Earth Observation and Geomatics Engineering 3(2) (2019) 92–98

EARTH OBSERVATION AND GEOMATICS ENGINEERING
website: https://eoge.ut.ac.ir

A spatial-spectral classification strategy for very high-resolution images using region covariance descriptors and multiple kernel learning algorithms

Saeid Niazmardi

Department of Surveying Engineering, Faculty of Civil and Surveying Engineering, Graduate University of Advanced Technology, Kerman, Iran

Article history:
Received: 16 January 2019, Received in revised form: 2 September 2019, Accepted: 20 September 2019

ABSTRACT

Extracting and modeling the spatial information content of very high resolution (VHR) images can dramatically increase the performances of urban area classification. However, extracting spatial features is a highly challenging task. During the years, several spatial feature extraction methods have been proposed, most of which are mainly designed for grayscale images. To use these methods for a multispectral image, usually, a dimensionality reduction step is required. As a result, these methods cannot optimally extract the spatial information contents of different bands of a multispectral image. To address this issue, we proposed the use of the region covariance descriptor (RCD) for spatial feature extraction from VHR images. The RCD features consider the covariance matrix of a local neighborhood of each pixel as the features. These features can model both the spatial information and the spectral relationship between bands. The RCD features lie in a Riemannian manifold, on which the common classification algorithms cannot be applied. To overcome this, we used Riemannian kernel functions. Also, we proposed a multiple kernel learning strategy for combining RCD and spectral features. The proposed strategy was evaluated for classifying a VHR image acquired over the urban area of Tehran, Iran. Furthermore, its obtained results were compared with those of ten other common spatial feature extraction methods. The results showed that the proposed classification strategy using the RCD features yielded at least 5% higher accuracies than the other feature extraction methods.

KEYWORDS

Covariance descriptors
Region covariance descriptor
Multiple kernel learning
High resolution image
Urban classification

1. Introduction

Generating up-to-date land cover maps of urban areas is of particular importance to different fields of urban management. These maps are usually generated through the classification of very high-resolution (VHR) images acquired over urban areas (Gerl et al., 2014; Shi et al., 2015; Tian et al., 2018). VHR images, due to their high spatial resolution, can provide better discrimination between different urban land-cover classes than any other remote sensing data modality.

Nevertheless, the classification of VHR images is a very challenging task because these images contain information about several undesired small-scale objects that reduce the separability between land cover classes (Demir & Bruzzone, 2016; Tuia et al., 2010). This issue can be overcome by extracting and using the spatial information of VHR images alongside their spectral information for classification. Thus, several spatial feature extraction methods, such as gray-level co-occurrence matrix, shape index, wavelet, and morphological attribute profiles (MAP), have been developed over the years (Audebert et al., 2018; Dalla Mura et al., 2010; Huang et al., 2007; Myint et al., 2004).

* Corresponding author

E-mail addresses: s.niazmardi@kgut.ac.ir

DOI: 10.22059/eoge.2020.285999.1057

Most of these methods, such as GLCM features and MAP, have been proposed for extracting the spatial features from grayscale images. There are two possible strategies for using these methods for multispectral images. The first one is to apply feature extraction methods to each band of multispectral images. This strategy provides a high-dimension redundant feature vector, which can reduce the performance of the classification. The second strategy is to use a dimensionality reduction method, such as principal component analysis, to map the image into a lower-dimension feature space, and then extract the spatial features from the first (or the first few) bands of the data in this feature space (Bernabé et al., 2013; Dalla Mura et al., 2010; Marpu et al., 2012). The main drawback of this strategy is its inability to model the spatial content of all spectral bands of the image. The same spatial feature may have different spectral responses in different bands. This spectral relation between bands, if accurately modeled, can be of great help for feature extraction (Fang et al., 2018).

In this paper, we proposed a novel feature extraction strategy for modeling both the spatial information of the data and the spectral relation between its spectral bands, using the region covariance descriptors (RCD). The RCD calculates the covariance matrix of the pixels in a region and uses it as features. Since this feature is calculated from the pixels in a region (or a local neighborhood), it can capture the spatial information of the images. Besides, since it uses all the spectral bands of the data, it also can capture the spectral characteristics of the image region from which the features are extracted. The RCD has shown very promising performances in different computer vision applications such as face recognition (Pang et al., 2008), action and gesture recognition (Sanin et al., 2013), image segmentation (Carreira et al., 2014), and emotion recognition (Liu et al., 2014). Recently the RCD has also been used for hyperspectral data classification in the remote sensing literature (Fang et al., 2018).

