A Review Paper on Object Detection for Improve the Classification Accuracy and Robustness using different Techniques

Divya Patel
Research Scholar, CSE Department
Parul Institute of Technolody,
Limda, Vadodara, India.

Pankaj Kumar Gautam
Asst. Prof., CSE Department
Parul Institute of Technolody,
Limda, Vadodara, India.

ABSTRACT
Object detection is a computer technology that connected to image processing and computer vision that deal with detecting instance objects of certain class in digital images and videos. Object detection is a challenging problem in vision based computer applications. It is used to identifying that whether in scene or image object is been there or not. In this review paper, we are going to present different techniques and methods for detecting or recognizing object with various benefits like efficiency, accuracy, robustness etc.

Keywords
Object detection, robustness.

1. INTRODUCTION
Generally, Object detection has applications in many areas of computer vision, including image fetching and video surveillance[1]. Well-researched domains of object detection include face detection and pedestrian detection. Good object detection system determined the presence or absence of objects in arbitrary scenes and be invariant to object scaling and rotation, the camera view point and changes environment. Address detection problem with different objectives, which are classified into two categories: 'specific' and 'conceptual'. The former involves detection of known objects and letter involves the detection of an object class or interested area. All object detection systems use models either explicitly or implicitly and allocate feature detectors based on these object models. The hypothesis formation and verification components vary in their importance in different approaches to object detection. Some systems use only hypothesis formation and then select the object with highest matching as the correct object. An object detection system must select correct tools and appropriate techniques for the processing. In the selection of appropriate methods for a particular application must been considered by many factors. An object detection system finds objects in the real world from an image of the world, using object models which are known a priori. This process is surprisingly tough.

2. BENIFITS OF TECHNIQUE
In this paper we are going to review different technique and approach for object detection which deal with problems occurring due to detection process and gives beneficial results. The different techniques will be discussed in next section.

Multi-component object detection method is good indistinguishable and powerful. Multi-class Hough transform approach improves the taxonomy of precision for detecting object. Latent Hough transform method can approve imperceptibility. In Boosted haar cascade technique object can be detected efficiently.

3. TECHNIQUES OF OBJECT DETECTION
3.1 Multi-component Object Detection Method
In this method two phases are there, 1. Training phase and 2. Detection phase[2]. In the training phase, a two-layer model is trained to capture and aggregate the components of an object category from data. Each first-layer model is a binary classifier trained with a seed and a list of aligned objects with the seed based on keypoint. A second-layer classifier takes the outputs of these component classifiers as input, and produces a final category-level classification score. In the detection phase, bounding boxes are generated for each image using selection scheme. After scoring these boxes with two-layer model, a non-maximum suppression is applied to produce final detection results. The component models which get are both easy to learn and highly able to recognize small differences. A second layer classifier is learned to sum the outputs of component models into final scores.

$$H(C,i) = \sum S(a) \times \frac{Area\,(B(a)\ \Omega\ B(C))}{Area(B(a))} \times 1(I(a) = I) \qquad \text{Eq. 1}$$

The algorithm achieves good imperceptibility and robustness for object detection[2].

3.2 Multi Class Hough Transform Approach
This approach is used for detecting scalable multi-class object[3]. Scalability of object detectors with respect to the number of classes is a very important matter for applications where many object classes need to be detected. The single-class detectors provide serial complexity for evaluation and the multi-class detectors constrain all objects at once, reduces detection accuracy. To overcome these limitations, a scalable multi-class detection approach is used which measures sub-linearly with the number of classes without decreasing the detection accuracy. As distributed discriminatory features are measured by learning all the classes and detection is also performed for all classes parallel, results in good classification, and increase the performances of the multi-class object detection.

For clustering, it transforms it into a symmetric dissimilarity matrix D by,

$$D = 1 - ½\,(S + S^T) \qquad \text{Eq. 2}$$

This approach also benefits from sharing features and an automatically built category taxonomy for robust scalability without degrading accuracy[3]. This algorithm improves the classification accuracy for detecting multi-class object.

3.3 Latent Hough Transform (LHT)

Latent hough transform based object detection method learn a codebook of voting elements such as the image features, heavy image blocks are so on which are extracted and matched in order to encode the location and scale of the object in the image[4]. This transformation allows partial observation of the training objects to carry a single object theorem and produces wrong positives by accumulating votes that are consistent in location but inconsistent in other properties like pose, color, shape or type. To overcome these drawbacks, the Hough transform is used with latent variables in order to enforce consistency among votes. Therefore, the votes which follow the assignment of the latent variables are considered to support a single hypothesis and Latent Hough Transform based training approach which has multiple weights assignments is applied for obtaining better detection accuracy. The approach can improve the imperceptibility and robustness very well.

3.4 Boosted Haar Cascade Technique

This technique is used for sliding window object detection without spatial clustering. This method draws attention to the fact that now a day sliding window object detection technique becomes remarkable due to its versatility with significant detection performance. This technique fails to achieve due to its insufficient precision and inaccurate localization[5].

