









Figure 5. Classification results of Potsdam

### 3.3 Analysis

An accuracy assessment was carried out. A sample-based error matrix is created and used for performing accuracy assessment. In GEOBIA, a sample refers to an object. The error matrix for the test area is shown in Figure 16. The user's accuracy, producer's accuracy, overall accuracy and Kappa coefficient is shown in Table 4.

Table 4. Overall accuracy.

	grassland	road	woodland	building	water	bareland	overall	UA
grassland	30	0	1	0	0	0	31	96.77
road	0	28	0	0	1	0	29	96.55
woodland	0	0	29	0	0	0	29	100
building	0	0	0	30	0	0	30	100
water	0	0	0	0	5	0	5	100
bareland	0	2	0	0	0	8	10	80
overall	30	30	30	30	6	8	134	
PA%	10	93.3	96.6	10	83.3	10		
	0	3	7	0	3	0		

OA=97.01%, Kappa=0.96

The overall accuracy is 97.01%, and the kappa coefficient is 0.96. Our method yields improvements as it depends on decision rule based on machine learning and human knowledge. This is based on the initial decision rules and the validation process, and some obvious classification errors may be corrected already within the following validation step.

### 4. CONCLUSION

This study has put forward a rule set building method for Land cover classification based on human knowledge and machine learning. The use of machine learning is to build rule sets effectively which will overcome the iterative trial-and-error approach. The use of human knowledge is to solve the shortcomings of existing machine learning method on insufficient usage of prior knowledge, and improve the versatility of rule sets. A two-step workflow has been introduced, firstly, an initial rule is built based on Random Forest and CART decision tree. Secondly, the initial rule is analyzed and validated based on human knowledge, where we use statistical confidence interval to determine its threshold. The test site is located in Potsdam City. We utilised the TOP,

DSM and ground truth data. The results show that the method could determine rule set for Land Cover classification semi-automatically, and there are static features for different land cover classes.

Nevertheless, the method is still in the process of development and improvement. Further in-depth studies may be required to (a) improve and refine rule set using human knowledge, (b) investigate the factors influencing classification, such as the spatial scale, the segmentation method employed, and the choice of samples, and (c) to investigate the automation of the method.

### ACKNOWLEDGEMENTS

This research was funded by: (1) the National Natural Science Foundation of China (Project Nos. 41371406, 41471299, and 41671440); (2) Central Public-interest Scientific Institution Basal Research Fund (Project Nos. 7771508, 777161101, and 777161102).

### REFERENCES

- Breiman, L., Cutler, A. Random Forests; Available online: [http://www.stat.berkeley.edu/users/breiman/RandomForests/cc\\_home.htm](http://www.stat.berkeley.edu/users/breiman/RandomForests/cc_home.htm) (accessed on 5 March 2017).
- Breiman, L., 2001. Random Forests. *Mach. Learn.* 45, 5-32.
- Chen X, Fang T, Huo H, Li D R., 2015. Measuring the Effectiveness of Various Features for Thematic Information Extraction From Very High Resolution Remote Sensing Imagery. *IEEE Transactions On Geoscience And Remote Sensing*, 53(9):4837-4851.
- Forestier G, Puissant A, Wemmert C, et al., 2012. Knowledge-based region labeling for remote sensing image interpretation. *Computers, Environment and Urban Systems*, 36(5): 470-480.
- ISPRS Test Project on Urban Classification, 3D Building Reconstruction and Semantic Labeling. <http://www2.isprs.org/commissions/comm3/wg4/2d-sem-label-potsdam.html>. (accessed on 5 March 2017).
- Rau J Y, Jhan J P, Rau R J., 2014. Semiautomatic Object-Oriented Landslide Recognition Scheme From Multisensor Optical Imagery and DEM. *Geoscience & Remote Sensing IEEE Transactions on*, 52(2):1336-1349.
- Verikas, A., Gelzinis, A., Bacauskiene, M. Mining data with random forests: A survey and results of new tests. *Pattern Recognit.* 2011, 44, 330-349.
- Yu L, Fu H H, Wu B, Clinton N, Gong P., 2016. Exploring the potential role of feature selection in global land-cover mapping. *International journal of remote sensing*, 37(23): 5491-5504.
- Ziaei Z, Pradhan B, Mansor S B., 2014. A rule-based parameter aided with object-based classification approach for extraction of building and roads from WorldView-2 images. *Geocarto International*, 29(5):554-569.