However, the spatial features extracted using the RCD lie in a Riemannian manifold (Rosipal & Krämer, 2005). A Riemannian manifold is a smooth manifold endowed with a metric. As a result, the common classification algorithms cannot be used for their classification, besides they cannot be used by feature fusion method alongside the spectral features for classification purposes, which lie in a Euclidean space. To address this issue, in this study, we propose to construct two different kernels using spectral (DN values of spectral bands) and RCD features and to combine these two kernels using multiple kernel learning algorithms. Using this strategy, the different space of the spectral and spatial feature no longer poses a problem.

The rest of this paper is organized as follows: the second section introduces the RCM for spatial feature extraction and the proposed strategy for the classification. The third section introduces the used data and the design of the experiments. The results are presented and discussed in the fourth section, and finally, we draw our conclusion in the fifth section.

2. Methodology

In this section, we initially provide some theoretical background on the region covariance descriptor and then present the proposed classification strategy.

2.1. Region covariance descriptors

Region Covariance Descriptor (RCD) adopts the covariance matrix, computed from the pixels belonging to a neighborhood of each pixel of an image, as a local descriptor or feature of that image. Since these features are extracted from local neighborhoods of the image, they can extract the spatial structure of the image. By considering the covariance matrix as the region descriptor, the RCD features also capture the statistical relation between different bands of pixels belonging to each region, without any assumption about their distributions (Liu et al., 2014).

Assume that the multispectral image with d spectral bands is denoted by I. For each pixel of this image, we can define a neighborhood R that contains N pixels $\mathbf{x}_i \in \mathbb{R}^d, i = 1,...,N$. Using these pixels, a $d \times d$ covariance matrix $\mathbf{C}_R$ can be calculated using:

$$\mathbf{C}_R = \frac{1}{N-1}\sum_{i=1}^{N}(\mathbf{x}_i - \boldsymbol{\mu})(\mathbf{x}_i - \boldsymbol{\mu})^T \quad (1)$$

Where $\boldsymbol{\mu}$ is the average of the feature vectors in the neighborhood R. The covariance is a symmetric positive definite (SPD) matrix, which lies in a Riemannian manifold space (Arsigny et al., 2007). Since the mathematical modeling in this space is not the same as the Euclidean space, the distance between SPD matrices, as the data points, cannot be estimated using the commonly used metrics such as the Euclidean distance (Li et al., 2013). To address this issue, one should use manifold-based distance metrics (Fang et al., 2018), one of which is called Log-Euclidean distance (LED), proposed in (Arsigny et al., 2007). LED estimates the distance between two covariance matrices such as $\mathbf{C}_1$ and $\mathbf{C}_2$ as:

$$d_{LED}(\mathbf{C}_1, \mathbf{C}_2) = \left\| \log(\mathbf{C}_1) - \log(\mathbf{C}_2) \right\|_F \quad (2)$$

where $\|.\|_F$ denotes the matrix Frobenius norm, and $\log(.)$ is the ordinary matrix logarithm operator. The Frobenius norm of a matrix is equal to the square root of the sum of absolute squares of its elements.

From the geometrical point of view, the logarithmic function maps a point from the manifold space into the tangent space (denoted by T_1) at the point of identity matrix I on the manifold. Then the distance is estimated in this tangent space (Wang et al., 2012).

It can be proven that by calculating the inner product in the tangent space, one can obtain a Riemannian kernel function for two covariance matrices $\mathbf{C}_1$ and $\mathbf{C}_2$ as:

$$\mathbf{K}_{\log}(\mathbf{C}_1,\mathbf{C}_2) = tr\left(\log(\mathbf{C}_1)\cdot\log(\mathbf{C}_2)\right) \tag{3}$$

This kernel is a parameter-free kernel, and usually, a radial base function (RBF) function of this kernel is considered, which is calculated for two covariance matrices $\mathbf{C}_1$ and $\mathbf{C}_2$ using (Vemulapalli et al., 2013):

$$\mathbf{K}_{\log}^{RBF}(\mathbf{C}_1,\mathbf{C}_2) = \exp\left(-\gamma\left\|\log(\mathbf{C}_1)-\log(\mathbf{C}_2)\right\|_F^2\right) \tag{4}$$

where γ is the kernel parameter. All the kernel-based classification algorithms can be used to classify the RCD features using this kernel function.

