The Egyptian Journal of Remote Sensing and Space Sciences xxx (xxxx) xxx

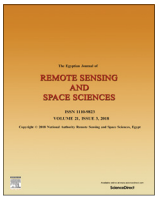
HOSTED BY

ELSEVIER

Contents lists available at ScienceDirect

The Egyptian Journal of Remote Sensing and Space Sciences

journal homepage: www.sciencedirect.com



Review Article

A review on graph-based semi-supervised learning methods for hyperspectral image classification

Shrutika S. Sawant, Manoharan Prabukumar *

SITE, VIT University, Vellore, India

ARTICLE INFO

Article history:
Received 24 May 2018
Accepted 5 November 2018
Available online xxxx

Keywords:
Hyperspectral images
Semi-supervised learning
Image classification
Graph-based learning

ABSTRACT

In this article, a comprehensive review of the state-of-art graph-based learning methods for classification of the hyperspectral images (HSI) is provided, including a spectral information based graph semi-supervised classification and a spectral-spatial information based graph semi-supervised classification. In addition, related techniques are categorized into the following sub-types: (1) Manifold representation based Graph Semi-supervised Learning for HSI Classification (2) Sparse representation based Graph Semi-supervised Learning for HSI Classification. For each technique, methodologies, training and testing samples, various technical difficulties, as well as performances, are discussed. Additionally, future research challenges imposed by the graph-based model are indicated.

Contents

1. Introduction

From the last decade, hyperspectral remotely sensed images are gaining a lot of attention due to their ability to carry a high volume of information about the object or scene. The advancement in imaging technology produces high dimensional images with hundreds of spectral bands (Melgani and Bruzzone, 2004). The rich amount of information helps HSI in performing a detailed analysis of the scene. Therefore, HSI is widely used in various applications, such as target detection (Yang et al., 2015; Zhang et al., 2014), land-cover classification for agricultural field analysis (Damodaran and Nidamanuri, 2014; Pinter et al., 2003), Mineral mapping, etc. Land cover classification is one of the active research topics, which have gained a lot of attention in the last decade (Hou et al., 2013). The traditional classification techniques cannot be applied directly to the HSI due to the following reasons, i) the traditional methods cannot accurately learn class conditional

Peer review under responsibility of National Authority for Remote Sensing and Space Sciences.

* Corresponding author.
E-mail address: mprabukumar@vit.ac.in (M. Prabukumar).

https://doi.org/10.1016/j.ejrs.2018.11.001
1110-9823/

Please cite this article as: S. S. Sawant and M. Prabukumar, A review on graph-based semi-supervised learning methods for hyperspectral image classification, The Egyptian Journal of Remote Sensing and Space Sciences, https://doi.org/10.1016/j.ejrs.2018.11.001

densities ii) A large number of training samples are necessary iii) Computational complexity of high dimensional data is very high (IDRISI Guide). This has fostered the idea of developing novel and precise techniques for HSI classification. HSI is classified as Supervised, Unsupervised, and Semi-supervised based on nature of available training samples. The supervised technique uses ground truth information (labeled data) for classification (Barisione and Solarna, 2016; Manoharan, 2018), whereas the unsupervised technique does not require any prior information (Villa and Chanussot, 2013).

Designing a classification system using HSI poses a few challenges: i) Availability of only a limited number of labeled data. As the dimensions of hyperspectral data are high, classification performance is limited by the availability of the number of samples. This is called the Hughes phenomenon (Richards and Richards, 1999). ii) Collection of labeled data is a highly time-consuming, expensive and a complicated process (Tan et al., 2014). However, unlabelled samples are largely available at no cost. This has motivated researchers to develop the semi-supervised learning (Shahshahani and Landgrebe, 1994) techniques, where unlabelled samples are added to the training data. Several strategies have been suggested by researchers (Di, 2011) in order for a competent development of semi-supervised methods (Sawant and Prabukumar, 2018). In the past decade, semi-supervised classification field has been emerging and enticing more attention due to its capability to solve pattern recognition problems (Zhang and Zhou, 2007; Sugiyama and Id, 2010). As per Fig. 1, semi-supervised learning generally comprise five various models namely, Generative model (Jin and Raich, 2012), Self-learning model (Li, 2013), Co-training model (Zhang, 2014), Transductive SVM (TSVM) learning model (Sun and Wang, 2014), and Graph-based learning model (Aydemir, 2017). A generative model is an old and very simple method but it assumes a certain type of distribution for unlabelled data (Jin and Raich, 2013). Self-learning and Co-training model considers a very few labeled samples, however, suffers from poor prediction (Li, 2013; Samiappan and Robert, 2015). TSVM model is less susceptible to over-fitting compared to self-learning and Co-training model, however, it is classifier dependent (Dalponte et al., 2015), whereas graph-based learning has shown better classification accuracy at the cost of more computational complexity and has become a gradually more active technique in the pattern recognition (Gu and Feng, 2012; Belkin and Niyogi, 2005; Camps Valls et al., 2007; Gomez Chova et al., 2008).

