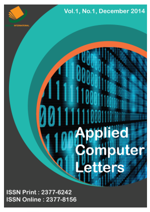
Applied Computer Letters (ACL)

DOI : http://doi.org/10.26480/acl.02.2017.16.19

ZIBELINE INTERNATIONAL PUBLISHING

ISSN: 2377-6242 (Print)
ISSN: 2377-8156 (Online)



SVDD ALGORITHM BASED ON NOISE COST FOR ANOMALY DETECTION IN HYPERSPECTRAL IMAGERY

Liyan Zhang*, Zhilin Liang, Xianling Zeng, Dandan Fu

Key Laboratory of resource Environment and Geographic Information System; Key Laboratory of 3-Dimensional Information Acquisition and Application, Capital Normal University, Beijing, China
****Corresponding Author's Email: zhangliyan010@126.com***

ARTICLE DETAILS

Article History:

Received 12 July 2017
Accepted 12 August 2017
Available online 8 October 2017

ABSTRACT

Anomaly detection algorithm based on Support Vector Data Description (SVDD) brings low detection rates due to background training samples being contaminated by anomalous data. To solve the problem, a new method based on SVDD with Noise Cost is proposed by introducing unbalanced data mining cost sensitive mind. This algorithm gives a different noise cost value to each background training samples through the neighbourhood clustering and then introduces the noise cost into SVDD to construct the SVDD hypersphere, thus making the classification interface more compact and improving the description ability of the anomaly and background value. At the same time, the sensitivity to the abnormal algorithm and the detection probability of the algorithm are greatly improved. Experimental results based on simulation data show that: compared to SVDD, this algorithm greatly reduces the false alarm rate, and improves the detection precision.

KEYWORDS

Noise Cost; Support Vector Data Description, Local Neighbourhood Clustering, Hyperspectral Imagery, Anomaly Detection

1. INTRODUCTION

Hyperspectral imaging spectrometer provides tens to hundreds of narrow-band (usually band-width<10nm) spectral information for each pixel, improving the land cover classification and identification abilities. With so many narrow-bands, hyperspectral images can distinguish different targets' subtle differences and this feature makes anomaly detection in hyperspectral imagery have significance application value in military affairs and civil use.

Based on a study, RX algorithm is a classical method for hyperspectral image anomaly detection. But at the higher complexity of background, it can't reflect the background pixel data distribution, and will lead to lower detection performance [1-5]. A researcher put forward LPD (Low Probability Detection) algorithm which is based on spectral mixing model [6]. LPD algorithm uses orthogonal vectors as the background spectrum in the low-frequency space. Thus, the influence of noise is great, and its missing rate is high [7]. In 2006, a researcher raised SVDD (Support Vector Data Description) algorithm which is driven by data [8]. According to research, being superior to the existing algorithms, it effectively reduces the false alarm rate and can detect abnormal target in hyperspectral imagery [9-11].

But in the process of SVDD sample training, all of the data points are equally treated, which makes abnormal pixel (perhaps noise), mixed in the algorithm training sample sensitive, and it inevitably affects the model accuracy in Hyperspectral Anomaly Detection, resulting in detection probability decreases [12-14]. To solve this problem, we introduce the cost sensitive thought and define noise cost for each data [15]. Then we put forward the SVDD algorithm based on noise cost for Anomaly detection in hyperspectral images. The possible noise samples will be given a small noise cost value, effectively reducing the impact of noise on the data set anomaly detection [16]. Here we use neighbourhood clustering method to determine the noise cost of training samples [17].

2. SVDD ALGORITHM BASED ON NOISE COST

The basic idea of SVDD (Amit Banerjee, 2006) is: constraining all the samples with common characteristics to a hypersphere and looking for a minimum closed hypersphere that can separate this class from other classes. Anomaly detection is essentially a class issue. Set up sample set $\boldsymbol{X}=\{\boldsymbol{x}_i, \ =1,\cdots,m\}$, $\boldsymbol{x}_i \in \boldsymbol{R}^n$, where m is the number of training sample sets. Hypersphere including sample sets, $=\{\boldsymbol{x}:\|\boldsymbol{x}-\boldsymbol{a}\|^2<R^2\}$ where R is radius, the center of the sphere is $\boldsymbol{a}$. Calculating minimum hypersphere enclosed sample sets is essentially a constrained optimization problem. On condition of all samples contained within the constraints of the ball, find the minimum radius of the ball. Formula (1) is as follows:

$$\min \quad f(R)=R \qquad s.t. \ \ (\boldsymbol{x}-\boldsymbol{a})(\boldsymbol{x}_i-\boldsymbol{a})^T-R^2\le 0, \qquad i=1,\cdots,m$$
(1)

