AIRPLANE DETECTION IN REMOTE SENSING IMAGES BASED ON OBJECT PROPOSAL

*Qinhan Luo, Zhenwei Shi**

Image Processing Center,School of Astronautics, Beihang University, Beijing 100191.
State Key Laboratory of Virtual Reality Technology and Systems, Beihang University, Beijing 100191.
Beijing Key Laboratory of Digital Media, Beihang University, Beijing 100191, PR China

ABSTRACT

Automatic detection of airplanes in remote sensing images (RSIs) remains a challenge. Its primary problem is how to locate the airplanes from the huge searching space of the image in an efficient way. In this paper, we utilize a simple but effective technology, Object Proposal, for airplane locating. The main objective of the technology is to generate a relatively small set of bounding boxes that most likely contain objects of interest. In our approach, a small set of bounding boxes that most likely contain the airplanes are first generated by the Object Proposal algorithm. Afterwards, a SVM classifier is trained on the HOG features to detect the airplanes. Finally, the trained object detector is applied to those bounding boxes instead of exhaustive search to complete the detection task. Experiments show that our Object Proposal method is effective in its ability of producing good quality proposals. It can be further utilized for the detection task to reduce the computation cost.

***Index Terms*—** airplane detection, Object Proposal, HOG-SVM

1. INTRODUCTION

Airplane detection in RSIs consists in determining whether a given image contains airplanes and if so, exporting the airplanes' bounding boxes. The task can be viewed as a binary classification problem, where each window of the image is assigned a target or non-target label. It is similar to the object detection task in natural images. However, the huge searching space of the remote sense image make it intolerable using the simple sliding window paradigm, in which a classifier tests for object presence in each candidate image window.

In early works [1, 2, 3, 4, 5, 6, 7, 8], this challenging task has been tackled with a two-stage paradigm which can be grouped as candidate region extraction and airplane classification. In the first stage, the candidate region is extracted utilizing various priori information of airplanes. Among these methods, some works [1, 2] consider airplanes to be salient in the images. Some works [3, 4, 5] formulate a simple hypothesis that airplanes have bright gray values. On these conditions, the threshold values are chosen manually to get the candidate region. Although these methods are effective in their provided images, these methods may lack flexibility and practicability. Apart from these methods, An et al. [6, 7, 8] cleverly takes advantage of the shape of airplanes. In the stage of locating airplanes the circle filter response is first computed, then the local maximums are taken as the centers of airplanes. The method show good performance on locating airplanes, but it cannot give out the objects' bounding boxes. In the second stage, decisions are made on each candidate region and their states are further determined by analyzing some discriminative features. In general, previous works of this stage mainly focus on template matching methods [2, 5] and machine learning based methods [1, 3, 4, 6, 7, 8].

General airplane detection should consider the complex environment and variations of airplanes' type, pose, and color. Although the above works give out some solutions, it is still unsolved to find a generally applicable location method which can give a relatively small set of bounding boxes that will not miss airplane targets in the remote sense image in the first stage.

Recently, Object Proposal [9, 10, 11, 13, 12], a new computer vision technology, is widely used for the task of object detection. The main interest of such methods is their ability to speed up recognition pipelines that make use of complex and expensive classifiers by considering only a few thousands of bounding boxes [9]. Its main objective is shown as in Fig. 1. Motivated by the success of Object Proposal technology in the natural images, it emerge as a fine solution for the task of airplane detection in RSIs.

In this paper we explore the applicability of Object Proposal in airplane detection of RSIs. We propose a simple class-specific object proposal strategy for airplane detection. Moreover we demonstrate that our proposal method is effective in producing good airplane proposals and can be further utilized for airplane detection task by cascade a simple classifier.

*The work was supported by the National Natural Science Foundation of China under the Grants 61273245, the Beijing Natural Science Foundation under the Grant 4152031, the funding project of State Key Laboratory of Virtual Reality Technology and Systems, Beihang University under the Grant BUAA-VR-16ZZ-03, and the Fundamental Research Funds for the Central Universities under the Grant YWF-14-YHXY-028 and YWF-15-YHXY-003.

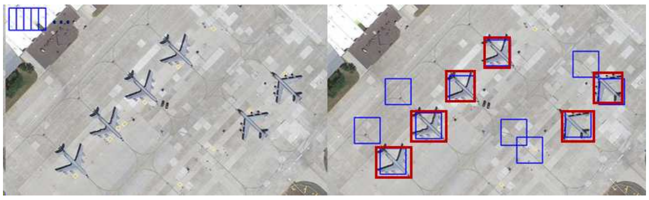

Fig. 1. Sliding a window for all possible locations, sizes, and aspect ratios is a challenge for the cost of computation. The proposal windows generated with our airplane specific object propose method provide high-quality coverage of airplanes in RSIs.

