ASIAN JOURNAL OF INFORMATION TECHNOLOGY
Journal Home Page: https://medwelljournals.com/journalhome.php?jid=1682-3915

Active Learning in Classification of Hyperspectral Imaging: A Review

[1]R. Elakkiya, [1]K. Thilagavathi and [2]A.Vasuki
[1]*Department of Electronics and Communication Engineering,*
[2]*Department of Mechatronics Engineering, Kumaraguru College of Technology, Coimbatore, India*

Key words: Hyperspectral image, active learning, classification, development, implemented, techniques

Corresponding Author:
R. Elakkiya
Department of Electronics and Communication Engineering, Kumaraguru College of Technology, Coimbatore, India

Page No: 173-179
Volume: 18, Issue 6, 2019
ISSN: 1682-3915
Asian Journal of Information Technology

Abstract: Hyperspectral images are used to characterize the objects with unprecedented accuracy of the data. The active learning aims at providing efficient training set by iterating the samples. This study reviews the concepts involved in active learning algorithm for classification of remote sensing image or hyperspectral image. The diversified vision of hyperspectral sensors was awakened with the latest development of remote sensing and geographical information. Imaging spectroscopy which is commonly known as hyperspectral remote sensing was recently inspected by researchers and scientists for exploring vegetations, minerals, etc. This hyperspectral imaging requires large data sets and new processing techniques. Several active learning algorithms are implemented in hyperspectral images for better classification and greater accuracy.

INTRODUCTION

Multispectral imaging was used in the classification of remote sensing images and satellite systems widely until 1960. It contains approximately 30 bands. Super spectral imagery is next to multispectral imagery as it contains more number of spectral bands. After the multispectral imagery came to an end hyperspectral imagery was used widely than the multispectral imagery as it had several advantages. It consists of hundreds of spectral bands. Hyperspectral remote sensing is processing of earth materials by means of continuous spectral bands. The hyperspectral imagery cube is a three dimensional cube consisting of two spatial dimensions and one spectral dimension. The hyperspectral imaging is the mapping of spectral signatures in images to specify land cover types. The information in hyperspectral images allows characterization, identification, distinction of sub pixels and also to classify images with high accuracy and robustness. The problem of curse in dimensionality reduction exists with less number of labeled samples and with high number of spectral signatures. Hyperspectral image classifiers have the ability to produce accurate land-cover maps with greater number of features, low-sized training datasets and high levels of spatial variability of the spectral signature. Support Vector Machine (SVM) classifier provides good generalization capabilities, especially in high dimensional space. Hyperspectral remote sensing systems provide additional discriminative features for classes that are spectrally similar, due to their high spectral resolution. By applying statistical learning model to spatial (Bruzzone and Carlin, 2006; Tuia *et al*., 2009) and spectral (Melgani and Bruzzone, 2004; Camps-Valls *et al*., 2004) resolution, the efficiency of remote sensing data can be improved.

Hyperspectral imaging: Hyperspectral imaging deals with gathering and processing the data across the electromagnetic spectrum. The hyperspectral Imagery cube is shown in Fig. 1. The spectral imaging divides the spectrum into number of bands. Although, the hyperspectral images contain more information than the regular RGB images, above ninety percent variance can be explained by a small portion of data. A lot of methods have been proposed to deal with hyperspectral data classification. The hyperspectral image classification is the processing step in remote sensing applications (Plaza *et al*., 2009; Cheng *et al*., 2015). The logistic

