X-ray Astronomical Point Sources Recognition Using Granular Binary-tree SVM

Zhixian Ma*, Weitian Li†, Lei Wang*, Haiguang Xu†, Jie Zhu*

*Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai, China
†Department of Physics and Astronomy, Shanghai Jiao Tong University, Shanghai, China
Email: {mazhixian, zhujie}@sjtu.edu.cn

Abstract—The study on point sources in astronomical images is of special importance, since most energetic celestial objects in the Universe exhibit a point-like appearance. An approach to recognize the point sources (PS) in the X-ray astronomical images using our newly designed granular binary-tree support vector machine (GBT-SVM) classifier is proposed. First, all potential point sources are located by peak detection on the image. The image and spectral features of these potential point sources are then extracted. Finally, a classifier to recognize the true point sources is build through the extracted features. Experiments and applications of our approach on real X-ray astronomical images are demonstrated. Comparisons between our approach and other SVM-based classifiers are also carried out by evaluating the precision and recall rates, which prove that our approach is better and achieves a higher accuracy of around 89%.

Index Terms—support vector machine (SVM); X-ray astronomical image; point sources; binary-tree; granule

I. INTRODUCTION

There are various astronomical objects that are radiating at different electromagnetic bands, from the long-wavelength radio band to the high-energy X-ray and Γ-ray observations. Among them, the X-ray observation is an important method to study the Universe, and have already revealed to us many exciting discoveries, such as the active galactic nucleus (AGN) and galaxy clusters filled with hot plasma. And many sources of our interest exhibit a point-like appearance. However, they are very far away and there are diffuse background radiation all over the sky. In addition, due to the imperfections of the instruments, the observed images are distorted, which can be described as the convolution with instrumental point spread function (PSF) [1]. It should be noted that the PSF varies across the instrument regions. Therefore, it is a challenge but of great importance to accurately detect these point sources (PS) from the faint observed images.

Since the astronomical sources are very distant and the X-ray radiation flux is very low, the instrument (e.g., CCD) is thus able to measure the position and energy of *every* incoming X-ray photon (i.e., an event). All the measured events are stored as an event table, from which the spacial image and/or the spectrum of a specific region can be extracted. Due to the very low radiation flux, the observed image suffers from significant Poisson noises [1], [2], which makes the PS detection very difficult and error-prone.

Recently, Masias *et al.* reviewed many source (both point-like and extended) detection methods [3]. And these methods can be roughly divided into three categories according their underlying techniques: (1) wavelet [1], [4]; (2) matched filters [5]; and (3) Bayesian techniques [2]. Although every method has its own strengthens and applicabilities, however, most of them only exploit the spacial information of the image, and just ignore the additional spectral information. We argue that the available spectral information should be utilized together with the spacial information, in order to further improve the source detection ability.

In this paper, we propose a new classifier based on the support vector machine (SVM) to recognize the potential PS, which takes advantage of available information from both the spacial and the spectral domain. All potential PS are first located by performing the peak detection on the spacial image. Then, both the spacial and spectral features are extracted and used to recognize the true PS by our SVM-based classifier.

The SVM has been proved as an excellent algorithm to solve the classification problems [6] and has been widely used in various areas [7]–[9]. Many algorithms based on the SVM are also proposed, such as Granular SVM (GSVM) [7], [8], and Binary-Tree SVM (BTS) [9]. Among them, the BTS is used for multi-class problems, while the GSVM can handle the imbalanced training sets well.

With regard to the X-ray astronomical images, they are usually very sparse, i.e., the samples of backgrounds are far more than the number of PS. Besides, the PS should be further divided into bright and faint classes, because the faint ones are almost as faint as the backgrounds. Consequently, a classifier to solve both the imbalanced training sets and multi-class classification problem is required. Thus we design our new granular binary-tree SVM (GBT-SVM) classifier by integrating both the GSVM and BTS.

The rest of the paper is organized as follows. In Sec. II, we describe the properties and spectral features of the point sources and backgrounds. In Sec. III, our proposed GBT-SVM classifier is explained in detail. After that, we briefly describe the peak detection approach used to locate the potential PS in Sec. IV. Then in Sec. V, experiments are carried out using the real X-ray astronomical images. Finally, we conclude in Sec. VI with some discussions and outlooks.

II. POINT SOURCE AND BACKGROUND PROPERTIES

The PS have different spacial and spectral properties compared to the backgrounds, thus they can be located in the
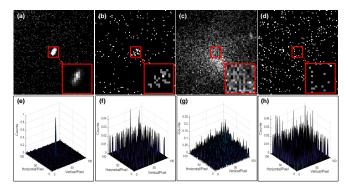
Fig. 1. Example images cropped from the Chandra observation of HCG 62 (observation ID: 921) showing the following 4 classes of objects: (a) bright PS, (b) faint PS, (c) bright background (i.e., extended source) and (d) faint background. And the bottom row shows the corresponding three-dimensional views.

image, and can be further recognized from backgrounds. In this section, we introduce both the spacial and spectral properties of the PS and backgrounds, and then define the features used for recognition.
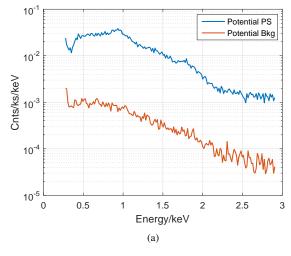
A. Spacial domain

A typical PS appears as an compact elliptical blob on the image that is brighter than its surrounding background. Its elliptical shape is due to the convolution with the instrumental PSF, which is generally elliptical [1]. And a PS usually have only one major peak at the centroid.

As for the background, it is generally uniformly distributed and obeys the Poisson distribution. There are many minor peaks within the background region without any major peak. However, there may exists extended source in the image which is much brighter than the plain background. Since it is also uniform on small scales of our interest, and it is more similar to the background compared to the PS, we regard the extended source as the *bright background*.

Therefore, we have following 4 class of objects in our work: (1) bright PS; (2) faint PS; (3) bright background (i.e., extended source) and (4) faint background. Fig. 1 shows an example of these 4 class of objects.

B. Spectral domain

A spectral describes the energy distribution of the X-ray photons that extracted from the region of interest, which can further reflect the internal physical processes of the origin source. As for the *Chandra*[1] X-ray observations, we extract the spectrum within energy range of 0.5 keV – 3.0 keV for this work. Generally, the PS is brighter than the background, thus it also has higher spectral intensities. In addition, the PS and background have different spectral shape due to there different radiation origins, as shown in Fig. 2(a). Therefore, the spectral properties can be utilized to recognize the PS from background.
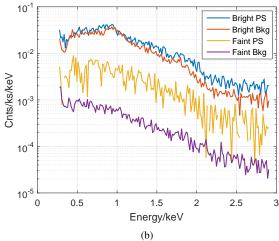
[1] *Chandra* X-ray Observatory: http://cxc.harvard.edu/

(a)

(b)

Fig. 2. Different spectral intensities and shapes among different class of PS and background. **(a)** Spectra of the potential PS and background (corresponding to the children in the 1st level of the binary tree as shown in Fig. 3); **(a)** Spectra of the 4 different classes of objects (corresponding to the children in the 2nd level of the binary tree).
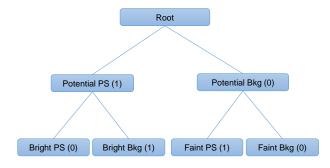

Fig. 3. A frame of the binary-tree structure SVM classifier.

C. Feature vector

In order to classify the 4 different classes of objects, we adopt a two-level binary tree for our GBT-SVM, as illustrated in Fig. 3.

The extracted spectrum of an object can be just used as the feature vector for our classifier, since it is already in vector

format. In addition, we define the following 4 spacial features: (1) counts per pixel (F_{cpp}); (2) peak-to-average ratio (F_{par}); (3) variance (F_{var}); and (4) number of peaks (F_{nop}). And the features (1)–(3) are defined as:

$$F_{\text{cpp}} = \frac{1}{mn} \sum_{i=1}^{m} \sum_{j=1}^{n} R(x_i, y_j), \tag{1}$$

$$F_{\text{par}} = \frac{R(x_p, y_p)}{F_{\text{cpp}}}, \tag{2}$$

$$F_{\text{var}} = \frac{1}{mn} \sum_{i=1}^{m} \sum_{j=1}^{n} [R(x_i, y_j) - F_{\text{cpp}}]^2, \tag{3}$$

where R represents the region of the object of interest, (x_p, y_p) is the coordinates of the peak, and m, n are the rows and columns of matrix R.

