Knowledge-Based Multi-sequence MR Segmentation via Deep Learning with a Hybrid U-Net++ Model

Jinchang Ren[1(✉)], He Sun[1], Yumin Huang[1], and Hao Gao[2]

[1] Department of Electronic and Electrical Engineering, University of Strathclyde, Glasgow, UK
{jinchang.ren,h.sun}@strath.ac.uk, yumin.huang.2016@uni.strath.ac.uk
[2] School of Mathematics and Statistics, University of Glasgow, Glasgow, UK
Hao.Gao@glasgow.ac.uk

Abstract. The accurate segmentation, analysis and modelling of ventricles and myocardium plays a significant role in the diagnosis and treatment of patients with myocardial infarction (MI). Magnetic resonance imaging (MRI) is specifically employed to collect imaging anatomical and functional information about the cardiac. In this paper, we have proposed a segmentation framework for the MS-CMRSeg Multi-sequence Cardiac MR Segmentation Challenge, which can extract the desired regions and boundaries. In our framework, we have designed a binary classifier to improve the accuracy of the left ventricles (LVs). Extensive experiments on both validation dataset and testing dataset demonstrate the effectiveness of this strategy and give an insight towards the future work.

Keywords: Cardiac image segmentation · Binary classifier · U-Net++

1 Introduction

Recent advances in the MRI technology have led to an effective way to manage the treatment plan for patients. Since the MR can provide a better enhancement in the infected area and highlight the illness part with special brightness, a variety of applications have been developed, such as the Carotid arterial plaque stress analysis [1], the tracking of myocardial deformation [2], etc. Among these tasks, the accurate segmentation of the MR image can help to extract the desired regions of interest (ROIs), which plays a key role in the clinical analysis. However, there are many challenging issues towards the performance of segmentation in MRI. Firstly, unlike the segmentation work in natural images, a more precise segmentation result is required in the MR image, even tiny errors may damage the result of further diagnose. Secondly, the manual segmentation of the desired boundaries is very time-consuming and error prone, which is not accessible in the practical applications. Several challenging examples are shown in the Fig. 1. What is more, it is crucial to fully utilize the provided multiple modalities of MR

M. Pop et al. (Eds.): STACOM 2019, LNCS 12009, pp. 280–289, 2020.
https://doi.org/10.1007/978-3-030-39074-7_30

images, which may improve the segmentation accuracy. Therefore, an effective and robust segmentation strategy is required, especially for the MS-CMRSeg Multi-sequence Cardiac MR Segmentation Challenge [3,4].

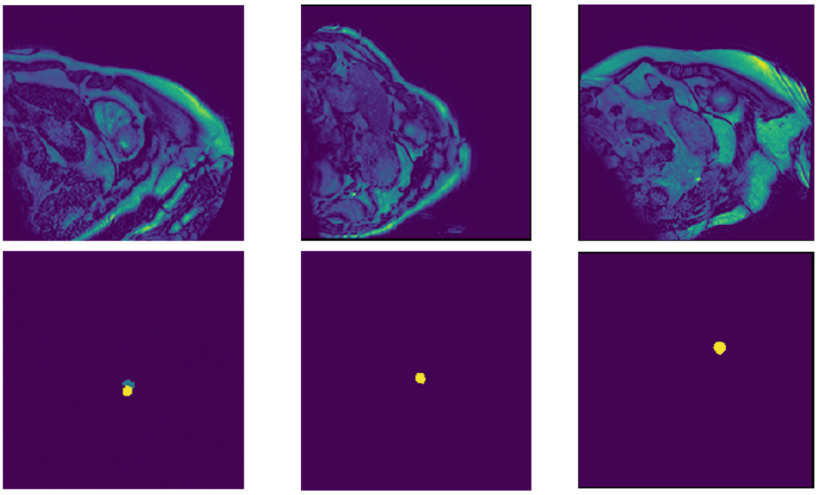

Fig. 1. Several examples of the challenging images.

In this paper, we propose an effective segmentation framework for the provided CMR images. To figure out the given task, we have analyzed the dataset and chosen a suitable deep learning framework first, i.e. the U-Net++ [4]. After that, some preprocessing techniques are considered before the U-Net++ to reduce the potential noise and improve the final performance. As the left vertical usually has a fixed shape in most of slices, we have designed a binary classifier module in the U-Net++ framework, which can significantly improve the accuracy of left ventricle.

