Category Independent Object Detection via Saliency Maps

Tomislav Puric
Zhiyong Wang, David Feng
School of Information Technologies
FACULTY OF ENGINEERING & INFORMATION TECHNOLOGIES

Motivation

- A plethora of images are taken everyday from numerous devices around the globe with little to no meaning of their content
- Humans can recognise a multitude of objects with minimal effort, despite various morphological representations. Computationally deriving such representations are not easily accomplished
- Most object detection methods provide category-dependent solutions [1] that achieve significant performance, however perform poorly to unknown categories due to their non-scalable characteristics
- Current progress such as [2] aim for category-independent methods for detecting and localising generic objects without specifying their categories, providing priors to improve object recognition algorithms
- Alexe et al. [2] work is limited to the bias of bounding box restrictions as they overlap surrounding background regions, impacting their 'objectness' measure.

Solution

- We take a category-independent approach combining segmentation and saliency cues as a prior
- Using saliency as a prior provides us with multiple region seeds rather than heuristic methods to evaluate key object regions in an over-segmented image
- We generate a binary segmentation of foreground (representing multiple generic objects of interest) and background (remaining regions)
- Our Approach consists of 4 main steps as illustrated in our pipeline (Fig 2.) and explained in the method section

Background

- Object Detection is the process of separating an image into two categories; foreground (objects) and background (non-objects) and delivering those object locations towards an object recognition pipelines)
- Labelling an image with is a difficult procedure due to nested, elongated and overlapping objects (Fig 1.)
- Current approaches are either category-dependent or category-independent

Fig 1. Image 000504 from the PASCAL dataset [3] and its labels demonstrates examples of dataset difficulties. Here the poster is ignored but the people in them are not

Dependent

- Category-dependent techniques such as [1] place emphasis on training data with features
- Features such as SIFT and HOG detectors are extracted from regions, and then classified for training purposes
- Up to 200 categories can be identified with high accuracy and efficiently once trained however they perform poorly given category-independent tasks.

Independent

- Category-independent research emphasises generic objects detection over categorical objects using image cues as seeds.
- Work in [2] demonstrates how objects are in well-defined closed boundaries, differ from their surroundings, and are salient
- Further work[4] provides segmented regions over bounding boxes for more suitable coordinate frames

