iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 3 Issue: VI Month of publication: June 2015

DOI:

International Journal for Research in Applied Science & Engineering Technology (IJRASET)

# Object Detection Using Diversified Search by Measuring Objectness of Image Window

Preeti Sinha[1], Toran Verma[2]
[1,2]*Department of Computer Science and Engineering, RCET, Bhilai.*

***Abstract— Object Detection and recognition is an important job in remote sensing and automation. Conventional object detection approaches are usually not sufficiently sturdy in addressing the variations of targets and typically suffer from restricted training samples. Search of possible object position in the image is also a restricting situation in object detection and needs to find a solution. This paper introduces a robust object detection approach which combines the diversified search approach and segmentation strategies. In proposed work, we tend to aim to capture all potential object locations. Rather than one technique to come up with potential object locations, we tend to diversify our search and use a range of complementary image partitioning to include all the possible conditions. As number of locations are reduced for search, this results in speeding up the searching process. Experimental results prove that the proposed method is efficient in reaching convergence quickly and gives better quantitative result.***
***Index Terms—Object Detection, LBP, HOG, SIFT .***

## I. INTRODUCTION

Object detection and recognition are the important job in the computer vision. An object detection system finds objects in the real world from an image of the world. Object detection can be defined as the in an exceedingly two-classification of object and non-object locations. Object detection algorithms generally use extracted options associated learning algorithms to acknowledge instances of an object class. It's unremarkably utilized in applications like image retrieval, security, surveillance, and automated systems. That while not segmentation is meted out over appropriate feature vectors extracted from image patches windows. Objects are standalone things with a well-defined boundary and center, such as cows, cars, and telephones, as opposed to amorphous background stuff, such as sky, grass, and road [10]. Detection of object from the complete image mainly depends upon the portioning if the images. An object in an image has close boundary in the space or different appearance from its surrounding or it is unique within an image [1].

Object detection sometimes leads to failure in situation like tiny object compared to image, very large object which contains most of the area in image, hierarchical objects and objects which have overlapping boundaries. In Figure salad and spoon is inside the bowel and the bowels are on the table. Therefore the category of the objects needs to be hierarchical. In Figure 1b the object boundaries are overlapped and the color of the background and object is same then different feature needs to be consider to get proper object region. In Figure 1c object in the image cover almost half of the image which creates problem in detection. In Figure 1d object (airplane) in the image is very small and almost escape from the detection system. Looking at the last two examples it can be say concluded that different size and shape of the detection window should be considered. Although objects will seem at a spread of locations and sizes inside a picture, some windows are more appropriate to cover object than others, even while not analyzing the component patterns within them. Challenges above can be overcome if two thing is considered first is segmentation and second is object window. Segmentation should be diversified that is different features need to be considered to create regions and different merging criteria. Object window is also very important should consider different shape and sizes of object window. This paper presents an adaptive method which include objectness measure and diversified search approach. objectness measure greatly reduce the number of evaluated windows evaluated by modern class-specific object detectors and it explicitly train it to distinguish windows containing an object from background windows. Objectness steers the localization algorithm towards objects rather than backgrounds [1]. Motivated by bottom-up segmentation, we tend to aim to take advantage of the structure of the image to come up with object locations. impressed by thorough search, we tend to aim to capture all

International Journal for Research in Applied Science & Engineering Technology (IJRASET)

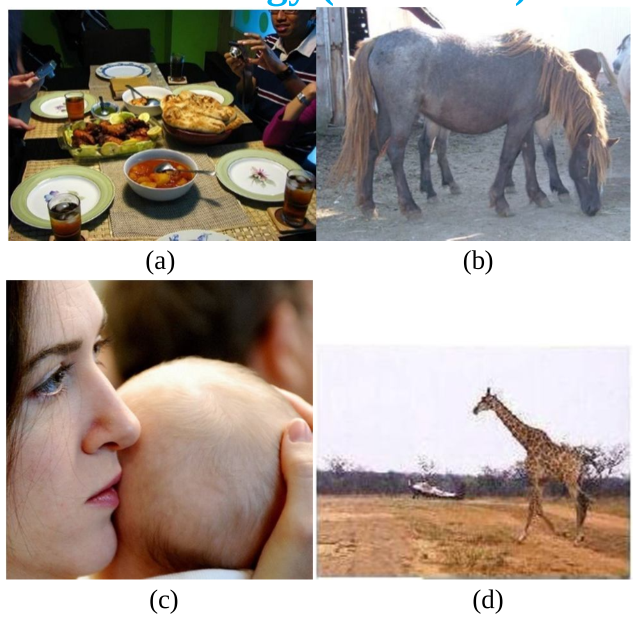



Fig. 1. Object detection Challenges (a) hierarchical objects (b) overlapping boundaries (c) very large object (d) tiny object