The rest of this paper is organized as follows. In Sect. 2, the utilized dataset is introduced, along with a brief summary of the Multi-sequence Cardiac MR Segmentation Challenge. The motivation and implementation of our designed framework is detailed in Sect. 3. In Sect. 4, the experimental results and analysis are presented and discussed. Finally, some concluding remarks are drawn in Sect. 5.

2 Dataset

In this paper, we have conducted experiments on the dataset from MS-CMRSeg 2019 Multi-sequence Cardiac MR (CMR) Segmentation Challenge [3,4]. The whole CMR data are come from 45 patients where each patient has been scanned by three CMR sequences, including the late gadolinium enhancement (LGE),

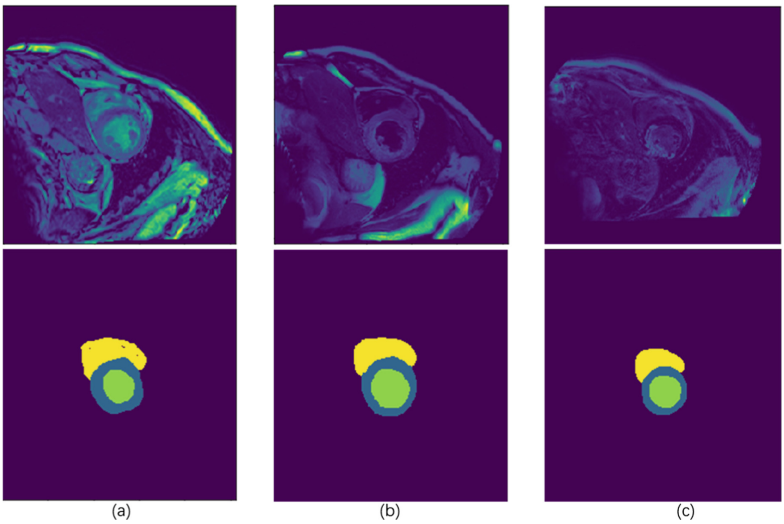

Fig. 2. Example of three different sequences from the same patient, the first row are the original images with false color and the second are the corresponding ground truth with false color, where the green, the cyan and the yellow pixels represents the left ventricle, the myocardium and the right ventricle, respectively: (a) bSSFP. (b) T2. (c) LGE (Color figure online)

T2 and balanced-Steady State Free Precession (bSSFP). In this challenge, the dataset is separated as training set and testing set. Four classes are labelled in the ground truth data, including the left ventricle, the right ventricle (RV), the myocardium and the background. The training set includes the LGE CMR images with ground truth of the first to the fifth patient, i.e. the validation dataset, the T2-weighted and bSSFP CMR image with corresponding ground truth for the first to the thirty-fifth patient. And for the last ten patients, only the T2-weighted and bSSFP CMR images are provided. For the testing dataset, the rest LGE CMR images are utilized for the final evaluation. For those three different sequences, each LGE CMR image usually consists of 10 to 18 slices, covering the main body of the ventricles, and the BSSFP image consist of 8 to 12 sequential slices which have been scanned at the diastolic end. Different from the other two sequences, T2 CMR image only has a small number of slices (3, 5, 6, 7 or 8 slices), which implies the constituted testing dataset has a relatively smaller size than the training set. In total, the number of slices from three sequences in the training dataset is less than 500 and the size of slices from these three sequences is inconsistent. Some examples from the training dataset are illustrated in Fig. 2.