Table 1. Quality of raw detection responses

Configuration	Recall	#FP/image	Localization
No filter, no grouping	95%	9.5	0.070, 0.055
No filter, grouping	95%	2.4	0.171, 0.163
filter, no grouping	91%	2.5	0.095, 0.063
filter, grouping	91%	0.35	0.150, 0.139

To overcome these problems, a detection approach is applied which prevent the requirement for spatial clustering of nearby detection responses such as traffic sign detection, where a super class of triangular warning signs can be considered. The main idea of the proposal is to omit blind spatial clustering of raw detections in order to preserve their localization accuracy[6]. It can detect the object efficiently while preserving the information.

4. OCCLUSION PATTERNS

Occlusion can be treat as just another source of noise instead, it include the occluder itself into the modeling, by mining characteristics, reoccurring occlusion patterns from annotated training data[7]. These patterns are then used as training data for dedicated detectors of varying experience.

4.1 Mining Occlusion Pattern

It mine occlusion patterns from training data by averaging fine-grained annotations in the form of 3D object bounding boxes and camera projection matrices that are readily available as part of the KITTI dataset[7].

Feature representation: It uses the following properties of occlusion patterns as features in their clustering: i) occlude left/right of occludee in image space, ii) occluder and occludee orientation in 3D object coordinates, iii) occlude is/is not itself occluded, iv) degree of occlusion of occludee.

Rule-based clustering: It found that a simple, greedy clustering scheme based on repeatedly splitting the training data according to fixed rules (e.g. based on assigning the viewing angle of the occluder to one of a fixed number of predetermined bins) resulted in sufficiently clean clusters.

4.2 Single Object Occlusion Pattern

A single-object class detector specifically trained to detect occluded objects from multiple viewpoints, occluded by various occluders[8]. In addition to the original components c $=1, \dots, C_{VISIBLE}$ that represent the appearances of instances of an object class of interest, it introduce additional mixture components dedicated to representing the different appearance of *occluded* objects of that class. In particular, it reserve a distinct mixture components, for each of the *occludee* members of clusters resulting from occlusion pattern.

4.3 Double-object Occlusion Pattern

A hierarchical double-object detector explicitly trained for accurate occluder/occludee bounding box localization[8]. While the single-object occlusion model has the potential to represent different occlusion patterns in the data, modeling occluder and corresponding occlude jointly suggests a potential improvement: the strong proof of the occluder should provide strong cues as to where to look for the occluded object. In these models occluder and occlude-occluded object are allowed to move w.r.t. a spatial models much like parts in the DPM.

5. CONCLUSION

In this paper we review different super-resolution based methods that can enhance efficiency and robustness. An object recognition system finds objects in the real world from an image. The object recognition problem can be defined as a labeling problem based on models of known objects. Formally, given an image containing one or more objects of interest and a set of labels corresponding to a set of models known to the system, the system should assign correct labels to regions, or a set of regions, in the image. For better result in occlusion patterns Shadow c-means approach can be used in future. It will very useful for better performance for object detecting.

6. REFERENCES

[1] Ashwin Deshpande, " Object Recognition Using Large Datasets".

[2] Chunhui Gu, Pablo Arbelaez, Yuanqing Lin, Kai Yu and Jitendra Malik, "Multi-Component Models for Object Detection", European Conference on Computer Vision, pp. 445-458, 2012.

[3] Nima Razavi, Juergen Gall and Luc Van Gool, "Scalable Multi-class Object Detection", IEEE Conference on Computer Vision and Pattern Recognition, pp. 1505-1512, 2011.

[4] Nima Razavi, Juergen Gall, Pushmeet Kohli, Luc Van Gool, "Latent Hough Transform for Object Detection", European Conference on Computer Vision, pp. 312-325, 2012.

[5] Bjorn Ommer and Jitendra Malik, “Multi-Scale Object Detection by Clustering Lines”, International Conference on Computer Vision, pp. 484-491, 2009.

[6] Sinisa Segvic, Zoran Kalafatic and Ivan Kovacek, “Sliding window object detection without spatial clustering of raw detection responses”, IFAC Symposium on Robot Control, pp. 114-119, 2012.

[7] A. Geiger, P. Lenz, and R. Urtasun. Are we ready for autonomous driving? the KITTI vision benchmark suite. In *CVPR*, 2012.

[8] Bojan Pepik, Michael Stark, Peter Gehler and Bernt Schiele, “Occlusion Patterns for Object Class Detection”, IEEE Conference on Computer Vision and Pattern Recognition, 2013.

[9] Xiaofeng Ren and Deva Ramanan, “Histograms of Sparse Codes for Object Detection”, IEEE Conference on Computer Vision and Pattern Recognition, pp. 3246-3253, 2013.

[10] Sushmita Mitra , Witold Pedrycz , Bishal Barman, “Shadowed c-means: Integrating fuzzy and rough clustering”, Elsevier