


Article

An Aircraft Object Detection Algorithm Based on Small Samples in Optical Remote Sensing Image

Ting Wang ^{1,*} , Changqing Cao ¹, Xiaodong Zeng ¹, Zhejun Feng ¹, Jingshi Shen ², Weiming Li ², Bo Wang ¹, Yuedong Zhou ¹ and Xu Yan ¹

¹ Xidian University, School of Physics and Optoelectronic Engineering, No. 2 Taibai South Road, Xi'an 710071, China; chqcao@mail.xidian.edu.cn (C.C.); xdzeng@xidian.edu.cn (X.Z.); zhjfeng@mail.xidian.edu.cn (Z.F.); wangbo599026@sina.com (B.W.); 15191903240@163.com (Y.Z.); yanxu_xy@163.com (X.Y.)

² Institute of Space Electronic Technology, Hangtian Road, Yantai 264670, China; jingshi_shen@163.com (J.S.); liweiming513@163.com (W.L.)

* Correspondence: tingw2016hg@126.com

Received: 21 July 2020; Accepted: 16 August 2020; Published: 20 August 2020



Abstract: In recent years, remote sensing technology has developed rapidly, and the ground resolution of spaceborne optical remote sensing images has reached the sub-meter range, providing a new technical means for aircraft object detection. Research on aircraft object detection based on optical remote sensing images is of great significance for military object detection and recognition. However, spaceborne optical remote sensing images are difficult to obtain and costly. Therefore, this paper proposes the aircraft detection algorithm, it can detect aircraft objects with small samples. Firstly, this paper establishes an aircraft object dataset containing weak and small aircraft objects. Secondly, the detection algorithm has been proposed to detect weak and small aircraft objects. Thirdly, the aircraft detection algorithm has been proposed to detect multiple aircraft objects of varying sizes. There are 13,324 aircraft in the test set. According to the method proposed in this paper, the f1 score can achieve 90.44%. Therefore, the aircraft objects can be detected simply and efficiently by using the method proposed. It can effectively detect aircraft objects and improve early warning capabilities.

Keywords: aircraft object detection; support vector machines; clustering; small samples learning

1. Introduction

At present, computer vision (CV) and natural language processing (NLP) and speech recognition are listed as three hotspot directions of artificial intelligence (AI). Computer vision can be used in many fields, such as detection of asteroids [1], meteors [2] and atmospheric events [3], self-driving [4], biomedical image detection [5], automated defect inspection [6], satellite image analysis [7], criminal analysis [8]. One of the classic applications is aircraft object detection. Object detection algorithm is used to identify which real-world objects [9] are and where the objects are in the image. For example, Figure 1 is an image in the Pascal VOC dataset [10], object detection algorithm needs to identify the bus and person in the image and locate them. Object detection algorithms can be used in a great number of fields, such as face detection, and shape detection [9]. Therefore, object detection has received great attention in the field of remote sensing, and it also faces many challenges [11].



Figure 1. An image in the Pascal VOC dataset [10], in which a bus and a person are marked by yellow boxes.

Traditional object detection has gone through several important phases. Viola and Jones use Harr-like features [12] to train strong classifiers in AdaBoost [13,14], and then implement face detection by using sliding window search strategy. Dalal and Triggs propose the local Histogram of Oriented Gradient (HOG) as the image feature, and support vector machine (SVM) as the classifier to detect pedestrians [15]. Felzenszwalb et al. [16,17] propose a multiscale Deformable Parts Model (DPM) method [18] based on the idea that most objects can be regarded as rigid bodies. Malisiewicz et al. adopt the integration idea in machine learning, and train an SVM classifier for each positive sample by using HOG features [19]. Ren et al. design the histogram of sparse code (HSC) based on sparse representation theory [20].

With the swift improvement of computer hardware performance, convolutional neural networks (CNN) [21] have once again attracted the attention of researchers. In the Large Scale Visual Recognition Challenge (ILSVRC) 2012 [22], Alex and Hinton used the AlexNet [23] to achieve a huge leap in detection performance on the ImageNet dataset [24]. Therefore, CNN has received extensive attention in the field of computer vision.

Since the advent of AlexNet, a large number of CNN-based object detection algorithms have appeared. In 2014, Ross Girshick et al. applied CNN to object detection and proposed R-CNN algorithm [25]. Firstly, a selective search algorithm was used to generate a series of candidate regions. Then CNN was used to extract these features in these regions. Finally, the SVM classifier was trained for classification.

In 2015, Kaiming He et al. proposed SPP-net [26], which solved the problem of input image size limitation by using spatial pyramid pooling, and mapped the candidate regions in the original image directly to the feature map. This makes SPP-net faster than R-CNN by dozens of times.

In 2015, Girshick proposed Fast R-CNN [27], which uses two parallel different fully connected layers to perform classification and positioning tasks separately. Ren et al. proposed the Faster R-CNN [28] algorithm, which uses neural networks to extract candidate regions, i.e., RPN networks.

