Target Detection using Machine Learning

Ahmed Mohsen Sabra, Mohamed Abdelhady

Military Technical College, Egypt, ahmedmohsensabra0@gmail.com, mohamedabdelhady58@gmail.com

Supervisor: *Col.Dr.* Ahmed T. Hafez, ACA Department

Military Technical College, Egypt, a.taimour@mtc.edu.eg

Supervisor: Col.Dr. *Essam H. Hamza, ACA Department*

Military Technical College, Egypt

Abstract -With the rapid technological development of various different satellite sensors, a huge volume of high-resolution image data sets can now be obtained and widely used in military and civilian fields. Detecting typical targets in satellite images is a challenging task due to the varying size, orientation and background of the target object. Traditional manually engineered features (i.e. HOG, Gabor feature and Hough transform, etc.) do not work well for massive high-resolution remote sensing image data. Thus, we are expected to find an efficient way to automatically learn the presentations from the massive image data and increase the computational efficiency of target detection. Robust and computationally efficient systems are required which can learn presentations from the massive satellite imagery. Comparing to the general objects in nature images, the edge information of targets in satellite images shows more distinctive and concise characteristics. This paper proposes a new target detection framework based on Edge Boxes and Convolutional Neural Networks (CNN). CNN can learn rich features automatically and is invariant to small rotation and shifts, has achieved state of-the-art performance in many image classification databases. Edge Boxes can generate a smaller set of object proposals based on the edges of objects. The proposed method can reduce the computational time of the detector. Moreover, CNN is invariant to minor rotations and shifts in the target object. Extensive experiments demonstrate that the proposed framework is effective in typical target detection systems

Keywords: Target Detection, Convolutional Neural Networks, Edge Boxes

I. Introduction

Automatic detection of military targets such as oil tanks, aircrafts, artillery, etc. in high resolution satellite imagery has great significance in military applications. With the rapid development of satellite imaging and geographic information systems, a large amount of high resolution images can be acquired effortlessly from Google Earth. The non-hyperspectral image data has been used in many civil and military applications. Various techniques and features have been proposed so far for automatic target detection in satellite imagery.

Zhang et al. [1] developed a hierarchical algorithm based on Adaboost classifier and HOG feature for detection of oil tanks.

Han et al. in [2] proposed a method based on graph search strategy and improved Hough Transform for detection of oil tanks in satellite imagery. Yildiz et al. in [3] employed Gabor feature and used SVM classifier to detect different aircrafts. Gabor filter is also employed by authors in [4,5] for road crack detection in aerial images and settlement zone detection is satellite images respectively.

Hsieh et al. in [6] employed Zernike moments, aircraft contour and wavelets and used SVM classifier for the detection of aircrafts in satellite images. Most of the methods discussed above use hand-crafted features and work effectively in their scenes only. Deep learning is a very effective method for learning optimum features directly from huge training dataset automatically. Now a day in numerous applications computer vision along with deep learning have outperformed humans. Furthermore, the use of Graphical Processing Units (GPUs) has decreased the training time of deep learning methods. Large databases of labelled data and pre-trained networks are now publicly available. The two popular models of deep learning are Deep Belief Network (DBN) [7] and Convolutional Neural Network (CNN) [8].

CNN is a modern deep learning method which is widely used for image recognition because it is invariant to small rotation and shifts [9]. DBN is a probabilistic generative model which is pre-trained as Restricted Boltzmann Machine layer by layer, and then finally tuned by back-propagation algorithm to become a classifier [9]. Chen et al. [10] employed object locating method along with DBN for aircraft detection in satellite images. Saliency has also been used for image classification by various researchers such as Li et al. [11] applied visual symmetry detection and saliency computation for aircraft and tank detection in satellite images. Zhang et al. [12] and Sattar et al. [13] employed saliency and used unsupervised learning for image classification.

Identification of fixation points, detection of image regions representing the scene and detection of dominant objects in a scene are the primary goals of saliency. Nevertheless, satellite images often contain several targets and correct localization of each target is required. Saliency cannot be directly employed for automatic target detection in satellite images and it needs the help of other methods such as symmetry detection.

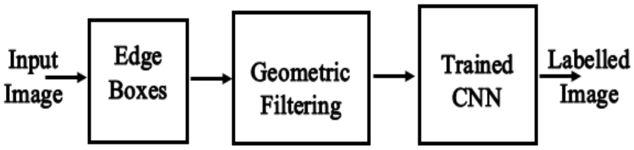

Figure 1. Conceptual level block diagram of the proposed target detection system

Figure 2. Examples from the tank dataset.

