SEMANTIC SEGMENTATION OF REMOTE SENSING IMAGERY USING OBJECT-BASED MARKOV RANDOM FIELD BASED ON HIERARCHICAL SEGMENTATION TREE WITH AUXILIARY LABELS

Lin He¹*, Zhaocong Wu², Yi Zhang², Zhongwen Hu³

¹ College of surveying and geo-informatics, Tongji University, 200092 Shanghai, China

² School of remote sensing and information engineering, Wuhan University, 430072 Wuhan, China ³ Geo-Environmental Monitoring of Coastal Zone of the Ministry of Natural Resources & Guangdong Key Laboratory of Urban Informatics & Shenzhen Key Laboratory of Spatial Smart Sensing and Services & Research Institute for Smart Cities, Shenzhen University, 518060 Shenzhen, China

Commission III, WG III/1

KEY WORDS: Hierarchical segmentation tree, Auxiliary label field, Object-based Markov random field, Multiscale segmentation, Remote sensing imagery, Semantic segmentation

ABSTRACT:

In the remote sensing imagery, spectral and texture features are always complex due to different landscapes, which leads to misclassifications in the results of semantic segmentation. The object-based Markov random field provides an effective solution to this problem. However, the state-of-the-art object-based Markov random field still needs to be improved. In this paper, an object-based Markov Random Field model based on hierarchical segmentation tree with auxiliary labels is proposed. A remote sensing imagery is first segmented and the object-based hierarchical segmentation tree is built based on initial segmentation objects and merging criteria. And then, the object-based Markov random field with auxiliary label fields is established on the hierarchical tree structure. A probabilistic inference is applied to solve this model by iteratively updating label field and auxiliary label fields. In the experiment, this paper utilized a Worldview-3 image to evaluate the performance, and the results show the validity and the accuracy of the presented semantic segmentation approach.

1. INTRODUCTION

Semantic segmentation is defined as a multi-label classification problem (Hu et al., 2018), which aims to assign category labels to each pixel on the image. It contains two tasks, image segmentation and target recognition. Semantic segmentation is of great significance to the understanding and analysis of the imagery and is widely applied in various fields, such as automatic driving, land-use and land-cover classification. Semantic segmentation of remote sensing imagery can be regarded as a typical multi-class image classification (Chen et al., 2013) by assigning predefined semantic classes to a remote sensing image. With the development of remote sensing satellite technology, the temporal, spatial and spectral resolution of remote sensing image have improved gradually, which makes the information obtained from remote sensing image more and more abundant. Due to a large number of categories in the natural scene, some show a high degree of similarity in the spectral or texture features, so how to accurately recognize and distinguish different landscapes is still a challenge in the research of remote sensing. The semantic segmentation of remote sensing image provides an effective solution for image retrieval and analysis.

There are two groups of semantic segmentation methods for the remote sensing imagery, traditional method and deep learning method (Hu et al., 2018). The deep learning method mainly utilizes Convolution Neural Network (CNN) (Pedro H. O., Collobert, 2013, Yu et al., 2018), Fully Convolutional Network (FCN) (Long et al., 2014, Chen et al., 2018) and other improved neural network methods (Paszke et al., 2016, Kampffmeyer et al., 2016) to extract features and obtain semantic segmentation results. Traditional methods include Support Vector Machine (SVM) (Gupta et al., 2013, Huang, Zhang, 2013), Random Decision Forest (RDF) (Gupta et al., 2013, Hermans et al., 2014), Markov Random Field (MRF) (Zheng, Wang, 2015, Zheng et al., 2017, Zheng et al., 2019) and Conditional Random Field (CRF) (Volpi, Ferrari, 2015, Thøgersen et al., 2016), etc.