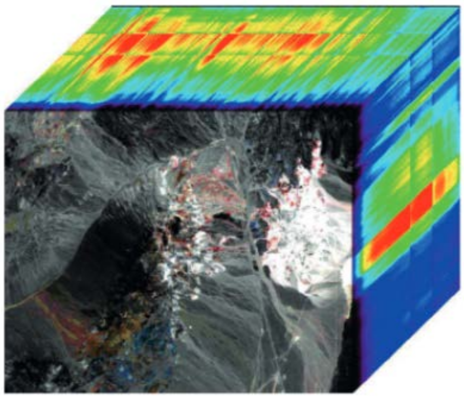

Fig. 1: Hyperspectral imagery cube from cuprite mining district scene 4, Nevada (Arslan *et al*., 2017)

regression, maximum likelihood and k nearest-neighbors are the classification methods and algorithms used in hyperspectral data classification. The majority of the above mentioned algorithms suffer from the "curse of dimensionality". Hughes phenomenon arises mainly due to increase in dimension of data for accuracy in classification and limited training samples (Chapelle *et al*., 2006). The goal of dimensionality reduction is to map high dimensional data into a low dimension while preserving the main features of the original data. Few classification methods were developed in order to reduce the dimension leading to less training samples of hyperspectral data.

SVM is an assuring classification method introduced for hyperspectral data classification. SVM exhibits low sensitivity to high dimensionality and is unlikely to suffer from the Hughes phenomenon. In most cases, SVM-based classifiers obtain better classification accuracy than the other widely used pattern recognition techniques. For a longer time, these classifiers were the state-of-the art methods. In recent years, the spatial information is growing vital for hyperspectral data classification. Spatial-spectral classification methods provide significant advantages in terms of performance parameters. To deal with spatial variability of spectral signature, some recent approaches try to incorporate spatial information based on the fusion of morphological information and original data followed by SVM as it provides good classification results. A new classification framework is proposed to exploit the spatial and spectral information using loopy belief propagation and active learning. In recent years, sparse representation-based methods have been widely used in many fields. The spatial-spectral kernel sparse representation is proposed to deal with hyperspectral data classification. SVM can generalize even with the limited number of samples (Mountrakis *et al*., 2011). SVMs and KFD (Kernel Fisher Discriminant analysis) approaches can be integrated as kernel methods framework. Kernel-based methods are based on mapping data from the original input feature space to a kernel feature space of higher dimensionality and then solving a linear problem in that space. These methods combine statistics and geometry in an effective way to interpret the learning algorithms (Camps-Valls and Bruzzone, 2005).

Active learning: The aim of Active Learning (AL) is to select informative samples from unlabeled data and to transfer the knowledge from the labeled data smoothly to the unlabeled data. It also minimizes the cost of hand labeling new samples. It can be explored for three central areas namely classification, parameter estimation and causal discovery. SVM classifiers have met with significant success in numerous real time classification tasks. However, they are typically used with a randomly selected training set. AL algorithm can be applied in the areas of text categorization, image retrieval and also to reduce the need for training data.

The aim of AL is to rank the learning set based on interest criterion or a heuristic which selects the most useful pixel to improve the model and also minimizes the number of training samples necessary to maintain discrimination capabilities (Pasolli *et al*., 2011). AL can effectively reduce labeling effort for remote sensing image classification. The primary idea of AL based methods is iteratively labeling in the unlabeled data, acquiring the true labels from the oracle (domain expert) and updating the current system according to the new data. In this case, the selection of informative samples in the unlabeled data and the construction of learning machine must be taken into consideration. AL can be applied to many realistic applications. For e.g., consider a new user in a movie recommender system with scarce preference information to whom recommendations can be improved by selecting several movies from rating (Houlsby *et al*., 2014). Another example in the field of medical imaging is collecting and selecting labels depending on the cost (Hoi *et al*., 2006). In AL for hyperspectral classification, unlabeled samples in the pool are usually evaluated and chosen in a greedy manner according to certain informative measures. Uncertainty sampling and query-by-committee are the techniques used to measure informativeness. In uncertainty sampling, a small set of labeled samples are used to train an initial system. The measurement of uncertainty can be measured by entropy or confidence score.