CNN is being used in various computer vision applications for the last two decades. The sliding window approach for target detection using CNN is very slow and contrast to mechanisms in the human vision system. Many objectiveness detection methods have been proposed to increase the

Computational efficiency of target detection, such as Edge Boxes [14], Binarized Normed Gradients (BING) [15] and Selective Search [16]. Selective search greedily merges low-level super-pixels to generate object proposals. Ross Girshick et al. [17] proposed the use of selective search along with CNN instead of sliding-window detector and achieved outstanding results on ILSVRC2013 detection dataset. BING generates object proposals based on binarized normed gradients. Edge Boxes uses the object boundaries in the image as a feature for proposing candidate objects [14].

Moreover, Edge Boxes are robust to varying size of objects. In satellite images, scale and orientation changes are the main characteristics of targets. Moreover, the edge information of targets in satellite imagery contains very prominent and concise attributes. The major challenges in target detection in satellite imagery include presence of targets in different sizes, different orientations and at very close locations. BING generates very loosely fitting proposals and thus is only suitable at low IoU. Selective search is relatively good in general object detection, but it is considerably slower and doesn't perform well if the size of objects is rather small.

Edge Boxes provides best tradeoff between speed and quality. An automatic target detection method based on Edge Boxes and Convolutional Neural Networks (CNN) is proposed in this paper. Fig. 1 illustrates the conceptual level block diagram of the proposed system. We use Edge Boxes to produce object proposals in the initial stage. The candidate objects proposed by Edge Boxes are filtered using some geometric checks while maintaining high recall rates. We then feed the potential object proposals to CNN for automatic feature extraction and classification.

Finally, the performance of our method is evaluated on a large military target dataset which contains tank and non-tank patches for training and test images as shown in Fig. 2. The target detection results using Edge Boxes and CNN are illustrated in Fig. 5. These methods are for panchromatic data. The proposed algorithm can be used for detection of any type of targets. However, we have only detected aircrafts so far due to availability of the aircraft dataset. The literature related to detection of any military target is of interest here. The proposed target detection system is explained in detail in Section II. The experimental analysis is presented in Section III followed by the results, conclusion and future prospects in Section IV and V, respectively.

II. TARGET DETECTION SYSTEM

The proposed framework for target detection system is shown in Fig. 1. We detect objects in input image using Edge Boxes and apply geometric checks to select military targets among the object proposals. A well-trained CNN is used to extract features of the proposed objects and classify them as aircraft or non-aircraft objects.

1) *Candidate Objects Proposal:*
Using Edge Boxes, the edge information of objects is very useful in remote sensing because it contains very prominent and concise attributes as shown. The Edge Boxes technique presented in [14] leverages the edge information to detect objects. In Edge Boxes, a single score from contours confined in a bounding box of the candidate object is calculated and edges with high affinity are grouped together using a greedy approach. The affinity between two groups is given by:

$$a(s_i.s_j) = \left|\cos(\theta_i - \theta_{ij})\cos(\theta_j - \theta_{ij})\right|^{\gamma} \quad (1)$$

Where s_i and s_j represents the pairs of groups, θ_i the mean orientation of a group s_i and the angle between mean positions of groups s_i and s_j is represented by $\theta_{i.j}$. The sensitivity of affinity to orientation variations is controlled by γ. The score of a bounding box of a candidate object is given by:

$$h_b = \frac{\sum_i (w_b(s_i) m_i)}{2(b_w + b_h)^k} \qquad (2)$$

Where width and height of a box are represented by b_w and b_h respectively, the sum of magnitude of all edges in a group s_i is represented by m_i. The value of $w_b(s_i) \in [0 \,.\, 1]$ indicates whether b contains s_i or not. To normalize the score, the magnitude of edges from box b^{in} centered in b is subtracted to improve the accuracy of Edge boxes:

$$h_b{}^{in} = h_b - \frac{\sum_{p \in b^{in}} (m_p)}{2(b_w + b_h)^k} \qquad (3)$$

Where b^{in} has height and width equal to $b_h/2$ and $b_w/2$ respectively.

2) *Candidate Objects Selection:*
The candidate objects proposed by Edge Boxes are very large in number for classification by CNN. We apply some geometric checks to discard the objects which are unlikely to be aircraft objects. The patches of aircraft in satellite images are generally small in size and square shaped, thus the objects with very large or very small area and high aspect ratio are discarded. The objects left behind in geometric filtering are passed to the CNN for automatic feature extraction and image classification.