2.2. Proposed classification strategy

The RCD features can locally capture the spatial information and the relationship between the bands. However, they cannot model the spectral information. In order to use both RCD and spectral features, we proposed a classification strategy based on multiple kernel learning (MKL) algorithms. In this strategy, one kernel is constructed using the spectral features and one using the RCD spatial features. These two kernels are then combined into a composite kernel using an MKL algorithm. The SVM is finally trained using the composite kernel.

The MKL algorithms are a category of kernel-based learning algorithms that use a combination of some precomputed basis kernels instead of a single one for training a kernel-based learning algorithm. Interested readers are referred to (Niazmardi et al., 2018; Niazmardi et al., 2017) for more information about the MKL algorithms.

For the proposed strategy, there are only two basis kernels that should be combined using the MKL algorithm, thus the composite kernel ($\mathbf{K}_C$) can be modeled as:

$$\mathbf{K}_c = \alpha\mathbf{K}_{spectral} + (1-\alpha)\mathbf{K}_{Spatial} \tag{5}$$

where, $\mathbf{K}_{Spatial}$ is the kernel that captures both the spatial information and the relationship between different bands of the data, and it is constructed using Eq. (4) considering the RCD features. $\mathbf{K}_{Spectral}$ is the kernel function constructed considering spectral features. α is the non-negative coefficient that should be estimated using the MKL algorithm. Figure 1 shows the flowchart of the proposed classification strategy.

3. Dataset and experimental setup

3.1. Dataset

The proposed classification strategy was used to classify a VHR image acquired by the GeoEye sensor from the urban area of Tehran in June 2012. The selected site included 900 × 900 pan-sharpened pixels with a spatial resolution of 0.5 m. Seven different land-cover classes characterized this dataset. The available samples for these classes were obtained using photo-interpretation and the available maps of the area. These samples were randomly divided into a set of 740 samples as the training set and a set of 3937 samples as the test set. Figure 2 shows the true color composite of this dataset. The land-cover classes and the available number of training and testing samples are provided in Table 1.

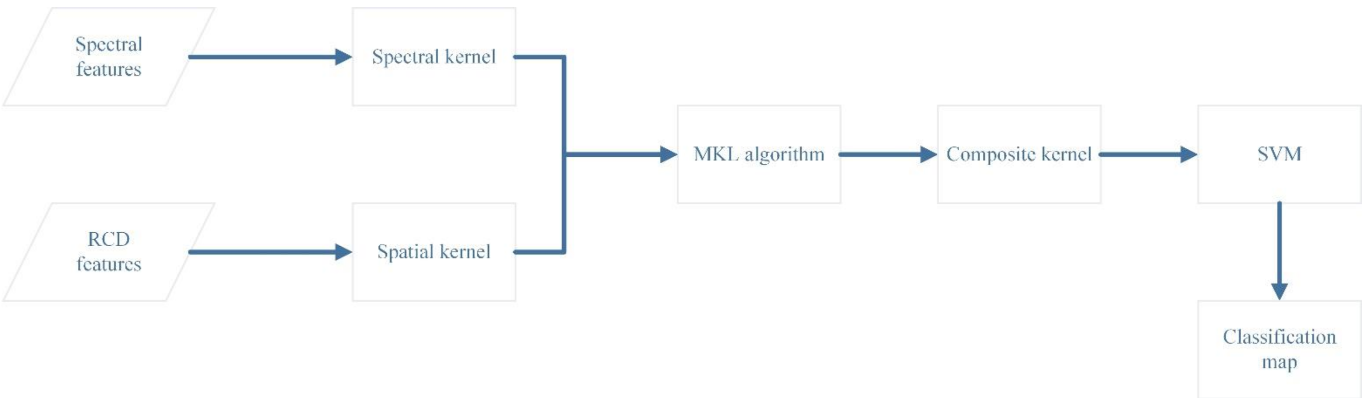

Figure 1. The flowchart of the proposed classification strategy.