In the process of SVDD sample training, all of the data points are equally treated. To solve this problem, we introduce the cost sensitive thought and define noise cost for each data. Then we put forward the SVDD algorithm based on noise cost. In the algorithm, we introduce a noise cost μ_i, $0 < \mu_i < 1$ training sample sets is $(x_1, \mu_1), \ldots, (x_i, \mu_i), \ldots, (x_m, \mu_m)$. Optimization objective function (1) is rewritten to formula (2):

$$\min \quad w(\xi_i, R, \boldsymbol{a}) = R^2 + C\mu_i \sum_{i=1}^{m} \xi_i \quad s.t. \quad \| \Phi(x_i) - \boldsymbol{a} \|^2 \le R^2 + \xi_i, \xi_i \ge 0, i = 1, \cdots, m \tag{2}$$

Where C is used to determine the penalty coefficient between hypersphere volume and the number of rejected targets, ξ_i is relaxation factor of building training samples in a hypersphere.

With Lagrange dual method, we transform optimization problem to dual function maximum optimization problem whose decision variables is Lagrange multiplier. Replace $\Phi(x)$ with X, and then construct Lagrange function:

$$L(R, \boldsymbol{a}, \boldsymbol{\alpha}, \beta) = R^2 - \sum_{i=1}^{m} \alpha_i \left\{ R^2 + \xi_i - \left\| \Phi(x_i) - \boldsymbol{a} \right\|^2 \right\} + C\mu_i \sum_{i=1}^{m} \xi_i - \sum_{i=1}^{m} \xi_i \beta_i \tag{3}$$

Where $\boldsymbol{\alpha} = \{\alpha_i \ge 0, i = 1, \cdots, m\}$, $\beta = \{\beta_i \ge 0, \ = 1, \cdots, m\}$ are Lagrange multipliers, C is a constant, $C\mu_i \sum_{i=1}^{m} \xi_i$ is penalty term. Calculate derivatives on both sides of the equation for R, $\boldsymbol{a}$ and ξ_i respectively. Make derivative 0, obtain the following equation:

$$\sum_i \alpha_i = 1, \quad \boldsymbol{a} = \sum_i \alpha_i \Phi(x_i), \quad \alpha_i = C\mu_i - \beta_i \tag{4}$$

$$\xi_i \beta_i = 0, \quad (\ ^2 + \xi_i - \left\| \Phi(x_i) - \boldsymbol{a} \right\|^2) \alpha_i = 0 \tag{5}$$

$$\max \quad Q(\alpha, \beta) = \sum_{i=1}^{m} \Phi(x_i) \bullet \Phi(x_i) \alpha_i - \sum_{i,j=1}^{m} \alpha_i \beta_j \Phi(x_i) \bullet \Phi(x_j) \quad s.t. \quad 0 \le \alpha_i \le C\mu_i, i = 1, 2, \ldots, m \tag{6}$$

According KKT conditions (Ben-Hur A, Horn D, 2002); we deduce formulas (5). From formulas (6), the original constrained optimization problem converts to a dual problem. Mapping function inner product operation can be expressed by the kernel function $K(x, y)$, $K(x, y) = \langle \Phi(x), \Phi(y) \rangle$. We select radial basis (RBF) function as the kernel function. Expressed as:

$$\mathrm{K(x, y)} = \exp\left(\frac{-\|x - y\|^2}{\sigma^2} \right) \tag{7}$$

The problem function is rewritten to formula (8):

$$\max \quad Q(\alpha, \beta) = \sum_{i=1}^{m} K(x_i, x_i) \alpha_i - \sum_{i,j=1}^{m} \alpha_i \beta_j K(x_i, x_j) \quad s.t. \quad 0 \le \alpha_i \le C\mu_i, i = 1, 2, \ldots, m \tag{8}$$

By analysing of formula (8), we can indicate that if $\xi_i > 0$, the sample points x_i will fall over the hyperspheres. While x_i satisfies $\beta_i = 0$, we call them Border Support Vector (BSV). If $0 < \alpha_i < C\mu_i$, we deduce $\xi_i = 0$ and $(R^2 - \left\| \Phi(x_i) - \boldsymbol{a} \right\|^2) = 0$. It illustrates that these points fall on the hypersphere and these points are called Support Vector (SV). The remaining sample points are inside the hypersphere.