potential object locations. Therefore, rather than employing a single sampling technique, we tend to aim to diversify the sampling techniques to account for as several image conditions as potential. Specifically, we tend to use a data-driven grouping based strategy wherever we tend to increase diversity by employing a style of complementary grouping criteria and a spread of complementary color areas with completely different properties. The set of locations is obtained by combining the locations of those complementary partitionings.

## II. RELATED WORKS

Previously various work has been done on object detection. Berg and Berg [11] realize painting pictures representative for an object category. Their approach returns pictures having an oversized, targeted object that is clearly separated from the background. These approaches don't appear appropriate for the PASCAL VOC 07 dataset wherever several are appeared in an image and that they are seldom principal.

Blaschko and Lampert [12] proposed victimisation the looks model to guide the search. This each alleviates the constraints of employing a regular grid, mounted scales, and glued ratio, whereas at constant time reduces the quantity of locations visited. This can be done by directly looking for the best window inside the image employing a branch and certain technique.

Felzenszwalb et al. [15] proposed a very successful sliding window approach for object detection. Their technique additionally performs an exhaustive search employing a linear SVM and HOG options. However, they hunt for objects and object components, whose combination leads to a powerful object detection performance.

Alexe, Deselaers and Ferrari [1] presented a generic objectness measure, quantifying how likely it is for an image window to contain an object of any class. They explicitly train it to distinguish objects with a well-defined boundary in space from amorphous background elements. The measure combines in a Bayesian framework several image cues measuring characteristics of objects, such as appearing different from their surroundings and having a closed boundary.

Volume 3 Issue VI, June 2015
ISSN: 2321-9653
International Journal for Research in Applied Science & Engineering Technology (IJRASET)

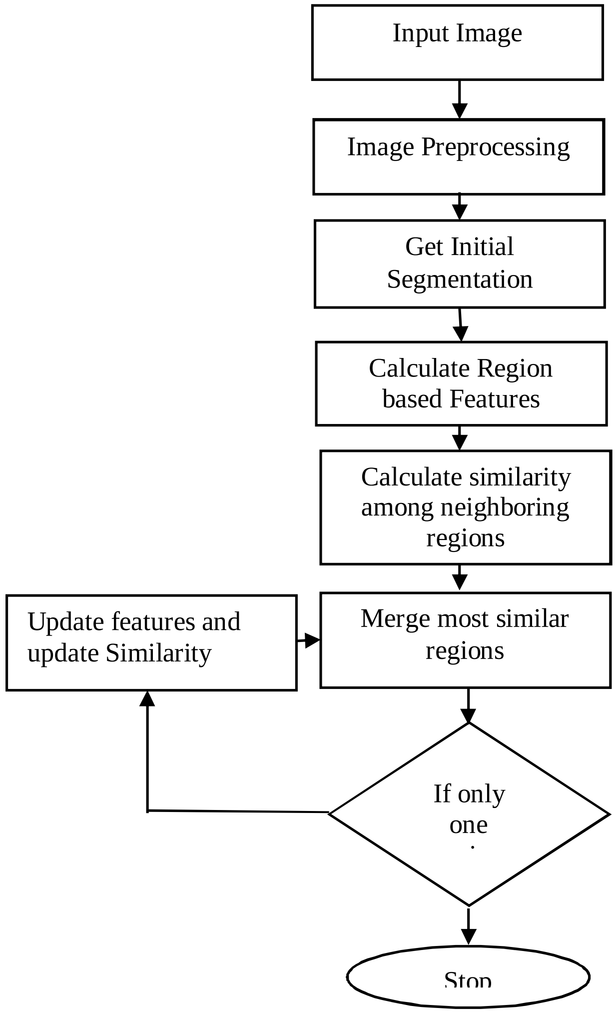



Fig. 2. Flowchart of Proposed approach

These include an innovative cue to measure the closed boundary characteristic. In experiments on the challenging PASCAL VOC 07 dataset. They show this new cue to outperform a state-of-the-art saliency measure, and the combined objectness measure to perform better than any cue alone. This paper is related to above work and also includes new strategies which will increase the efficiency of the object detection system. We include diversified searching for generic objects. Inclusion of objectness measure increases the speed up the convergence.

