An Approach to Detect Cavities in X-Ray Astronomical Images Using Granular Convolutional Neural Networks

Zhixian MA†(a), Student Member, Jie ZHU†, Weitian LI††, and Haiguang XU††, Nonmembers

SUMMARY Detection of cavities in X-ray astronomical images has become a field of interest, since the flourishing studies on black holes and the Active Galactic Nuclei (AGN). In this paper, an approach is proposed to detect cavities in X-ray astronomical images using our newly designed Granular Convolutional Neural Network (GCNN) based classifiers. The raw data are firstly preprocessed to obtain images of the observed objects, i.e., galaxies or galaxy clusters. In each image, pixels are classified into three categories, (1) the faint backgrounds (BKG), (2) the cavity regions (CAV), and (3) the bright central gas regions (CNT). And the sample sets are then generated by dividing large images into subimages with a window size according to the cavities’ scale. Since the number of BKG samples are far more than the other types, to achieve balanced training sets, samples from the major class are split into subsets, i.e., granule. Then a group of three-convolutional-layer granular CNN networks without subsampling layers are designed as the classifiers, and trained with the labeled granular sample sets. Finally, the trained GCNN classifiers are applied to new observations, so as to estimate the cavity regions with a voting strategy and locate them with elliptical profiles on the raw observation images. Experiments and applications of our approach are demonstrated on 40 X-ray astronomical observations retrieved from Chandra Data Archive (CDA). Comparisons among our approach, the β-model fitting and the Unsharp Masking (UM) methods were also performed, which prove our approach was more accurate and robust.

key words: X-ray astronomical image, cavity, detection, granular convolutional neural network (GCNN)

1. Introduction

During the last few decades, telescopes and instruments are launched and constructed, by which researches on the Universe can be conducted. Among them, the study of the central black holes and their host galaxies becomes a hot topic. However, the black holes are usually unobservable because nearly nothing can escape from their huge gravitational field. Thus, a probe is required to detect them, and the Active Galactic Nuclei (AGN) is a perfect one [1].

The AGN is a compact region at the galaxy center, and always has bright luminosity. It radiates highly energetic jets, blows and heats the gas in the interstellar medium around it, and generates bubbles or cavities. This is believed as the result of the accretion of mass by the central black holes in the galaxy [1]. As for such phenomena, they can be observed at the corresponding frequency bands with their energy. For instance, the jets are salient on the radio observations, and the cavities are at the X-ray frequency band. However, the radio telescopes are suffered from low resolutions and interferences from instruments and human activities, thus detection of cavities at the X-ray band becomes a good choice, especially for the observations from the famous high resolution Chandra X-ray Observatory (CXC).

There are three main types of methods to detect cavities in the X-ray astronomical images, and they are the (1) visually detection [2], (2) β-model fitting [3]–[6], and (3) the unsharp masking based approaches [4], [6]–[8]. As for the visually detections, they rely on the observers’ experience, and usually fail at insignificant cavities. With regard to the approach β-model fitting, it fits the central region of the object by a two dimensional function with an elliptical plane view, and subtracts the fitted pattern from the raw images. After subtraction, the cavities are more salient on the residual images. However, there may exist spurious structures, since the objects’ centers are sometimes irregular and asymmetric, and the β function is symmetric. And for the UM methods, the image sharpening is applied. They convolve the raw image with two Gaussian kernels of different variances, and make subtraction or division between the two convolved images, so as to improve the contrast of the target structures [9]. Though widely applied, the UM methods still have a deficit that they are sensitive to noise [6].

Machine learning based methods have been widely applied in many areas, especially for image classification, object detection, and image segmentation. Among them, the Convolutional Neural Network (CNN) based models have been famous in recent years, because the networks can not only perform as classifiers, but also be perfect feature extractors [10]. For instance, Lecun et al. designed the LeNet to recognize handwritten digits [11], Krizhevsky et al. constructed the AlexNet to classify images on very large scale ImageNet datasets [12], and etc. In this work, we apply the CNNs to automatically learn intrinsic features of the cavity regions, segment them from their neighbors, and thus locate them on the raw X-ray astronomical images. The proposed networks combined the three-convolutional-layer LeNet structure [11] with the unpooling strategy proposed by Khalaf et al. in their study on retinal vessel segmentations [13].