The main difference between the SVDD algorithm and SVDD based on the cost of noise is that the Lagrange multipliers in the dual problem have different upper bound. The upper bound is 1 in SVDD algorithm. However, the upper bound of Lagrange multiplier is dynamically determined by the noise cost of each sample in SVDD based on the cost of noise.

3. NOISE COST MODEL BASED ON NEIGHBOURHOOD CLUSTERING

As SVDD anomaly detection is a partial anomaly detection algorithm, we use Neighbourhood Clustering Segmentation method (Derong CHEN, Liyan ZHANG, 2007) to replace the KNN method and use spectral similarity instead of the KNN distance as the metric value.

For sample points (x_i , x_j) their spectral similarity can be represented by the cosine of spectral angle, defined as formula (9). Where (x_i , x_j) represent pixel spectrum vectors. For each sample point x_i, find the set S_i^k, here k is the number of its nearest neighbour points. According to formula (9), the average distance between x_i and each pixel in the set S_i^k can be defined as formula (10):

$$d(x_i, x_j) = \langle x_i, x_j \rangle \Big/ \sqrt{\langle x_i, x_i \rangle} \sqrt{\langle x_j, x_j \rangle} \tag{9}$$

$$d_i = \frac{1}{k} \sum \left(\langle x_i, x_j \rangle \Big/ \sqrt{\langle x_i, x_i \rangle} \sqrt{\langle x_j, x_j \rangle} \right) \tag{10}$$

Take the figure 1 for example, for the points x_1 and x_6, we can calculate $S_1^4 = \{x_2, x_3, x_4, x_5\}$, $S_6^4 = \{x_7, x_8, x_9, x_{10}\}$ by formula (9).

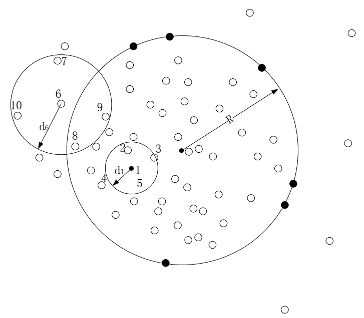


Figure 1: The feature space sample distribution

By formula (10), we obtain the average distance between the point and the set. As it can be predicted from Figure 2, and $d_1 < d_6$.Based on the assumption above, we can build a model between noise cost μ_i and the distance d_i. Define $d_{\max}$, $d_{\min}$ as the maximum and minimum values of d_i respectively. Noise cost model is defined as formula (15). Where $\sigma < 1$ and f is a control parameter.

$$d_{\max} = \max(d_i \mid x_i \in X), d_{\min} = \min(d_i \mid x_i \in X) \tag{14}$$

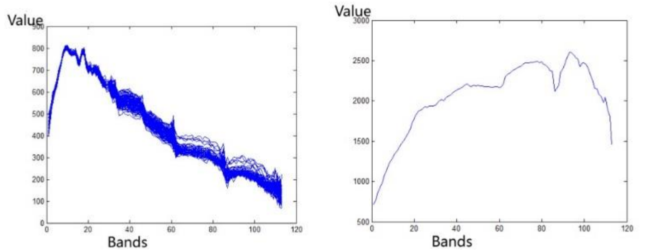
$$\mu_i = 1 - (1 - \sigma)\left(\frac{d_i - d_{\min}}{d_{\max} - d_{\min}}\right)^f \tag{15}$$

4. EXPERIMENT AND ANALYSIS

To validate SVDD algorithm based on noise cost performs better than SVDD, we can demonstrate it by the following experiment.

4.1 Emulation Data Experiment

Emulation data include background pixels and singular pixels. As shown in Figure 2 (a), the spectrums of airport are background pixels. And in Figure 2 (b), the spectrums of planes are abnormal pixels. Image size is 200 × 200, the number of bands to 113. Figure 3 is band 14 image of the airport emulation data.



(a) Background pixel spectral curves (b) Singular pixels spectral curves

Figure 2: Spectral curves of simulation data

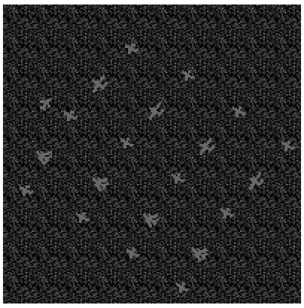
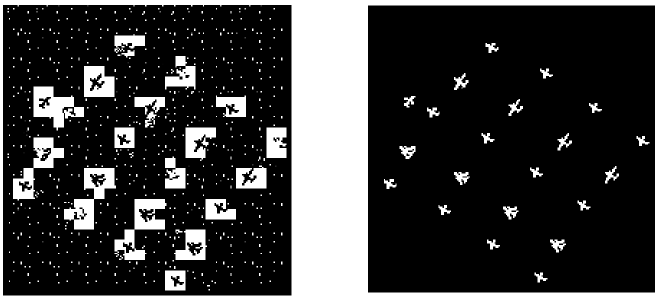