Since the extraction of the candidate region is time-consuming and complicated, some researchers directly find the position and category information of the object of interest on the input image. The most representative algorithms are YOLO [29] and Single Shot Multi-Box Detector (SSD) [30].

In 2018, Yin Zhou proposed a generic 3D detection network, i.e., VoxelNet [31]. It can perform feature extraction and bounding box prediction into a single stage, end-to-end trainable deep network.

Although deep learning has achieved great success in object detection, the detection performance will be significantly reduced in the case of insufficient datasets. Generally, spaceborne optical remote sensing images are difficult to obtain and costly. Therefore, this paper studies the aircraft object detection algorithm with small samples.

The circle frequency filter algorithm is small in calculation, good in saliency objects detection. However, the performance of circle frequency filter greatly reduces when detecting small objects and multiple objects. Therefore, this article proposed the fusion algorithm to improve the circle frequency filter detection accuracy with small samples. The experimental results show that the proposed method can correctly detect 12,215 aircraft on test set, accounting for 91.68% of the total number of aircraft in the test set. From the data analysis, it can be seen that in the case of small samples, the proposed method has better detection performance than Faster R-CNN, and the detection speed is faster.

2. Theoretical Analysis and Methods

Generally, the grayscale, shape, size and shadow of aircraft objects in remote sensing images are used to distinguish aircraft from other interfering objects. Usually aircraft objects are brighter than the background, and most aircraft have similar shapes that they have four parts of the aircraft nose, left wing, tail and right wing. As shown in Figure 2a, the portion where the aircraft fuselage and the wingspan intersect may form a rectangle whose length and width are the widths of the fuselage and the wing, respectively. If the center of gravity of the rectangle is the center of the circle, draw a circle is with a diameter larger than the width of the fuselage and smaller than the length of the fuselage and wingspan, as shown in Figure 2b. The pixel points on the circumference are extracted clockwise, and the gray value of the pixels on the circumference will show the characteristic of “bright-dark-bright-dark-bright-dark-bright-dark” [32]. As shown in Figure 2c, a bar graph is made by taking 601 pixels on the circumference as the horizontal axis and the gray value at each pixel as the vertical axis. However, when the center of the circle is not at the center of the aircraft, it does not have such a characteristic. Therefore, based on this feature, aircraft and non-aircraft can be distinguished.

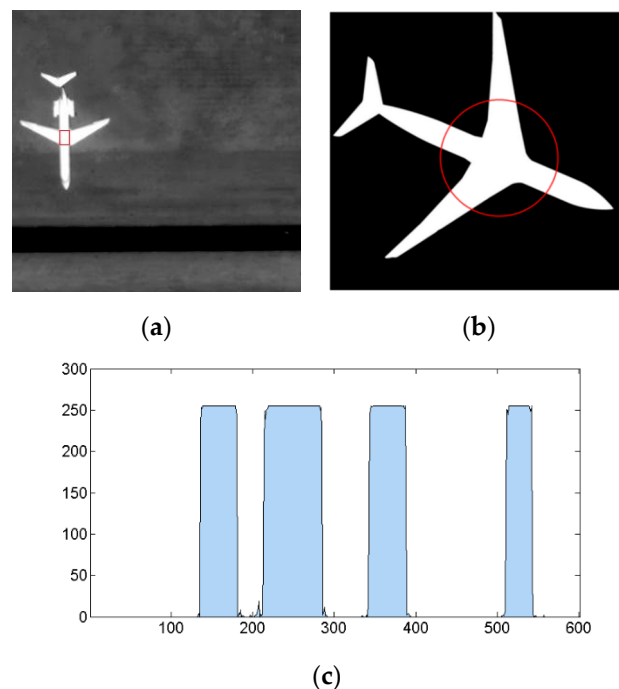


Figure 2. (a) The rectangle of the intersection of the aircraft fuselage and the wingspan. (b) The circumference on the aircraft, the size of the original image is 1024×958 , the radius of the circle is 200 pixels. (c) The gray value curve on the circumference, the horizontal axis is 601 pixels, and the vertical axis is the gray value at each pixel point.

Let f_k ($k = 0, 1, 2, \dots, N - 1$) denote the gray value on the circumference centered on (i, j) and the radius is R . Make a discrete Fourier transform on this one-dimension array and obtain

$$F = \sum_{k=0}^{N-1} f_k e^{-j \frac{2\pi}{N} kn} \quad (k = 0, 1, 2, \dots, N - 1), \quad (1)$$

Square the above equation to get the amplitude value.

$$F^2 = \left(\sum_{k=0}^{N-1} f_k \cos \frac{2\pi}{N} kn \right)^2 + \left(\sum_{k=0}^{N-1} f_k \sin \frac{2\pi}{N} kn \right)^2, \quad (2)$$

where n is a constant used to control the number of periods of the sine and cosine functions for the discrete Fourier transform.