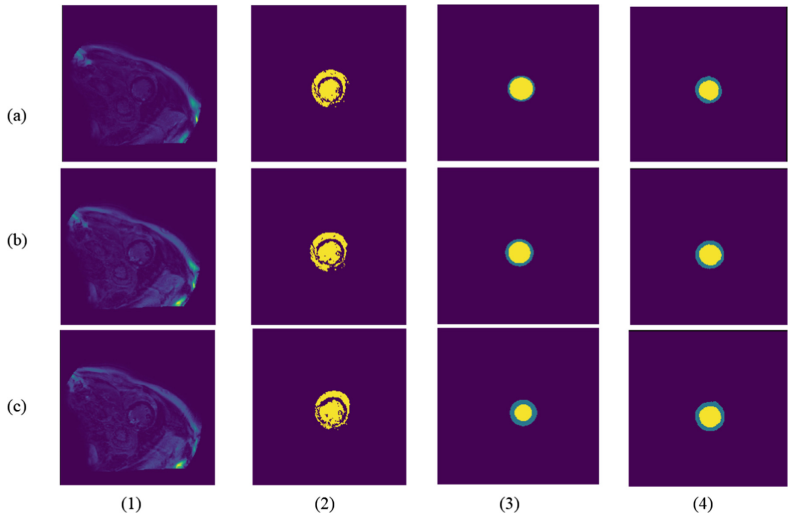

Fig. 3. Three classes segmentation results (a)–(c) by using K-means++ with Hough transform. (1) original images. (2) Results from K-means with Hough transform. (3) Results from further Hough transform. (5) Ground truth. The segmentation accuracy for left ventricle and myocardium after further Hough transform are 0.7443, 0.8677 and 0.5053, respectively.

3 The Proposed Framework

3.1 Motivation

In this section, our designed model will be discussed, which includes our motivation and implementation. Currently, numerous methods have been proposed for the segmentation problem, which can be classified as traditional methods and deep learning-based methods. To design a more robust framework, we have investigated both above methods and tried to combine them together.

First, some traditional image segmentation algorithms have been considered, including the Hough transform [5], the watershed [6], the active contours [7], the level set [8], and the K-means++ [9], etc. The Hough transform can achieve a better performance in left ventricle and myocardium, but it heavily relies on the chosen parameter. The watershed also suffers this dilemma and some redundant boundaries could be generated if parameters are not selected appropriately. For both active contours and the level set methods, they require some initial set of points to evolve the final boundary, which is not efficient in dealing with the given task. Although most of the traditional methods are not favorable, we have found out that the fusion of the binary K-means++ and Hough transform might be useful for the segmentation of left ventricles and myocardium, some results from a three classes (background, left ventricle and myocardium) classifier based on K-means++ and Hough transform are shown in the Fig. 3. What is more,

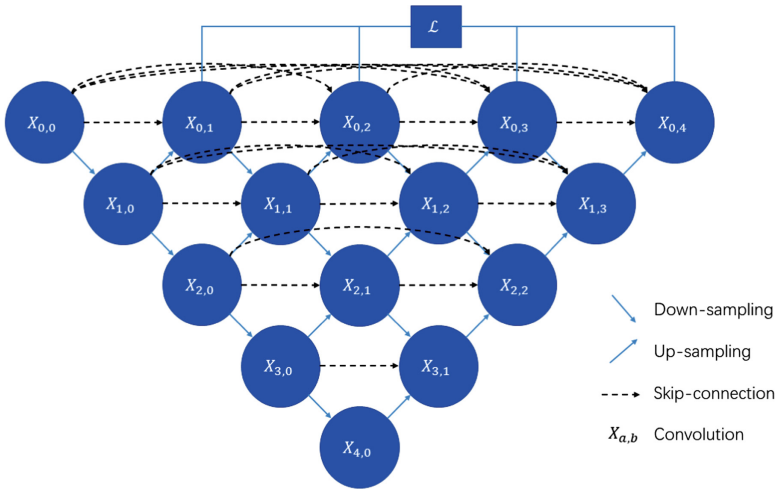

Fig. 4. The structure of U-Net++.

the binary classification demonstrates its superiority in the segmentation of left vertical, which can increase the accuracy.

In the last few years, the U-Net [10] framework has drawn great attention in the medical image segmentation area due to its efficiency and simple framework. With its skip connection architecture, the U-Net can better capture both local and global features, which fulfils the requirement of the medical imaging segmentation. According to the nested and dense skip connection, the U-Net++ is proposed to promote the performance and it has been proved to be a robust model. Compared to the U-Net, the U-Net++ can increase the accuracy from 1% to 5% in the validation dataset. Furthermore, the U-Net++ does not suffer from the heavy computational burden, which gives us a better option to investigate more. The structure of U-Net++ is shown in the Fig. 4. For increasing the segmentation accuracy, we have attempted to exploit some popular techniques, like data augmentation, etc. However, these techniques could not improve the accuracy significantly with less computational burden, which is not supportive for this challenge. For the data augmentation problem, the details are discussed in the experimental results.