Figure 2. The true color composite of the used data dataset.

Table 1. The land-cover classes and the number of samples used for training and testing the algorithm.

No	Land-Cover Class	Number of training samples	Number of testing samples
1	Building with white roof	124	625
2	Building with dark roof	101	636
3	Industrial Building	125	262
4	Asphalt	98	595
5	Soil	119	600
6	Vegetation	91	646
7	Shadow	82	573
	Total	**740**	**3937**

3.2. Experimental setup

In this section, we performed two experiments to evaluate the performance of the proposed classification strategy. The first experiment aimed at evaluating the performance of the proposed classification strategy for urban area classification. In this experiment, we used the Gaussian kernel function for spectral features and the kernel function of Eq. (4) for RCD features. Both these kernels had one free kernel parameter which was selected with 5-fold cross-validation from the set $\{10^{-2}, 10^{-1}, ..., 10^{4}\}$. To study the effects of the size of the neighborhood window on the classification performance we changed it in range $\{11, 13, 15, 17, 19, 21\}$. Although all the MKL algorithms can be exploited to estimate the value of the coefficient (α in Eq. (5)), in this experiment we used a trial-and-error method. We adopted the values between 0.1 and 0.95, with a step size increment of 0.05 as the coefficient. The RCD features were extracted for each pair of window size and coefficient. Also, the performance of each pair was evaluated using the overall accuracy.

In the second experiment, the performance of the RCD features was compared with that of other spatial feature extraction techniques. For this purpose, we adopted extended attribute profiles (EAPs). We considered four different attributes and used 15 threshold values to calculate each of them. The used attributes included area (calculated using the threshold values equally spaced in the range 50 and 1000), the moment of intertie (calculated using the threshold values equally spaced in the range 0.1 and 1), diagonal of the bounding box (calculated using the threshold values equally spaced between 1 and 100), and standard deviation (thresholds are estimated adopting the method proposed in (Marpu et al., 2012)). Besides, we also considered the mean and entropy values extracted from GLCM (considering 9×9 neighborhood) and four different Gabor filters (with the orientations of 0, $\pi/4$, $\pi/2$ and $3\pi/4$) as the spatial features for comparison.

4. Results and discussion

Figure.3 shows the overall accuracies obtained from the classification of the dataset that adopted different values for the size of the neighborhood window and the coefficient. The highest accuracy was 87.20 %, obtained by setting the size of the neighborhood window and the coefficient respectively to 19 and 0.95. It can also be observed that the results are less affected by the size of the neighborhood window than by the coefficient. This is probably due to the considered values as the size of the neighborhood, which is relatively homogenous. Nevertheless, adopting a higher value as the coefficient always resulted in better classification performance.

Table 2 shows the obtained classification performance of the features for comparison. Besides, we also reported the obtained results of the proposed strategy considering the RCD features. In this table, the EAPs that considered area, moment of inertia, standard deviation, and diagonal of the bounding box were respectively denoted as Area, Inertia, STD, and DB.