Then, we combine the above 4 spacial features with the spectral features to derive the final feature vector to be used in our GBT-SVM classifier:

$$F = (F_{\text{spec}}, F_{\text{cpp}}, F_{\text{par}}, F_{\text{var}}, F_{\text{nop}}), \tag{4}$$

where F_{spec} is the spectral feature of the object.

III. GBT-SVM MODEL

To tackle the multi-class problem described above, we design a new SVM-based classifier, namely GBT-SVM, which consists of three submodels: (1) classifier for the potential PS and background; (2) classifier for bright PS and background; and (3) classifier for faint PS and background. In this section, we first briefly introduce the basically SVM model of our classification problem, then describe the strategy to deal with the imbalanced dataset by adopting the thinking of granular computing and sampling. Finally the GBT-SVM model is explained.

A. Basic SVM model

The spacial and spectral features of each PS and background object are extracted as described in Sec. II-C, and are represented as

$$S = \{(\boldsymbol{x_i}, y_i)\}, i = 1, \cdots, M, \tag{5}$$

where S is the sample feature set, $\boldsymbol{x_i}, y_i$ are the ith sample's feature vector and its classification label, and M is the amount of samples.

Then the classifier is defined as follows,

$$f(\boldsymbol{x}) = \boldsymbol{w}^T \phi(\boldsymbol{x}) + b, \tag{6}$$

where $\boldsymbol{w}$ is the weight vector, and b is the bias, and $\phi(\boldsymbol{x})$ is a mapping function that maps the feature vector $\boldsymbol{x}$ to a higher dimensional linear space for easier classification.

And the solution of parameters $\boldsymbol{w}$ and b can be derived by solving the following optimization problem, where the soft margin strategy with loss function is also considered (see Eq. 7).

$$\min_{\boldsymbol{w}, b} \quad \frac{1}{2} \|\boldsymbol{w}\|^2 + C \sum_{i=1}^{M} \xi_i$$
$$\textbf{s.t.} \quad y_i(\boldsymbol{w}^t \phi(\boldsymbol{x_i}) + b) \geq 1 - \xi_i,$$
$$\xi_i \geq 0, i = 1, 2, \cdots, m, \tag{7}$$

where C is the trade-off between the structural risk (target) and the empirical risk (miss probability) [10]. $\xi_i, i = 1, 2, \cdots, M$ are the slack variables, which are defined as:

$$\xi_i = \max(0, 1 - y_i(\boldsymbol{w}^T \phi(\boldsymbol{x_i}) + b)) \tag{8}$$

To solve Eq.7, the Lagrange function is defined.

$$L(\boldsymbol{w}, b, \boldsymbol{\alpha}, \boldsymbol{\xi}, \boldsymbol{\mu}) = \frac{1}{2} \|\boldsymbol{w}\|^2 + C \sum_{i=1}^{M} \xi_i$$
$$+ \sum_{i=1}^{M} \alpha_i (1 - \xi_i - y_i(\boldsymbol{w}^T \phi(\boldsymbol{x_i}) + b))$$
$$- \sum_{i=1}^{M} \mu_i \xi_i, \tag{9}$$

where α_i, μ_i are the Lagrange multipliers. Let the partial derivation of L with respect to parameters $\boldsymbol{w}, b, \xi_i$ be zero, then the above equation is rearranged to its dual form:

$$\max_{\boldsymbol{\alpha}} \quad \sum_{i=1}^{M} \alpha_i - \frac{1}{2} \sum_{i=1}^{M} \sum_{j=1}^{M} \alpha_i \alpha_j y_i y_j \phi(\boldsymbol{x_i})^T \phi(\boldsymbol{x_j}),$$
$$\textbf{s.t.} \quad \sum_{i=1}^{M} \alpha_i y_i = 0,$$
$$0 \leq \alpha_i \leq C, i = 1, 2, \cdots, M. \tag{10}$$

Then we solve Eq. 6 by calculating $\alpha_i, i = 1, 2, \cdots, M$, and the result is:

$$f(\boldsymbol{x}) = \sum_{i=1}^{M} \alpha_i y_i [\phi(\boldsymbol{x_i})^T \phi(\boldsymbol{x})] + b, \tag{11}$$

where $\phi(\boldsymbol{x_i})^T \phi(\boldsymbol{x})$ can be calculated by the kernel function $\kappa(\boldsymbol{x_i}, \boldsymbol{x_j})$. In this work, the RBF (Radical Basis Function) kernel is used [10].