Query-by-committee also starts with labeled data. Multiple distinct models are trained and requested to vote on labels of the unlabeled samples. Despite pre-existing extensive studies of active learning, the works on hyperspectral image classification began only in recent years. Recent studies show that the SVM-based AL methods are more preferred as its generalization capability under small training set and complex high dimensional data conditions are feasible. Multi Class-

Level Uncertainty (MCLU) method was introduced wherein the uncertainty is measured by the decisions from support vector machines in one-against all multi class architecture. Samples that are closer to the separating hyperplane are considered to be more uncertain (Settles, 2010). The key components in active learning are an uncertainty measure, an information density measure and an adaptive combination framework. The uncertainty sampling is based on uncertainty of information that selects the uncertain samples to label the data (He *et al*., 2014). To overcome the drawbacks of uncertainty sampling information, density measure can be performed as it helps to shrink the training sets (Li and Guo, 2013). The two dimension criterion considers sample and label dimensions where it selects the most informative labels to reduce the uncertainty (Qi *et al*., 2008).

Active learning based classification: An integrated algorithm which suited well to the problems with very few training samples was proposed in Indian Pines and Pavia data sets with the help of Loopy Belief Propagation (LBP) based method which estimates the conditional marginals for the classification of spectral-spatial hyperspectral images and to collect the marginals by using the AL algorithm efficiently to exploit the spectral and spatial information from the hyperspectral data (Li *et al*., 2012).

One spatial and two local spectral embedding methods in conjunction with the SVM classifier, was implemented along with a Radial Basis Function (RBF) kernel which showed excellent performance on AVIRIS and hyperion hyperspectral data when compared to random sampling. In spectral domain, the manifold learning is applied to search intrinsic low-dimensional structures of hyperspectral data (Di and Crawford, 2011a, b). Manifold learning provides higher classification accuracies and improved representation relative to linear dimensionality reduction methods (Bachmann *et al*., 2005; Kim and Crawford, 2010).

The Simple Linear Iterative Clustering (generally abbreviated as SLIC) method was initially introduced to generate superpixels. Then, the proposed active learning algorithm is presented in the context of superpixels to guarantee the diversity of selected informative samples from unlabeled data pools. Superpixel-based active learning is abbreviated to SPAL. Superpixels of the image have advantages in adherence to image boundaries, which is usually a preprocessed step for image segmentation. Together with speed and memory efficiency, this simple linear iterative clustering algorithm is a good choice to generate superpixels for image segmentation. Accordingly, the SLIC method is adopted to generate superpixels which can represent better images compared to other superpixel methods. The SLIC algorithm is similar to k-means clustering algorithm. The process comprises of assigning and updating steps in an iterative manner. In the assignment step, each pixel should be assigned to its nearest cluster based on a similarity measure. Then the clusters are updated in terms of their member pixels in the updating step. The process is repeated until it reaches convergence. The key issue is to define a proper similarity measure for the algorithm ("Superpixel-based active learning for the classification of hyperspectral images School of Computer Science, Shanghai Key Laboratory of Data Science Key Laboratory for Information Science of Electromagnetic Waves (MoE) State Key Laboratory of Satellite," n.d.). Multiclass classification problems can be sorted using one against-all architecture based on support vector machine classifier (Patra and Bruzzone, 2011). Using 'informative' data points gives better learning rates rather than updating the classifier with randomly chosen data points from the new area. MacKay proposed an active learning algorithm which increases the information gain from a user-defined variable on adding the new data point. This technique is exactly turned out to be well utilizing random points, batch semi-supervised methods and entropy-based active learning method. The area of investigation can be increased when more hyperspectral data are readily available without any restriction (Rajan *et al*., 2006).