Compared with the deep learning method, the application of the traditional method in remote sensing imagery semantic segmentation is more conducive to analyze the physical information of landscapes. MRF is an effective method for semantic segmentation in traditional methods. One of the classic problems of the MRF is the Maximum a Posterior (MAP) (Kollar, 2014) estimation of the state vector. Based on the given parameters of model and data sets, the most likely state of the sequence is estimated through the posterior distribution. In the application of image semantic segmentation, the algorithm obtains the optimal result by solving MAP estimation of the label field. In semantic segmentation, the MRF model makes full use of spatial information constraints, which is suitable for spectral and texture processing of remote sensing imagery.

The classical method of semantic segmentation based on the MRF model is Pixel-based Markov Random Field (PMRF) (Geman, S, 1984, Besag, 1993). This method is defined on pixels and is utilized to measure the similarity between pixels. The most advantage of the model is its regular context, which is convenient for spatial relationship description and model solution. However, due to the improvement of spatial resolution of remote sensing imagery, the classical model is not suitable for

^{*}Corresponding author

capturing complex macroscopical features, and the calculation is time-consuming (Zheng, Wang, 2015). The Multi-resolution Markov Random Field (MRMRF) (Noda et al., 2002, Zheng et al., 2010) defined on the image pyramid structure extends the classic PMRF model. The method improves the computational efficiency and extends the descriptive spatial features to a certain extent, but it is still a pixel level Markov Random Field. With the application and development of Object-Based Image Analysis (OBIA) in remote sensing image, the Objectbased Markov Random Field (OMRF) (Yu, Clausi, 2008, Wang, Zhang, 2009, Blaschke, 2010, Zhang et al., 2017) has been widely used. A image is divided into several over-segmented regions, then the relationship between regions is expressed as Region Adjacency Graph (RAG), and finally the semantic segmentation is completed by the MRF model. The OMRF model is defined based on RAG, which breaks through the limitation of the pixel level model in the description of spatial features, and can better capture the macroscopical features of the image.

The OMRF used in the semantic segmentation of remote sensing imagery still faces many challenges. For example, there will be some misclassification if the OMRF does not fully emphasize spatial relationships. On the contrary, if spatial relations are overemphasized, the OMRF model will lead to oversmooth results. Hierarchical processing provides an effective strategy for spatial relationship analysis (Marfil, Bandera, 2015). With the application of multi-scale strategy and object-based method in remote sensing imagery segmentation, the OMRF and the MRF model based on hierarchical information is widely used in the process of remote sensing imagery semantic segmentation. Zheng et al. (Zheng et al., 2019) established the probability graph with a multilayer structure, integrates pixel label and object label to build the hybrid label field, and utilized joint distribution to capture the isotropy of the same layer and the anisotropy of different layers. In the iterative update of hybrid label field, multiple-granularity is integrated and interaction between different granularity layers is achieved. In the other article (Zheng et al., 2017), they defined two auxiliary labels in different categories, and built the conditional probability distribution of label field and auxiliary label fields.

Inspired by the interaction between layers based on the conditional probability distribution model (Zheng et al., 2017), this paper establishes the Object-based Markov Random Field based on hierarchical segmentation tree with auxiliary labels (OMRF-HA) to realize the semantic segmentation. In the hierarchical segmentation tree, there is a good corresponding relationship between objects in different layers, hence the classification reference of the segmented objects between adjacent layers is more reliable. When the model updates the label field and auxiliary label fields for the iterative probabilistic inference, the auxiliary label fields will have a more positive impact on the label field.

The remaining of this paper is organized as follows: Section 2 introduces the details of the proposed approach. The experimental results, analysis and evaluation are presented in Section 3. Finally, the conclusion is given in Section 4.

2. METHODOLOGY

This paper proposes an object-based Markov random field model with auxiliary label fields based on hierarchical segmentation tree structure. This algorithm first segments the remote sensing imagery into objects and establishes the object-based hierarchical segmentation tree. Based on the hierarchical tree structure, the OMRF model with auxiliary label is built. The result of semantic segmentation is obtained by iterative updating the label field and auxiliary label fields. Post-processing is applied to remove noises. An illustration is shown in Figure 1.



Figure 1. Flow diagram of Object-based Markov Random Field based on hierarchical segmentation tree with auxiliary labels (OMRF-HA).