With regard to the samples for training the CNN model, we split the full image into patches, and form three categories, (1) the faint backgrounds (BKG), (2) the cavity regions (CAV), and (3) the bright central gas regions (CNT),
according to their luminosity and density. This strategy can not only benefit the detection of cavities, but also provide a large scale of dataset to train the networks.

The cavity and central gas regions usually take up few parts in the images when compared to the backgrounds, which leads to imbalanced dataset and harms the models’ performance. To handle this, many solutions were proposed at the data or classification levels [14], [15]. As for the data level based approaches, they downsample the majority class, or oversample the minority class. The famous approaches are the downsampling based TLINK [15], and oversampling method SMOTE [16]. However, the downsampling may exclude potentially useful data, while training on the oversampled minority dataset may lead to overfitting [15]. With regard to the algorithm level, many approaches were proposed by attaching cost adjusting functions at the classifier outputs [15], or ensemble subclassifiers training on the readjusted datasets [17]. In this work, we take advantage of granular thinking proposed by Tang et al., which separates outputs [15], or ensemble subclassifiers training on the reconstituted training sets [18]. This thinking considers both the data and algorithm levels, which has been successfully applied in our previous work on point sources detection in the X-ray astronomical images [19].

The rest of the paper is organized as follows. In Sect. 2 we explain the cavities in X-ray astronomical images with examples, and the previous cavity detection approaches. After that, in Sect. 3, the preprocessing of the raw data, and preparing of subimage samples are introduced. Then, in Sect. 4, the granular convolutional neural networks of our approach is described, as well as locations of detected cavity regions in Sect. 5. Experiments and comparisons are carried out in Sect. 6 with dicussions. Finally, we conclude in Sect. 7 with outlooks.

2. Cavity and Previous Works

To design cavity detection approach, a better understanding of it is required. In this section, we first explain the properties of the cavities, and then the previous cavity detection approaches.

2.1 Properties of Cavity

In general, cavities exist pairwisely and symmetrically around the objects’ center, where the AGNs locate. Compared with the central gas regions, the cavities are with sparser gas density, and fainter luminosities. In addition, there are also fainter backgrounds outside them, thus the cavities are embedded between the central gas and the backgrounds. By means of these, it is possible to differ the cavities from their neighbors by their density and luminosities.

As a gas bubble blown by the jets from the central black hole, a cavity in the X-ray astronomical image is defined as an elliptical region [1], [4], [6]. And the mathematical model of a cavity is as follows,

$$\begin{align*}
  (x, y) &\leq (x_c, y_c) + \left(\cos \theta - \sin \theta \right) \left( a \cdot \cos \phi - b \cdot \sin \phi \right),
\end{align*}$$

where \((x, y)\) is the coordinate of a pixel in the cavity region, \((x_c, y_c)\) is the center coordinate, \(\theta\) is the orientation angle, and \(a, b\) represent the semi-major and semi-minor axes. \(\phi\) is the free parameter which varies from 0 to \(\pi\).

2.2 Previous Works

There are many works attempt to detect X-ray astronomical cavities. As introduced in Sect. 1, we separate them into three types, (1) visually detection, (2) \(\beta\)-model fitting, and (3) the unsharp masking based approaches. The latter two are widely applied in recent works.

As for the \(\beta\)-model fitting, the \(\beta\) function is as follows,

$$I_\beta(x_r, y_r) = A \left[ 1 + \left( \frac{r}{r_0} \right)^\beta \right],$$

where \(I_\beta\) represents the \(\beta\) fitted image, and \((x_r, y_r)\) is the coordinate of the pixel, which is calculated by Eq. (3). \(A\) is the value w.r.t. the center pixel \((x_c, y_c)\), \(\beta\) is the spectrum index, and \(r = x_r^2 + y_r^2 \cdot (a^2/b^2)\) represents the radial distance from \((x, y)\) to the center, in which \(a\) and \(b\) are the semi-major and semi-minor axes.

$$\begin{align*}
  (x_r, y_r) &= \left(\cos \theta - \sin \theta \right) \left( x - x_c \right) - \left(\sin \theta \cdot \cos \theta \right) \left( y - y_c \right).
\end{align*}$$

Thus, \((A, r_0, a, b, \beta, \theta)\) are the parameters to be fitted. After that, the fitted \(\beta\) image is subtracted from the raw image to generate the residual image.