## III. RESEARCH METHODOLOGY

### A. Object Detection

Flow chart presented in the Figure 2 shows proposed approach. Input image is a normal color image which needs an object detection process. Then preprocessing of input image is done. In preprocessing step the image is converted into the HSV (hue-saturation-value) image. Image segmentation is done using the graph based segmentation method proposed by Felzenszwalb and Huttenlocher in [14]. Let $G = (V, E)$ be an undirected graph with vertices $v_i \in V$ , the set of elements to be segmented, and edges $(v_i, v_j) \in E$corresponding to pairs of neighboring vertices. Each edge $(v_i, v_j) \in E$ has a corresponding weight$w(v_i, v_j)$, which is a non-negative measure of the dissimilarity between neighboring elements $v_i$ and $v_j$. In the case of image segmentation, the elements in $V$ are pixels and the weight of an edge is some measure of the dissimilarity between the two pixels connected by that edge (e.g., the difference in intensity, color, motion, location or some other local attribute)[14]. In the graph-based approach, a segmentation S is a

Volume 3 Issue VI, June 2015
ISSN: 2321-9653
**International Journal for Research in Applied Science & Engineering Technology (IJRASET)**

partition of $V$ into components such that each component (or region) $C \in S$ corresponds to a connected component in a graph $G' = (V, E')$, where $E' \subseteq E$. In other words, any segmentation is induced by a subset of the edges in E. There are different ways to measure the quality of a segmentation but in general we want the elements in a component to be similar, and elements in different components to be dissimilar.

Bottom-up merging is good and general approach to segmentation [6, 13], therefore we have used it for proposed search. As a result of the method of grouping itself is stratified, we will naturally generate locations in all scales by continued the grouping method till the complete image becomes one region. This satisfies the condition of capturing all scales. We use a greedy rule to iteratively merge regions together: Firstly the similarities between all neighboring regions are calculated. Then most similar regions are sorted along, and new similarities are calculated between the ensuing region and its neighbors. The method of grouping the foremost similar regions is recurrent till the complete image becomes one region. We first take four features for each region which will be used to calculate similarity between regions; criteria to merge that regions. These four features are color, texture, size and fill. We calculate one dimensional color histogram for each color channel for each region. Color feature for each region $r_i$ is represented as $C_i = \{C_i^1, \ldots\ldots, C_i^n\}$, where n=75 which represents dimensionality 25 for each color channel. The color histograms are normalized using the $L_1$ norm. We represent texture using fast SIFT-like measurements because SIFT works for material recognition [2]. Gaussian derivatives in eight orientations has been taken using s = 1 for each color channel. For each orientation for each color channel we extract a histogram using a bin size of 10. This leads to a texture histogram $T_i = \{t_i^1, \ldots\ldots, t_i^n\}$, for three color channels for each region $r_i$ with dimensionality n = 240 when three color channels are used. Texture histograms are normalized using the $L_1$ norm. Size features are calculated for each regions $r_i$ which will help to merge smaller regions. Last feature which we will calculate is the fill feature. This feature help to merge the regions which fits into other. Next step is calculation of similarity measure. We outline four complementary, fast-to-compute region based similarity measures. These measures are vary between 0 and 1 that facilitates blend of those similarity measures. These four similarity measure defined by given four equations:

$$S_{color} = \sum_{k=1}^{n} \min(C_i^k, C_j^k)$$

$$S_{texture} = \sum_{k=1}^{n} \min(t_i^k, t_j^k)$$

$$S_{size} = 1 - \frac{size(r_i) + size(r_j)}{size(image)}$$

$$S_{fill} = 1 - \frac{size(BB_{ij}) - size(r_i) - size(r_j)}{size(image)}$$

Where $size(image)$ is the size of image pixels and $BB_{ij}$ is the bounding box around two test regions. Finally combined similarity will be calculated using the formula $S = a_1 S_{color} + a_2 S_{texture} + a_3 S_{size} + a_4 S_{fill}$