The gray value curve on the circumference has 4 peaks and 4 valleys, which is similar to the sine and cosine curves of 4 cycles. So, if four cycles of sine and cosine functions are selected in Equations (1) and (2) (that is, $n = 4$), the amplitude value of the circumferential frequency filter [32] in the center of the aircraft is large.

Since the non-aircraft center point usually does not have the special properties, the value of the circle frequency filter [32] is small. The amplitude value is the square of the value obtained by the Fourier transform, therefore, the amplitude values on the center of the aircraft and the non-aircraft center will vary greatly. Therefore, it is easy to distinguish between aircraft centers and non-aircraft centers.

2.1. Weak and Small Object Detection Algorithm

Our dataset has 419 images, with a size of 305×296 . It contains a total of 2175 aircraft. In order to better verify the effect of the proposed method, the aircraft dataset of RSOD-Dataset [33] and UCAS_AOD [34] are added to our dataset. RSOD-Dataset included 4993 aircraft in 446 images, with a size of 1044×915 . UCAS_AOD dataset includes 7482 aircraft in 1000 images. Images in RSOD-Dataset and UCAS_AOD dataset both have three channels, therefore, they need to be converted to grayscale. Among the dataset, 2841 aircraft are weak and small objects. They are either smaller than 10×10 , or the grayscale value of them is almost equal to that of the ground background. The aircraft in each image is manually marked in the form of (x, y, r) , where (x, y) is the center of the circle with the characteristics of four peaks and four valleys, and r is the radius. The dataset is divided into a training set and a test set, of which the training set accounts for about 10% and the test set 90%. Training set contains 1326 aircraft. In order to compare the test results with Faster R CNN, our dataset is annotated according to the label form of the PASCAL VOC2007 [10] dataset.

The circle frequency filter detection algorithm is used to detect the saliency object, and the detection accuracy is high. The amplitude value of each point on the circumference is calculated and normalized to between 0 and 255. According to multiple experiments, the detection effect is better when the threshold is set to 80. Therefore, the gray value greater than 80 is set to 255, and the other gray values are set to 0. As shown in Figure 3b, circles are drawn on the original image centering on pixel points with a gray value of 255 in the filter result image, and the circumference is just at the center of the airplane.

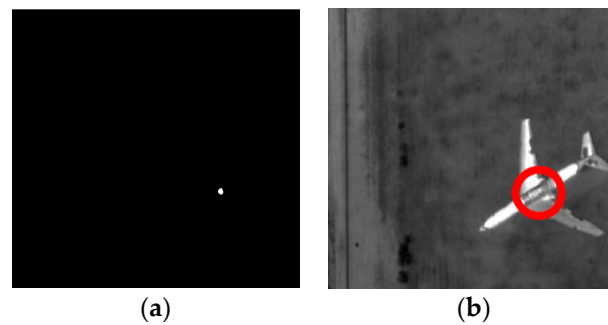


Figure 3. The result graph after circle frequency filtering; (a) Threshold segmentation map; (b) Circles with radii of the circumference centered on the points in (a) is marked on the original image.

As shown in Figure 4a,b there are a total of six aircraft in the picture. Only two aircraft can be correctly detected. As shown in Figure 4c,d the circle frequency filter cannot detect aircraft. Therefore, the circle frequency filter is not suitable for weak and small aircraft object detection. Based on this problem, the circle frequency filter detection algorithm is improved to specifically detect weak and small aircraft objects.

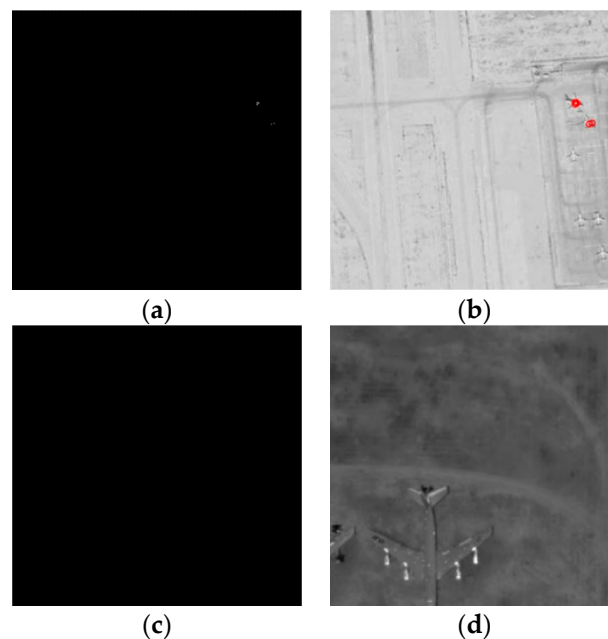


Figure 4. The result graph of weak and small objects; (a,b) small object; (c,d) weak objects.

The circle frequency filter detection algorithm has the following two reasons for the inaccurate detection of weak aircraft objects. Firstly, the difference between the gray value of the aircraft object and the background is too small, resulting in a mixture of the background and the object, which is difficult to distinguish. Secondly, the edge of the aircraft is blurred, and the gray value on the circumference does not show the characteristics of four peaks and four valleys. So, it can't determine whether it is an aircraft object.