The three main steps involved in semi supervised classification using active learning are; semisupervised learning, spectral unmixing and active learning. After the k-means clustering operation, the resulting end member Mixture-Tuned Matched Filtering (MTMF) based unmixing is done. An unsupervised clustering method k-means is used on the available labeled samples to solve the problems obtained by end member extraction algorithms. It is proved that MTMF can outperform other techniques. The classification probabilities and spectral unmixing abundance are combined in the active learning step to learn the most informative unlabeled samples. On the other hand in order to select the most informative unlabeled samples, two kinds of criteria are considered. From the classification probabilities viewpoint, it means to consider both the boundary regions between two classes and the regions with large number of boundaries. The spectral unmixing consists of two main endmembers. It finds the misclassified samples easily. The semi supervised classification using active learning is shown in Fig. 2 (Sun *et al*., 2016).

The HSeg algorithm is a segmentation hierarchy that contains different level of details of the image. It is necessary to select a single optimum segmentation level from the hierarchy. It is based on three key steps; Pixelwise classification, region classification which is obtained by assigning every spatially connected region from the segmentation result to the most frequently occurring class within the region. The segmentation result at every hierarchical level is evaluated and the level that gives the highest classification accuracy is selected. The

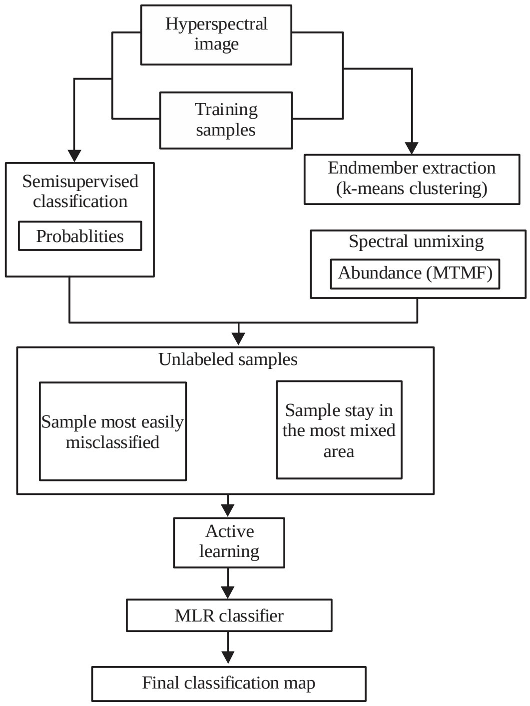

Fig. 2: Semisupervised classification with active learning (Sun *et al*., 2016)

steps involved in spectral-spatial classification of hyperspectral imagery are as follows; identification of the subspaces (views), back-end classifier, AL query criterion, decision fusion strategy. To improve the efficiency of every view and promote the diversity across different views, spatial features are incorporated based on three steps.

Run HSeg separately for each view, extract the spatial features from HSeg derived segmentation for each view independently, in each view, the original spectral features and the derived spatial features are connected in the form of stacked vector. The ensemble AL approach yields better classification performance compared to single view AL approaches on a widely used hyperspectral imagery dataset (Zhang and Crawford, 2015). The spatial and spectral features are connected together as a stacked vector is a segmentation approach and the training set is extended using Semi Supervised Learning (SSL) (Zhang and Chen, 2002). Figure 3 represents the hyperspectral classification involved in supervised classification (Qiu *et al*., 2017).

In MLR via. variable splitting and augmented Lagrangian (LORSAL) algorithm, the classification stage assigns a specific label to each pixel at the final stage. In this stage, the MLR classifier is used to obtain image of the class labels. To ensure the effectiveness of the feature-driven AL, two assessment indices to guide the feature selection was introduced. The first one