Among the numerous semi-supervised techniques, the graph-based semi-supervised technique is gaining a lot of attention due to its ability to achieve a satisfactory classification performance. These methods help the construction of a graph from both labeled and unlabeled data by utilizing a graph Laplacian regularizer to smooth the classification function with respect to the data manifold. The effect of Laplacian regularizer depends upon how the graph is constructed. The construction of the graph structure involves two-step procedure: determination of graph adjacency relationship and calculation of the graph weights. In this review article, the focus is on the graph-based semi-supervised learning techniques due to its excellence in practice involving the review of the techniques recommended by researchers in the last decade and suggest future research directions.

The remainder of the paper is structured as follows: Section 2 highlights a brief overview of the graph learning models. Section 3 deals with the classification methods that use the spectral information only. Section 4 discusses techniques which consider both spectral and spatial information. Section 5 provides a summary of the literature survey, followed by the indication of future challenges.

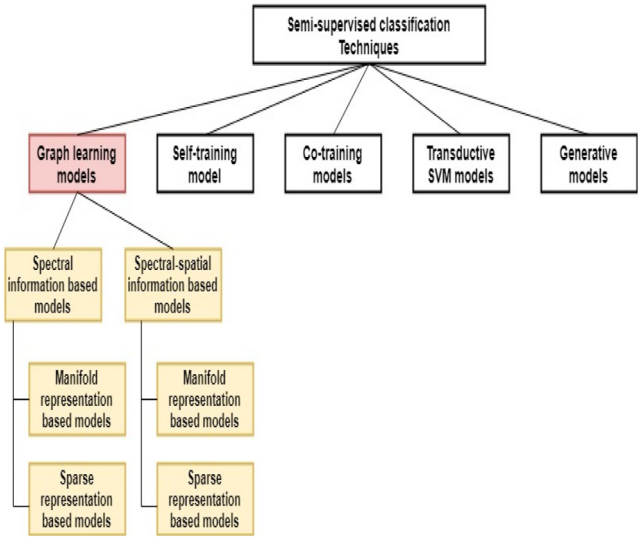


Fig. 1. Hierarchical structure of semi-supervised classification techniques.

2. Graph-based Semi-supervised learning

In the graph-based methods, label information of each sample is propagated to its neighboring sample until a global stable state is reached on the complete dataset. Here, a graph is constructed with nodes and edges, where the nodes are specified by unlabelled and labeled samples, whereas the edges specify the similarities among the labeled as well as unlabelled samples. Here the label of each data sample is progressed to its neighboring points (Sheikhpoura and Sarrama, 2017). The graph structure is represented as $G = (V, E)$, where a set of vertices V which signify both the labeled and unlabeled data samples and a set of edges E which denote the similarities among the labeled as well as unlabelled samples from the dataset (Fig. 1). In Fig. 2(a) two shaded circles are the initially labeled vertices (±1), while the white nodes represent unlabelled samples. The similarity between samples is represented by the thickness of the edges. Fig. 2(b) shows that the graph technique classifies the unlabelled samples according to the weighted distance. The two clusters are formed, even though samples are connected by thin edges.

Consider the HSI dataset which is represented as, $X = [x_1, x_2, \ldots, x_M]$, where $x_i \in F^N$, F is feature vector, M is the total number of pixels in the HSI and N is the total number of spectral bands or the feature dimension. Let, L = {1, ..., l} be the labeled samples corresponding to labels y1, ..., yl and U = {l + 1, ..., l + u} be the unlabelled samples. The process of construction of the graph involves two stages (Cheng et al., 2010): the initial phase involves graph adjacency matrix construction and second phase deals with graph weight calculation. The graph adjacency matrix is constructed either by the

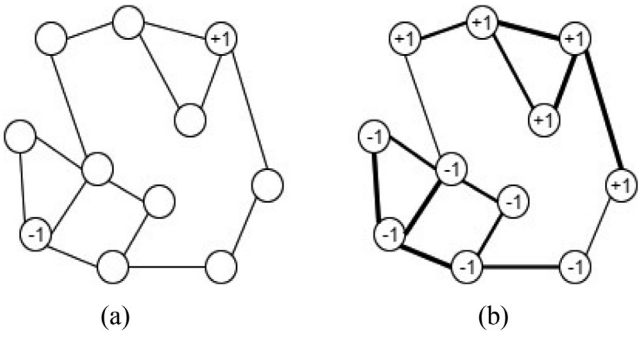


Fig. 2. Graph-based classification of dataset: (a) Before classification (b) After classification.