2.1 Object-based Hierarchical Segmentation Tree

Considering the homogeneity of regions and edge features, the watershed algorithm (Li et al., 2010) is utilized to divide the remote sensing image I into initial over-segmented regions. The initial segmentation set is recorded as R and the number of objects is n. Each segmentation region is recorded as $R_i(i = 1, 2, ..., n)$ and there is no intersection between divided objects. The initial segmentation regions are gradually merged to build a hierarchical segmentation tree (Wu et al., 2019) by using a merging criteria (Hu et al., 2013) combining color and textural features with spatial constraint. The level set of the hierarchical segmentation tree is H, the number of levels is m and the level is recorded as $H_j(j = 1, 2, ..., m)$.



Figure 2. The hierarchical segmentation tree and the corresponding objects.

Each node on the hierarchical segmentation tree corresponds to each segmentation object on the remote sensing imagery. The

nodes of H_1 corresponds to the objects of the initial segmentation result. Each node in the upper layer corresponds to a segmentation object obtained after j - 1 times merging. As shown in Figure 2, (b-1)-(b-3) corresponds to the segmentation image at node i-iii respectively. The position of the segmented object(I) corresponding to the upper node i on the image will fully cover the segmented objects (II and III) corresponding to the lower nodes ii and iii. After m - 1 merging, the initial segmentation objects will be merged into one object covering the whole image, corresponding to the node on the top level of the hierarchical segmentation tree.

2.2 OMRF Model

The OMRF model first initializes the label field X and assigns a random label field $X_i (i = 1, 2, ..., n)$ for each region R_i . $x = \{x_i | i = 1, 2, ..., n\}$ is marked as the realization of X. MRF model uses the MAP estimation to transform the problem of semantic segmentation into the optimal realization of the given observation image (Xia et al., 2006). The optimal result of semantic segmentation is obtained through the estimation of P(x|I) for the optimal realization \hat{x} . The posterior probability is solved by Bayesian formula and the optimal realization (Zheng, Wang, 2015) is simplified as follows:

$$\hat{x} = \operatorname*{arg\,max}_{x} P(x|I)$$

$$= \operatorname*{arg\,max}_{x} \frac{P(I|x) \cdot P(x)}{P(I)} = \operatorname*{arg\,max}_{x} P(I|x) \cdot P(x)$$
(1)

In the Equation 1. the likelihood function P(x|I) shows the conditional probability of the image which belongs to a realization x. The joint distribution P(x) is utilized to model the spatial interactions between objects and possesses the Markovianity property in the MRF model.

$$P(x_i|x_j, \forall j \in V/\{v_i\}) = P(x_i|x_j, \forall j \in N_i)$$
(2)

where N_i is the adjacent objects set of R_i . According to the Hammersley–Clifford theorem (Kollar, 2014), P(x) obeys the Gibbs distribution,

$$P(x) = \frac{1}{Z} \exp(-U(x)) \tag{3}$$

where, Z is the normalizing constant $Z = \sum_{x} exp(-U(x))$, and U(x) denotes the energy function which is equal to the sum of the clique potentials $U(x) = \sum_{c \in C} V_c(x)$ over all possible cliques C. Based on multilevel logistic(MLL) model defines $V_c(x)$:

$$V(x_i, x_j) = \begin{cases} -\beta, if \ x_i = 0\\ \beta, otherwise \end{cases}$$
(4)

 I_i denotes a vector composed by the spectral value of pixels in R_i . $P(I_i|x_i)$ is assumed to obey the Gaussian distribution. μ_l and $\sum_l (l \in \{1, 2, ..., k\})$ denote the mean and variance, respectively.

$$P(I_i|x_i = l) = \prod_{I_i \in I} P(I_i|x_i = l)$$

=
$$\prod_{I_i \in I} (2\pi)^{-n/2} |\Sigma_l|^{-1/2}$$
(5)
$$\times \exp[-\frac{1}{2} (I_i - \mu_i)^T \Sigma_l^{-1} (I_i - \mu_i)]$$