With regard to unsharp masking, it is a famous image sharpening technique that can increase the contrast of the target scale objects. It convolves the raw image with filters to generate blurred (i.e., unsharp) images as masks, which is then divided or subtracted from the raw image. In Shin et al.’s work [6], they adopted two Gaussian smooth kernels whose \(\sigma\) are 2 and 10 pixels, and the UM image was then obtained by dividing the large scale smoothed image over the small scale image.

On the \(\beta\) fitted residual image and the UM image are cavities then visually located with ellipses. Although both of the above approaches can improve contrast of cavity regions in the images, the visually locating of cavities may still lead biases. As a result, an automatic cavity detection and location approach is required.

3. Data Preparation

To detect the cavities, the raw observed images should be preprocessed so as to prepare sample sets for training the classifiers. In this section, we firstly introduce the procedure of preprocessing, and then the labeling strategy of patchwise sample preparation.

3.1 Preprocessing of Image

The raw images of the galaxies or galaxy clusters are re-
Subimages are defined as follows, as the overlapped pixels between two subimages, and the retrieved from the Chandra Data Archive (CDA), and reprocessed using the CIAO (Chandra Interactive Analysis of Observations) v4.9 by the official guide [20]. Since the cavities usually locate around the central of the objects, to shorten computation time, we only crop the central regions with size of 200 × 200 pixels from the raw images of size around 1400 × 1400 pixels. After that, the small scale point sources (PS) on them are detected by the CIAO tool wavdetect, visually checked, and filled by the CIAO tool dmfilth [19]. Then, the PS removed images are smoothed with a small σ Gaussian filter to avoid the noise. Finally, we obtain the preprocessed images after normalization.

In Fig. 1, we illustrate an example of the preprocessing on object HCG 62 (ObsID: 10462). Figure 1 (1) is the raw observation image, Fig. 1 (2) is the cropped image and Fig. 1 (3) is the preprocessed with cavity regions marked.

3.2 Sample Preparation

Since the CNN based classifier is trained under supervision, thus the samples of cavities and non-cavity regions with labels are required. In this work, we define a strategy that splits the 200 × 200 preprocessed image into small patches (subimages), and label them with a mask.

There are three types of structures in the preprocessed image, (1) the faint backgrounds (BKG), (2) the cavity regions (CAV), and (3) the bright central gas regions (CNT). Thus, we define a mask to label the regions. It is an image as the same size as the preprocessed image, in which pixels in the faint backgrounds are set as zeros, pixels in the cavity regions are set as ones, and the bright gas central pixels are with twos (see Fig. 1 (4) for details).

As for a single sample or patch, in our strategy, it is a subimage from the preprocessed image and labeled with a mask. Denote $w$ as the window size of the patch, and $p_o$ as the overlapped pixels between two subimages, and the subimages are defined as follows,

$$I_m^{p_o} = \begin{pmatrix}
I_{p_o+1, j, p_o+1} & \cdots & I_{p_o+1, j, p_o+w} \\
\vdots & \ddots & \vdots \\
I_{p_o+w, j, p_o+1} & \cdots & I_{p_o+w, j, p_o+w}
\end{pmatrix}, \tag{4}
$$

where $I_m^{p_o}$ is the patch matrix, i.e., the subimage, $p_d = w - p_o$ is the shifting pixels between two subimages, and $m = i * p_d + j$ is the corresponding index. $i = 0, 1, \ldots, M_r - 1$ and $j = 0, 1, \ldots, M_c - 1$, where $M_r, M_c$ are the number of patches per row and column.

The number of subimages $M$ for one observation is calculated by,

$$M = M_r \cdot M_c = (\lfloor \frac{N_r - w - 1}{p_d} \rfloor + 1) \cdot (\lfloor \frac{N_c - w - 1}{p_d} \rfloor + 1), \tag{5}\text{where } N_r \text{ and } N_c \text{ are row and column sizes of the preprocessed image, and the symbol } \lfloor \cdot \rfloor \text{ means rounding.}
$$

With regard to the label of the sample, it is calculated as follows,

$$L_m = L[\max(N_{BKG}^m, N_{CAV}^m, N_{CNT}^m)] \tag{6}\text{Where } L_m, m = 1, 2, \ldots, M \text{ represents the label of sample } m, \text{ and } N_{BKG}^m, N_{CAV}^m, N_{CNT}^m \text{ are the number of pixels belonging to the three region types. } L(\cdot) \text{ means the corresponding label of this type, which belongs to } [0, 1, 2].$$

In this work, we carefully filtered 40 observations from Shin et al.’s [6] and Dong et al.’s [4] works. In summary, the observations are with an average exposure time of 46.23 ks, and possesses around two cavities in each. The smallest scale of those cavities is around 10 px, thus we set window size $w$ of the samples as 10. The overlapped pixels $p_o$ is set as 5, which is adjustable in real applications. And for a 200 × 200 image, there are 1521 samples of it.

4. Design of the Classifier

After preparation of the samples, in this section, the cavity detection classifiers that use the convolutional neural networks are designed. We firstly introduce the classifiers’ architectures, and then the handling of imbalanced data.

4.1 Classifier Architecture

The designed GCNN classifier is as Fig. 2 illustrated, in
The frame of the designed Granular Convolutional Neural Network (GCNN) based classifier. It consists of five layers, in which are three convolutional layers without pooling layer, a fully connected layer, and a softmax output layer with three units. The three units or classifications are, (1) the faint backgrounds (BKG), (2) the cavity regions (CAV), and (3) the bright central gas regions (CNT).

Table 1  The structure of the designed CNN based cavity classifier. $C_i, i = 1, 2, 3$ represent the three convolutional layers.

<table>
<thead>
<tr>
<th>Layer</th>
<th>structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input</td>
<td>a set of images of size $10 \times 10$</td>
</tr>
<tr>
<td>$C_1$</td>
<td>kernel size: $2 \times 2$, 15 kernels</td>
</tr>
<tr>
<td>$C_2$</td>
<td>kernel size: $3 \times 3$, 15 kernels</td>
</tr>
<tr>
<td>$C_3$</td>
<td>kernel size: $4 \times 4$, 15 kernels</td>
</tr>
<tr>
<td>Fully connected (FC)</td>
<td>a vector with 240 nodes</td>
</tr>
<tr>
<td>Output</td>
<td>a softmax layer of 3 units</td>
</tr>
</tbody>
</table>

which there are three convolutional layers for hierarchical feature extraction, and a fully connected layer for label estimating. After each layer, the rectified linear unit (ReLU) is formed as the activation function, which is non-saturating and can increase the convergence speed [12], [21], [22]. The number of kernels, as well as the kernel sizes of the convolutional layers are listed in Table 1.

Since our target is to locate the cavity regions from their neighbors, and the regions are shift variant. To avoid introducing biases, though widely applied in many CNN based networks, we do not cascade subsampling or pooling layers after the convolutional layers and follow the structure proposed by Khalaf et al. [13]. In addition, the sample patches defined in Sect. 3.2 are only with tens of pixels, the computational expenses may not be significantly affected by removing the subsampling layers.

4.2 Imbalanced Dataset

Since the background samples are far more than the other classes, which leads to imbalanced datasets. To handle it, we divide the majority background samples into multiple subsets (i.e., granule), in which the number of samples are approximately equal to the minority cavity samples and the central gas samples [18], [19].

Denote the $N_{\text{sub}}$ as the number of subsets of the majority class, and it is defined as follows,

$$N_{\text{sub}} = \left\lfloor \frac{N_{\text{maj}}}{N_{\text{min}}} \right\rfloor,$$

where $N_{\text{maj}}$ and $N_{\text{min}}$ are the major and minor sample sets, and $\lfloor \cdot \rfloor$ means rounding down.

Since the sample patches are generated from different observations with different exposure times and angular distances, we do not separate the major set by randomly selecting samples. Instead, we uniformly downsample the major sets so as to cover all the observations in each granula (see Eq. (8)).

$$S^{k}_{\text{maj}} = \{S^{k+i\cdot N_{\text{sub}}}_{\text{maj}}\},$$

$$i = 1, 2, \ldots, N_{\text{min}}, k = 1, 2, \ldots, N_{\text{sub}},$$

where $S^{k}_{\text{maj}}, k = 1, 2, \ldots, N_{\text{sub}}$ means the subsample sets. With regard to each one, it combines a subset of major samples and the whole cavity and central gas samples.

Finally, the label of a single sample is decided after estimating it with all the sub-classifiers, and a voting strategy is applied, by which the label obtaining most of the tickets wins.

5. Location of Cavity Regions

After training of the classifiers, the cavities can be detected and located on the raw images. In this section, an algorithm that locates the detected cavities is proposed.

As for a new observation, it should be split into patches with overlaps as Sect. 3.2 defined, and input into the GCNN classifiers. Then the potential cavities can be located by a reunited image called the reunited mask image. In this image, as the same as the mask image defined in Sect. 3.2, pixels in cavity regions possess value one, but pixels belonging to the rest two categorizes are all set as zero.

Since the patches are overlapped, labels of pixels in the overlapped regions should be particularly considered. In this algorithm, the estimated label of a single patch is reset with Boolean value and spreaded pixelwisely. This means if this subimage was estimated as CAV, then pixels in it would be set with Boolean value 1 (i.e., True), and if estimated as BKG or CNT, pixels would be set with Boolean 0 (i.e., False). After that, the reunited mask image is obtained by combining the reset patches, in which labels of the overlapped pixels are calculated by binary addition (See Alg. 1.
Algorithm 1 Reunited mask image construction algorithm

1: **Input**: patches $I_w$ and estimated labels $L_m$, $m = 0, \ldots, M - 1$
2: **Input**: $w, p_d, N_p, \text{and} \ N_w$
3: **Initialize reunited mask image**: $I_0 = \text{bool}(zeros(N_p, N_w))$
4: $M_0 = \text{round}(N_p - w - 1/p_d) + 1$
5: $M_1 = \text{round}(N_w - w - 1/p_d) + 1$
6: for $i = 0 : M_0 - 1$
7: for $j = 0 : M_1 - 1$
8: if $L_i == 1$ then
9: \[ I_{iM_0}[1:w,1:w] = \text{bool}(1) \]
10: else
11: \[ I_{iM_0}[1:w,1:w] = \text{bool}(0) \]
12: end if
13: \[ I_{iM_0}[1:p_d : i \cdot p_d + w, j \cdot p_d : j \cdot p_d + w] += I_{iM_1} \]
14: end for
15: end for
16: **Output**: the reunited mask image $I_0$

for details).

On the reunited mask image, the connected domains [23], i.e. cavity regions, are then detected, and parameters w.r.t. the cavities defined by Eq. (1) are estimated. Finally, the obtained locations of the cavities can be marked on the raw images.

6. Experiments and Result

Experiments on real X-ray astronomical observations were carried out to demonstrate the performance of our proposed GCNN classifiers. And comparisons of our approach with $eta$-model fitting and unsharp masking are performed, as well as the methods to handle imbalanced dataset.

As introduced in Sect. 3, 40 X-ray observations were selected. In each of them, regions of the cavities were located by combination of Shin et al.’s [6], Dong et al.’s [4] results and the experts’ visually checking. And regions faint backgrounds and the bright central gas regions were decided according to the fitted $eta$ function of the observations. We randomly selected 35 observations to generate the training and validation datasets, and the rest 5 were as the test dataset to evaluate performance of our proposed approach. The corresponding numbers of BKG, CAV, and CNT samples were 45617, 6114, and 9109 respectively. Thus 7 granular subclassing sets were formed (see Eq. (7)).

The CNN based granular subclassifiers were formed as the same as Fig. 2 defined, and trained by the training subsets above. In this work, as for each subclassifier, the granular subset was randomly divided into training and validation sets with a ratio of 4 : 1. 500 times of training epochs were conducted, and batch gradient descent was applied to adjust the network parameters with cross entropy loss. The computational time of each epoch was around 1.33 seconds. Thus training for all of the 7 GCNNs cost around 1.29 hours.