Please cite this article as: S. S. Sawant and M. Prabukumar, A review on graph-based semi-supervised learning methods for hyperspectral image classification, The Egyptian Journal of Remote Sensing and Space Sciences, https://doi.org/10.1016/j.ejrs.2018.11.001

k-nearest neighbor or by ε nearest neighbor. For calculation of graph weights, one of the following equation is used.

i). The Gaussian similarity function which is represented as:

$$g(x_i, x_j) = \exp\left(-\frac{\| x_i - x_j \|^2}{2\sigma^2}\right) \quad (1)$$

where factor σ controls the width of the neighborhoods.

ii). The inverse Euclidean distance function is given as:

$$g(x_i, x_j) = \| x_i - x_j \|^{-1} \quad (2)$$

From the above steps, x_i and x_j have an associated weight w_{ij}. If samples are not connected then, $w_{ij} = 0$. Finally, the weight matrix W is calculated among all labeled and unlabeled samples in order to perform the classification. The normalized graph Laplacian is defined as,

$$L = I - D^{(-1/2)} W D^{(-1/2)} \quad (3)$$

where, D is a diagonal matrix with degrees d_1, d_2, …., d_N and $d_i = \sum_{j=i}^{n} w_{ij}$.

An important property of Laplacian is given as,

$$F'LF = \frac{1}{2}\sum_{i,j=1}^{n} w_{ij}\left(\frac{f_i}{\sqrt{d_i}} - \frac{f_j}{\sqrt{d_j}}\right)^2 \quad (4)$$

where F is a vector and f_i is its element.

The objective function formulated next is to be minimized in order to classify the data.

Recently the graph based techniques have become the topic of interest for researchers due to the sparseness properties, association with kernel methods, solid mathematical foundation, and better performance. In the next section, we will discuss some of the graph based methods used for HSI classification.

3. Spectral information based graph Semi-supervised learning for HSI classification

HSI provides rich spectral information about the scene as the information is collected in narrow spectral bands. In the HSI classification method, the pixels are assigned to the class in accordance with their spectral signature. So the spectral information plays a vital role in the analysis of HSI. In the past decade, many graph-based semi-supervised learning methods for HSI classification has been evolved, these have considered either manifold or sparse representation of the data structure.

3.1. Manifold representation based graph Semi-supervised learning for HSI classification

The HSI data usually resides on a nonlinear sub manifold causing inefficiency of linear algorithms (Bachmann et al., 2005). Manifold learning-based algorithms have been applied to the HSI for the exploration of the nonlinear structure of HSI data (Lunga et al., 2014; Du and Zhang, 2014). In the past years, a variety of manifold learning based algorithms have been developed enabling the discovery of the nonlinear manifold structure concealed in the high- dimensional data. These include Isomap (Lunga et al., 2014), Locally Linear Embedding (LLE) (Roweis and Saul, 2000), and Laplacian Eigenmaps (LE) (Belkin and Niyogi, 2003).

In semi-supervised learning, the graph constructed from labeled and unlabeled samples is used for manifold representation. Graph-based semi-supervised learning techniques utilize a graph Laplacian regularizer to make the classification function smooth with respect to the data manifold, a low-dimensional subspace on which the high-dimensional data actually reside. So to explore valuable properties of manifold learning, Ma and Crawford (2015) have constructed a graph using the manifold learning method. A study of the semi-supervised learning and manifold learning has been undertaken for finding the relationship between the non-linear data points. The weight matrix is used for determination of the local properties of data points and their neighborhoods. The relationship between semi-supervised learning and local manifold learning has helped the discovery of a graph structure which does effective classification of the structure. The ability of the local manifold learning to preserve local geometry of each neighborhood has helped a tremendous increase in classification accuracy tremendously.

Cheng and Zhu (2016), have proposed the classification of HSI with the active learning method by combining discriminant analysis and robust regression. The robust regression is used for the extraction of abundant information from labeled and unlabeled data. Then, finally, the discriminant analysis is used for improvement of the separability between the classes by exploiting the pairwise constraint term. With these constraints, samples which belong to the same class tend to be close and those from different class go away.