Where $a_i$ is having value 0 or 1 depending upon it is used or not. Next step is the merging of two most similar regions and make bounding box on the resulting region depending on the objectness of the bounding box. Duplicate bounding boxes will also be removed by passing only one bounding box to next iterative level. Then in the next step the system will calculate features of the new regions of that iteration. This is achieved with the fast to compute mathematical formulas that inherit the properties of uncombined regions. Features will be calculated using similarity measures of uncombined regions as:

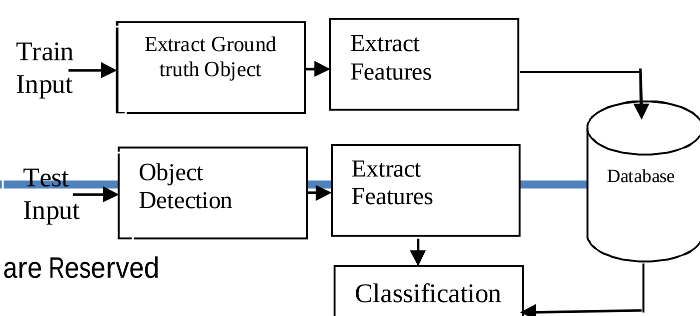
$$size(r_t) = size(r_i) + size(r_j)$$

$$texture(r_t) = texture(r_i) + texture(r_j)$$

The complete process iteratively executes until only one region is left and all the boxes around objects will be added in each iteration.

### *B. Object Recognition*

Block diagram of generalized object recognition system is shown in the Figure 3.



Volume 3 Issue VI, June 2015
ISSN: 2321-9653
International Journal for Research in Applied Science & Engineering Technology (IJRASET)

Two blocks needs to be explained one is feature extraction and classification. Feature extraction for object recognition is a crucial step and two features are standard histograms of oriented gradients (HOG) [5] and bag-of-words [6, 7]. HOG has been shown to be successful in combination with the partbased model by Felzenszwalb et al. [2] but bag of words is more appropriate with proposed object detection task because it is powerful feature[8,9]. Classifier for object recognition can be Support Vector Machine with a histogram intersection. . SVM evaluates more relevant information in a convenient way [11]. As with any supervised learning model, first train a support vector machine, and then cross validate the classifier.

## IV. RESULT AND DISCUSSION

We presented a diversified and fast to compute object detection approach. We tested our proposed object detection method and it gives satisfactory result. To evaluate the quality of our object detection we used the Average Best Overlap (ABO) used in [3]. The Average best overlap was calculated using the formula

$$ABO = \frac{1}{|G^c|} \sum_{g_i^c \in G^c} \, \sum_{l_j \in L} maxoverlap(g_i^c, l_j)$$

Where $g_i^c \in G^c$ represents groudtruth and $l_j \in L$ is object hypothesis. We tested our proposed algorithm on PASCAL VOC 2007.

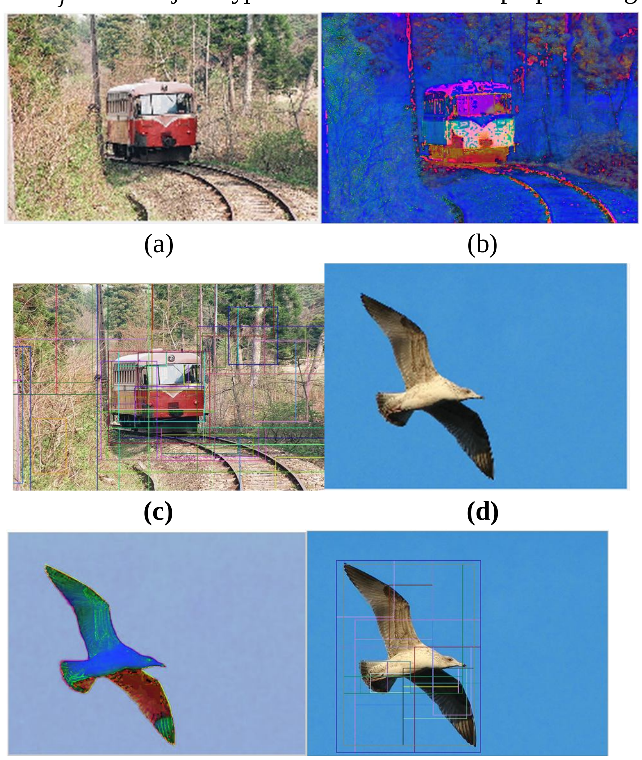



Fig.4. Input Image, HSV Image and Output object detection frames

Result of proposed method is compared with Felzenszwalb et al. [14].