B. Imbalanced dataset

Since the amount of faint background samples is far more than the point and extended sources, which leads to an imbalanced sample set. Tang *et al.* proposed a method namely Granular SVM (GSVM), which separates the majority sample set into multiple subsets (i.e., information granules) so as to achieve a better hyperplane [8], [11]. In this work, we take advantage of this method to deal with the imbalanced sample sets.

Taking the classification between the potential PS and background as an example, it compares sample amounts among the classes and choose the larger class as the major class. In this

example, the potential background is the major class, while the potential PS is the minor one. Then the number of granular sets is determined by:

$$N_{\mathrm{gra}} = \lfloor \frac{N_{\mathrm{maj}}}{N_{\mathrm{min}}} \rfloor, \tag{12}$$

where $\lfloor \cdot \rfloor$ means rounding down.

After that, the major sample set is separated into N_{gra} subsets. In this work, the samples are acquired from different observations of different exposure times. Thus, we take advantages of Tang *et al.*'s methods of under-sampling [8], and generate the granular sets by uniformly sampling on the major sample sets so as to cover all the observations.

$$S_{\mathrm{maj}}^{k} = \{S_{\mathrm{maj}}^{k+iN_{\mathrm{gra}}}\},$$
$$i = 1, 2, \cdots, N_{min}, k = 1, 2, \cdots, N_{gra}. \tag{13}$$

Finally, there generates N_{gra} sub training sets. As for each subset, it combines a granular of major samples and the whole minor samples, and respective SVM model can be trained and obtained. With those submodels, the class label of a sample can be decided by a voting strategy, i.e., the label with most tickets wins.

C. GBT-SVM

We build our GBT-SVM classifier by combining the basic SVM and the granular sets. Our classifier has a binary tree structure, and the whole model is divided into three submodels. In each submodel, the major sample is evaluated and separated into granular subsets, and the number of subsets is calculated by Eq. 12. Then the classifier is obtained as follows,

$$C = C_1 \oplus (C_{2\mathrm{L}} \cup C_{2\mathrm{R}}), \tag{14}$$

where C_1 is the GSVM classifier of 1$^{\mathrm{st}}$ level (i.e., potential PS and background), $C_{2\mathrm{L}}$ and $C_{2\mathrm{R}}$ are the GSVM classifiers of leaves in our binary-tree model. And total number of submodels in the classifier is:

$$N_C = N_{\mathrm{gra}}^{1} + N_{\mathrm{gra}}^{2\mathrm{L}} + N_{\mathrm{gra}}^{2\mathrm{R}}. \tag{15}$$

In addition, the 4 classes of objects are encoded based on their labels so as to estimate whether a sample is PS or background. We use three bits to encode the information of each object. The first two bits are labels of the two levels in our binary tree as Fig. 3 shows. And the third bit is the decision label which is the *XOR* of the first two bits (See Tab. I).

The GBT-SVM classification algorithm is displayed in Alg. 1.

TABLE I
CODE TABLE FOR THE LEAVES IN THE BINARY-TREE STRUCTURE.

Region type	Label$_1$	Label$_2$	Decision label
Bright PS	1	0	1
Bright Bkg	1	1	0
Faint PS	0	1	1
Faint Bkg	0	0	0