3.2. Sparse representation based graph Semi-supervised learning for HSI classification

In recent times, the benefits of compressed sensing have led to the suggestions of several sparse representations (SR) based methods for graph based semi-supervised learning (Cheng et al., 2009; Yan and Wang, 2009; He et al., 2856). Sparse representation based graphs can help getting the adjacency relationship among the samples and weights simultaneously and automatically. Also, this graph has natural discriminating power which increases performance. Many authors have proposed various versions of the SR methods.

Tan and Zhou (2015), have constructed block sparse graph for HSI classification using discriminant analysis. Unlabelled data samples are picked using sparse representation and regularized collaborative representation. Regularized collaborative representation improves the accuracy of representation. Sparse representation has helped improvement to discriminant capability.

Luo and Huang (2016), have used the manifold learning based on sparse representation for the construction of the graph. The use of sparse coefficient of labeled samples helps construction between the class graph and within the class graph, whereas using the sparse coefficient of unlabelled samples, the unsupervised graph is constructed. All these graphs are used for generation of the projection matrix for feature extraction. Because of the sparse based manifold structure of the data, the discrimination ability of class has increased, and the algorithm has achieved very good accuracy compared to other methods. The computational cost of the algorithm depends on the number of spectral bands as well as the number of labeled and unlabeled samples.

Morsier and Borgeaud (2016), have proposed a graph representation with kernel low rank and sparse subspace clustering for the classification of HSI assuming that data lies on the union of manifolds. This algorithm has been of help in the construction of the graph in an unsupervised manner and performs the classification in an unsupervised as well as semi-supervised fashion. The highly connected graph with a low-rank structure which provides the smoother kernel distortion helps the algorithm outperforming the competing methods that have been trained on the traditional KNN classifier.

Wang et al. (2016), have proposed a non-negative sparse semi-supervised learning method for addressing the ill-posed problem. This method involves two stages: (1) analysis of the discriminant behavior of labeled samples for assessment of the separability

Please cite this article as: S. S. Sawant and M. Prabukumar, A review on graph-based semi-supervised learning methods for hyperspectral image classification, The Egyptian Journal of Remote Sensing and Space Sciences, https://doi.org/10.1016/j.ejrs.2018.11.001

between samples; (2) construction of the non-negative sparse graph based on unlabelled samples by adding regularization term which then extracts the precise information. This block non-negative sparse method effectively utilizes the unlabelled samples which have achieved better accuracy in comparison to the conventional sparse representation.

Even though SR based approaches have presented an impressive performance for HSI classification, some of them have failed to consider the structure of the class underlying dataset. The performance could be improved with consideration of the probabilistic class structure by exploiting partial label information. Yuanjieshao and Nongsang (2017), have developed a sparse representation graph based semi-supervised learning model. Probabilistic class structure estimates the probability relationship between each data sample point and each class of entire data. Probabilistic class structure formulates the sparse structure enhancing the discriminative ability of graph structure. The estimation of class distribution and addition of regularization with class structure increases makes the system slower.

4. Spectral-spatial information based graph Semi-supervised learning for HSI classification

Most of the graph based semi-supervised HSI classification methods use only spectral information neglecting the spatial information. The spectral information alone is not useful for discriminating different classes (Yang et al., 2014). Recently, many articles have referred to the excellent classification performance through exploitation of the spatial neighborhood information along with the spectral information. These methods are also categorized on the basis of manifold or sparse representation.

4.1. Manifold representation based graph Semi-supervised learning for HSI classification

Luo and Liao (2016), have proposed a graph based model where both spatial and spectral information is considered for the improvement of the classification performance. Graph adjacency is calculated by measuring the similarity between two spatial neighborhoods which is termed as the sum of minimum distance (SMD) and local manifold learning (LML) is combined to decide the edge weights. The combination of SMD and LML vectors preserves the local properties of the pixel. Spatial information is used for the identification of the similar neighbors. The combination of SMD and LML preserves the local geometry of both spectral neighborhood and spatial neighborhood. The algorithm outperforms over other methods, but its performance has degraded with few labeled samples.