2. OUR APPROACH

The method proposed in this paper builds on the work of [10, 11], where an objectness value, defined as how likely a detection window covers a generic object, is first computed, then the windows with high scores will be extracted as the propose windows.

Our method is based on the fact that most airplanes in the RSIs appear big differences in size, color, orientation but share a common character that they all have well-defined closed boundaries. We observe that the windows which contain an airplane can be discriminated from the windows without an object by the simple gradient feature when resized into a small fixed size (e.g. 16×16). The algorithm is divided into two main steps: proposal generation, proposal verification.

2.1. Proposal generation

In this stage, the goal is to extract the windows that most likely contain airplanes. A score, representing how likely the window contains the airplane, is first computed in each window of the given image. Then the window with high score will be extracted as the propose windows. For the necessity to consider all the windows. It confines the use of the complicate features in the image. Like [11], we only use the simple norm gradient feature to measure the window. We learn a model from the samples of airplanes in a fixed small size (e.g.16×16) in advance. To solve the variation of targets size, the input image is then resized into different scales so that the targets in the resized image may correspond to the learned samples. Thus the score is defined as:

$$\mathbf{s}_l = \langle \mathbf{w}, \mathbf{g}_l \rangle, \tag{1}$$

$$l = (i, x, y), \tag{2}$$

Where,$\mathbf{s}_l, \mathbf{g}_l, i, (x, y)$ are filter score, norm gradient feature of a window, size and position of a window. The windows with high scores in each scale will be extracted as the proposal window.

2.1.1. *Learning the model*

In this stage, we collect 3340 airplanes (indicated by the bounding boxes) with various sizes, orientations and colors from the google earth. Each sample is resized into a small size (e.g.16 $\times$ 16) and the simple norm gradient feature is computed.

In the processing of learning, we learn the model using the one-class SVM [14] tentatively. During our training only the positive samples, airplane, are considered. The motivate is that the background takes a dominate role in the searching space of remote sensing image and has a mixed and disorder distribution in the feature space. We tend to extract the windows most likely containing the airplanes despite the potential false-alarms it may produce. Our implementation is builded on the libSVM [14]. The learned $\mathbf{w}$ (Fig. 2), seems to represent some appearance of the airplanes.

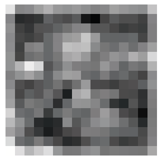

Fig. 2. The $\mathbf{w}$ learned by the one-class SVM.

2.1.2. *Generating the proposal windows*

After a filter implementation using the learned $\mathbf{w}$, we get a score response in each resized image (Fig. 3). The value represents the possibility of containing an airplane in its corresponding window. Using non-maximal suppression (NMS) [15], we select a set of proposals in each scale. The algorithm is described in Alg. 1. Where $\mathbf{G}$ is the norm gradient of the image, k, m, n represent the image's scale change, half the size of $\mathbf{w}$, the numbers of proposals we hope to extract. The $imfliter, Inf$ follow the MATLAB function. The $MaximunmAt$ function returns the position of the maximum in the response of the filtered image.

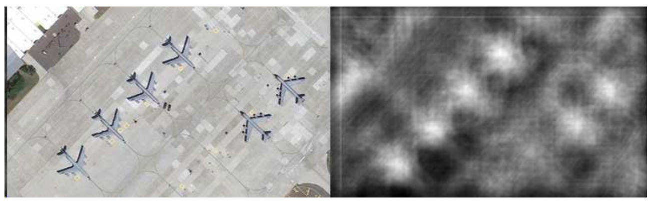

Fig. 3. The filter response. The window containing the airplanes have higher response than the background.

2.2. Proposal verification

To further detect the airplanes, we design a proposal verification stage. In this stage,we train a linear SVM on HOG fea-

Algorithm 1 Proposal generating algorithm

Input : $\mathbf{G}, \mathbf{w}, k, m, n$
Output : Box
$\mathbf{E} \leftarrow imfilter(\mathbf{G}, \mathbf{w})$;
  For $i = 1 : n$
    $(x_i, y_i) \leftarrow MaximumAt(E)$;
    $Box(i,:) \leftarrow (y_i - m, x_i - m, y_i + m, x_i + m)$;
    $E(x_i - m : x_i + m, y_i - m : y_i + m) = -Inf$;
  end For
  $Box = k * Box$;
Return Box;

tures. For the changes in orientation and size, we divide the training samples into 32 orientations and the same size(e.g. 40 $\times$ 40). In each orientation, a SVM is learned. Then the proposals generated from the first stage will be removed if have negative response in all the 32 classifiers. Part of our training data is shown in Fig. 4.