The training process was run on a computer with an Intel Core i5 3.10 GHz CPU, 8G DDR3 RAM and an NVIDIA GT730 1G GPU. And our code was guided by NVIDIA CUDA documentation [24] and written according to two Python packages namely Theano and Lasagne.

Estimated labels of the split test subimages were generated by the trained GCNN classifiers, thus we can locate the cavities on the 5 test observations. The reunited mask images were recovered as Alg. 1 introduced, and the respective connected domains were detected. Corresponding parameters of the elliptical cavities were then estimated. Finally, we mark the results on the raw observations.

Since the main aim is to locate the cavity regions in the X-ray observations, the performance of the proposed approach can be assessed both at piecewise and pixelwise. As for the piecewise, the errors among parameters of estimated cavities and their nearest correct cavities are calculated. We denote them as $(E_R, E_{\text{can}}, E_{\text{maj}}, E_{\text{min}}, E_{\text{rot}})$, which means errors of cavity numbers, center distances, semi-major and semi-minor axes, and the rotation angles.

With regard to pixelwise performance, three indices namely accuracy, sensitivity, and specificity are applied [13], which are defined as follows,

\[ R_{\text{acc}} = \frac{TP}{(TP + FN)} \]
\[ R_{\text{sen}} = \frac{TP}{(TN + TP + FN)} \]
\[ R_{\text{spe}} = \frac{TN}{(TN + FP)} \]
\[ R_{\text{acc}} = \frac{(TP + TN)}{N} \]

where $N$ represents the total number of pixels in an observation. $N_{TP}$ and $N_{TN}$ are the truly detected cavity and non-cavity pixels. $N_{FP}$ and $N_{FN}$ are the misclassified cavity and non-cavity pixels.

To compare the proposed approach with the other two methods, cavities were also detected by them. As for the $eta$-model fitting approach, the $eta$ functions of the five test observations were estimated, and the residual maps were generated as well. With regard to the unsharp masking, following Shin et al.’s [6] work, the adopted smooth kernels of the small and large scale are set to 2 and 10 pixels. Then the UM images are obtained by dividing the large scale smoothed images over the small scale images. For both of them, cavities are then visually detected and marked with ellipses.

In addition, comparisons of the proposed granular-based imbalanced data handling strategy with the other three approaches namely SMOTE [16], Randomly Down-Sampling (RDS) [15], and Threshold Adjusting (TA) were also demonstrated, so as to prove the efficiency of our approach. For the SMOTE and RDS, they only did data adjusting, and the classifiers were as the same as the GCNN. As for the TA, the proportions of CAV, CNT, and BKG samples were considered and attached as a cost function at the output of the CNN network.

The results of the 5 test observations by the proposed approach are displayed in Fig. 3, as well as the compared two methods. The performances of them were also assessed and listed in Table 2, as well as results of comparisons on the imbalanced data handling approaches in Table 3.

Since that, some summaries are obtained as follows,

- In all of the five test observations, detected cavities located near the objects’ central regions, and existed pair-
Fig. 3  The test results of proposed cavity detection approach, the $\beta$-model fitted and the unsharp masking method. From left to right are objects NGC 1316 (ObsID: 2022), NGC 4104 (ObsID: 6939), NGC 4552 (ObsID: 2072), NGC 4636 (ObsID: 323), and PKS 2153-69 (ObsID: 16084). From up to bottom are the raw cropped 200×200 images, the mask images, the reunited mask images, the preprocessed images, the $\beta$-model fitted residual images, and the unsharp masked images. Detected cavities by our approach, $\beta$-model fitting, and UM are marked with green, red, and yellow ellipses.

- In summary, the proposed approach automatically detected most of the cavities, and achieved mostly the best performances than the other methods both at pixelwise and piecewise, especially for the sensitivity and accuracy.
- Though the specificities of all the three methods were over 95%, the sensitivities were not satisfied, especially for the object NGC 4104. In our view, it is because the cavity region taken up very small parts, thus the posi-
Table 2 Evaluations and comparisons among the proposed GCNN approach, β-model fitting, and Unsharp Masking (UM). $R_{sen}$, $R_{spe}$, and $R_{acc}$ represent the three pixelwise assessment indices namely sensitivity, specificity, and accuracy.