3) *Feature Extraction and Classification using CNN:*
The Convolutional Neural Network (CNN) is a modern deep learning method which is being widely used for image analysis tasks such as image classification and object detection and segmentation. Krizhevsky et al. [18] achieved excellent recognition rates on Large Scale Visual Recognition Challenge dataset using standard backpropagation for training a deep CNN. A CNN consists of several layers: convolutional, activation and pooling layers in alternation followed by a fully connected layer that produces the output. Unlike typical neural networks, only a small region of input neurons known as Local Receptive Field (LRF) is connected to the hidden neurons. LRF is translated across the image using convolution to map the input to hidden neurons. The hidden layers in CNN learn to detect different features in an image. The weights and biases for all neurons is a hidden layer are the same. Thus, all hidden neurons detect the same features such as edges and blobs in different regions of an image, making the CNN tolerant to translation of objects in an image. Activation transforms the output of each neuron by using activation functions such as Rectified Linear Unit (ReLU) which maps the output of a neuron to the highest positive value, or if the output is negative, ReLU maps it to zero. Pooling reduces the dimensionality of the feature map by condensing the output of small regions of neurons into a single output, thus simplifying the following layers and reducing the number of parameters to learn. The final layer connects the neurons from the last hidden layer to the output neurons which produce the final output. The class probabilities are determined by the value of each node in the final layer

III. SIMULATION

Dataset Specifications:
We use the publicly available military target dataset. The dataset contains 500 tank patches, 5000 non-tank patches. Some patches and test images are shown in Fig. 2. We resize the patches to 32×32 and use them to train the CNN. Simulation was done using PYCHARM with PYTHON 3.9 language as shown in Fig.3

Figure 3 Data set collection

Training:
By using The Convolutional Neural Network (CNN) algorithm the code used to train data set shown in Fig. 4

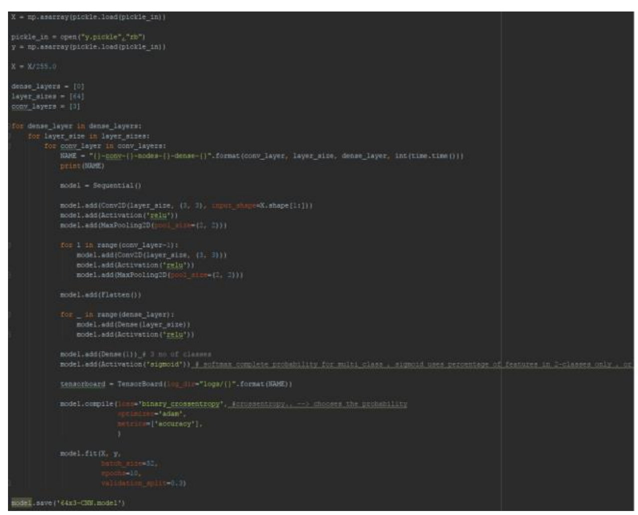

Figure 4 Training data set

Evaluation Metrics:

We use precision and recall as our evaluation metrics. They are defined as:

$$Precision = \frac{Number\ of\ Detected\ Targets}{Number\ of\ Detected\ Objects} \quad (4)$$

$$Recall = \frac{Number\ of\ Detected\ Targets}{Number\ of\ Targets} \quad (5)$$

The tank dataset contains 1000 images. Some examples are shown in Fig. 2(c). Edge boxes method is used to generate object proposals which are filtered using geometric checks as discussed in Section II. For each detected object in the input image, the trained CNN is used to predict whether it is aircraft or a non-aircraft object. The results show that the proposed system reaches high value of precision and recall both respectively on the complete dataset and achieves a high average precision and recall of respectively on the complete dataset. Excellent precision and recall rates have been obtained in images with unvarying illumination and less noise. Relatively lower recall rates are obtained in images where the aircraft objects coincide or overlap with other objects. Lower precision rates with high recall rates are observed in images containing a lot of objects similar to tank objects.

IV. RESULTS

It was found that by using CNN algorithm higher efficiency was achieved of target detection as shown in Fig. 5

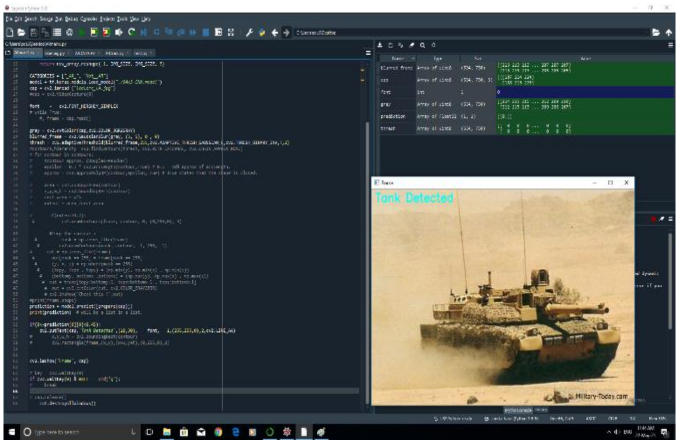

Figure 5. Proposed detection program.

The experimentation results clearly show the effectiveness of the proposed target detection system in complex scenes.