Algorithm 1 GBT-SVM decision algorithm

1: **Input:** $Sample$ and **Load:** $Classifiers$
2: $Predict_1 = zeros(1, N_{\mathrm{gra}}^{1})]$
3: **for** $i = 1 : N_{\mathrm{gra}}^{1}$ **do**
4:     $Predict_1(i) = predict(Sample, Classifiers.L_1(i))$
5: **end for**
6: $Label_1 = vote(Predict_1)$
7: **if** $Label_1 == 1$ **then**
8:     $Predict_{2\mathrm{L}} = zeros(1, N_{\mathrm{gra}}^{2\mathrm{L}})]$
9:     **for** $j = 1 : N_{\mathrm{gra}}^{2\mathrm{L}}$ **do**
10:         $Predict_{2\mathrm{L}}(j) = predict(Sample, Classifiers.L_{2\mathrm{L}}(j))$
11:     **end for**
12:     $Label_2 = vote(Predict_{2\mathrm{L}})$
13: **else**
14:     $Predict_{2\mathrm{R}} = zeros(1, N_{\mathrm{gra}}^{2\mathrm{R}})]$
15:     **for** $j = 1 : N_{\mathrm{gra}}^{2\mathrm{R}}$ **do**
16:         $Predict_{2\mathrm{R}}(j) = predict(Sample, Classifiers.L_{2\mathrm{R}}(j))$
17:     **end for**
18:     $Label_2 = vote(Predict_{2\mathrm{R}})$
19: **end if**
20: $FinalLabel = xor(Label_1, Label_2)$
21: **if** $FinalLabel == 1$ **then**
22:     **Output:** The sample is a point source.
23: **else**
24:     **Output:** The sample is a background.
25: **end if**

IV. POTENTIAL PS LOCALIZATION

In this section, the approach employed to locate all the potential PS in described. A PS generally has a major peak on the image, as mentioned in Sec II-A, so it is intuitive to adopt the peak detection method.

The background noise is suggested to subject to the Poisson distribution [2]. And for the parameter λ in such distribution, it can be estimated without bias using the average of the group of samples [12]:

$$\hat{\lambda} = \bar{I} = \frac{\sum_{i=1}^{M} \sum_{j=1}^{N} I_{i,j}}{MN}, \tag{16}$$

where $\hat{\lambda}$ is the estimated value, and I is the image matrix.

Thus, the raw image can be preprocessed to reduce the noise. The parameter λ, is estimated and set as a threshold, pixels with counts less than which are set to zero (See Eq. 17).

$$I_r(i,j) = \begin{cases} 0, & I(i,j) \le \hat{\lambda}; \\ I(i,j), & I(i,j) > \hat{\lambda}, \end{cases} \tag{17}$$

where $\{(i,j); i = 1, \cdots, M; j = 1, \cdots, N\}$ are the pixel coordinates in the image I, and I_r is the noise-reduced image.

Finally, peaks in the preprocessed image are located and listed as potential point sources. In our work, to reduce the complexities, we do not detect peaks by covering all pixels. Instead, pixels of the two dimensional matrix are sorted in descending order, and only pixels with values greater than a threshold are extracted as PS centers.

However, as for the extended sources (i.e., bright backgrounds), they often have spread bright pixels. Thus local maxima in these regions are not as significant as point sources. To solve this problem, the neighbors around a maximum pixel

are considered. If there are pixels whose values equal or approach to the maximum, it will be eliminated from the PS list.

V. EXPERIMENTS AND RESULT

Experiments on real X-ray astronomical observations were carried out to demonstrate the performance of our proposed GBT-SVM classifier. And comparisons of our approach with GSVM, and BTSVM were also performed.

A. Datasets

All the datasets in our work were obtained from the *Chandra Data Archive*[2] and processed using the CIAO[3] software v4.4 by following the official guide [13]. After manually filtering, 25 observations were selected, among which 20 observations were chosen as the training datasets while the remaining 5 were used as the testing datasets.

Each observation averagely has 20 point sources and the average exposure time is about 41.62 ks. To evaluate the performance of our approach, all point sources in the raw images were detected with *wavdetect* [1] provided by the CIAO software and then visually checked. They were then set as the *reference group*.

For each training observation, we randomly selected 150 faint backgrounds, 30 bright backgrounds, and the bright and faint PS were manually distinguished according to their brightness and spectrum in the reference group. Then the feature vectors of each sample were generated as explained in Sec. II-C.

As for the testing dataset, the potential point sources were detected with peak detection method as described in Sec. IV, and then the corresponding features were extracted.