<table>
<thead>
<tr>
<th>Name (ObsID)</th>
<th>Approaches</th>
<th>$R_{sen}$ (%)</th>
<th>$R_{spe}$ (%)</th>
<th>$R_{acc}$ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NGC 1316 (2022)</td>
<td>β-model</td>
<td>28.59</td>
<td>98.57</td>
<td>96.79</td>
</tr>
<tr>
<td></td>
<td>UM</td>
<td>32.52</td>
<td>98.67</td>
<td>96.99</td>
</tr>
<tr>
<td></td>
<td>GCNN</td>
<td><strong>73.48</strong></td>
<td><strong>99.28</strong></td>
<td><strong>98.79</strong></td>
</tr>
<tr>
<td>NGC 4104 (6939)</td>
<td>β-model</td>
<td>3.40</td>
<td>99.97</td>
<td>98.57</td>
</tr>
<tr>
<td></td>
<td>UM</td>
<td>10.77</td>
<td>99.99</td>
<td>99.57</td>
</tr>
<tr>
<td></td>
<td>GCNN</td>
<td>10.95</td>
<td>99.98</td>
<td><strong>99.68</strong></td>
</tr>
<tr>
<td>NGC 4552 (2072)</td>
<td>β-model</td>
<td>59.89</td>
<td>98.57</td>
<td>97.85</td>
</tr>
<tr>
<td></td>
<td>UM</td>
<td>89.38</td>
<td>99.37</td>
<td>99.17</td>
</tr>
<tr>
<td></td>
<td>GCNN</td>
<td><strong>94.38</strong></td>
<td>99.03</td>
<td>98.96</td>
</tr>
<tr>
<td>NGC 4636 (323)</td>
<td>β-model</td>
<td>40.35</td>
<td>69.39</td>
<td>68.56</td>
</tr>
<tr>
<td></td>
<td>UM</td>
<td>36.20</td>
<td>69.39</td>
<td>67.73</td>
</tr>
<tr>
<td></td>
<td>GCNN</td>
<td><strong>95.45</strong></td>
<td>86.76</td>
<td><strong>88.63</strong></td>
</tr>
<tr>
<td>PKS 2153-69</td>
<td>β-model</td>
<td>80.40</td>
<td>99.85</td>
<td>99.66</td>
</tr>
<tr>
<td></td>
<td>UM</td>
<td>22.79</td>
<td>99.81</td>
<td>97.24</td>
</tr>
<tr>
<td></td>
<td>GCNN</td>
<td><strong>86.79</strong></td>
<td>99.57</td>
<td>99.50</td>
</tr>
</tbody>
</table>

Table 3 Comparisons among the proposed GCNN, β-model fitting, and the UM. The five indices are the errors of cavities $E_n$, center distances $E_{cn}$, semi-major and semi-minor axes $E_{maj}, E_{min}$, and the rotation angles $E_{rot}$.

<table>
<thead>
<tr>
<th>Approaches</th>
<th>$E_n$</th>
<th>$E_{cn}$ (px)</th>
<th>$E_{maj}$ (px)</th>
<th>$E_{min}$ (px)</th>
<th>$E_{rot}$ (°)</th>
</tr>
</thead>
<tbody>
<tr>
<td>β-model</td>
<td>$1.00 \pm 0.71$</td>
<td>$0.40 \pm 0.55$</td>
<td>$13.83 \pm 16.15$</td>
<td>$16.65 \pm 17.11$</td>
<td>$8.84 \pm 8.92$</td>
</tr>
<tr>
<td>UM</td>
<td>$32.43 \pm 18.24$</td>
<td>$93.45 \pm 13.46$</td>
<td>$38.06 \pm 21.17$</td>
<td>$93.72 \pm 0.19$</td>
<td>$2.36 \pm 11.84$</td>
</tr>
<tr>
<td>SMOTE</td>
<td>$50.75 \pm 30.01$</td>
<td>$96.54 \pm 5.60$</td>
<td>$95.44 \pm 7.05$</td>
<td>$95.67 \pm 6.69$</td>
<td>$95.57 \pm 7.05$</td>
</tr>
<tr>
<td>RDS</td>
<td>$48.19 \pm 29.74$</td>
<td>$0.49 \pm 0.24$</td>
<td>$96.89 \pm 5.42$</td>
<td>$97.32 \pm 4.80$</td>
<td>$97.18 \pm 4.78$</td>
</tr>
<tr>
<td>TA</td>
<td>$72.21 \pm 35.35$</td>
<td>$93.23 \pm 13.35$</td>
<td>$92.29 \pm 13.30$</td>
<td>$92.29 \pm 13.30$</td>
<td>$92.29 \pm 13.30$</td>
</tr>
<tr>
<td>GCNN</td>
<td>$72.21 \pm 35.35$</td>
<td>$93.23 \pm 13.35$</td>
<td>$92.29 \pm 13.30$</td>
<td>$92.29 \pm 13.30$</td>
<td>$92.29 \pm 13.30$</td>
</tr>
</tbody>
</table>