3.2 Implementation

After the introduction of background of this challenge, the implementation of our proposed framework is presented in this subsection. Since the desired ROIs are located at the center of the slice and the size of each slice is inconsistent, we have employed a pre-processing step to deal with such inconsistency. We have resized all slices from the training dataset into the same size as both of their height and width are set to be 256. To remove the potential noise from the background, we have cropped the resized slices and keep the central part of each slice, where

the final height and width are set as 128. After the pre-processing step, we have trained two classifiers simultaneously with the training set. As discussed above, the binary classifier aiming to distinguish the left ventricle and the rest classes are advantageous, we have trained one binary classifier to find the left ventricle. Another four-classes segmentation model has been trained to classify all four classes in each slice. In the final, we combine the results from these two classifiers simply by considering the classification map from the binary model as the priority. The classification map generated from the four class classifier is filtered by the classification map generated from the binary classifier. The diagram of our framework is presented in the Fig. 5.

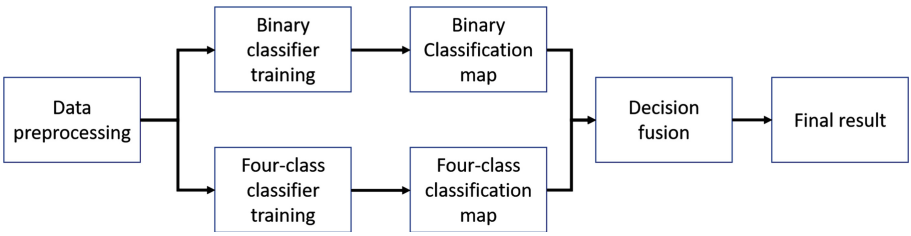

Fig. 5. The flowchart of our proposed framework.

Table 1. The mean and standard deviation of the Dice score on three classes.

	Left ventricle	Myocardium	Right ventricle
Dice	0.757 ± 0.127	0.470 ± 0.117	0.539 ± 0.151

4 Results

4.1 Parameter Settings

In this part, the utilized parameters in our proposed framework will be given in detail. The optimal parameters are chosen through the performance on validation dataset. For the training epochs, it is set to 600, and the batch size equals to 16. We have used the Adam optimizer with a learning rate of 1e-6. For achieving better performance, we have initialized the weight of loss of both classifiers, the binary model is set to be 1:6 for the background and the left ventricle and the four classes model is 1:6:6:6 where the background is set to be 1 and the rest are 6. Generally, we have only utilized the training dataset to train our model and the validation dataset to prove the efficiency of our approach. However, for the purpose of achieving better performance on the challenge, we have utilized the validation dataset to train the model in the last submission. For the hardware requirement, all our experiments and the training process of our proposed framework are conducted on the Google Colab with TESLA K80 (12 GB).

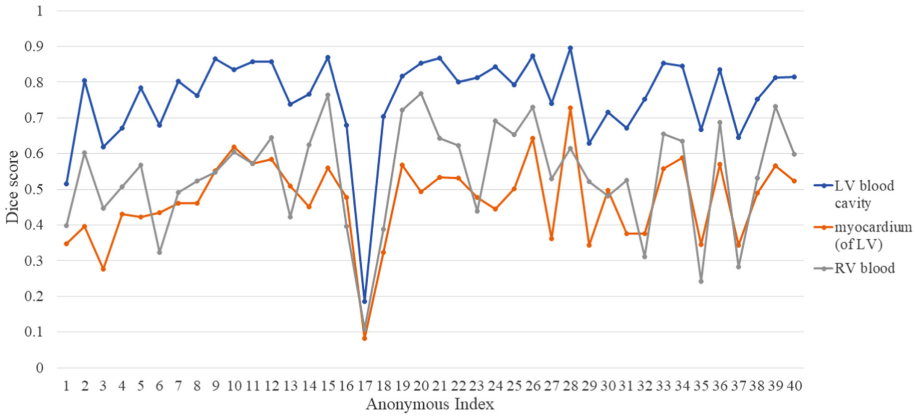

Fig. 6. The Dice score from the organizers.

4.2 Performance of Test Dataset

To evaluate the performance our proposed framework, we have tested our proposed framework on the challenge test dataset and received the feedback from organizers, including the Dice score, the Jaccard index, the Surface and Hausdorff distance. The result about the Dice score is shown in Fig. 6 and corresponding mean and standard derivation are depicted in the Table 1. From the result, it can be clearly seen that the accuracy of the left ventricles is much better than the rest, which proves that the binary classifier is suitable for the segmentation of left ventricles.