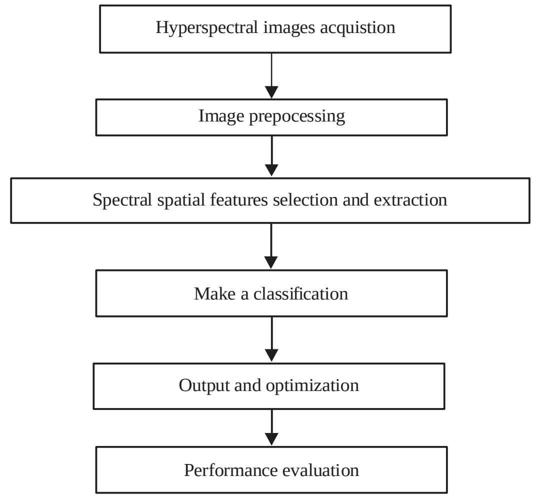

Fig. 3: Hyperspectral image classification (Qiu *et al*., 2017)

is the overall error probability and the other is the Fisher ratio. These two indices can roughly assess the discriminativeness of the constructed feature space. Therefore, the useless features can be removed ahead of looping. The representative and discriminative information are gradually improved by active learning. Meanwhile at each iteration, the samples are treated as candidate pool for active learning (Liu *et al*., 2016, 2017).

Binary feedback active learning involves two steps; a query image is selected from the pool; a sample image is selected from a known category and used with the query image (Joshi *et al*., 2013). A search algorithm and a criterion function are the feature selection techniques where solutions for different feature selection problems can be generated using search algorithm (Serpico and Bruzzone, 2001). In the supervised classification only labeled data is used to train the classifier. The unsupervised method groups pixel with similar spectral characteristics (means, standard deviations, etc.,) into unique clusters based on certain criteria. While in semi-supervised classification both labeled and unlabeled data can be used to train the classifier (Sabale and Jadhav, 2015). In case of remote sensing image classification, active learning can be effectively used in the spatial domain (Stumpf *et al*., 2013). The kernel-based method was performed initially to reduce the curse of dimensionality following a semi-supervised approach, which exploits the wealth of unlabeled samples in the image and naturally gives relative importance to the labeled ones through a graph-based methodology.

Finally, it incorporates contextual information through a full family of composite kernels. As the graph method relies on inverting a huge kernel matrix formed by both labeled and unlabeled samples, nystrom method was introduced in the formulation to speed up the classification process (Camps-Valls *et al*., 2007). Two algorithms the Extended Morphological Profile-Kernel Principal Component Analysis (EMP-KPCA) and the Multiple Spectral-Spatial Classifier-Minimum Spanning Forest (MSSC-MSF) provide better performance in classification accuracies (Fauvel *et al*., 2012). An active learning algorithm based on weighted incremental dictionary learning is proposed for various applications. The AL algorithm selects training samples with two selection criteria, namely representative and uncertainty. This algorithm trains a deep network efficiently by selecting the training samples iteratively (Liu *et al*., 2017). Initially, the SVM is performed with a limited number of training samples and then a 3×3 window of majority voting algorithm is introduced. Secondly, a testing step is implemented in order to calculate the classification error by subtracting the classification rate of current iteration with the previous iteration. If the classification error is less than predefined threshold, then the data is good for mapping (thematic mapping), else remove the non-correct pixels by comparing each pixel with the markers. The corrected pixels are taken as training pixels for the next iteration. This algorithm solve problems with very few training samples available and performs well than the ISVM and the other existing approaches over the considered analysis scenario (AVIRIS Indian Pines) (Baassou *et al*., 2013). Using spatial-spectral label propagation the Semi-supervised classification can be done for hyperspectral images (Wang *et al*., 2014).

Discovering Representativeness and Discriminativeness by Semi Supervised Active Learning (DRDbSSAL) algorithm with active learning settings, supervised clustering approach and classifier are adopted to assign pseudolabels for the unlabeled data to enhance the performance of active learning. The representative and discriminative information are gradually improved by active learning. The proposed method provides a novel way to exploit the representative and discriminative information from the unlabeled data without a tradeoff parameter. This follows clustering approach DRDbSSAL that mines both representative and discriminative information by assigning pseudolabels to the unlabeled data with a supervised clustering technique and classifiers, thereby improving the final classification results with the labeled data and the unlabeled data with the pseudolabels (Wang *et al*., 2017).