Figure 3: 14nd band imagery of simulation data

In the experiment, we select that the size of background window is 15×15-7×7and kernel function parameter $\sigma = 23$ simulation data for testing. We use both SVDD algorithm and SVDD algorithm based on noise cost to do the experiment on the simulation data. Figure 4(a) and figure4 (b) is the result of SVDD and SVDD algorithm based on noise cost.



(a)SVDD (b) SVDD Based on Noise Cost

Figure 4: Results of the two methods on simulation data

4.2 Real Hyperspectral Images Experiment

To further examine the effectiveness of the algorithm, we perform experiment on 224-band AVRIS hyperspectral imagery. Its spatial resolution is 20×20m. There are 13 abnormal targets. It is shown that band 5 and abnormal target distribution in figure 5.

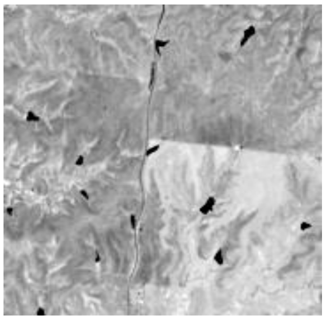
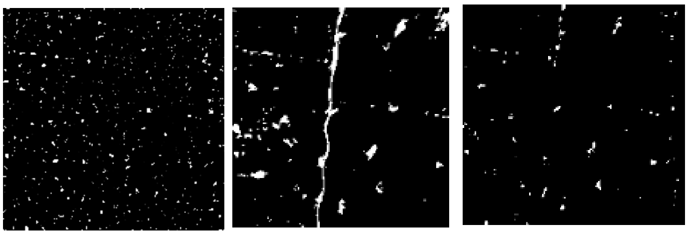


Figure 5: The fifth band image of AVIRIS Data

SVDD, LPD and SVDD algorithm based on noise cost are respectively used to do the experiment on AVRIS. In the experiment, we select that the size of background window is 15×15-7×7. In other word, the inner window is 7×7 and the external is 15×15. Traverse the image in the unit of seven pixels row by row.



(a) SVDD (b) LPD (c) SVDD Based on Noise Cost

Figure 6: Results of the three methods on AVIRIS data

The result is shown in Figure 6. From Figure 6, we conclude that comparing to SVDD and LPD, SVDD Based on Noise Cost reduces the influence of error samples and have a higher detection rate. However, on the condition that the anomaly targets are relatively weak, a handful of background pixels do not participate in the construction of training samples. They may be erroneously detected as abnormal targets and mainly scatter around the weak abnormal pixels.

Algorithm characteristic curve is used to describe the relationship between detection probability P_d and false alarm probability P_f on the condition of different detection thresholds. It provides a quantitative analysis for the algorithm detection performance. Figure 7 is ROC curves of the three methods.

It provides a quantitative analysis for the algorithm detection performance. Figure 7 is ROC curves of the three methods.

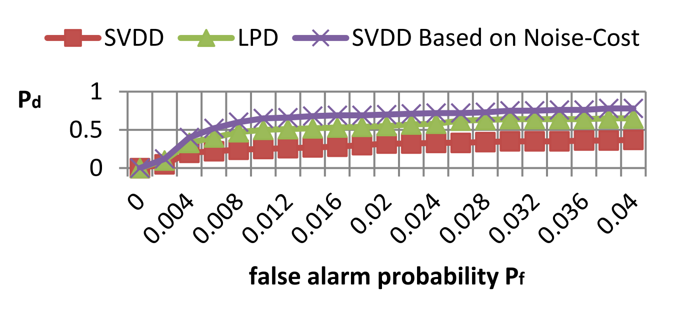


Figure 7: ROC curves of the three methods

In Figure 7, we can see that SVDD Based on Noise Cost not only improves the detection probability P_d , but also reduces the false alarm probability P_f . In addition, the classified boundaries of different hyperspheres are more compact in the feature space. The SVDD model based on noise cost is more sensitive to abnormal pixels. It can be concluded that SVDD Based on Noise Cost is superior to SVDD and LPD anomaly detection algorithm.