Usually, most of the grayscale values on the aircraft object are not much different, and the background grayscale values are not much different too. We can use this feature to solve the above two problems, that is, to square each point on the image matrix. It is assumed that the grayscale value of a point x on the aircraft target is 9, and the grayscale value of the background point y adjacent to x is 4, and the difference in grayscale value is 5. Only the circle frequency filter detection algorithm cannot distinguish between the object and the background, so that such a weak object cannot be detected.

Therefore, we can square the two grayscale values, and the grayscale values of the x and y points are 81 and 16, respectively. Then the gray values sampled on the circumference normalize to make the difference between the data more obvious. Based on the characteristics of the aircraft object, the data can be determined whether it is an aircraft object.

As shown in Figure 5, weak and small aircraft can be detected using the proposed algorithm. In Figure 5b, the circumferential radius of the detected aircraft is only 2 pixels, and the minimum difference between objects and background in Figure 5d is only 1 pixel. In Figure 5a,b, some of the detection circumferences are not located at the center of the aircraft. Therefore, mean shift clustering is used to cluster the detection points to the center of the aircraft.

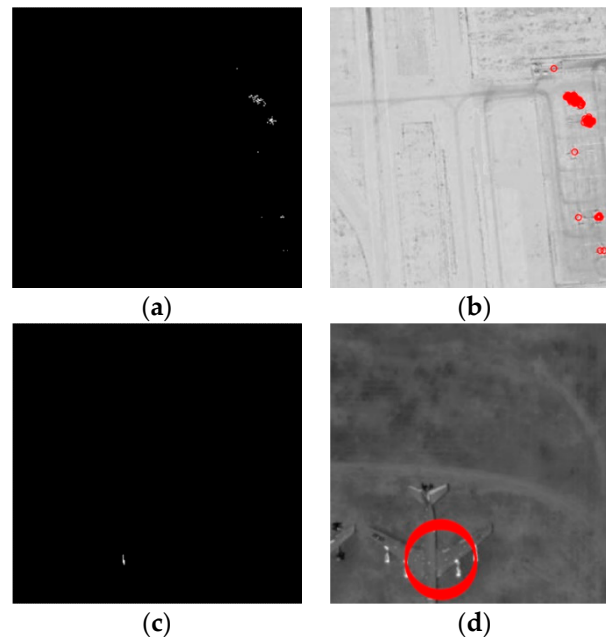


Figure 5. The result graph of proposed method. (a,c) Threshold segmentation map; (b,d) Circles with radii of the circumference centered on the points in (a) and (c) is marked on the original image.

2.2. Fusion Algorithm

As shown in Figure 5, the use of weak and small object detection algorithms to detect aircraft objects will mark multiple bounding boxes near the center of the aircraft. Since the bounding boxes are concentrated near the center of the aircraft, only one bounding box can be left on each aircraft by using the mean shift cluster.

2.2.1. Clustering Algorithm

In order to discover interesting groups, cluster analysis is applied to a variety of scientific areas [35]. In this paper, mean shift cluster algorithm is used to cluster points which are the result of weak and small object detection algorithm. Mean Shift, a method of nonparametric probability density estimation, was originally proposed by Fukunaga et al. [36]. Yizong Cheng [37] defined a family of kernel functions, which made the contribution of drift to the mean drift vector different with the difference of sample and drift point; secondly, he set a weight coefficient, which made the importance of different sample points different.

Let X be the n -dimensional Euclidean space R^n . Then, donate the i th component of $x \in X$ by x_i , and the norm of $x \in X$ is a non-negative number $\|x\|$. A function $K: X \rightarrow R$ is said to be a kernel if there exist a profile $k: [0, \infty] \rightarrow R$

$$K(x) = k(\|x\|^2), \quad (3)$$

where k is nonnegative, nonincreasing and piecewise continuous.

Let K be a kernel and $\omega: S \rightarrow (0, \infty)$ a weight function. Then, with the kernel K at $x \in X$, the sample mean is:

$$m(x) = \frac{\sum_{s \in S} K(s-x)\omega(s)s}{\sum_{s \in S} K(s-x)\omega(s)}, \quad (4)$$

Let T be a finite set (cluster centers). The evolution of T in the form of iterations $T \leftarrow m(T)$ is called a mean shift algorithm.

Let X be the d -dimensional Euclidean space, \mathbb{R}^d . Donate the i th component of $x \in X$ by x_i , and the mean vector at $x \in X$ is:

$$M_h(x) = \frac{1}{k} \sum_{x_i \in S_h} (x_i - x), \quad (5)$$

and S_h refers to a spherical field with a radius with h .

$$S_h(x) = \left\{ y \mid (y-x)(y-x)^T \leq h^2 \right\}, \quad (6)$$

With the kernel function and sample weight, the mean vector at $x \in X$ is

$$M_h(x) = \frac{\sum_{x_i \in S_h} K\left(\frac{x_i-x}{h}\right)\omega(x_i)(x_i-x)}{\sum_{x_i \in S_h} K\left(\frac{x_i-x}{h}\right)\omega(x_i)}, \quad (7)$$

where $\omega(x_i)$ is the weight of the sample x_i , and $\sum_{x_i \in X} \omega(x_i) = 1$.