The goal of AL is to obtain satisfactory classification performance with fewer labeled samples comparative to conventional passive learning wherein the training set is often selected randomly or manually without interaction with the classifier. It can be roughly categorized as approaches involving uncertainty sampling, query by committee, margin sampling and expected model change (Di and Crawford, 2011a, b). Spatial preprocessing can be done for classifying the hyperspectral images using MLR (Nidhin Prabhakar *et al*., 2015). Collaborative Active and Semi-Supervised Learning (CASSL) involves an effective pseudolabel verification procedure that provides collaborative labeling method along with human experts. The classifiers improve classification performance by acquiring the label pixels carefully (Wan *et al*., 2014). AL can be performed with stacked autoencoders for classifying hyperspectral images based on neural network. The differences between active learning and semi supervised learning that addresses the classification problems by few training samples was proposed by Progressive Semi Supervised Support Vector Machine-Diversity (PS^3VM-D) technique which allows iterative SSL approach to find the distribution of classes and to reach convergence in lesser iterations with respect to the standard algorithm. It can be concluded that the AL techniques are effective and can be readily used in operational applications. SSL techniques yet require additional works to be adopted to relate their convergence properties and also to further investigate the validation procedures. AL and SSL paradigms can be effectively combined to define learning algorithms that exploit both labeled and semi labeled samples in training phase and for selection of new samples to be labeled by the user (Persello and Bruzzone, 2014).

CONCLUSION

This study represents a rich and healthy research community developing new heuristics for active learning in remote sensing and signal processing. Active learning has strong potential for remote sensing data processing. The efficient number of training samples is required for processing large number of digital images. The combination of AL with other algorithms can readily enhance the image classification accuracy. New problems are being tackled with active learning algorithms, guaranteeing the efficient processing of hyperspectral data. It includes the algorithms that are used to obtain better identification and also can improve the accuracy of the data.

REFERENCES

Arslan, O., O. Akyurek and S. Kaya, 2017. A comparative analysis of classification methods for hyperspectral images generated with conventional dimension reduction methods. Turk. J. Electr. Eng. Comput. Sci., 25: 58-72.

Baassou, B., H. Mingyi, M.I. Farid and M. Shaohui, 2013. Hyperspectral image classification based on iterative support vector machine by integrating spatial-spectral information. Proceedings of the 2013 IEEE International Symposium on Geoscience and Remote Sensing (IGARSS'13), July 21-26, 2013, IEEE, Melbourne, Australia, pp: 1023-1026.

Bachmann, C.M., T.L. Ainsworth and R.A. Fusina, 2005. Exploiting manifold geometry in hyperspectral imagery. IEEE Trans. Geosci. Remote Sens., 43: 441-454.

Bruzzone, L. and L. Carlin, 2006. A multilevel context-based system for classification of very high spatial resolution images. IEEE. Trans. Geosci. Remote Sens., 44: 2587-2600.

Camps-Valls, G. and L. Bruzzone, 2005. Kernel-based methods for hyperspectral image classification. IEEE. Trans. Geosci. Remote Sens., 43: 1351-1362.

Camps-Valls, G., L. Gomez-Chova, J. Calpe-Maravilla, J.D. Martin-Guerrero, E. Soria-Olivas, L. Alonso-Chorda and J. Moreno, 2004. Robust support vector method for hyperspectral data classification and knowledge discovery. IEEE. Trans. Geosci. Remote Sens., 42: 1530-1542.

Camps-Valls, G., T.V.B. Marsheva and D. Zhou, 2007. Semi-supervised graph-based hyperspectral image classification. IEEE. Trans. Geosci. Remote Sens., 45: 3044-3054.

Chapelle, O., M. Chi and A. Zien, 2006. A continuation method for semi-supervised SVMs. Proceedings of the 23rd International Conference on Machine Learning (ICML'06), June 25-29, 2006, Pittsburgh, Pennsylvania, USA., pp: 185-192.

