A COARSE-TO-FINE MODEL FOR AIRPLANE DETECTION FROM LARGE REMOTE SENSING IMAGES USING SALIENCY MODLE AND DEEP LEARNING

Z. N. Song[1], H. G. Sui[2*],

[1] Remote Sensing and Information Engineering. College of Wuhan University, Wuhan, China – biao.da@163.com

[2,*] State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan, China – haigang_sui@263.net

ICWG ll/III

ABSTRACT:

High resolution remote sensing images are bearing the important strategic information, especially finding some time-sensitive-targets quickly, like airplanes, ships, and cars. Most of time the problem firstly we face is how to rapidly judge whether a particular target is included in a large random remote sensing image, instead of detecting them on a given image. The problem of time-sensitive-targets target finding in a huge image is a great challenge: 1) Complex background leads to high loss and false alarms in tiny object detection in a large-scale images. 2) Unlike traditional image retrieval, what we need to do is not just compare the similarity of image blocks, but quickly find specific targets in a huge image. In this paper, taking the target of airplane as an example, presents an effective method for searching aircraft targets in large scale optical remote sensing images. Firstly, we used an improved visual attention model utilizes salience detection and line segment detector to quickly locate suspected regions in a large and complicated remote sensing image. Then for each region, without region proposal method, a single neural network predicts bounding boxes and class probabilities directly from full images in one evaluation is adopted to search small airplane objects. Unlike sliding window and region proposal-based techniques, we can do entire image (region) during training and test time so it implicitly encodes contextual information about classes as well as their appearance. Experimental results show the proposed method is quickly identify airplanes in large-scale images.

1. INTRODUCTION

1.1 Background

With the rapid development of remote sensing observation technologies, we have entered an era of remote sensing big data (Li, 2017). Due to the large volume of images and datasets, time-sensitive-targets searching of remote sensing images is very difficult and urgent. Most of time the problem firstly we face is how to rapidly judge whether a particular target is included in a large random remote sensing image, instead of detecting them on a given image. Nevertheless, the problem of time-sensitive-targets target finding in a huge image is a great challenge: 1) Complex background leads to high loss and false alarms in tiny object detection in a large-scale images. 2) Unlike traditional image retrieval, what we need to do is not just compare the similarity of image blocks, but quickly find specific targets in a huge image.

1.2 Related Work

As discussed above, A few works have been done on detecting target from optical satellite images. For a given image, the target detection task can be simply described as "where is the target" (Li, 2011). In object detecting, we get regions of interest (ROI) which may have targets in the first step likewise. Considering the methods used in optical satellite images target detection, many researchers like R-CNN (Girshick, 2015) use region proposal methods to first generate potential bounding boxes in an image and then run a classifier on these proposed boxes. After classification, post-processing is used to refine the bounding boxes, eliminate duplicate detections, and rescore the boxes based on other objects in the scene (Girshick, 2014). These complex pipelines are slow and hard to optimize because each individual component must be trained separately.

Other works reframe object detection as a single regression problem, straight from image pixels to bounding box coordinates

and class probabilities. Using that system, you only look once (Redmon, 2016) at an image to predict what objects are present and where they are. While it can quickly identify objects in images it struggles to precisely localize some objects, especially small ones.

1.3 Work Flow

In order to overcome such problems, this paper tried to explore a new way of thinking for airplane detecting in huge images. The work flow of this method is shown in Figure 1.
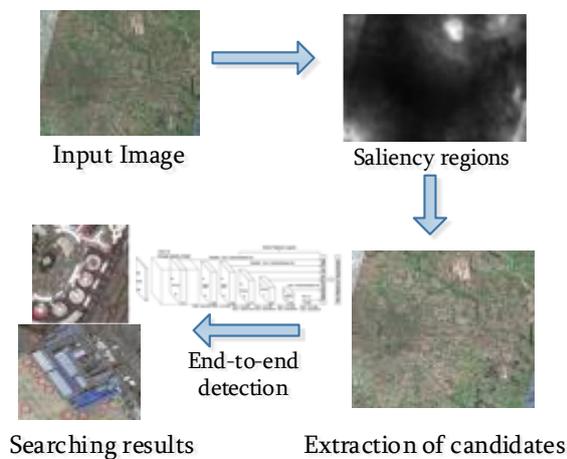
Figure 1. The work flow structure

Firstly, different from traditional studies, this proposed model simply and directly distinguishes each image regions which may contain airplane targets or not in large images, without other pre-processing or huge number of siding windows. In our model, inspired by studies of visual attention, we go further by using combined salience detection and line segment detector to predict the whole image areas which may contain airplanes.