In the process of the proposed algorithm, firstly, calculate the mean vector $M_h(x)$ between a random point A in the sample and its surrounding radius with h . Then, the next drift direction of this point is calculated, that is $A = M_h(x) + A$. At the new center A , the previous steps repeat to iterate out the new center. When the point no longer moves, it forms a cluster with the surrounding points, and all the points encountered during this iteration should be grouped into this cluster. Calculate the distance between this cluster and the other cluster, and merge into the same cluster if the distance is less than the threshold, and form a new cluster if it's not satisfied. Finally, repeat the previous steps until all points in the sample are accessed.

According to our experience, h take to be twice the radius of the circumference. After clustering the filtered results on Figure 5, the resulting graph is shown in Figure 6. Compared with other algorithms, Mean Shift can automatically discover cluster structure without specifying the number of clusters.

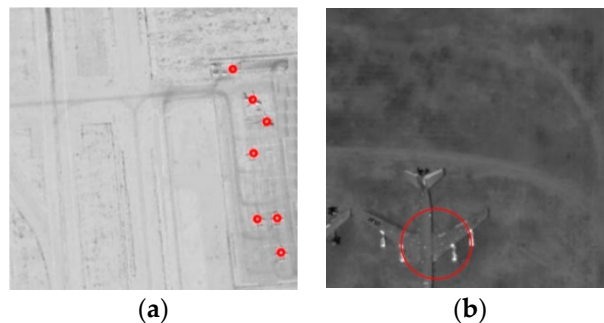


Figure 6. Clustering result graph. (a) Multiple small objects; (b) Weak object.

2.2.2. False Alarms Elimination

There are not only airplanes in the original images, but also many buildings around the airport. Detecting aircraft only using circle frequency filter can produce many false alarms. Therefore, support vector machine (SVM) is used to eliminate false alarms. Firstly, the dataset is divided into a training set

and a test set. The training set is used to fit the SVM model, and the test set is used to finally evaluate the performance and classification ability of the model. The training set includes 66,300 negative samples and 13,260 positive samples. Then, the SVM model is trained on the training set. Finally, the classification ability of the model was evaluated on the testing set.

3. Results

In many images, aircraft vary in size. Therefore, we use the circle frequency filter first to detect the large and saliency aircraft, then deduct the part of these aircraft on the original image, and finally, the small aircraft. When the weak and small aircraft is included in images, the improved circle frequency filter algorithm is used for detection. The trained SVM model is used to classify the test results, and then the targets classified as non-aircraft are removed. The fusion algorithm flow is shown in Table 1.

Table 1. The fusion algorithm flow.

The Proposed Algorithm
Input: Test dataset images Execute: <ol style="list-style-type: none"> (1) According to the statistical radius, the circumference of the corresponding radius is taken for the pixels on each image. (2) Calculate the magnitude of the gray value on the circumference. (3) According to the amplitude value and the special characteristics of the aircraft to determine whether it is an aircraft. (4) According to the weak and small object detection algorithm, process data and calculate the magnitude to determine whether it is an aircraft. (5) Remove some false alarms by clustering algorithm. (6) Use the trained SVM to eliminate a false alarm.
Output: Images contain the location of the aircraft

In order to compare the detection results of the proposed method and Faster R CNN [28], we train Faster R CNN on our training set and test it. A pre-trained backbone network VGG16 model on the Imagenet [38] dataset is adopted. In order to ensure that there is no over-fitting, the maximum number of iterations is 2000. Part of the image detection map is shown in Figure 7.

There are 13,324 aircraft in the test set, which is recorded as P. A total of 18,812 samples were detected using the proposed algorithm. Using the SVM model to eliminate false alarms, 13,687 objects were obtained. Among them, 12,215 samples can be correctly detected, denoted as TP (True Positive). 1109 samples were classified as negative sample. A sample of 1472 non-aircraft objects was incorrectly judged by the improved method and recorded as FP (False Positive). A total of 1109 aircraft targets were not be detected, and they were denoted as FN (False Negative). Therefore, TN (True Negative) is 0.

Faster R CNN detects a total of 13,687 objects. Among them, 11,335 samples can be correctly detected. 4361 non-aircraft objects was incorrectly judged by Faster R CNN and recorded as FP. A total of 1989 aircraft objects were not be detected, and they were denoted as FN. TN is 0.

According to the actual test results of the dataset, precision, accuracy, recall, missing alarm and false alarm are shown in Table 2. 12,215 aircraft can be correctly detected by using the proposed algorithm. It can be seen from both the visualization results and some evaluation indicators that the proposed method is better than Faster R CNN on our dataset. In the case of small samples, the detection accuracy of the object detection method based on a convolutional neural network is greatly reduced. On the test set, Faster R CNN takes about 6.383 s to detect an image, while the proposed algorithm only takes 2.951 s. Therefore, the proposed algorithm can be adopted to efficiently detect aircraft objects.