B. Evaluation strategy

The accuracy R_A of our PS recognition approach is defined by the combination of true positive (TP) and false negative (FN) measurements [3], [14]:

$$R_A = \frac{N_{\text{TP}} + N_{\text{FN}}}{N_S}, \tag{18}$$

where N_S is the number of samples, N_{TP} represents the true PS our approach recognized, and N_{FN} is the number of true backgrounds discarded.

In addition, to evaluate the generalization abilities of our GBT-SVM classifier, and compare it with other SVM based methods, two famous measurements are utilized, i.e., the *precision* and *recall* rates [15]. As for a good or generalized classifier, both of the precision and recall rates should be large enough. The two measurements are defined as:

$$P = \frac{\overline{TP}}{\overline{TP} + \overline{FP}}, \tag{19}$$

$$R = \frac{\overline{TP}}{\overline{TP} + \overline{FN}}, \tag{20}$$

TABLE II
RESULTS OF THE POTENTIAL PS LOCATED BY THE PEAK DETECTION APPROACH. N_{ref} IS NUMBER OF TRUE PS ACCORDING TO THE REFERENCE GROUP. N_T AND N_F ARE THE NUMBER OF TRUE PS AND FALSE PS DETECTED BY OUR APPROACH AND FALSE DETECTED PS, RESPECTIVELY.

Name (ObsID)	N_{ref}	N_T	N_F	Accuracy (%)
3C 186 (9774)	42	40	13	75.47
MACS J2140.2-2339 (4974)	15	15	4	78.95
NGC 6482 (3218)	13	13	7	65.00
NGC 7619 (3955)	21	20	4	83.33
RCS J1107.3-0523 (5825)	25	25	2	92.59

TABLE III
COMPARISONS AMONG THE CLASSIFICATION RESULTS OF GSVM, BT-SVM, AND GBT-SVM.

Name (ObsID)	Approaches	P (%)	R (%)	Accuracy(%)
3C 186 (9774)	GSVM	61.90	72.22	33.96
	BT-SVM	73.81	93.94	62.26
	GBT-SVM	81.40	87.50	**75.47**
MACS J2140.2-2339 (4974)	GSVM	81.25	92.86	73.68
	BT-SVM	81.25	92.86	73.68
	GBT-SVM	82.35	93.33	**78.95**
NGC 6482 (3218)	GSVM	65.00	100.00	65.00
	BT-SVM	68.42	92.86	70.00
	GBT-SVM	81.25	76.47	**85.00**
NGC 7619 (3955)	GSVM	83.33	100.00	83.33
	BT-SVM	83.33	100.00	83.33
	GBT-SVM	90.91	90.91	**91.67**
RCS J1107.3-0523 (5825)	GSVM	95.24	95.24	77.78
	BT-SVM	92.31	100.00	88.89
	GBT-SVM	96.00	100.00	**92.59**

where P and R are precision and recall rates, respectively; $\overline{TP}$, $\overline{FP}$, $\overline{TN}$, and $\overline{FN}$ are the true positive, false positive, true negative and false negative detections averaged over all the tree-structured submodels, respectively.

C. Experiments and comparisons

Above all, the peak detection approach described in Sec. IV were applied to generate potential PS lists in the X-ray images. And the results were listed in Tab II). Compared with the reference PS, our peak detection algorithm located nearly all of the true point sources, but with some spurious detections.

Therefore, we apply our GBT-SVM classifier to recognize the true PS and to discard the spurious PS for each test observations. The classifier was already trained with the training datasets.

The results are displayed in Tab. III. And the GSVM and BT-SVM classifiers were also applied to recognize the true PS for comparisons. It can be found that, our proposed GBT-SVM classifier achieved the highest detection accuracy and thus had the best performance among all of the methods. Besides, the

precision and recall measurements of the GBT-SVM were also very promising.

In addition, as for GSVM and BTSVM, the accuracy rates of some observations are even less than the accuracy rates only after peak detection. In our opinion, it is because some true point sources were wrongly classified as bright backgrounds, as well as the spurious PS were misjudged as point sources.

Finally, we also combined all the 5 test observations and carried out the comparison, and the detection accuracy of GSVM, BT-SVM and GBT-SVM were 75.56%, 80% and 88.89%, respectively.