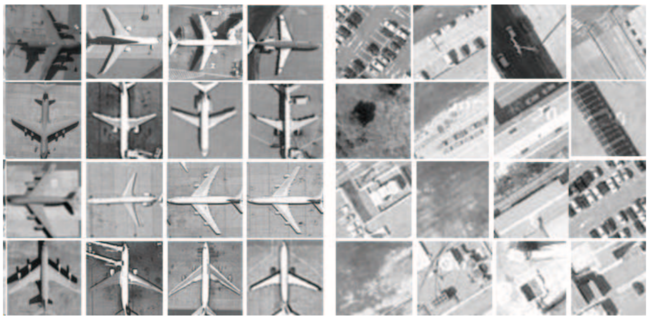

Fig. 4. Part of the data for training the SVM.

3. EXPERIMENTS AND RESULTS

In this section, we use two experiments to evaluate the performance of our method. The test data is collected from the google earth. It includes 20 images of complex airport scenes. The image size varies from 1000 $\times$ 1000 to 2000 $\times$ 2000.

3.1. Performance of the Object Proposal

The evaluation framework we used is the standard for Object Proposals methods [9]. It is based on the analysis of the detection recall achieved by a given method under certain conditions. Recall is calculated as $recall = \frac{N_d}{N_{gt}}$, where N_d represents number of the detected airplanes,N_{gt} represents the total number of airplanes. The propose windows will be considered as true only when it have a large value (e.g. 0.5) of the intersection over union (IoU) with the Ground Truth. As shown in Fig. 5. Our method can get a high $recall$ just considering the 200 proposals in each size.

By contrast, we adopt the same strategy as [3] in which the bounding boxes are extracted from the segmented region

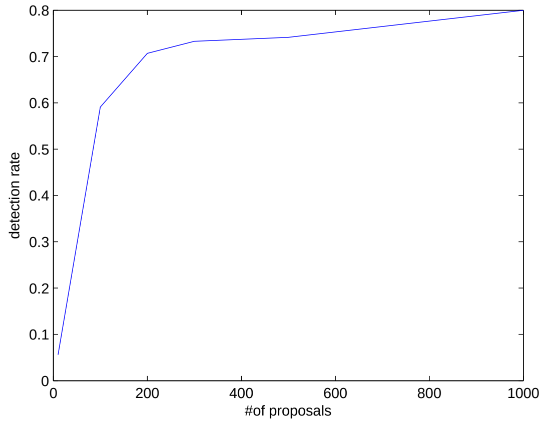

Fig. 5. Performance of our method at 0.5 IoU.

to state the effects of the Object Proposal method. In our experiments, we utilize the algorithm in [13]. Then we get the $recall = 0.2$ on the whole data, such method may be confronted with problems when the airplane' size become large and the aircraft-body'color is different from that of the wings. Fig. 6 show such a case.

Fig. 6. Examples approach [3] cannot solve. The candidate region extracted from segment is difficult on how to combine the airplane-body with the wings.

3.2. Efficiency of the detector

To evaluate the effects of Object Proposal on the detector. We use a sliding-window paradigm. At first, we choose 100 proposal windows using our method in a suitable scale of the image, then we detect the airplanes on these windows. On the other hand, we slide windows to detect the airplanes. These two approaches behave the same on the detection results but Object Proposal really have huge advantage on the computation cost for the target detection task (Tab. 1). The result is shown in Fig. 7.

Table 1. Computation Time.