Gustavo and Marsheva (2007), have presented the graph based composite kernel model for semi-supervised learning. Along with the spectral information, spatial variability of the spectral signature is also exploited. The family of the composite kernel is used for the integration of the contextual information in the classifier. As the method invokes large kernel matrices, Nystrom optimization algorithm (Williams and Seeger, 2001) is used for reducing the computational complexity. The Nystrom method is one of the effective methods for spectral decomposition of the large kernel matrices. The method has shown good accuracy and robustness in an ill-posed situation but the thematic map obtained from this technique is noisy and is sensitive to the parameter alpha.

Gao and Ji (2014), have proposed a novel framework for the performance of HSI classification by using two layer graph-based learning to overcome the challenges of limited data and compound distribution of the classes. In the first layer, the simple graph structure is constructed using unsupervised learning. In the second layer, a hypergraph is constructed using semi-supervised learning, where each vertex represents one pixel and the hyper edge is constructed by using the pair-wise pixel relevance obtained from the first layer. This method outperforms over the state-of-art methods but at the cost of computational complexity.

Martinez-Uso et al. (2014), have proposed the transductive approach for graph-based semi-supervised learning using the probabilistic relaxation theory. This technique uses contextual information for multi-class assignment of labels. Here, the probabilistic relaxation theory is used for assigning a label to the nodes based on the initial contextual information in an iterative manner. The performance is tested in the noise-free channel, so there are chances of performance degradation in presence of significant noise contents.

Wang and Hao (2014), have classified new samples by constructing the spectral-spatial graph which randomly selects unlabelled samples from the spatial information. 2D Gabor filter is used for the extraction of the spatial data which constructs the spatial graph to increase classification accuracy. The authors have also proposed a new adaptive method to find the width of Gaussian function for a spectral-spatial graph. Hence, labels are propagated from labeled to unlabelled samples, on the basis of the spectral-spatial graph. With the incorporation of the spatial information, this method has shown greater accuracy compared to the state of art methods but it takes more computation time for label allocation.

Caol and Chen (2016), Ma and Andong (2016), have proposed a novel deep classification framework by incorporating both the spatial and spectral information. The algorithm works in two steps: graph-based spatial fusion and Convolutional Neural Network (CNN). Initially, the spatial fusion technique is used to extract the spatial spectral features. In the second step, the fused features of the input are used for learning the CNN and via deep hierarchy, with convolutional and pooling layers it understands the spectrum efficiently from the fused input features. Thus, a relationship is formed between the spectral and spatial features and the class distribution which has increased the classification performance.

4.2. Sparse representation based graph Semi-supervised learning for HSI classification

Chen et al. (2017) have presented the double sparse graph discriminant analysis for the dimensionality reduction of HSI in a semi-supervised manner. The positive and negative relationships among the data points are used for forming the double sparse graph, where positive sparse graph reflects intra-class association among the data points while the negative sparse graph indicates inter-class association among the data points. To explore the precise discriminant information, the spatial information along with k nearest neighborhood is utilized to select unlabelled samples which are denoted as pseudo labels.

Xue and Du (2017), have presented a method based on sparse regularization for more accurate classification. This method is an inductive method which can predict the unseen data. Semi-supervised learning is done by adopting the sparse graph regularization (SGR) and the total variation sparse graph regularization (TVSGR) which is an extension of SGR. Here spatial information is considered for the formation of TVSGR. And, finally to get more smooth classification map, graph cut method is integrated with the above referred sparse methods. Though this method is robust to noise and free from tuning of complex parameters, computational complexity is still high.

Aydemir (2017) has developed a novel method by using graph-based learning. Initial samples are selected using subtractive clustering which has provided more informative and quality samples compared to the traditional k means clustering. Decision level

Please cite this article as: S. S. Sawant and M. Prabukumar, A review on graph-based semi-supervised learning methods for hyperspectral image classification, The Egyptian Journal of Remote Sensing and Space Sciences, https://doi.org/10.1016/j.ejrs.2018.11.001

combination of the results obtained by kernel sparse representation and SVM classifier is used for the classification which yields the higher accuracy. Furthermore, this algorithm also uses the spatial information by incorporating the contextual data about the surrounding neighbors of the pixel which has largely affected on classification accuracy.

Fred et al. (2003), have proposed a locality constraint low-rank method by incorporating the contextual information during the feature extraction. This locality constraint low-rank method helps good preservation of the spatial-spectral information. A hypergraph model is constructed for the exploration of the high order relationship among the samples, instead of the conventional graph model. Then the semi-supervised hypergraph learning method is performed on the basis of the low-rank representation model for HSI classification.

5. Discussion and conclusion

Designing an efficient classification system using HSI is very important since HSI reflects a huge data while the classification performance is limited by the availability of the number of labeled samples. In this paper, an overview of some recent articles have been presented to address the challenges and methodologies of the current system. In this section, we summarize this work by outlining the research challenges faced by the graph-based semi-supervised classification system.