V. CONCLUSION

Automatic target detection in aerial imagery has great significance in military applications. In our work, we propose the use of Edge Boxes algorithm for object detection and CNN for classification in satellite images. The Edge Boxes technique leverages the edge information to detect objects. Edge Boxes is robust to varying size of objects. CNN effectively learns optimum features directly from huge amount of data automatically. Moreover, CNN is invariant to minor rotations and shifts in the target object. Encouraging experimental results have been obtained on a large dataset. The high precision and recall rates show the optimum performance and robustness of our system in complex scenes.

In our future work, we will try to improve the performance of our system and lower the computational cost. We will also apply it in other areas where target detection is used.

VI. REFERENCES

[1] L. Zhang, Z.W. Shi, X.R. Yu, “A hierarchical oil depot detector in highresolution images with false detection control”, 7th International. Congress on Image and Signal Processing (CISP), 2014, pp. 530-535

. [2] X.W. Han, Y.L. Fu, G. Li, “Oil Depots Recognition Based on Improved Hough Transform and Graph Search”, Journal of Electronics & Information Technology, Vol. 33, No.1, pp. 66-72, 2011.

[3] C. Yildiz, E. Polat, “Detection of stationary aircrafts from satellite images”, IEEE 19th Conference on Signal Processing and Communications Applications (SIU), 2011, pp. 518-521.

[4] H.A. Khan, M. Salman, S. Hussain, K. Khurshid, “Automation of Optimized Gabor Filter Parameter Selection for Road Cracks Detection”, International Journal of Advanced Computer Science and Applications, Vol. 7. No. 3, pp. 269-275, 2016.

[5] H. Iftikhar, K. Khurshid, “Fusion of Gabor Filter and Morphological Operators for the Detection of Settlement Zones in Google Earth Satellite Images”, IEEE International Conference on Signal and Image Processing Applications (ICSIPA), Nov 2011.

[6] J.W. Hsieh, J.M. Chen, C.H. Chuang, K.C. Fan, “Aircraft type recognition in satellite images”, IEE Proceedings - Vision, Image and Signal Processing, Vol. 152, No.3, pp. 307-315, 2005.

[7] G.E. Hinton, S. Osindero, Y.W. Teh, “A fast learning algorithm for deep belief networks”, Neural Computation 2006, Vol. 18, No.7, pp.1527-1554, 2006.

[8] Y. Lecun, L. Bottou, Y. Bengio, P. Haffner, "Gradient-based learning applied to document recognition," in Proceedings of the IEEE, vol. 86, no. 11, pp. 2278-2324, Nov 1998.

[9] G.E. Hinton, “A practical guide to training restricted boltzmann machines”, Neural Networks: Tricks of the Trade - Springer, 2012, pp.599-619.

[10] X.Y. Chen, S.M. Xiang, C.L. Liu, C.H. Pan, “Aircraft Detection by Deep Belief Networks”, The 2nd IAPR Asian Conference on Pattern Recognition, 2013, pp. 54-58.

[11] W. Li, S.M. Xiang, H. B. Wang, C.H. Pan, “Robust airplane detection in satellite images”, 18th IEEE International Conference on Image Processing (ICIP), pp. 2821-2824, 2011.

[12] F. Zhang, B. Du and L. Zhang, "Saliency-Guided Unsupervised Feature Learning for Scene Classification," in IEEE Transactions on Geoscience and Remote Sensing, Vol. 53, No. 4, pp. 2175-2184, April 2015.

[13] S. Sattar, H. A. Khan, K. Khurshid, “Optimized Class-Separability in HyperSpectral Images”, 36th IEEE International Geo-science and Remote Sensing Symposium (IGARSS), July 2016

[14] C. L. Zitnick, P. Dollar, “EdgeBoxes: Locating object proposalsfrom edges”, European Conference on Computer Vision Zurich, pp. 391-405, September 2014.

[15] M.M. Cheng, Z.M. Zhang, W.Y. Lin, P. Torr, “BING: Binarized normed gradients for objectness estimation at 300fps”, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp.3286-3293, 2014.

[16] J.R. R. Uijlings, K.E. A. van de Sande, T. Gevers, A.W.M. Smeulders, “Selective search for object recognition”, International Journal of Computer Vision - Springer, Vol. 104, No.2, pp.154-171, 2013.

[17] R. Girshick, J. Donahue, T. Darrell, J. Malik, “Rich feature hierarchies for accurate object detection and semantic segmentation”, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp.580-587, 2014.

[18] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet Classification with Deep Convolutional Neural Networks”, The 26th Annual Conference on Neural Information Processing Systems (NIPS), pp. 1097-1105, 2012.

[19] A.J. Chen, J.Z. Li, “Automatic recognition method for quasicircular oil depots in satellite remote sensing images”, OptoElectronic Engineering, Vol. 33, No.9, pp. 96-100, 2006.

[20] P. Dollar, C.L. Zitnick, “Fast edge detection using structuredforests”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. PP, No.99, pp.1, 2014.