2.3 OMRF-A Model

OMRF-A model (Zheng et al., 2017) defines two auxiliary labels $AX^1 = \{AX_i^1 | 1 \le i \le n\}$ and $AX^2 = \{AX_i^2 | 1 \le i \le n\}$ to participate in probability inference. The optimal realization \hat{x} of the label field X is given by:

$$\begin{aligned} \hat{x}|(AX^{1} = ax^{1}, AX^{2} = ax^{2}) \\ &= \underset{x \in \Omega}{\arg \max} P(X|I, ax^{1}, ax^{2}) \\ &= \underset{x \in \Omega}{\arg \max} \frac{P(X|I, ax^{1}, ax^{2})}{P(I|ax^{1}, ax^{2})} \\ &= \underset{x \in \Omega}{\arg \max} \frac{P(I|X, ax^{1}, ax^{2})P(X|ax^{1}, ax^{2})}{P(I|ax^{1}, ax^{2})} \\ &= \underset{x \in \Omega}{\arg \max} P(I|X, ax^{1}, ax^{2})P(X|ax^{1}, ax^{2}) \\ &= \underset{x \in \Omega}{\arg \min} E(I|X, ax^{1}, ax^{2}) + E(X|ax^{1}, ax^{2}) \end{aligned}$$
(6)

OMRF-A model proposes an approximate method to simplify the solution of MAP, which will realize the iterative local optimization.

$$\hat{x}_{i}|(AX^{1} = ax^{1}, AX^{2} = ax^{2})$$

$$= \underset{x_{i} = l, l \in \Lambda}{\arg\min} \{-\frac{1}{2} \ln(\det |\Sigma_{l}|)$$

$$-\frac{1}{2}(I_{i} - \mu_{l})^{T} \Sigma_{l}^{-1}(I_{i} - \mu_{l})$$

$$+ \beta \cdot [\sum_{j \in N_{i}} V(x_{i} = l, x_{j})$$

$$+ V'(x_{i} = l, ax_{i}^{1}) + V'(x_{i} = l, ax_{i}^{2})]\}$$
(7)

OMRF-A model develops a semi-supervised method to iteratively optimize auxiliary realizations.

$$a\hat{x}_{i}^{u}|(X = \hat{x}, AX^{3-u} = ax^{3-u})$$

$$= \underset{ax_{i}^{m}u=l, l \in \Phi^{u}}{\arg\min} \{-\frac{1}{2}\ln(\det |\Delta_{l}^{u}|)$$

$$-\frac{1}{2}(I_{i} - \Theta_{l}^{u})^{T}(\Delta_{l}^{u})^{-1}(I_{i} - \Theta_{l}^{u})$$

$$+ \frac{k}{k_{u}'}\beta \cdot [\sum_{j \in N_{i}} V(ax_{i}^{u} = l, ax_{j}^{u})$$

$$+ V_{A}'(ax_{i}^{u} = l, ax_{i}^{3-u}) + V_{A}'(ax_{i}^{u} = l, \tilde{x}_{i})]\}$$
(8)

where Θ_l^u and Δ_l^u are the mean and variance of Gaussian distribution for auxiliary label field $ax_l^u (l \in \Theta^u, u = 1, 2)$.

2.4 OMRF-HA Model

OMRF-HA model selects segmentation objects for the label field and auxiliary label fields based on hierarchical segmentation tree structure. The proposed model chooses a layer in hierarchical segmentation tree as the label field, and the upper and lower levels near it as the auxiliary label fields. If we build too many layers in the hierarchical segmentation tree, there will be little difference between the label field and auxiliary label fields, which will weaken the constraint of auxiliary label fields to the label field. On the contrary, if the number of layers is small, there will be great differences between adjacent levels. Once the regions of different categories are merged in the highlevel, large errors will be produced in the corresponding auxiliary label, which will also have a negative impact on the label

This contribution has been peer-reviewed.