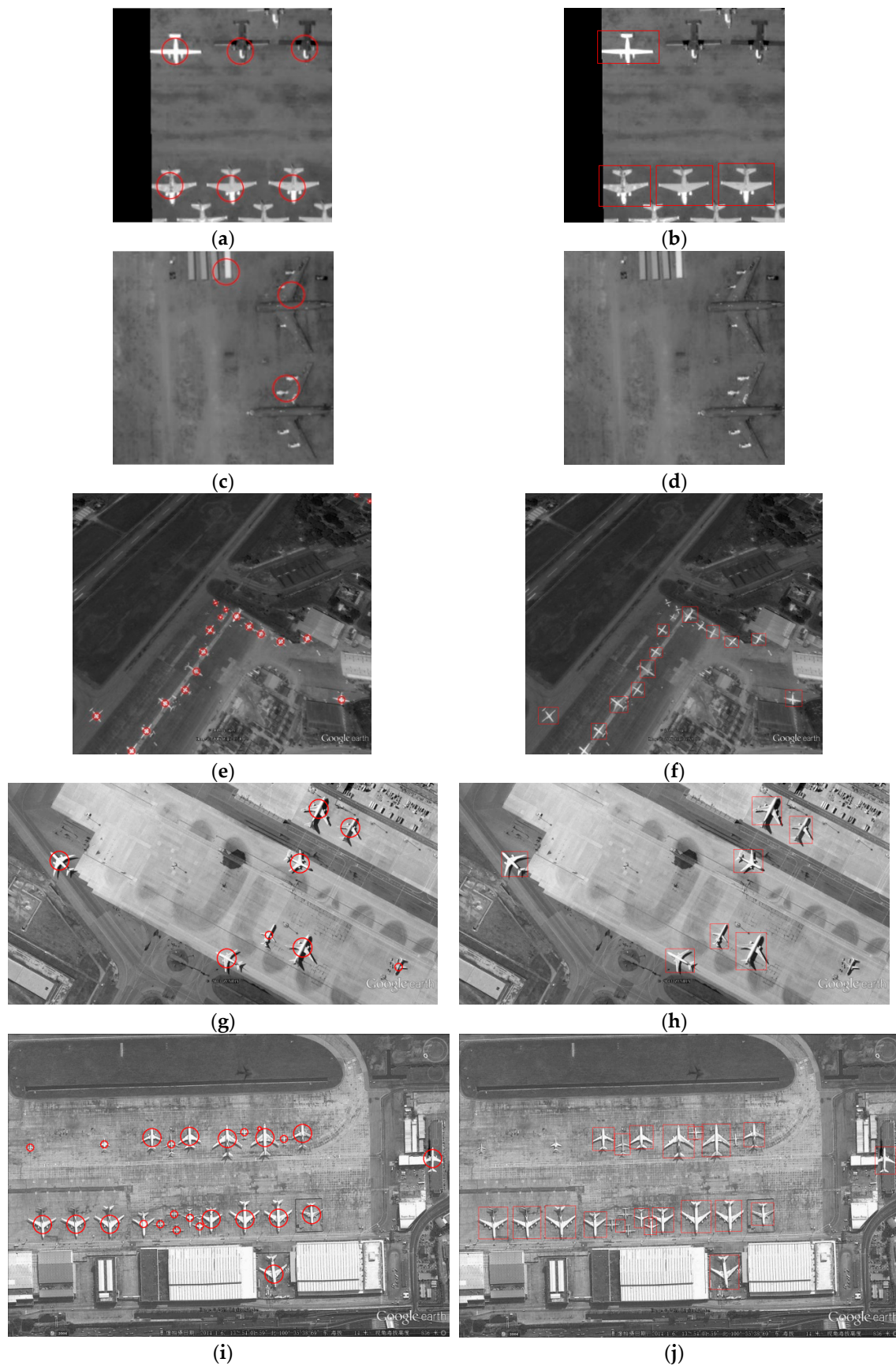


Figure 7. (a,c,e,g,i) are the detection results of the proposed method; (b,d,f,h,j) are the detection results of Faster R CNN.

Table 2. Some evaluation indicators.

Evaluation Indicator	The Proposed Algorithm	Faster R CNN
Precision	89.25%	72.22%
Accuracy	82.56%	64.09%
Recall	91.68%	85.07%
Missing Alarm	8.32%	14.93%
False Alarm	10.75%	27.78%
F-measure	90.44%	78.12%

4. Conclusions

There are 13,324 aircraft in the test set, and 12,215 aircraft can be detected by using the proposed method. The recall is 91.68%, the precision is 89.25%, and the accuracy is 82.56%. Experiments show that the proposed method is simple and effective for detecting aircraft objects. Especially for small samples learning, the proposed algorithm can accurately detect aircraft objects, which greatly saves costs for the research. The research on aircraft object detection has significant value for improving early warning capability and counterattack capability. However, about 9% aircraft can't be detected by using our algorithm, mainly because the center of the circle on the aircraft don't have the characteristics of four peaks and four valleys. To address this problem, in the next research, a new algorithm will be proposed to detect the remaining aircraft.