Cheng, G., J. Han, L. Guo, Z. Liu, S. Bu and J. Ren, 2015. Effective and efficient midlevel visual elements-oriented land-use classification using VHR remote sensing images. IEEE. Trans. Geosci. Remote Sens., 53: 4238-4249.

Di, W. and M.M. Crawford, 2011a. Active learning via multi-view and local proximity co-regularization for hyperspectral image classification. IEEE. J. Sel. Top. Signal Process., 5: 618-628.

Di, W. and M.M. Crawford, 2011b. View generation for multiview maximum disagreement based active learning for hyperspectral image classification. IEEE. Trans. Geosci. Remote Sens., 50: 1942-1954.

Fauvel, M., Y. Tarabalka, J.A. Benediktsson, J. Chanussot and J.C. Tilton, 2012. Advances in spectral-spatial classification of hyperspectral images. Proc. IEEE., 101: 652-675.

He, T., S. Zhang, J. Xin, P. Zhao and J. Wu *et al*., 2014. An active learning approach with uncertainty, representativeness and diversity. Sci. World J., Vol. 2014, 10.1155/2014/827586

Hoi, S.C., R. Jin, J. Zhu and M.R. Lyu, 2006. Batch mode active learning and its application to medical image classification. Proceedings of the 23rd International Conference on Machine Learning (ICML'06), June 25-29, 2006, Pittsburgh, Pennsylvania, USA., pp: 417-424.

Houlsby, N., J.M. Hernandez-Lobato and Z. Ghahramani, 2014. Cold-start active learning with robust ordinal matrix factorization. Proceedings of the 31st International Conference on Machine Learning (ICML'14) Vol. 32, January 2014, Beijing, China, pp: 766-774.

Joshi, A.J., F. Porikli and N.P. Papanikolopoulos, 2013. Scalable active learning for multiclass image classification. Patt. Anal. Mach. Intelli. Trans., 34: 2259-2273.

Kim, W. and M.M. Crawford, 2010. Adaptive classification for hyperspectral image data using manifold regularization kernel machines. IEEE. Trans. Geosci. Remote Sens., 48: 4110-4121.

Li, J., J.M. Bioucas-Dias and A. Plaza, 2012. Spectral-spatial classification of hyperspectral data using loopy belief propagation and active learning. IEEE. Trans. Geosci. Remote Sens., 51: 844-856.

Li, X. and Y. Guo, 2013. Adaptive active learning for image classification. Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition, June 23-28, 2013, IEEE, Portland, Oregon, USA., pp: 859-866.

Liu, C., L. He, Z. Li and J. Li, 2017. Feature-driven active learning for hyperspectral image classification. IEEE. Trans. Geosci. Remote Sens., 56: 341-354.

Liu, P., H. Zhang and K.B. Eom, 2016. Active deep learning for classification of hyperspectral images. IEEE. J. Sel. Top. Applied Earth Obs. Remote Sens., 10: 712-724.

Melgani, F. and L. Bruzzone, 2004. Classification of hyperspectral remote sensing images with support vector machines. IEEE Trans. Geosci. Remote Sens., 42: 1778-1790.

Mountrakis, G., J. Im and C. Ogole, 2011. Support vector machines in remote sensing: A review. ISPRS J. Photogramm. Remote Sens., 66: 247-259.

Pasolli, E., F. Melgani, D. Tuia, F. Pacifici and W.J. Emery, 2011. Improving active learning methods using spatial information. Proceedings of the 2011 IEEE International Symposium on Geoscience and Remote Sensing, July 24-29, 2011, IEEE, Vancouver, Canada, pp: 3923-3926.

Patra, S. and L. Bruzzone, 2011. A batch-mode active learning technique based on multiple uncertainty for SVM classifier. IEEE. Geosci. Remote Sens. Lett., 9: 497-501.

Persello, C. and L. Bruzzone, 2014. Active and semisupervised learning for the classification of remote sensing images. IEEE. Trans. Geosci. Remote Sens., 52: 6937-6956.