After getting the suspicious areas, we focus on finding out the salient regions containing airplanes by using a single end-to-end network to detect objects from these regions rapidly.

2. PREDICTION OF CANDIDATE REGIONS

Airport is an artificial target, and it has obvious structural characteristics, moreover, airport is a prerequisite for the existence of an aircraft, in most cases. So we can check the airport and then look for the plane. The background color and texture of airport area are uniform, and the edge features are few and irregular. Visual attention model can adequately simulate the

characteristics of human eyes and quickly find the most "saliency" goals in the scene because it makes full use of a variety of characteristic information such as intensity, color and direction of an image (Yao, 2015). As saliency maps computed from enhanced contrast based filtering for salient region detection provide a coarse indication of the structure in the visual contents, we use line segment detector (LSD) (Grompone, 2010) as saliency map feature extractor witch complement more detailed and effective features to support the target/non-target classification task. Figure 2 shows an example of airplane candidate regions extraction.

(a). The original image   (b). Salient regions
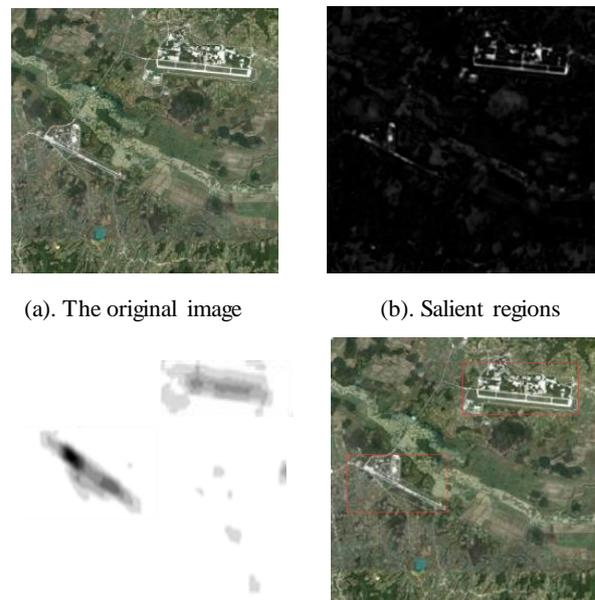
(c).Line-probability distribution   (d).Airdrome detection

Figure 2. Airplane candidate regions extraction

3. ARIPLANES DETECTION

After getting the suspicious areas, we unify the separate components of object detection into a single neural network. Our network uses features from the entire image to predict each bounding box. It also predicts all bounding boxes across all classes for an image simultaneously. This means our network reasons globally about the full image and all the objects in the image. The design referencing YOLO enables end-to-end training and real time speeds while maintaining high average precision. During multi-scale training we use sum of squared error loss, and predict an objectness score for each bounding box using logistic regression. Experimental results show the proposed method is quickly (less than 1 minute) identify airplanes in large-scale (larger than 20000 * 20000) images.

Figure 3. Airplane candidate regions detecting results.

4. CONCLUSION

In this paper, an efficient algorithm for detecting airfields and aircraft using salient suspicious extraction and end-to-end deep learning is proposed. Experimental results show the proposed method is quickly detect airplanes in large-scale images. However, for the purpose of practical application of object detection, what we have done is just show the suspicious object candidates, we have to classify single targets respectively, and this is mainly work we will study in the future.

REFERENCES

Zhifeng Xiao, Yang Long, Deren Li. High-Resolution Remote Sensing Image Retrieval Based on CNNs from a Dimensional Perspective. *Remote Sens.*, 2017, 9(7), 725.

Zhicheng Li, Laurent Itti, 2011. Saliency and gist features for target detection in satellite images. *IEEE Trans. Image Processing,* 20(7), pp. 2017 - 2029.

Ross Girshick. Fast R-CNN. *Comp. Vis. and Pat. Recog.* 2015. abs/1504.08083, 2, 5, 6, 7.

R.Girshick, J.Donahue, T.Darrell, and J.Malik. Rich feature hierarchies for accurate object detection and semantic segmentation. *Comp. Vis. and Pat. Recog.* 2014, 1, 4, 7.

Grompone v G R, Jakubowicz J, Morel J M. LSD: a fast line segment detector with a false detection control. *IEEE Trans. on Patt. Analys. & Mac. Intel.* 2010, 32(4):722.

Xiwen Yao, Junwei Han, Lei Guo, Shuhui Bu. A coarse-to-fine model for airport detection from remote sensing images using target-oriented visual saliency and CRF. *Neurocomp.* 2015, 162–172.