set of large-scale images (larger than 4500×3500). In addition, we measure the segmentation accuracy with a dataset contains 100 training images and 69 test images. The size of the images is varying from the scale of 600×600 to 2000×2000 . The segmentation ground truth is annotated by several experts. Note that we use this dataset for the segmentation accuracy evaluation as the image size of this dataset is adaptable to the competing muscle image segmentation methods.

Efficiency evaluation

To evaluate the efficiency, we build a small cluster using 8 Linux machines, each with 6 cores (Intel i7@3.60GHz \times 6) and 32 GB RAM. Each core is treated as a independent computing unit (worker node). In total we construct a cloud cluster with 48 nodes and 256 GB RAM.

The parallelism of the proposed method has two levels: data level parallelism using cloud computing and model level parallelism using multi-cores. Based on our observation, there is a trade-off between the tile size and the number of tiles (each tile is a task distributed to a worker node in the cluster). Given a test image, the more tiles we have, then the smaller tile size we obtain. If the tile size is too small, the computation duty of a worker node is too slight to maximize the performance of the multi-core parallel hierarchical tree region selection algorithm. Meanwhile, a large number of tiles would bring too much data communication cost. On the other hand, our model level parallelism may have resource (cores of each machine) conflicts with data level parallelism. Practically we use only 2 cores of each machine as worker nodes in the cluster, and thus in total we use a maximum number of 16 worker nodes.

In Fig. 4a, we visualize the time cost using different number of worker nodes in the cluster with a 4600×3800 test image. As we can see, as the number of nodes increases, the time cost drops dramatically. We can achieve a significant speed improvement when the number of node increasing from 1 to 8, but the time decreasing is not obvious from 9 to 12. This is attributed to the trade-off between the size and the number of image tiles, and the data communication overhead. The time cost for data communication will gradually increase as the tile size decreases. In Fig. 4b, we compare the time cost between the Spark based parallel mode and the standalone mode. We can obtain more than 10 times speedup with $5 \times (5000 \times 4000)$ image size.

Segmentation performance

To evaluate segmentation performance, we report precision, recall and F_1 -score, which is defined as

$$\begin{aligned} Precision &= \frac{|S \cap G|}{|S|}, & Recall &= \frac{|S \cap G|}{|G|}, \\ F_1\text{-score} &= 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall}, \end{aligned} \quad (4)$$

where S is the segmented cell region and G is the corresponding groundtruth cell region. $|\cdot|$ means the area of the region. Since the evaluation is cell-wised, for each test image, precision and recall is computed by averaging all cell evaluation results.

Figure 5 shows some the segmentation results, where the test images exhibit significant variations on cell sizes, shapes and appearances. It is clear that the proposed algorithm can accurately segment out most of the individual cell, which demonstrates the robustness of our proposed method. Figure 6 shows the precision-recall curve of our method. Our proposed method can preserve high precisions at recalls in a large range, which means that our method is capable to preserve and segment most of the cells in muscle images.

We compare the proposed parallel muscle image segmentation algorithm with two state-of-the-art image segmentation algorithms: 1) gPb [29], which is an edge-based image segmentation algorithm and has been widely used in the image segmentation field. The major drawback is its low efficiency, which takes about 300s for a 1000×1000 test image; 2) Isoperimetric graph partition (ISO) [37], which produces high quality segmentations as a spectral method with improved speed and stability. In Table 1, the proposed method outperforms the comparative

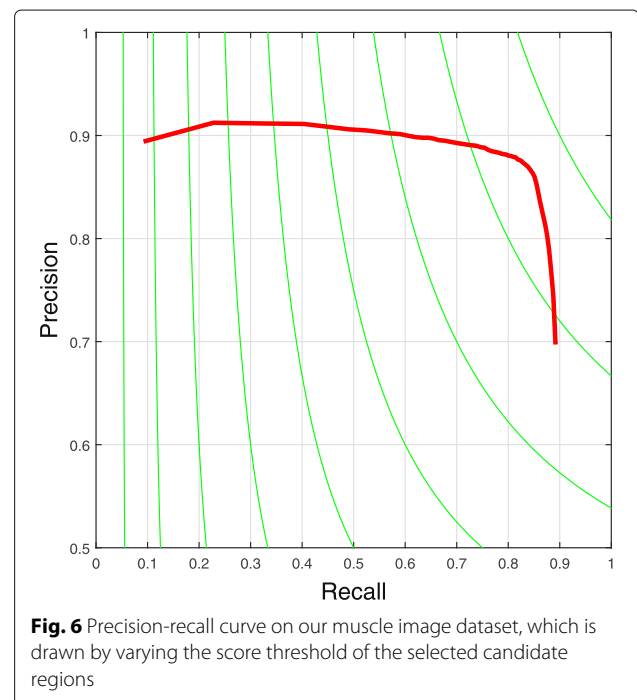


Table 1 The comparison results of state-of-the-art image segmentation algorithms

Method	F ₁ -score (%)		Precision (%)		Recall (%)	
	mean	std	mean	std	mean	std
ISO [37]	80.50	0.0993	89.88	0.0589	74.29	0.1369
gPb [29]	79.04	0.0780	91.23	0.0515	70.11	0.0962
Proposed method	84.61	0.0134	85.99	0.0035	85.11	0.0181