Figure 3. The hierarchical segmentation tree(a) and The corresponding segmentation images(b) for label fields.

field. In this paper, we select an appropriate segmentation layer and control the number of segmentation layers to obtain a satisfied label field and auxiliary label fields. As shown in Figure 3, the red lines i, ii and iii in (a) are shown as the selected layers corresponding to the segmentation images (b-1), (b-2) and (b-3), respectively. The black dotted lines are shown as the corresponding relations of the orange objects between these layers. (b-1) and (b-3) are selected as auxiliary label fields and (b-2) are selected as the label field. Our approach uses the same number of classes for three fields. Pixels in auxiliary label field layers and the label field layer get labels from the same label set $\Theta = \{1, 2, \ldots k\}$, (in the Figure 3, k = 3).

For the auxiliary label field corresponding to highest level in three selected layers, each object is randomly assigned at frist, that is, the initial value of the pixel in each object is the same. The other auxiliary label field and the label field are marked as same value with it. The optimal realizations of the label field and the auxiliary label fields are calculated iteratively according to Equation 7 and Equation 8, respectively. Due to the weakening of randomness for MRF by object-based method, in the process of iteration, pixel-based operation is carried out. The semantic segmentation results are finally produced after postprocessing.

3. EXPERIMENT

Sub images captured from a Worldview-3 image were used to evaluate the effectiveness of the proposed approach. The Worldview-3 image was acquired on 5 January 2017, and it covers part of urban areas in Shenzhen City and Hong Kong City, China, with different types of land covers, like residential areas, roads, mangroves and water. It has multispectral band (1.2m/pixel) with eight spectral channels (coastal blue, blue, green, yellow, red, red edge, near-infrared 1 and near-infrared 2). 3339 pixels were labeled identified manually as seven categories (artificial structure(including ships, cars and buildings), shadow, road, lawn, wood, water, mud).

Through experiments for remote sensing images, OMRF-HA model could not only highlight the details on the image, but also remove noises to some extent. We compared the semantic segmentation results of PMRF, OMRF and OMRF-HA models before post-processing with the same bands, iteration times and classification numbers.

As shown in Figure 5, Compared with the classification of PMRF on the shadow of buildings, OMRF and the proposed method could get shadow with more complete shape, and the extraction effects of vegetation and shadow are also better. Compared with OMRF, PMRF and OMRF-HA are more complete on road classification with less noises. In the red box at the lower-left corner, PMRF and OMRF-HA could show the differences between shadows of wood and water, while PMRF produces more noises.

The experimental results show that the average accuracy of the method is 90.44% for the classification of the artificial structure. In Figure 6(a) and Figure 6(b), there is a large homogeneous building area, so that the classification accuracy of the artificial structure is high (98.48%). In Figure 6(c), although the buildings are dense and incoherent, they are arranged regularly and the difference of spectral values is small with the accuracy of classification results as 92.45%. In Figure 6(d), the spectral values of buildings at the sunny side and shade are quite different and ships may carry mud and sand in Figure 6(c), and thus the classification precision of artificial structure are lower which are 89.93% and 72.13%, respectively. The spectral texture values of mud and water are relatively single, so they can be distinguished accurately (97.67% for mud and 94.20% for water).

The results of lawn and wood show well. In the Figure 6(b), the lawn in the middle of the road can also be correctly classified. The average accuracies of these two classes are all more than 90% on test images, and the highest accuracy of wood is 99.33%. However, at the edge of woods, with the decrease of the leaf density, the spectral and texture characteristics decrease as well, so that the misclassification appears.

Compared with other categories, shadow and road detection results show common. In Figure 6(e), there are many obvious characteristics of woods under the shadow which causes many errors. Some misclassifications of road and shadow appear in Figure 6(b). However, it's remarkable that the method can distinguish roads with long and thin shape in Figure 6(d), and also get results of road detection with less noises in in Figure 6(c).