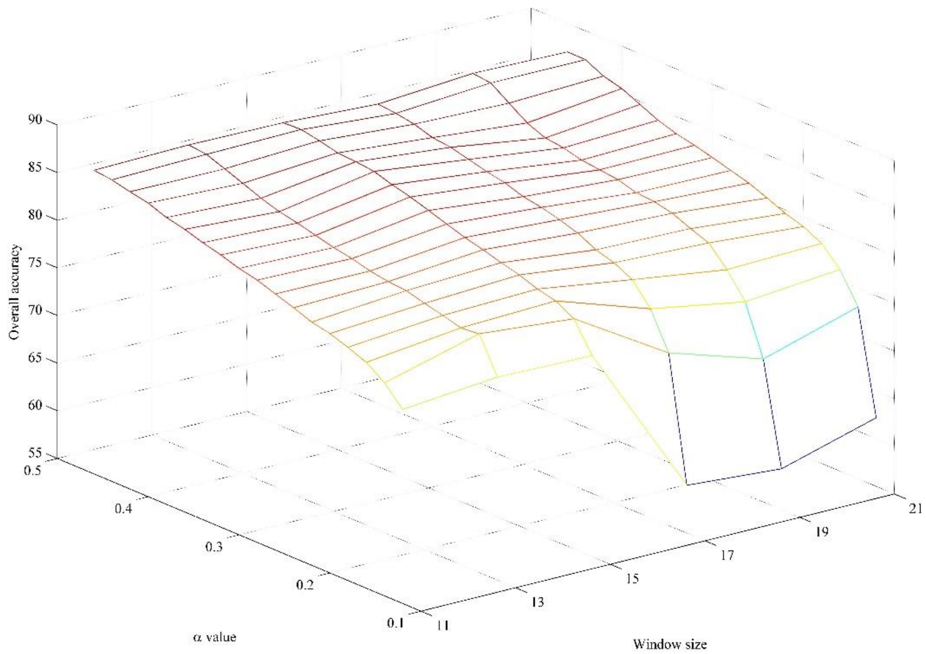

Figure 3. The classification performance of the proposed strategy considering different values for coefficient and the size of the sliding window.

Table 2. The classification performance of different spatial features.

Class No.	*RCD*	*Gabor*				*GLCM*		*EAP*			
		0	$\pi/4$	$\pi/2$	$3\pi/4$	*Mean*	*Entropy*	*Area*	*Inertia*	*STD*	*DB*
1	86.88	82.56	94.88	82.56	90.40	99.04	87.68	95.84	87.52	80.96	87.52
2	86.01	79.56	74.53	79.56	77.83	61.48	20.28	81.60	66.35	59.59	66.35
3	87.02	62.59	80.92	62.59	78.63	59.54	25.95	62.98	81.67	97.71	81.68
4	74.12	72.60	54.79	72.60	66.05	56.64	12.27	32.94	40.84	73.61	40.84
5	99.83	91.50	92.33	91.50	92.17	70.50	48.17	66.67	92.16	100	92.17
6	85.14	72.44	71.52	72.44	80.96	86.99	68.11	71.21	71.82	73.06	71.83
7	91.62	91.97	73.82	91.97	86.74	84.99	37.17	81.85	95.81	91.80	95.81
OA	**87.20**	80.31	77.32	80.31	82.09	75.57	44.70	71.32	76.00	81.00	76.00
Kappa	**0.85**	0.77	0.73	0.76	0.79	71.28	35.06	0.67	0.72	0.77	0.72

As the table shows, the proposed classification strategy yielded the highest overall accuracy of the classification (87.20 %). As for other features, the Gabor filter with an orientation of $3\pi/4$ and EAP with considering the area attribute provided accuracy values of 82.09% and 81%, respectively, which were ranked second and third best spatial features.

The higher performance of the proposed strategy is due to its ability to model the spatial features better. Also, the use of the MKL method makes it possible for a proper combination of both spatial and spectral features.

Regarding the class accuracies, for most classes, the proposed classification strategy showed an acceptable performance. The high performance of this strategy is more evident through comparing the classes such as asphalt and building with dark roofs, for which all the other spatial features showed very poor performances. For a better visual comparison, the classification map obtained from the proposed strategy is shown in Figure 4.

5. Conclusion

In this paper, a new strategy was proposed for the classification of urban areas using VHR images. The proposed strategy takes advantage of the region's covariance descriptor for characterizing the spatial information and multiple kernel learning algorithms to combine spatial and spectral features properly. The proposed strategy has two open parameters, namely the size of the neighborhood window and the coefficient, whose effects on the classification performance were studied in our experiments. The obtained results substantiated the effectiveness of the proposed strategy since it outperformed ten commonly used spatial feature extraction methods in the literature. Besides, the results showed that the size of the neighborhood from which the RCD features were extracted had less effect on the classification performance than the coefficient criterion. Although the proposed strategy showed a very promising performance considering the RBF function of the Riemannian kernel function, the performances of other Riemannian kernel functions should be studied as well. Additionally, the performance of RCD features in extracting spatial information from other remote sensing data modalities should be investigated further.