4.3 Experimental Analysis

In this part, we will present the related experiments to discuss some interesting issues and give some further comparison about our framework.

In our model, we have defined a pre-processing step and designed a binary classifier. Therefore, it is crucial to prove the significance of these two ideas. Therefore, we have done a comparison between our proposed methods and one four classes classifier framework on the validation dataset, which is depicted as 'Binary+Four' and 'Only Four'. Besides, we have also implemented the data augmentation step to inspect its function. The rest parameters are kept the same as above. The performance is shown in the Table 2, which includes the 'ROI'accuracy, the total accuracy and the training time. The 'ROI' accuracy is calculated only from the left ventricle, myocardium and right ventricle pixels, which is interpreted in Fig. 8. It can be recognized that after the combination of binary and four-class classifiers, the accuracy can be improved by 2%. As seen in Table 2, the effect of data augmentation can be noticed, where the accuracy is increased by 9% with more training samples from the data augmentation. Although data augmentation can generate more training samples and improve the accuracy, its huge computational burden seems impractical, especially for the limited time challenge or real-time applications (Fig. 7).

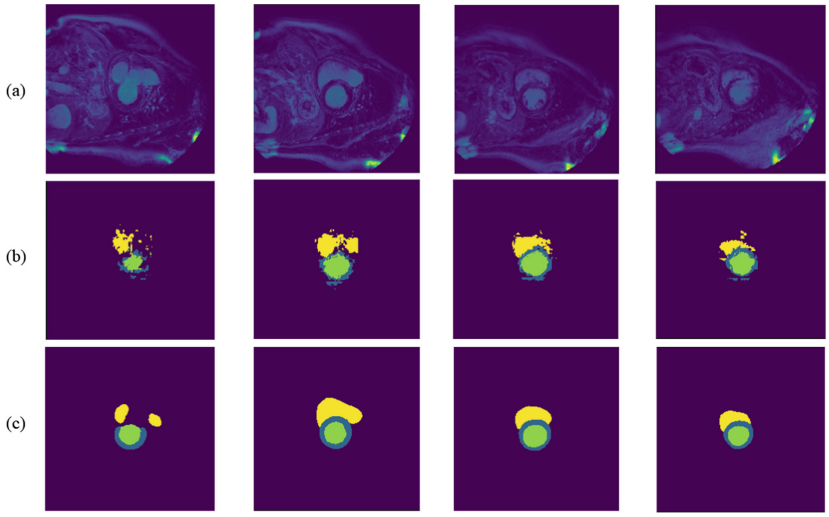

Fig. 7. Some results from the validation dataset. (a) original images. (b) segmented mask. (c) Ground truth.

Table 2. The mean and standard deviation of the Dice score on three classes.

Model	Augmentation	ROI accuracy	Total accuracy	Training time (h)
Binary+Four	Yes	0.6789	0.9803	18
Binary+Four	No	0.5893	0.9764	1.8
Only Four	Yes	0.6572	0.9802	9
Only Four	No	0.5707	0.9762	0.9

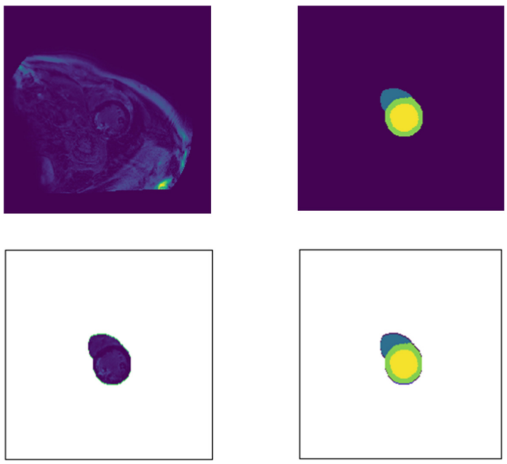

Fig. 8. The example of defined 'ROI'. Left top: the original image. Right top: ground truth of original image. Left bottom: cropped 'ROI'. Right bottom: ground truth of cropped 'ROI'.