3.1 Conclusion

This paper establishes the Object-based Markov Random Field based on hierarchical segmentation tree with auxiliary labels (OMRF-HA) to realize the semantic segmentation. This algorithm first segments the remote sensing imagery into objects and establishes the object-based hierarchical segmentation tree. Based on the hierarchical tree structure, the objectbased Markov Random Field model with auxiliary label is built. Then the iterative updating process of Markov random fields is implemented. The results are finally obtained, after the postprocessing combining geometric morphology, spectral features and texture features.

We build the label field layer and auxiliary label field layers on the hierarchical segmentation tree which can enhance the positive effect of auxiliary field labels on the label field layer via the constraint relationship of objects on the hierarchical segmentation tree. Considering the influence of object-based segmentation method on the reduction of sample size for Markov Random Field, this paper combines the PMRF with the OMRF semantic segmentation. Through experiments for seven different categories landscapes on remote sensing image, the performances of the method for high-resolution remote sensing image semantic segmentation were analyzed and compared. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLIII-B3-2020, 2020 XXIV ISPRS Congress (2020 edition)



Figure 4. False color composite images with sample points.



Figure 5. Comparison of semantic segmentation results before post-processing. (a) False color composite image. (b) Results of PMRF. (c) Results of OMRF. (d) Results of OMRF-HA.



Figure 6. The results of experiments. (a)-(e) False color composite image. (a-1)-(e-1) The results of the proposed method.

ACKNOWLEDGEMENTS

This work was jointly supported by the National Key R&D Program of China (No. 2017YFC0506200), by the National Natural Science Foundation of China (NSFC) (No. 41501369 and No.41871227)

REFERENCES

Besag, J., 1993. Statistical analysis of dirty pictures**. *Journal of Applied Statistics*, 20(5-6), 63-87.

Blaschke, T., 2010. Object based image analysis for remote sensing. 65(1), 2-16.

Chen, G., Zhang, X., Wang, Q., Fan, D., Gong, Y., Zhu, K., 2018. Symmetrical Dense-Shortcut Deep Fully Convolutional Networks for Semantic Segmentation of Very-High-Resolution Remote Sensing Images. *IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing*, 11(5), 1633-1644.

Chen, K., Jian, P., Zhou, Z., Guo, J., Zhang, D., 2013. Semantic Annotation of High-Resolution Remote Sensing Images

	Precision(a)	Precision(b)	Precision(c)	Precision(d)	Precision(e)	Average
Artificial structure	0.9948	0.9821	0.9245	0.8993	0.7213	0.9044
Shadow	0.8902	0.8818	0.8000	0.9909	0.6262	0.8378
Road	0.7624	0.8538	0.9720	0.7000	Λ	0.8220
Lawn	λ	0.9286	λ	0.9394	0.9000	0.9227
Wood	0.9350	0.9933	0.9762	0.8000	0.8396	0.9088
Water	λ	λ	0.9172	0.9344	0.9744	0.9420
Mud	λ	λ	λ	١	0.9767	0.9767

Table 1. The evaluation results of experiments.

via Gaussian Process Multi-Instance Multilabel Learning. *IEEE Geoscience & Remote Sensing Letters*, 10(6), 1285-1289.

Geman, D., S, G., 1984. Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 6(5-6), 721-741.

Gupta, S., ArbelPerceptual organization and recognition of indoor scenes from rgb-d images. *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on.*

Hermans, A., Floros, G., Leibe, B., 2014. Dense 3d semantic mapping of indoor scenes from rgb-d images. 2014 IEEE International Conference on Robotics and Automation (ICRA).

Hu, Y., Chen, Z., Lin, W., 2018. Rgb-d semantic segmentation: A review. 2018 IEEE International Conference on Multimedia & Expo Workshops (ICMEW).

Hu, Z., Wu, Z., Qian, Z., Qian, F., Xu, J., 2013. A Spatially-Constrained Color–Texture Model for Hierarchical VHR Image Segmentation. *IEEE Geoscience & Remote Sensing Letters*, 10(1), 120-124.

Huang, X., Zhang, L., 2013. An SVM Ensemble Approach Combining Spectral, Structural, and Semantic Features for the Classification of High-Resolution Remotely Sensed Imagery. *IEEE Transactions on Geoscience & Remote Sensing*, 51(1), 257-272.