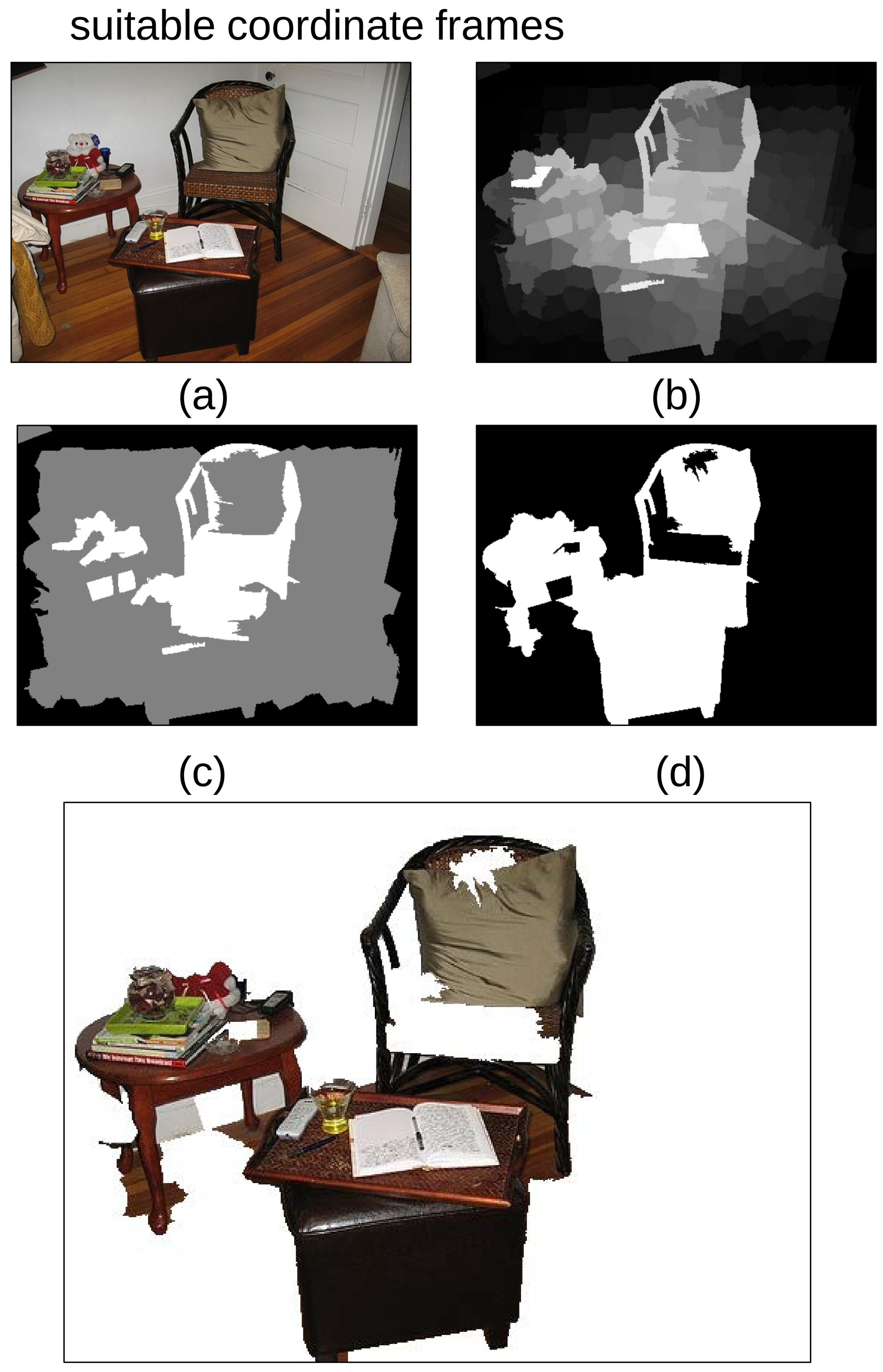

Fig 2. Pipeline approach (a) input, (b) saliency map, (c) region features for GMM [black = background, white = foreground], (d) result of ranking algorithm with GMM classification (e) output

Method

Step 1: Prior

- With many segmentation algorithms to pick, we select SLIC due to its efficiency, control over the number of regions and the overall compactness. SLIC is also the precursor to our saliency map prior.
- We follow the bottom-up saliency approach of absorbing Markov Chains[5], which places focus on category-independent objects and considers the appearance, divergence and spatial distributions of salient objects.
- Step 1 is illustrated in Fig 2 (b), providing us with an over-segmented saliency map

Step 2: Feature Extraction

- Features are extracted using CEDD[8] in order to provide us with a 24-bin histogram

Step 3: GMM Classification

- The 24 features from the selected border regions represented in black (Fig 2c.) are fed into a Gaussian Mixture Models (GMM) in order to provide us with possible foreground regions to extend our manifold-ranking implementation in step 4. The GMM can be written as:

$$\mathbf{P(\theta|x)} = \sum_{i=1}^{k} (\phi_i N(\mu_i, \Sigma_i))$$

Step 4: Manifold Ranking

- With promising foreground regions generated from Step 3 as a prior, a one-class classification ranking algorithm [6] is derived to improve the binary segmentation. The ranking algorithm can be written as

$$\mathbf{f^*} = (\mathbf{I} - \alpha\mathbf{S})^{-1}\mathbf{y},$$

- Where **I** is an identity matrix, $\alpha = 1/(1+\mu)$ and **S** is the normalised Laplacian matrix, $\mathbf{S} = \mathbf{D}^{-1/2}\mathbf{W}\mathbf{D}^{-1/2}$.
- Results are produced as a binary map (Fig 2. d, e)

Results

- Experiments were performed on the PASCAL VOC 07 dataset[3] and is compared with [7].
- PASCAL VOC 07 only contains 210 images out of the 4952 that have their ground truth segmentations labelled. This sample was used in the analysis and results are shown in Fig 4.

	Ours	[7]
Precision	0.66	0.64
Recall	0.37	0.18

Fig 3. Recall is drastically improved whilst Precision is on par with [7]

- Further analysis on the remaining images with bounding box labels, [4] and the MSRA dataset are currently in process

References

[1] R. B. Girshick, J. Donahue, T. Darrell, and J. Malik. Rich feature hierarchies for accurate object detection and semantic segmentation. arxiv:1311.2524 [cs.CV], 2013.
[2] B. Alexe, T. Deselaers, and V. Ferrari. What is an object? In CVPR, pages 73–80, 2010
[3] M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman. The PASCAL Visual Object Classes Challenge. In IJCV, pages 303-338, 2010.
[4] X. Li, Y. Li, C. Shen, A. R. Dick, and A. Hengel. Contextual hypergraph modelling for salient object detection. ICCV, 2013
[5] Jiang, B., Zhang, L., Lu, H., Yang, C., Yang, M.H.: Saliency detection via absorbing markov chain. In: ICCV, 2013
[6] C. Yang, L. Zhang, H. Lu, X. Ruan, and M.-H. Yang. Saliency detection via graph-based manifold ranking. In CVPR, 2013.
[7] S. A. A. Shah, M. Bennamoun, F. Boussaid, and A. A. El-Sallam. Automatic object detection using objectness measure. In 1st International Conference on Communications, Signal Processing, and their Applications (ICCSPA), pages 1-6, 2013.
[8] S. A. Chatzichristofis and Y. S. Boutalis. Cedd: Color and edge directivity descriptor: A compact descriptor for image indexing and retrieval. In A. Gasteratos, M. Vincze, and J. K. Tsotsos, editors, ICVS, volume 5008 of Lecture Notes in Computer Science, pages 312–322. Springer, 2008.