5. CONCLUSION

SVDD Algorithm brings low accuracy and detection rate. To solve this problem, we introduce the cost-sensitive ideology and put forward SVDD Algorithm Based on Noise Cost for Anomaly detection in hyperspectral imagery. In the algorithm, background training samples are given different noise-cost values. We introduce noise cost into SVDD to rebuild hypersphere in this way. Then, perform experiment on simulation data and real hyperspectral data AVRIS. The results show that:

1. Reduce the probability of abnormal pixel mixed into training samples
2. Enhance the algorithm sensitivity to abnormal pixels
3. Improve the hyperspectral imagery abnormality detection accuracy.
4. Reduce the false alarm rate.
5. The classified boundaries are more compact
6. Reduce the computation time.

Disclosure statement: No potential conflict of interest was reported by authors.

Funding: This work is supported by the national natural science foundation of China (41201075) and the commission of science and technology project (KM201210028012).

REFERENCES

[1] Reed, I.S., Yu, X. 1990. Adaptive Multiple-band CFAR Detection of an Optical Pattern with Unknown Spectral Distribution. IEEE Transactions on Acoustics Speech and Signal Processing, 38 (10), 1760-1770.

[2] Banerjee, A., Burlina, P., Diehl, C. 2006. A support vector method for anomaly detection in hyperspectral imagery. Journal of Geoscience and Remote Sensing, IEEE Transactions on, 44 (8), 2282-2291.

[3] Khazai, S., Homayouni, S., Safari, A. 2011. Anomaly detection in hyperspectral images based on an adaptive support vector method. Journal of Geoscience and Remote Sensing Letters, IEEE, 8 (4), 646-650.

[4] Shi, Z.W., Wu, J., Yang, S. 2012. RX and its variants for anomaly detection in hyperspectral images. Journal of Infrared and Laser Engineering, 41 (3), 796-802.

[5] Liu, M., Du, X.P., Sun, J. 2013. Anomaly Detection Based on Improved Orthogonal Subspace Projection Algorithm in Hyperspectral Imagery. Journal of Academy of Equipment, 23 (4), 92-96.

[6] Harsanyi, J.C. 1993. Detection and Classification of Subpixel Spectral Signatures in Hyperspectral Image Sequences Ph.D. Thesis. Baltimore: University of Maryland Baltimore County.

[7] Zhang, L.Y., Chen, D.R., Tao, P. 2007. Anomaly Detection for Hyperspectral Imagery Based on Vertex Component Analysis. Journal of Astronautics, 28 (5), 1262-1265.

[8] Banerjee, A. 2006. A Support Vector Method for Anomaly Detection in Hyperspectral Imagery. IEEE Transactions on Geoscience and Remote Sensing, 44 (8), 2282-2291.

[9] Chen, D.R., Zhang, L.Y., Tao, P. 2007. Xuping CAO. Support Vector Data Description for Anomaly Detection in Hyperspectral Imagery Combined with Neighbouring Clustering Segmentation. Journal of Astronautics, 28 (3), 767-771.

[10] Chen, D.R., Gong, J.L., Chen, Q. 2008. Support Vector Data Description for Fast Anomaly Detection in Hyperspectral Imagery Based on Sample Segmentation. Acta Armamentarii, 29 (9), 1049-1053.

[11] Banerjee, A., Burlina, P., Meth, R. 2007. Fast hyperspectral anomaly detection via SVDD Image Processing. ICIP 2007. IEEE International Conference on. IEEE, 4, IV-101-IV-104.

[12] Hen, D.R., Gong, J.L., He, G.L., Cao, X.P. 2010. A RFS-SVDD Algorithm for Hyperspectral Global Anomaly Detection. Journal of Astronautics, 31 (1), 228-232.

[13] Pu, X.F., Lei, W.H., Tang, J.J. 2010. Anomaly Detection for Hyperspectral Image Based on SVDD with Negative Examples. Opto-Electronic Engineering, 37 (012), 83-87.

[14] Zheng, E.H. 2006. Cost Sensitive Data Mining Based on Support Vector Machines: Theories and Applications. Zhejiang University.

[15] Ben-Hur, A., Horn, D., Siegelmann, H.T. 2002. Support vector clustering. The Journal of Machine Learning Research, 2, 125-137.

[16] He, S.Y., Zhang, L.Y. 2008. Research on Image Compression Method Based on Anomaly Detection. Chinese Journal of Electron Devices, 2008, 31 (2), 410, 413.

[17] Jie, L.I., Zhao, C.H., Mei, F. 2010. Detecting Hyperspectral Anomaly by Using Background Residual Error Data. Journal of Infrared Millimeter Waves, 29 (2), 150-155.