Kampffmeyer, M., Salberg, A. B., Jenssen, R., 2016. Semantic segmentation of small objects and modeling of uncertainty in urban remote sensing images using deep convolutional neural networks. 2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW).

Kollar, D., 2014. Probabilistic Graphical Models.

Li, D., Zhang, G., Wu, Z., Yi, L., 2010. An Edge Embedded Marker-Based Watershed Algorithm for High Spatial Resolution Remote Sensing Image Segmentation. *IEEE Trans Image Process*, 19(10), 2781-2787.

Long, J., Shelhamer, E., Darrell, T., 2014. Fully Convolutional Networks for Semantic Segmentation. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 39(4), 640-651.

Marfil, R., Bandera, A., 2015. *Hierarchical segmentation of range images inside the combinatorial pyramid.*

Noda, H., Shirazi, M. N., Kawaguchi, E., 2002. MRF-based texture segmentation using wavelet decomposed images. *Pattern Recognition*, 35(4), 771-782.

Paszke, A., Chaurasia, A., Kim, S., Culurciello, E., 2016. ENet: A Deep Neural Network Architecture for Real-Time Semantic Segmentation.

Pedro H. O., P., Collobert, R., 2013. Recurrent Convolutional Neural Networks for Scene Parsing. *IDIAP RESEARCH RE-PORT*.

Thøgersen, M., Escalera, S., Gonzàlez, J., Moeslund, T. B., 2016. Segmentation of RGB-D indoor scenes by stacking random forests and conditional random fields. *Pattern Recognition Letters*, 80, 208-215.

Volpi, M., Ferrari, V., 2015. Semantic segmentation of urban scenes by learning local class interactions. 2015 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW).

Wang, X., Zhang, X., 2009. A new localized superpixel markov random field for image segmentation. 2009 IEEE International Conference on Multimedia and Expo.

Wu, Z., He, L., Hu, Z., Zhang, Y., Wu, G., 2019. Hierarchical Image Segmentation Evaluation using Scale-sets Hierarchy. *IEEE Journal of Selected Topics in Applied Earth Observations* & *Remote Sensing*, 12(8), 2718-2727.

Xia, G. S., He, C., Sun, H., 2006. An unsupervised segmentation method using markov random field on region adjacency graph for sar images. *Radar, 2006. CIE '06. International Conference on.*

Yu, B., Yang, L., Chen, F., 2018. Semantic Segmentation for High Spatial Resolution Remote Sensing Images Based on Convolution Neural Network and Pyramid Pooling Module. *IEEE Journal of Selected Topics in Applied Earth Observations* & Remote Sensing, 11(9), 3252-3261.

Yu, Q., Clausi, D. A., 2008. IRGS: Image Segmentation Using Edge Penalties and Region Growing. 30(12), 2126.

Zhang, X., Xiao, P., Feng, X., 2017. Toward combining thematic information with hierarchical multiscale segmentations using tree Markov random field model. *Isprs Journal of Photogrammetry & Remote Sensing*, 131, 134-146.

Zheng, C., Qin, Q., Liu, G., Hu, Y., 2010. Image segmentation based on multiresolution Markov random field with fuzzy constraint in wavelet domain. *Iet Image Processing*, 6(3), 213-0.

Zheng, C., Wang, L., 2015. Semantic Segmentation of Remote Sensing Imagery Using Object-Based Markov Random Field Model With Regional Penalties. *IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing*, 55(5), 1924-1935. Zheng, C., Wang, L., Chen, X., 2019. A Hybrid Markov Random Field Model With Multi-Granularity Information for Semantic Segmentation of Remote Sensing Imagery. *IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing*, 12(8), 2728-2740.

Zheng, C., Yun, Z., Wang, L., 2017. Semantic Segmentation of Remote Sensing Imagery Using an Object-Based Markov Random Field Model With Auxiliary Label Fields. *IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing*, 55(5), 3015-3028.