Author Contributions: Conceptualization, T.W.; Formal Analysis, X.Z. and Y.Z.; Investigation, C.C. and W.L.; Methodology, T.W. and B.W.; Software, Z.F.; Validation, T.W. and J.S.; Visualization, X.Y.; Writing—Review & Editing, T.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Fundamental Research Funds for the Central Universities (K5051399208) and Major Instruments of the Ministry of Science and Technology (2012YQ12004702) and the 111 Project (B17035).

Conflicts of Interest: The authors declare no conflict of interest.

References

- Tommy, G.; Robert, J.; Larry, D.; Steve, C.; Matthew, H.; Timothy, S. The pan-STARRS synthetic solar system model: A tool for testing and efficiency determination of the moving object processing system. *Publ. Astron. Soc. Pacific* **2011**, *123*, 423–447. [\[CrossRef\]](#)
- Trigo-Rodriguez, J.M.; Madiedo, J.M.; Gural, P.S.; Castro-Tirado, A.J.; Llorca, J.; Fabregat, J.; Vitek, S.; Pujols, P. Determination of meteoroid orbits and spatial fluxes by using high-resolution all-sky CCD cameras. *Earth Moon Planets* **2008**, *102*, 231–240. [\[CrossRef\]](#)
- Serge, S.; Iacovella, F.; van der Velde, O.; Montanyà, J.; Füllekrug, M.; Farges, T.; Bór, J.; Georgis, J.-F.; NaitAmor, S.; Martin, J.-M. Multi-instrumental analysis of large sprite events and their producing storm in southern France. *Atmos. Res.* **2014**, *135–136*, 415–431. [\[CrossRef\]](#)
- Shitao, C.; Songyi, Z.; Jinghao, S.; Badong, C.; Zheng, N. Brain-inspired cognitive model with attention for self-driving cars. *IEEE Trans. Cogn. Dev. Syst.* **2019**, *11*, 13–25. [\[CrossRef\]](#)
- Olaf, R.; Philipp, F.; Thomas, B. U-net: Convolutional networks for biomedical image segmentation. In *Lecture Notes in Computer Science, Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Munich, Germany, 5–9 October 2015*; Springer: Berlin/Heidelberg, Germany, 2015. [\[CrossRef\]](#)
- Shuang, M.; Hua, Y.; Zhouping, Y. An unsupervised-learning-based approach for automated defect inspection on textured surfaces. *IEEE Trans. Instrum. Meas.* **2018**, *67*, 1266–1277. [\[CrossRef\]](#)
- Logar, T.; Bullock, J.; Nemni, E.; Bromley, L.; Quinn, J.A.; Luengo-Oroz, M. PulseSatellite: A tool using human-AI feedback loops for satellite image analysis in humanitarian contexts. *AAAI* **2020**, *34*, 13628–13629. [\[CrossRef\]](#)
- Pompílio, A.; Jefferson, F.; Luciano, O. Multi-perspective object detection for remote criminal analysis using drones. *IEEE Geosci. Remote Sens. Lett.* **2020**, *17*, 1283–1286. [\[CrossRef\]](#)