detected a large cavity locating at the northwest. However, this may be a spurious detection. Because there were point sources in this area (see the raw image of NGC 1316 in Fig. 3), and removing of them in the processing procedure could leave false structures.

- There were some misclassified small scale regions in NGC 4636, which led to the largest errors on detected cavity numbers of our approach. In our view, this can be solved by visually checking according to the properties of cavities on the detected results. In addition, this could also be solved by training the GCNN networks with larger datasets, i.e., more observations with correct recognized cavities.

- The granular-based imbalanced data handling strategy outperformed the other approaches both at accuracy and robustness. In addition, the performance of TA and RBS were also better than β-model fitting and UM, which proved the CNN based classifiers are effective.

### 7. Conclusion

In this paper, we propose an approach to automatically detect cavities in the X-ray astronomical images. The raw observations are firstly preprocessed and separated into subimage patches to generate the samples, as well as the labels by generated mask images. After that, to handle the imbalanced dataset, the majority background samples are uniformly divided into subsets to form training granule. Correspondingly, a group of three-convolutional-layer granular CNN (GCNN) networks without pooling layers are constructed as the cavity classifiers. Then the subclassifiers are trained and tested by real X-ray observations.

The results and comparisons of our approach to the other usual cavity detection approaches namely β-model fitting, and unsharp masking are presented in Sect. 6, which highlight that our approach is accurate and robust, especially for observations with irregular and asymmetrical centers, and low SNRs. We also compared the proposed imbalanced data handling strategy with the other sampling and cost-sensitive based approaches. It shows that our approach achieved significantly higher performances, which prove that our GCNN based classifiers are accurate and suitable for the X-ray cavity detection cases.

In the future work, we are going to detect more objects with cavities by the proposed approach, and train the GCNN classifiers iteratively. And the detections of other structures like the central gas profile, and point-like sources are also our interest.

### Acknowledgments

This work is supported by the National Natural Science Foundation of China (grant Nos. 61371147 and 11433002).

### References

Zhixian Ma received his B.S. degree in telecommunication engineering in 2014 from Xidian University, Xi’an, China. He is a Ph.D. candidate in Shanghai Jiao Tong University, China, he works in Department of Electronic Engineering, Shanghai Jiao Tong University, China. His research interests include image processing, signal processing, etc.

Jie Zhu received his Ph.D. degree in information system from Shanghai Jiao Tong University, China. Since 1999, he works as a professor in Department of Electronic Engineering, Shanghai Jiao Tong University, China. He was the chairman of IEICE Shanghai section. His research interests include speech signal processing, multimedia system, etc.

Weitian Li received his B.S. degree in applied physics in 2013 from Shanghai Jiao Tong University, China. He is currently working towards a Ph.D. degree in the Department of Physics and Astronomy in Shanghai Jiao Tong University. His research interests include evolution of galaxy clusters in X-ray and low-frequency radio bands, and astronomical image processing.
Haiguan Xu received his Ph.D. degree in theoretical physics in 1998 from Shanghai Jiao Tong University, China. He also received a Ph.D. degree in astrophysics in 1997 from the University of Tokyo, Japan. Since 2002, he is a professor in the Department of Physics and Astronomy in Shanghai Jiao Tong University. His research interests include evolution of galaxy clusters, large-scale structures of the Universe, and low-frequency radio astronomy.