VI. CONCLUSION

In this paper we propose a new point source recognition approach for the X-ray astronomical observations. The potential PS are first located by the peak detection approach, and then a newly designed Granular Binary-tree SVM (GBT-SVM) classifier is trained to recognize the true PS with the spurious PS discarded. Our approach not only utilize the spacial features, but also fully exploit the available spectral features of the PS and background. And the comparison results presented in Sec. V highlight that our approach is accurate and has good generalization abilities.

We also compare our classification approach to the other SVM-based approaches, i.e., GSVM and BT-SVM. It shows that our classifier achieved a significantly higher detection accuracy, which proves that our GBT-SVM classifier is accurate and is very suitable for the X-ray PS recognition.

In the further work, we are planning to find approaches to accurately determine the outlines of the PS, so that we are able to better analyze these objects.

ACKNOWLEDGMENT

REFERENCES

[1] P. E. Freeman, V. Kashyap, R. Rosner, and D. Q. Lamb, "A wavelet-based algorithm for the spatial analysis of poisson data," *The Astrophysical Journal*, vol. 138, pp. 185–218, 2002.

[2] F. Guglielmetti, R. Fischer, and V. Dose, "Background-source separation in astronomical images with bayesian probability theory — i. the method," *Monthly Notices of the Royal Astronomical Society*, vol. 396, no. 1, pp. 165–190, 2009.

[3] M. Masias, J. Freixenet, X. Lladó, and M. Peracaula, "A review of source detection approaches in astronomical images," *Monthly Notices of the Royal Astronomical Society*, vol. 422, no. 2, pp. 1674–1689, 2012.

[4] P. S. Broos, L. K. Townsley, E. D. Feigelson, K. V. Getman, F. E. Bauer, and G. P. Garmire, "Innovations in the analysis of chandra-acis observations," *The Astrophysical Journal*, vol. 714, no. 2, p. 1582, 2010.

[5] A. Vikhlinin, W. Forman, C. Jones, and S. Murray, "Rosat extended medium-deep sensitivity survey: Average source spectra," *The Astrophysical Journal*, vol. 451, p. 564, 1995.

[6] C. Cortes and V. Vapnik, "Support-vector networks," *Machine learning*, vol. 20, no. 3, pp. 273–297, 1995.

[7] Y. Tang, B. Jin, Y. Sun, and Y.-Q. Zhang, "Granular support vector machines for medical binary classification problems," in *Computational Intelligence in Bioinformatics and Computational Biology, 2004. CIBCB'04. Proceedings of the 2004 IEEE Symposium on*. IEEE, 2004, pp. 73–78.

[8] Y. Tang and Y.-Q. Zhang, "Granular svm with repetitive undersampling for highly imbalanced protein homology prediction," in *Granular Computing, 2006 IEEE International Conference on*. IEEE, 2006, pp. 457–460.

[9] B. Fei and J. Liu, "Binary tree of svm: a new fast multiclass training and classification algorithm," *Neural Networks, IEEE Transactions on*, vol. 17, no. 3, pp. 696–704, 2006.

[10] C.-C. Chang and C.-J. Lin, "Libsvm: a library for support vector machines," *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 2, no. 3, p. 27, 2011.

[11] Y. Tang, Y.-Q. Zhang, N. V. Chawla, and S. Krasser, "Svms modeling for highly imbalanced classification," *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, vol. 39, no. 1, pp. 281–288, 2009.

[12] K. Tinmmermann and R. Nowak, "Multiscale modeling and estimation of poisson processes with application to photon-limited imaging," *IEEE Transactions on Information Theory*, vol. 45, no. 3, pp. 846–862, April 1999.

[13] A. Fruscione, J. C. McDowell, G. E. Allen, N. S. Brickhouse, D. J. Burke, J. E. Davis, N. Durham, M. Elvis, E. C. Galle, D. E. Harris *et al.*, "Ciao: Chandra's data analysis system," in *SPIE Astronomical Telescopes+ Instrumentation*. International Society for Optics and Photonics, 2006, pp. 62 701V–62 701V.

[14] T. Fawcett, "An introduction to roc analysis," *Pattern recognition letters*, vol. 27, no. 8, pp. 861–874, 2006.

[15] P. Harrington, *Machine learning in action*. Manning, 2012.