Plaza, A., J.A. Benediktsson, J.W. Boardman, J. Brazile and L. Bruzzone *et al*., 2009. Recent advances in techniques for hyperspectral image processing. Remote Sens. Environ., 113: S110-S122.

Prabhakar, T.N., G. Xavier, P. Geetha and K.P. Soman, 2015. Spatial preprocessing based multinomial logistic regression for hyperspectral image classification. Procedia Comput. Sci., 46: 1817-1826.

Qi, G.J., X.S. Hua, Y. Rui, J. Tang and H.J. Zhang, 2008. Two-dimensional active learning for image classification. Proceedings of the 2008 IEEE International Conference on Computer Vision and Pattern Recognition, June 23-28, 2008, IEEE, Anchorage, Alaska, USA., pp: 1-8.

Qiu, Q., X. Wu, Z. Liu, B. Tang and Y. Zhao *et al*., 2017. Survey of supervised classification techniques for hyperspectral images. Sens. Rev., 37: 371-382.

Rajan, S., J. Ghosh and M. Crawford, 2006. An active learning approach to knowledge transfer for hyperspectral data analysis. Proceedings of the 2006 IEEE International Symposium on Geoscience and Remote Sensing, July 31-August 4, 2006, IEEE, Denver, Colorado, USA., pp: 541-544.

Sabale, S.P. and C.R. Jadhav, 2015. Hyperspectral image classification methods in remote sensing-a review. Proceedings of the 2015 International Conference on Computing Communication Control and Automation, February 26-27, 2015, IEEE, Pune, India, pp: 679-683.

Serpico, S.B. and L. Bruzzone, 2001. A new search algorithm for feature selection in hyperspectral remote sensing images. IEEE. Trans. Geosci. Remote Sens., 39: 1360-1367.

Settles, B., 2010. Active learning literature survey. Mach. Learn., 15: 201-221.

Stumpf, A., N. Lachiche, J.P. Malet, N. Kerle and A. Puissant, 2013. Active learning in the spatial domain for remote sensing image classification. IEEE. Trans. Geosci. Remote Sens., 52: 2492-2507.

Sun, Y., X. Zhang, A. Plaza, J. Li, I. Dopido and Y. Liu, 2016. A new semi-supervised classification strategy combining active learning and spectral unmixing of hyperspectral data. Proceedings of International Conference on High-Performance Computing in Geoscience and Remote Sensing (SPIE’10007), October 24, 2016, Edinburgh, UK., pp: 1-8.

Tuia, D., F. Pacifici, M. Kanevski and W.J. Emery, 2009. Classification of very high spatial resolution imagery using mathematical morphology and support vector machines. IEEE. Trans. Geosci. Remote Sens., 47: 3866-3879.

Wan, L., K. Tang, M. Li, Y. Zhong and A.K. Qin, 2014. Collaborative active and semisupervised learning for hyperspectral remote sensing image classification. IEEE. Trans. Geosci. Remote Sens., 53: 2384-2396.

Wang, L., S. Hao, Q. Wang and Y. Wang, 2014. Semi-supervised classification for hyperspectral imagery based on spatial-spectral label propagation. ISPRS. J. Photogram. Remote Sens., 97: 123-137.

Wang, Z., B. Du, L. Zhang, L. Zhang and X. Jia, 2017. A novel semisupervised active-learning algorithm for hyperspectral image classification. IEEE. Trans. Geosci. Remote Sens., 55: 3071-3083.

Zhang, C. and T. Chen, 2002. An active learning framework for content-based information retrieval. IEEE Trans. Multimedia, 4: 260-268.

Zhang, Z. and M.M. Crawford, 2015. An ensemble active learning approach for spectral-spatial classification of hyperspectral images. Proceedings of the 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS’15), July 26-31, 2015, IEEE, Milan, Italy, pp: 4963-4966.