Method	Windows	Time(s)
Sliding Windows	321984	878
Our Method	100	0.615

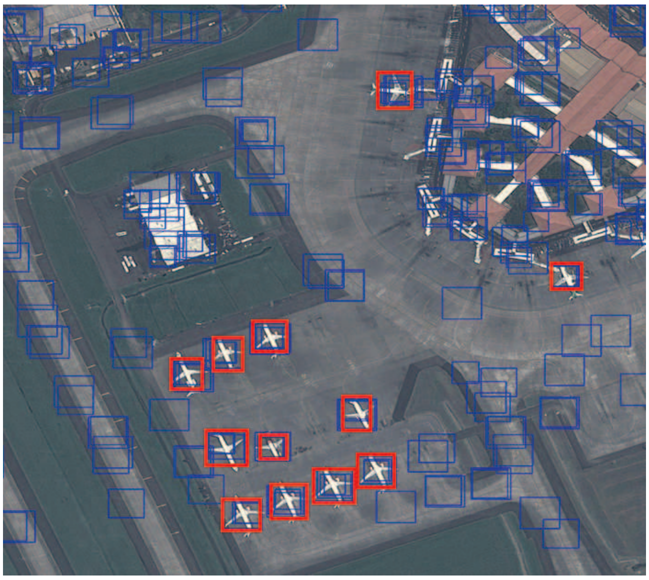

Fig. 7. A result of detection. Blue box: the proposal windows; Red box: the proposal windows after verification.

4. CONCLUSION

We propose a new method for locating the airplanes in RSIs based on Object Proposal. This proposed method can give out a small set of bounding box that most likely contain airplanes in the RSIs and show effective in the complex scenes. Based on the Object Propose technology we train a HOG-SVM classifier to further detect airplanes. This approach show good performance on reducing the computation for the task of target detection.

5. REFERENCES

[1] J. W. Han,P .C. Zhou, and J. Wu, "Efficient, simultaneous detection of multi-class geospatial targets based on visual saliency modeling and discriminative learning of sparse coding," *Journal of Photogrammetry and Remote Sensing*. vol.89,pp. 37-48,2014.

[2] W. Li, S. M. Xiang, H. B. Wang and C. H. Pan,"Robust airplane detection in satellite images," *International Conference on Image Processing*, pp. 2821-2824, 2011.

[3] S. K. Bo, and Y. J. Jing, "Region-based airplane detection in remotely sensed imagery," *Image and Signal Processing*, vol. 4, pp. 1923-1926, 2010.

[4] X. Y. Cheng, S. M. Xiang, C. L. Liu and C. H. Pan,"Aircraft Detection by Deep Belief Nets," *2013 2nd IAPR Asian Conference on Pattern Recognition*, pp.54-58, 2013.

[5] H. Sun, X. Sun, H. Q. Wang, Y. Li and X. J. Li "Automatic Target Detection in High-Resolution Remote Sensing Images Using a Contour-Based Spatial Model," *IEEE Geoscience and Remote Sensing Letters*, vol. 9, no. 5, pp. 886-890, 2012.

[6] Z. Y. An, Z. W. Shi, X. C. Teng, X. R. Yu and W. T "An automated airplane detection system for large panchromatic image with high spatial resolution," *Optik - International Journal for Light and Electron Optics*, vol. 125, pp. 2768-2775, 2014.

[7] L. Liu and Z. W. Shi, "Airplane detection based on rotation invariant and sparse coding in remote sensing images," *Optik - International Journal for Light and Electron Optics*, Vol. 125, Issue 18, pp. 5327-5333, September 2014.

[8] Z. Y. An and Z. W. Shi, " An airplane detection method for panchromatic image," *2013 IEEE China Summit and International Conference on Signal and Information Processing (ChinaSIP)*, pp. 189-192, July 2013.

[9] J. Hosang, R. Benenson, P. Dollar and B. Schiele "What makes for effective detection proposals?," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp.1-1, 2015.

[10] B. Alexe, T. Deselarers, and V. Ferrari,"What is an object?,"*IEEE,CVPR*, pp. 73-80, 2010.

[11] M. M. Cheng, Z. M. Zhang, W. Y. Lin and P. Torr, "BING: Binarized Normed Gradients for Objectness Estimation at 300fps," *IEEE,CVPR*, pp. 3286-3293, 2014.

[12] J. R. Uijlings, K. E. van de Sande, T. Gevers, and A. W. Smeulders, "Selective Search for Object Recognition," *IJCV*, vol. 104, no. 2, pp. 154-171, 2013.

[13] C. L. Zitnick and P. Dollar, "Edge Boxes: Locating Object Proposals from Edges," *ECCV*,2014

[14] R.-E. Fan, K.-W. Chang, C.-J. Hsieh, X.-R. Wang, and C.-J. Lin, "Liblinear: A library for large linear classification," *The Journal of Machine Learning Research*,vol. 9, pp 1871-1874, 2008

[15] A. Neubeck and L. Van, "Efficient Non-Maximum Suppression," *18th International Conference on Pattern Recognition*,vol. 3, pp 850-855, 2006