5.1. Summary of survey work

The survey work will be summarized as:

- Existing methods assume that the data as single manifold or well separated multiple manifolds. So when a graph is constructed with KNN edges or ε edges, the nearby nodes are strongly connected and have similar labels. However, in high dimensional data, distribution of data forms overlapping of multiple manifolds. Though single manifold follows label smoothness assumption, multiple manifolds fail to satisfy this assumption. Therefore, existing graph-based methods will not achieve the finest performance.
- Most of the algorithms involves the construction of the adjacency graphs using the k nearest neighbors (KNN) method, but the drawback of using this nearest neighborhood criterion is that it does not obtain sufficient discriminant data.
- Most of the algorithms use spectral as well as spatial information which has increased the classification accuracy tremendously at cost of more computational time.
- HSI data contains redundant information of a huge volume which is not useful for the data processing as it results only in additional computational burden. Therefore, band reduction is necessary for HSI classification. Instead of using all available bands of HSI data, a few bands are to be selected with most discriminative information which in turn reduces the computational burden in the training phase.
- For overcoming the above-mentioned drawbacks, some researchers have proposed sparse representation based learning models. Sparse representation has shown good discriminating capability. But the computation of the sparse coefficient takes more time and computational complexity is still remain a challenge.

Therefore, in order to address the above challenges, there is a need to develop a sparse representation based spectral-spatial graph learning algorithm by incorporating an effective optimization algorithm. There is also the need to select the most useful bands for enabling reduction in the computational burden through use of the optimization algorithm. The challenges presented in this section suggests the better efficiency of classification systems for HSI is remaining as a very active area for research.

Conflict of interest

None.

Acknowledgment

The authors thank Vellore Institute of Technology, Vellore for providing a VIT seed grant for carrying out this research work and the Council of Scientific & Industrial Research (CSIR), New Delhi, India for the award of CSIR-SRF.

References

Aydemir, M., 2017. Semi-supervised hyperspectral image classification using small sample sizes. IEEE Geosci. Remote Sens. Lett. 14 (5), 621–625.
Bachmann, C.M., Ainsworth, T.L., Fusina, R.A., 2005. Exploiting manifold geometry in hyperspectral imagery. IEEE Trans. Geosci. Remote Sens. 43 (3), 441–454.
Barisione, F., Solarna, D., 2016. Supervised classification of thermal infrared hyperspectral images through Bayesian, markovian, and region-based approaches. IGARSS IEEE.
Belkin, M., Niyogi, P., 2003. Laplacian eigenmaps for dimensionality reduction and data representation. Neural Comput. 15 (6), 1373–1396.
Belkin, M., Niyogi, P., 2005. Semi-supervised learning on manifolds. Mach. Learn. J. 56, 209–239.
Camps Valls, G., Bandos Marsheva, T., Zhou, D., 2007. Semi-supervised graph-based hyperspectral image classification. IEEE Geosci. Remote Sens. 45 (10), 3044–3054.
Caol, J., Chen, Z., 2016. Graph-based Deep convolutional networks for hyperspectral image classification. IEEE, IGARSS. 978-1-5090-3332-4/16.
Chen, P., Jiao, L., Liu, F., Zhao, J., Zhao, Z., Liu, S., 2017. Semi-supervised double sparse graphs based discriminant analysis for dimensionality reduction. Pattern Recogn. 61, 361–378.
Cheng, H., Liu, Z., Yang, J., 2009. Sparsity induced similarity measure for label propagation. Proceedings of the IEEE 12th International Conference Comput. Vis., Kyoto, Japan, pp. 317–324.
Cheng, B., Yang, J., Yan, S., Fu, Y., Huang, T., 2010. Learning with L1-graph for image analysis. IEEE Trans. Image Process. 19 (4), 858–866.
Cheng, G., Zhu, F., 2016. Semi-supervised Hyperspectral Image Classification via Discriminant Analysis and Robust Regression. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 9 (2), 595–608.
Dalponte, M., Liviu, T., 2015. Semi-supervised SVM for individual tree crown species classification. In: International Society for Photogrammetry and Remote Sensing. Elsevier, pp. 77–87.
Damodaran, B., Nidamanuri, R., 2014. Dynamic linear classifier system for hyperspectral image classification for land cover mapping. IEEE J. Selected Topics Appl. Earth Observ. Remote Sens. 7, 2080–2093.
Di, W., 2011. Active learning via multi-view and local proximity co-regularization for Hyperspectral image classification. Signal Process. IEEE 5, 618–628.
Du, B., Zhang, L., . A discriminative metric learning based anomaly detection method. IEEE Trans. Geosci. Remote Sens. 52 6 (844–6), 857.
Fred, A., Joseph, W., Jonathan, F., 2003. Comparison of airborne hyperspectral data and EO-1 hyperion for mineral mapping. IEEE Trans. Geosci. Remote Sens. 41 (6), 1388–1400.
Gao, Y., Ji, R., 2014. Hyperspectral image classification through bilayer graph-based learning. IEEE Trans. Image Process. 23 (7), 2769–2778.
Gomez Chova, L., Camps Valls, G., Munoz Mari, J., Calpe, J., 2008. Semi-supervised image classification with Laplacian support vector machines. IEEE Geosci. Remote Sens. Lett. 5 (3), 336–340.
Gu, Y., Feng, K., 2012. L1-graph semi-supervised learning for hyperspectral image classification. Proc. IEEE IGARSS, 1401–1404.
Gustavo, C., Marsheva, T., 2007. Semi-supervised graph-based hyperspectral image classification. IEEE Trans. Geosci. Remote Sens. 45 (10), 3044–3054.
He, R., Zheng, W., Hu, B., Kong, X., 2856. Non-negative sparse coding for discriminative semi-supervised learning. In: Proceedings of the 24th IEEE Conference Comput. Vis. Pattern Recognition. Colorado Springs, USA, pp. 2849–2856.
Hou, B., Zhang, X., Ye, Q., Zheng, Y., 2013. A novel method for hyperspectral image classification based on Laplacian Eigen map pixels distribution flow. IEEE J. Selected Topics Appl. Earth Observ. Remote Sens. 6 (3), 1602–1618.
GIS and Image Processing. IDRISI Guide, Volume 1.
Jin, G., Raich, R., 2012. On surrogate supervision Multiview learning. Machine Learning for Signal Processing (MLSP), IEEE International Workshop, pp. 1–6.
Jin, G., Raich, R., Miller, 2013. A generative semi-supervised model for Multiview learning when some views are label-free. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 3302–3306.