5 Conclusion

In this paper, we have proposed a U-Net++ based framework for the MRI segmentation, especially for the MS-CMRSeg Multi-sequence Cardiac MR Segmentation Challenge. Although the performance on the right ventricle and myocardium are not good enough, we have obtained better performance on the left ventricle with the binary classifier module, which can give an insight about the Cardiac segmentation task.

Due to the time limitation of this challenge, there are still some unsolvable issues about the Cardiac MR segmentation. In the future, we will focus on the design of a more flexible model for the Cardiac MR segmentation, which can fully utilize the multiple sequence data and capture more 3D information. What is more, some multi-stage methods will also be investigated [11–16] and more reliable features will be extracted [17–20] to remove the noise and further improve the accuracy.

References

1. Gao, H., et al.: Carotid arterial plaque stress analysis using fluid-structure interactive simulation based onin-vivomagnetic resonance images of four patients. J. Biomech. **42**, 1416–1423 (2009)
2. Schuster, A., et al.: Cardiovascular magnetic resonance myocardial feature tracking concepts and clinical applications. Circ. Cardiovasc. Imaging **9**(4), e004077 (2016)
3. Zhuang, X.: Multivariate mixture model for cardiac segmentation from multi-sequence MRI. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, Athens Greece, pp. 581–588 (2016)
4. Zhuang, X.: Multivariate mixture model for myocardial segmentation combining multi-source images. IEEE Trans. Pattern Anal. Mach. Intell. (2018). https://doi.org/10.1109/TPAMI.2018.2869576
5. Ballard, D.H.: Generalizing the Hough transform to detect arbitrary shapes. Pattern Recogn. **13**(2), 111–122 (1981)
6. Meyer, F.: Topographic distance and watershed lines. Signal Process. **38**(1), 113–125 (1994)
7. Chan, T.F., Vese, L.A.: Active contours without edges. IEEE Trans. Image Process. **10**(2), 266–277 (2001)
8. Sussman, M., Smereka, P., Osher, S.: A level set approach for computing solutions to incompressible two-phase flow. J. Comput. Phys. **114**(1), 146–159 (1994)
9. Arthur, D., Vassilvitskii, S.: K-means++: the advantages of careful seeding. In: Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms, pp. 1027–1035. Society for Industrial and Applied Mathematics (2007)
10. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) MICCAI 2015. LNCS, vol. 9351, pp. 234–241. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24574-4_28
11. Ijitona, T.B., et al.: SAR sea ice image segmentation using watershed with intensity-based region merging. In: 2014 IEEE International Conference on Computer and Information Technology, Xi'an, pp. 168–172 (2014)

12. Ren, J., et al.: Effective SAR sea ice image segmentation and touch floe separation using a combined multi-stage approach. In: 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Milan, pp. 1040–1043 (2015)
13. Xie, X., et al.: Automatic image segmentation with superpixels and image-level labels. IEEE Access **7**, 10999–11009 (2019)
14. Huang, H., et al.: Combined multiscale segmentation convolutional neural network for rapid damage mapping from postearthquake very high-resolution images. J. Appl. Remote Sens. **13**(2), 022007 (2019)
15. Sun, G., et al.: Dynamic post-earthquake image segmentation with an adaptive spectral-spatial descriptor. Remote Sens. **9**(9), 899 (2017)
16. Hwang, B., et al.: A practical algorithm for the retrieval of floe size distribution of Arctic sea ice from high-resolution satellite Synthetic Aperture Radar imagery. Elem. Sci. Anth. **5**, 38 (2017)
17. Han, J., et al.: Background prior-based salient object detection via deep reconstruction residual. IEEE Trans. Circuits Syst. Video Technol. **25**(8), 1309–1321 (2015)
18. Han, J., et al.: Object detection in optical remote sensing images based on weakly supervised learning and high-level feature learning. IEEE Trans. Geosci. Remote Sens. **53**(6), 3325–3337 (2015)
19. Cheng, G., et al.: Effective and efficient midlevel visual elements-oriented land-use classification using VHR remote sensing images. IEEE Trans. Geosci. Remote Sens. **53**(8), 4238–4249 (2015)
20. Fang, L., et al.: Classification of hyperspectral images by exploiting spectral-spatial information of superpixel via multiple kernels. IEEE Trans. Geosci. Remote Sens. **53**(12), 6663–6673 (2015)