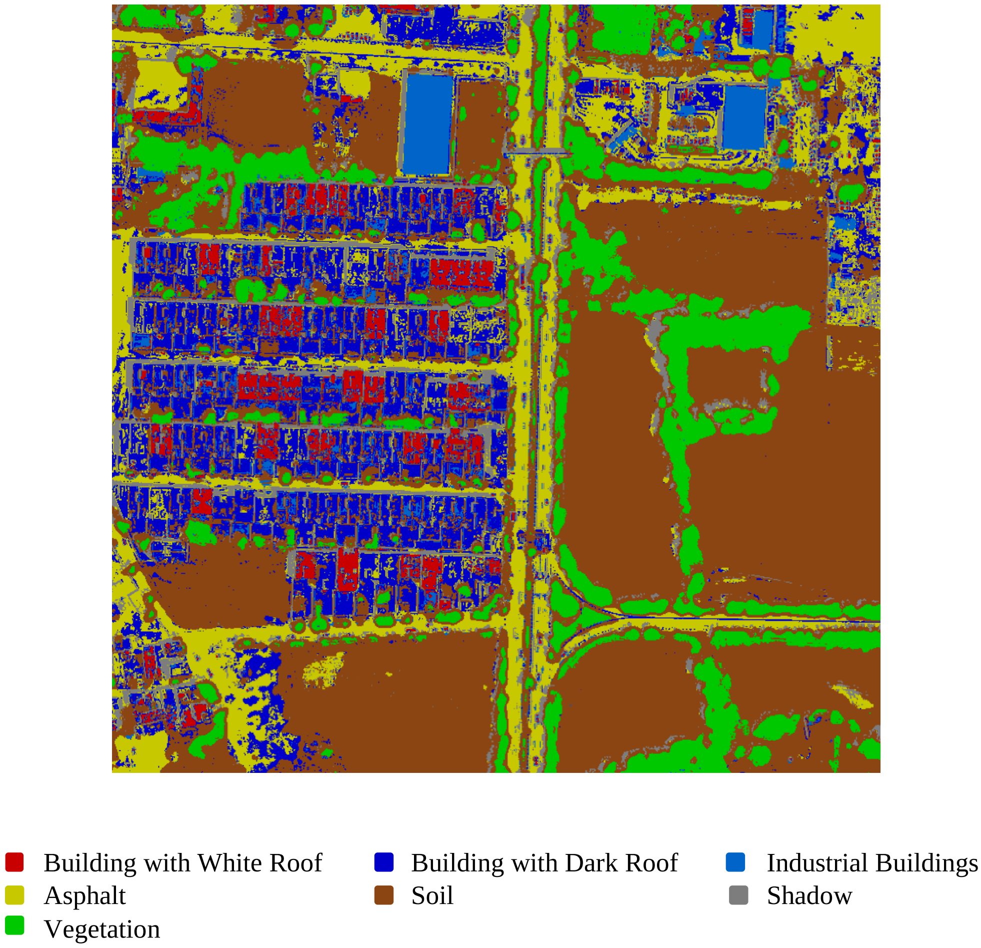

Figure 4. Classification map obtained using the proposed classification strategy by setting the window size and coefficient respectively to 19 and 0.95.

References

Arsigny, V., Fillard, P., Pennec, X., & Ayache, N. (2007). Geometric means in a novel vector space structure on symmetric positive-definite matrices. *SIAM journal on matrix analysis and applications*, *29*(1), 328-347.

Audebert, N., Le Saux, B., & Lefèvre, S. (2018). Beyond RGB: Very high resolution urban remote sensing with multimodal deep networks. *ISPRS Journal of Photogrammetry and Remote Sensing*, *140*, 20-32.

Bernabé, S., Marpu, P. R., Plaza, A., Dalla Mura, M., & Benediktsson, J. A. (2013). Spectral–spatial classification of multispectral images using kernel feature space representation. *IEEE Geoscience and Remote Sensing Letters, 11*(1), 288-292.

Carreira, J., Caseiro, R., Batista, J., & Sminchisescu, C. (2014). Free-form region description with second-order pooling. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 37*(6), 1177-1189.