Our results reported here are computed by setting the score threshold of the selected candidate regions to 0.26 (see Fig. 6). Our proposed method shows significantly higher recall than others. These entries in boldface are means the best results

segmentation approaches. Although gPb performs a high precision, it exhibits very low recall. Compared with these algorithms, our algorithm achieves largely improved recall while exhibits significantly improved running time cost.

Conclusion

In this paper, we propose a parallel approach for fast and accurate H&E stained skeletal muscle image segmentation using cloud computing and multi-core programming, which can provide a high throughput solution for computer-aided muscle image analysis with significantly reducing the labor efforts. Specifically, we present a novel muscle image segmentation framework and demonstrate its accessibility to be parallelized. Then a data parallel approach is proposed to accelerate the proposed segmentation method in a master-worker parallelism manner based on the Spark cloud computing platform. To further maximize the computational efficiency on each worker node, we propose to a new strategy to parallelize our proposed hierarchical tree inference algorithm for region selection using multi-core techniques. Experimental results indicate a more than 10 times speed improvement compared with the standalone mode of the proposed segmentation method. Moreover, the comparison results with several competing methods demonstrate the superior performance of the proposed method on our H&E skeletal muscle image dataset.

Endnote

¹<https://github.com/pdollar/toolbox/tree/master/channels>

Abbreviations

COPD: Chronic obstructive pulmonary disease; CRF: Conditional random field; H&E: Haematoxylin and eosin; ISO: Isoperimetric graph partition; OWT-UCM: Oriented watershed transform and ultra-metric contour map; RDDs: Resilient distributed datasets; RF: Random forest; SRF: Structured random forest

Acknowledgements

We would like to thank all study participants.

Funding

This work was supported by the National Key R&D Program of China under grant 2017YFB1002504, National Natural Science Foundation of China (No. 81727802) and National Natural Science Foundation of China (No. 61701404)

Availability of data and materials

The data that support the findings of this study are available from cooperative institution Muscle Miner but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of cooperative institution Muscle Miner.

Authors' contributions

LC and ZZ conceived of the study as the principle investigator. JF and LY helped draft the manuscript and complete the study design and technical details, and participated in the experiment design. All authors read and approved the final manuscript.

Ethics approval and consent to participate

Not applicable.

Consent for publication

The authors have obtained consent to publish from the participant (or legal parent or guardian for children) to report individual patient data.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹Department of Information Science and Technology, Northwest University, Xi'an, China. ²Department of Computer and Information Science and Engineering, University of Florida, Gainesville, FL, USA.

Received: 23 April 2018 Accepted: 7 March 2019

Published online: 28 March 2019

References

- Fry CS, Lee JD, Mula J, et al. Inducible depletion of satellite cells in adult, sedentary mice impairs muscle regenerative capacity without affecting sarcopenia. *Nat Med.* 2015;21(1):76.
- Lawlor MW, Viola MG, Meng H, et al. Differential muscle hypertrophy is associated with satellite cell numbers and Akt pathway activation following activin type IIb receptor inhibition in Mtm1 p. R69C mice. *Am J Pathol.* 2014;184(6):1831–42.
- Mula J, Lee JD, Liu F, et al. Automated image analysis of skeletal muscle fiber cross-sectional area. *Am J Physiol Heart Circ Physiol.* 2012.
- Liu F, Mackey AL, Srikuea R, et al. Automated image segmentation of haematoxylin and eosin stained skeletal muscle cross-sections. *J Microsc.* 2013;252(3):275–85.
- Janssens T, Antanas L, Derde S, et al. CHARISMA: An integrated approach to automatic H&E-stained skeletal muscle cell segmentation using supervised learning and novel robust clump splitting. *Med Image Anal.* 2013;17(8):1206–19.
- Su H, Xing F, Lee JD, et al. Learning based automatic detection of myonuclei in isolated single skeletal muscle fibers using multi-focus image fusion. In: 2013 IEEE 10th International Symposium on Biomedical Imaging. IEEE; 2013. p. 432–5.
- Xie Y, Xing F, Kong X, et al. Beyond classification: structured regression for robust cell detection using convolutional neural network. In:

- International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham: Springer; 2015. p. 358–65.
8. Xing F, Yang L. Fast cell segmentation using scalable sparse manifold learning and affine transform-approximated active contour. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham: Springer; 2015. p. 332–9.
 9. Nguyen BP, Heemskerk H, So PTC, et al. Superpixel-based segmentation of muscle fibers in multi-channel microscopy. *BMC Syst Biol*. 2016;10(5):124.
 10. Bova N, Gál V, Ibáñez Ó, et al. Deformable models direct supervised guidance: A novel paradigm for automatic image segmentation. *Neurocomputing*. 2016;177:317–33.
 11. Klemencič A, Kovačič S, Pernuš F. Automated segmentation of muscle fiber images using active contour models. *Cytom J Int Soc Anal Cytol*. 1998;32(4):317–26.
 12. Liu F, Xing F, Zhang Z, et al. Robust muscle cell quantification using structured edge detection and hierarchical segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham: Springer; 2015. p. 324–31.
 13. Yang L, Qi X, Xing F, et al. Parallel content-based sub-image retrieval using hierarchical searching. *Bioinformatics*. 2013;30(7):996–1002.
 14. Yang L, Kim H, Parashar M, et al. High Throughput Landmark Based Image Registration Using Cloud Computing. *MICCAI2011-HP/DCI*. 2011;38–47.
 15. Ghaznavi F, Evans A, Madabhushi A, et al. Digital imaging in pathology: whole-slide imaging and beyond. *Annu Rev Pathol Mech Dis*. 2013;8:331–59.
 16. Van Aart E, Sepasian N, Jalba A, et al. CUDA-Accelerated Geodesic Ray-Tracing for Fiber Tracking. *Int J Biomed Imaging*. 2011;2011:698908.
 17. Kagadis GC, Kloukinas C, Moore K, et al. Cloud computing in medical imaging. *Med Phys*. 2013;40(7):070901.
 18. Yang L, Qi X, Xing F, et al. Parallel content-based sub-image retrieval using hierarchical searching. *Bioinformatics*. 2013;30(7):996–1002.
 19. Qi X, Wang D, Rodero I, et al. Content-based histopathology image retrieval using CometCloud. *BMC Bioinformatics*. 2014;15(1):287.
 20. Donoser M, Schmalstieg D. Discrete-continuous gradient orientation estimation for faster image segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2014. p. 3158–65.
 21. Arbeláez P, Pont-Tuset J, Barron JT, et al. Multiscale combinatorial grouping. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2014. p. 328–35.
 22. Su H, Xing F, Kong X, et al. Robust cell detection and segmentation in histopathological images using sparse reconstruction and stacked denoising autoencoders. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham: Springer; 2015. p. 383–90.
 23. Liu X, Song M, Tao D, et al. Semi-supervised node splitting for random forest construction. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2013. p. 492–9.
 24. Breiman L. Random forests. *Mach Learn*. 2001;45(1):5–32.
 25. Lim JJ, Zitnick CL, Dollár P. Sketch tokens: A learned mid-level representation for contour and object detection. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2013. p. 3158–65.
 26. Dollár P, Zitnick CL. Structured forests for fast edge detection. In: Proceedings of the IEEE international conference on computer vision. 2013. p. 1841–8.
 27. Dollár P, Zitnick CL. Fast edge detection using structured forests. *IEEE Trans Pattern Anal Mach Intell*. 2015;37(8):1558–70.
 28. Dollár P, Tu Z, Perona P, et al. Integral channel features. 2009;91:1–11.
 29. Arbeláez P, Maire M, Fowlkes C, et al. Contour detection and hierarchical image segmentation. *IEEE Trans Pattern Anal Mach Intell*. 2011;33(5):898–916.
 30. Roerdink JBTM, Meijster A. The watershed transform: Definitions, algorithms and parallelization strategies. *Fundam Informaticae*. 2000;41(1, 2):187–228.
 31. Liu F, Xing F, Yang L. Robust muscle cell segmentation using region selection with dynamic programming. In: 2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI). IEEE; 2014. p. 521–4.
 32. Arteta C, Lempitsky V, Noble JA, et al. Learning to detect cells using non-overlapping extremal regions. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Berlin: Springer; 2012. p. 348–56.
 33. Lafferty JD, McCallum A, Pereira FCN. Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data. Morgan Kaufmann Publishers Inc.; 2001. p. 282–9.
 34. Lempitsky V, Vedaldi A, Zisserman A. Pylon model for semantic segmentation. In: Advances in neural information processing systems. 2011. p. 1485–93.
 35. Zaharia M, Chowdhury M, Franklin MJ, et al. Spark: Cluster computing with working sets. *HotCloud*. 2010;10(10-10):95.
 36. Zaharia M, Chowdhury M, Das T, Dave A, Ma J, McCauley M, Franklin MJ, Shenker S, Datasets RD. A Fault-Tolerant Abstraction for In-Memory Cluster Computing. *Ion Stoica: NSDI*; 2012. p. 12.
 37. Grady L, Schwartz EL. Isoperimetric graph partitioning for image segmentation. *IEEE Trans Pattern Anal Mach Intell*. 2006;28(3):469–75.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