Please cite this article as: S. S. Sawant and M. Prabukumar, A review on graph-based semi-supervised learning methods for hyperspectral image classification, The Egyptian Journal of Remote Sensing and Space Sciences, https://doi.org/10.1016/j.ejrs.2018.11.001

Li, Jun, Jose, M., 2013. Semi-supervised hyperspectral image classification using soft sparse multinomial logistic regression. Geosci. Remote Sens., IEEE Lett. 10, 318–322.
Lunga, D., Prasad, S., Crawford, M., Ersoy, O., 2014. Manifold-learning-based feature extraction for classification of hyperspectral data: a review of advances in manifold learning. IEEE Signal Process. Mag. 31 (1), 55–66.
Luo, F., Huang, H., 2016. Semi-supervised sparse manifold discriminative analysis for feature extraction of hyperspectral images. IEEE Trans. Geosci. Remote Sens. 54 (10), 6197–6211.
Luo, R., Liao, W., 2016. Feature extraction of hyperspectral images with semi-supervised graph learning. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 9 (9), 4389–4400.
Ma, L., Andong, L., 2016. Graph-based semi-supervised learning for spectral-spatial Hyperspectral image classification. Elsevier 83, 133–142.
Ma, L., Crawford, M., 2015. Local-manifold-learning-based graph construction for Semi-supervised Hyperspectral image classification. IEEE Trans. Geosci. Remote Sens. 53 (5), 2832–2844.
Prabukumar, M., Sawant, S., 2018. Band clustering using expectation – maximization algorithm and weighted average fusion-based feature extraction for hyperspectral image classification. J. Appl. Remote Sens. 12 (4), 046015.
Martinez-Uso, A., Pla, F., Sotoca, J., 2014. Modelling contextual constraints in probabilistic relaxation for multi-class semi-supervised learning. Knowl.-Based Syst. 66, 82–91.
Melgani, F., Bruzzone, L., 2004. Classification of hyperspectral remote sensing images with support vector machines. IEEE Trans. Geosci. Remote Sens. 42–8, 1778–1790.
Morsier, F., Borgeaud, M., 2016. Kernel low-rank and sparse graph for unsupervised and semi-supervised classification of hyperspectral images. IEEE Trans. Geosci. Remote Sens. 54 (6), 3410–3420.
Pinter, P., Hatfield, J., Schespers, J., Barnes, E., Moran, M., 2003. Remote sensing for crop management. ISPRS J. Photogramm. Remote Sens. 69, 647–664.
Richards, J.A., Richards, J., 1999. Remote Sensing Digital Image Analysis. Springer.
Roweis, S.T., Saul, L.K., 2000. Nonlinear dimensionality reduction by locally linear embedding. Science 290 (5500), 2323–2326.
Samiappan, S., Robert, J., 2015. Semi-supervised co-training and active learning framework for Hyperspectral image classification. IGARSS IEEE.
Sawant, S.S., Prabukumar, M., 2018. "Semi-supervised techniques based hyperspectral image classification: A survey,". 2017 Innovations in Power and Advanced Computing Technologies. i-PACT.
Shahshahani, B.M., Landgrebe, D.A., 1994. The effect of unlabelled samples in reducing the small sample size problem and mitigating the Hughes phenomenon. IEEE Trans. Geosci. Remote Sens. 32 (5), 1087–1095.