Dalla Mura, M., Benediktsson, J. A., Waske, B., & Bruzzone, L. (2010). Morphological attribute profiles for the analysis of very high resolution images. *IEEE Transactions on Geoscience and Remote Sensing, 48*(10), 3747-3762.

Demir, B., & Bruzzone, L. (2016). Histogram-Based Attribute Profiles for Classification of Very High Resolution Remote Sensing Images. *IEEE Transactions on Geoscience and Remote Sensing, 54*(4), 2096-2107.

Fang, L., He, N., Li, S., Plaza, A. J., & Plaza, J. (2018). A new spatial–spectral feature extraction method for hyperspectral images using local covariance matrix representation. *IEEE Transactions on Geoscience and Remote Sensing, 56*(6), 3534-3546.

Gerl, T., Bochow, M., & Kreibich, H. (2014). Flood damage modeling on the basis of urban structure mapping using high-resolution remote sensing data. *Water, 6*(8), 2367-2393.

Huang, X., Zhang, L., & Li, P. (2007). Classification and Extraction of Spatial Features in Urban Areas Using High-Resolution Multispectral Imagery. *IEEE Geoscience and Remote Sensing Letters, 4*(2), 260-264. https://doi.org/10.1109/LGRS.2006.890540

Li, P., Wang, Q., Zuo, W., & Zhang, L. (2013). Log-Euclidean kernels for sparse representation and dictionary learning. Proceedings of the IEEE International Conference on Computer Vision,

Liu, M., Wang, R., Li, S., Shan, S., Huang, Z., & Chen, X. (2014). Combining multiple kernel methods on riemannian manifold for emotion recognition in the wild. Proceedings of the 16th International Conference on multimodal interaction,

Marpu, P. R., Pedergnana, M., Dalla Mura, M., Benediktsson, J. A., & Bruzzone, L. (2012). Automatic generation of standard deviation attribute profiles for spectral–spatial classification of remote sensing data. *IEEE Geoscience and Remote Sensing Letters, 10*(2), 293-297.

Myint, S. W., Lam, N. S.-N., & Tyler, J. M. (2004). Wavelets for urban spatial feature discrimination. *Photogrammetric Engineering & Remote Sensing, 70*(7), 803-812.

Niazmardi, S., Demir, B., Bruzzone, L., Safari, A., & Homayouni, S. (2018). Multiple kernel learning for remote sensing image classification. *IEEE Transactions on Geoscience and Remote Sensing, 56*(3), 1425-1443.

Niazmardi, S., Safari, A., & Homayouni, S. (2017). A novel multiple kernel learning framework for multiple feature classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 10*(8), 3734-3743.

Pang, Y., Yuan, Y., & Li, X. (2008). Gabor-based region covariance matrices for face recognition. *IEEE Transactions on Circuits and Systems for Video Technology, 18*(7), 989-993.

Rosipal, R., & Krämer, N. (2005). Overview and recent advances in partial least squares. International Statistical and Optimization Perspectives Workshop" Subspace, Latent Structure and Feature Selection",

Sanin, A., Sanderson, C., Harandi, M. T., & Lovell, B. C. (2013). Spatio-temporal covariance descriptors for action and gesture recognition. 2013 IEEE Workshop on applications of Computer Vision (WACV),

Shi, H., Chen, L., Bi, F.-k., Chen, H., & Yu, Y. (2015). Accurate urban area detection in remote sensing images. *IEEE Geoscience and Remote Sensing Letters, 12*(9), 1948-1952.

Tian, T., Li, C., Xu, J., & Ma, J. (2018). Urban area detection in very high resolution remote sensing images using deep convolutional neural networks. *Sensors, 18*(3), 904.

Tuia, D., Ratle, F., Pozdnoukhov, A., & Camps-Valls, G. (2010). Multisource composite kernels for urban-image classification. *IEEE Geoscience and Remote Sensing Letters, 7*(1), 88-92.

Vemulapalli, R., Pillai, J. K., & Chellappa, R. (2013). Kernel learning for extrinsic classification of manifold features. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,

Wang, R., Guo, H., Davis, L. S., & Dai, Q. (2012). Covariance discriminative learning: A natural and efficient approach to image set classification. 2012 IEEE Conference on Computer Vision and Pattern Recognition,