9. Apoorva, R.; Mohana, M.; Pakala, R.; Aradhya, H.V.R. Object detection algorithms for video surveillance applications. In Proceedings of the International Conference on Communication and Signal Processing (ICCSP), Chennai, India, 3–5 April 2018.
10. Everingham, M.; Eslami, S.M.A.; van Gool, L.; Williams, C.K.I.; Winn, J.; Zisserman, A. The pascal visual object classes challenge: A retrospective. *Int. J. Comput. Vis.* **2014**, *111*, 98–136. [\[CrossRef\]](#)
11. Yang, X.; Fu, K.; Sun, H.; Sun, X.; Yan, M.; Diao, W.; Guo, Z. Object detection with head direction in remote sensing images based on rotational region CNN. In Proceedings of the IGARSS 2018—2018 IEEE International Geoscience and Remote Sensing Symposium, Valencia, Spain, 22–27 July 2018.
12. Viola, P.; Jones, M. Rapid object detection using a boosted cascade of simple features. In Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Kauai, HI, USA, 8–14 December 2001.
13. Hiromoto, M.; Sugano, H.; Miyamoto, R. Partially parallel architecture for adaBoost-based detection with haar-like features. *IEEE Trans. Circuits Syst. Video Technol.* **2009**, *19*, 41–52. [\[CrossRef\]](#)
14. Zhu, J.; Rosset, S.; Zou, H.; Hastie, T. Multi-class adaboost. *Stat. Interface* **2006**, *2*, 349–360.
15. Dalal, N.; Triggs, B. Histograms of oriented gradients for human detection. In Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Diego, CA, USA, 20–25 June 2005.
16. Felzenszwalb, P.; McAllester, D.; Ramanan, D. A discriminatively trained, multiscale, deformable part model. In Proceedings of the 2008 IEEE Conference on Computer Vision and Pattern Recognition, Anchorage, AK, USA, 23–28 June 2008.
17. Felzenszwalb, P.; Girshick, R.; McAllester, D.; Ramanan, D. Object detection with discriminatively trained part-based models. *IEEE Trans. Pattern Anal. Mach. Intell.* **2010**, *32*, 1627–1645. [\[CrossRef\]](#) [\[PubMed\]](#)
18. Felzenszwalb, P.; Girshick, R.; McAllester, D. Cascade object detection with deformable part models. In Proceedings of the 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Francisco, CA, USA, 13–18 June 2010.
19. Malisiewicz, T.; Gupta, A.; Efros, A. Ensemble of exemplar-SVMs for object detection and beyond. In Proceedings of the 2011 International Conference on Computer Vision, Barcelona, Spain, 6–13 November 2011.
20. Ren, X.; Ramanan, D. Histograms of sparse codes for object detection. In Proceedings of the 2013 IEEE Conference on Computer Vision and Pattern Recognition, Portland, OR, USA, 23–28 June 2013.
21. LeCun, Y.; Boser, B.; Denker, J.S.; Henderson, D. Backpropagation applied to handwritten zip code recognition. *Neural Comput.* **1989**, *1*, 541–551. [\[CrossRef\]](#)
22. Berg, A.; Deng, J.; Li, F.F. Large Scale Visual Recognition Challenge. 2012. Available online: www.image-net.org/challenges (accessed on 1 July 2020).
23. Krizhevsky, A.; Sutskever, I.; Hinton, G. ImageNet classification with deep convolutional neural networks. In Proceedings of the Advances in Neural Information Processing Systems 25, Lake Tahoe, CA, USA, 3–6 December 2012.
24. Deng, J.; Dong, W.; Socher, R. Image Net: A large-scale hierarchical image database. In Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 20–25 June 2009.
25. Girshick, R.; Donahue, J.; Darrell, T.; Malik, J. Region-based convolutional networks for accurate object detection and segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2016**, *38*, 142–158. [\[CrossRef\]](#) [\[PubMed\]](#)
26. He, K.; Zhang, X.; Ren, S.; Sun, J. Spatial pyramid pooling in deep convolutional networks for visual recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **2015**, *37*, 1904–1916. [\[CrossRef\]](#) [\[PubMed\]](#)
27. Girshick, R. Fast r-cnn. In Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV), Santiago, Chile, 7–13 December 2015.
28. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 1137–1149. [\[CrossRef\]](#) [\[PubMed\]](#)
29. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You only look once: Unified, real-time object detection. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016.
30. Liu, W.; Anguelov, D.; Erhan, D.; Szegedy, C.; Reed, S.; Fu, C.-Y.; Berg, A.C. SSD: Single shot multibox detector. In Proceedings of the European Conference on Computer Vision, Amsterdam, The Netherlands, 8–16 October 2016.

31. Zhou, Y.; Tuzel, O. VoxelNet: End-to-end learning for point cloud based 3D object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018.
32. Hongping, C.; Yi, S. Airplane detection in remote-sensing image with a circle-frequency filter. In Proceedings of the SPIE—The International Society for Optical Engineering, Wuhan, China, 4 January 2006.
33. Long, Y.; Gong, Y.; Xiao, Z.; Liu, Q. Accurate object localization in remote sensing images based on convolutional neural networks. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 2486–2498. [[CrossRef](#)]
34. Haigang, Z.; Xiaogang, C.; Weiqun, D.; Fu, K.; Ye, Q.; Jiao, J. Orientation robust object detection in aerial images using deep convolutional neural network. *IEEE Int. Conf. Image Process.* **2015**. [[CrossRef](#)]
35. Caruso, G.; Gattone, S.A.; Balzanella, A.; Di Battista, T. Models and theories in social systems. Studies in systems, decision and control. In *Cluster Analysis: An Application to a Real Mixed-Type Data Set*; Flaut, C., Hořková-Mayerová, Š., Flaut, D., Eds.; Springer: Berlin/Heidelberg, Germany, 2019; Volume 179, pp. 525–533.
36. Fukunaga, K.; Hostetler, L. The estimation of the gradient of a density function, with applications in pattern recognition. *IEEE Trans. Inf. Theory* **1975**, *21*, 32–40. [[CrossRef](#)]
37. Cheng, Y. Mean shift, mode seeking, and clustering. *IEEE Trans. Pattern Anal. Mach. Intell.* **1995**, *17*, 790–799. [[CrossRef](#)]
38. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet classification with deep convolutional neural networks. *Commun. ACM* **2017**, *60*, 84–90. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).