Sheikhpoura, R., Sarrama, M., 2017. A Survey on semi-supervised feature selection methods. Pattern Recogn. 64, 141–158.
Sugiyama, M., Id, T., 2010. Semi-supervised local Fisher discriminant analysis for dimensionality reduction. Mach. Learn. 78, 35–61.
Sun, Z., Wang, C., 2014. Semi-supervised classification for Hyperspectral Imagery with transductive multiple-kernel learning. Geosci. Remote Sens., IEEE Lett. 11, 1991–1995.
Tan, K., Zhou, S., 2015. Semi-supervised discriminant analysis for hyperspectral imagery with block-sparse graph. IEEE Geosci. Remote Sens. Lett. 12 (8), 1765–1769.
Tan, K., Li, E., Du, Q., Du, P., 2014. An efficient semi-supervised classification approach for hyperspectral imagery. ISPRS J. Photogramm. Remote Sens. 97 (11), 36–45.
Villa, A., Chanussot, J., 2013. Unsupervised methods for the classification of hyperspectral images with low spatial resolution. Pattern Recogn. 46, 1556–1568.
Wang, X., Gao, Y., Cheng, Y., 2016. A non-negative sparse semi-supervised dimensionality reduction algorithm for hyperspectral data. Neurocomputing 188, 275–283.
Wang, L., Hao, S., 2014. Semi-supervised classification for hyperspectral imagery based on spatial-spectral Label Propagation. Elsevier 97, 123–137.
Williams, Christopher, Seeger, Matthias, 2001. Using the Nystrom method to speed up kernel machines. Proc. NIPS.
Xue, Z., Du, P., 2017. Sparse graph regularization for hyperspectral remote sensing image Classification. IEEE Trans. Geosci. Remote Sens. 1–16.
Yan, S., Wang, H., 2009. Semi-supervised learning by sparse representation. Proceedings of the SIAM International Conference Data Mining, SIAM, Sparks, NV, USA, pp. 792–801.
Yang, S., Shi, Z., Tang, W., 2015. Robust hyperspectral image target detection using an inequality constraint. IEEE Trans. Geosci. Remote Sens. 53, 3389–3404.
Yang, L., Yang, S., Jin, P., Zhang, R., 2014. Semi-supervised hyperspectral image classification using a spatio-spectral Laplacian support vector machine. IEEE Geosci. Remote Sens. Lett. 11 (3), 651–655.
Yuanjieshao, A., Nongsang, 2017. Probabilistic class structure regularized sparse representation graph for semi-supervised Hyperspectral image classification. Pattern Recogn. 63, 102–114.
Zhang, X., 2014. Modified co-training with spectral and spatial views for semi-supervised hyperspectral image classification. Appl. Earth Observ. Remote Sens. IEEE 7, 2044–2055.
Zhang, L., Tao, D., Huang, X., Du, B., 2014. Hyperspectral remote sensing image subpixel target detection based on supervised metric learning. IEEE Trans. Geosci. Remote Sens. 52, 4955–4965.
Zhang, D.Q., Zhou, Z.H., 2007. Semi-supervised dimensionality reduction. Proceeding 7th SIAM International Conference Data Mining, pp. 387–393.

Please cite this article as: S. S. Sawant and M. Prabukumar, A review on graph-based semi-supervised learning methods for hyperspectral image classification, The Egyptian Journal of Remote Sensing and Space Sciences, https://doi.org/10.1016/j.ejrs.2018.11.001