Table 1: Result Comparison

| S. N. | Method | ABO |
|---|---|---|
| 1 | Felzenszwalb et al. [12] | 0.829±0.052 |
| 2 | Single Strategy | 0.690±0.171 |
| 3 | Proposed Method | 0.889±0.029 |

**International Journal for Research in Applied Science & Engineering Technology (IJRASET)**

## V. CONCLUSION

In this paper, we introduce an adaptive object detection approach which uses diversified search strategy and objectness measure. The result shows proposed method is better than the method with which it is compared. Our strategy can without much of a stretch be stretched out to different measurements and utilized for computational improvement and improvement in result. Proposed method can be improved by used better features. Research in this bearing is under way and will be introduced in a future report.

## REFERENCES

[1] B. Alexe, T. Deselaers, and V. Ferrari. Measuring the objectness of image windows. IEEE transactions on Pattern Analysis and Machine Intelligence, 2012. 3, 8, 10, 13.
[2] C. Liu, L. Sharan, E.H. Adelson, and R. Rosenholtz. Exploring features in a bayesian framework for material recognition.In Computer Vision and Pattern Recognition 2010. IEEE, 2010.
[3] I. Endres and D. Hoiem. Category independent object proposals. In ECCV, 2010. 2, 3, 6, 8, 9, 10, 11, 13
[4] Rajesh Yakkundimath, Jagdeesh D. Pujari and Abdulmunaf S. Byadgi, “Classification of Fungal Disease Symptoms affected on Cereals Using Color Texture Features”, International Journal of Signal Processing, Image Processing and Pattern Recognition, 2013, pp. 321-330
[5] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In CVPR, 2005. 1, 2, 3, 5
[6] G. Csurka, C. R. Dance, L. Fan, J.Willamowski, and C. Bray. Visual categorization with bags of keypoints. In ECCV Statistical Learning in Computer Vision, 2004.
[7] J. Sivic and A. Zisserman. Video google: A text retrieval approach to object matching in videos. In ICCV, 2003.
[8] H. Harzallah, F. Jurie, and C. Schmid. Combining efficient object localization and image classification. In ICCV, 2009.1, 2, 3, 5, 6, 8
[9] [17] C. H. Lampert, M. B. Blaschko, and T. Hofmann. Efficient subwindow search: A branch and bound framework for object localization. TPAMI, 31:2129–2142, 2009. 2, 5
[10] G. Heitz and D. Koller. Learning spatial context: Using stuff to find things. In ECCV, 2008.
[11] T. L. Berg and A. Berg. Finding iconic images. In CVPR Internet Vision Workshop, 2009.
[12] C. H. Lampert, M. B. Blaschko, and T. Hofmann. Efficient subwindow search: A branch and bound framework for object localization. TPAMI, 31:2129–2142, 2009. 2, 5
[13] G. Wang and D. Forsyth. Joint learning of visual attributes, object classes and visual saliency. In ICCV, 2009.
[14] P. F. Felzenszwalb and D. P. Huttenlocher. Efficient Graph- Based Image Segmentation. IJCV, 59:167–181, 2004. 1, 3, 4, 5, 7
[15] P. F. Felzenszwalb, R. B. Girshick, D. McAllester, and D. Ramanan. Object detection with discriminatively trained part based models. TPAMI, 32:1627–1645, 2010. 1, 2, 3, 5, 6, 8, 11, 12, 13.
[16] J. Platt, “Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods,” Adv. Large Margin Classifiers, vol. 10, no. 3, pp. 61–74, Oct. 1999.
[17] H. Bay, T. Tuytelaars, and L. Van Gool, “SURF: Speeded up robust features,” in Proc. Eur. Conf. Comput. Vis., 2006, pp. 404–417.
[18] H. Tao, H. S. Sawhney, and R. Kumar, “Object tracking with Bayesian estimation of dynamic layer representations,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 24, no. 1, pp. 75–89, Jan. 2002.

10.22214/IJRASET

INDEX COPERNICUS

45.98

IMPACT FACTOR:
7.129

TOGETHER WE REACH THE GOAL

IMPACT FACTOR:
7.429


# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY