CATEGORY INDEPENDENT OBJECT PROPOSALS USING QUANTUM SUPERPOSITION

Junaid Malik, Caglar Aytekin, Moncef Gabbouj

Laboratory of Signal Processing, Tampere University of Technology, Tampere, Finland

ABSTRACT

Object proposals improve the efficiency of object detection by providing probable locations of objects in an image. Most of the state-of-the-art object proposal methods employ a supervised approach and learn object features from ground truth annotations. We present a novel unsupervised approach for generating object proposals that is based on the human visual system and quantum mechanical principles. Despite of being devoid of any learnt priors pertaining to objects in images, the proposed method is shown to yield competitive results with supervised approaches.

Index Terms— object detection, object proposals, quantum superposition, category-independence.

1 INTRODUCTION

Object recognition and detection are important tenets in the pursuit of gaining scenic understanding from digital images. The aim in object recognition is to predict the presence or absence of an object of a particular class in an image whereas object detection refers to estimating the exact location of the object along with its class [1]. Object detection pipelines [2]–[4] used to adopt a sliding window approach to generate candidate windows spread over the entire image at possibly different scales. Each window is evaluated and classified separately by the detection algorithm. Such an exhaustive search strategy makes the whole process computationally expensive and prohibits the use of more sophisticated detection algorithms [5].

Object proposal generators provide a pool of probable locations in an image which have a high likelihood of covering an object. Applied as a precursor to detection, object proposals narrow down the number of possible object locations that are evaluated, thus relieving the computational burden on the detector. Contrary to detection, object proposals are desired to be category-independent to ensure their utility for any detection pipeline. Detection frameworks of [6]–[11] have successfully used object proposals, exhibiting state-of-the-art performance on ImageNet[12] and PASCAL [1] datasets.

Current object proposal generation schemes can be divided into two categories on the basis of their output: window-based and region-based [5]. Window-based methods evaluate rectangular windows based on their *objectness*, i.e., the likelihood of enclosing an object in the image. Windows with a high enough score are treated as detection proposals. Alexe et al. [13] proposed an objectness measure based on boundary characteristics and saliency of windows. Rahtu et al. [14] improved the results in [13] by presenting more efficient measures to sample windows from an image and exploiting object boundary characteristics. BING [15] uses normed gradients of resized detection windows as features to identify windows containing objects. Edge Boxes [16] groups edges obtained using structured forest [17] and defines the objectness of a window as being proportional to the fraction of edges contained in it.

Region-based methods produce proposals as pixel-wise segments instead of bounding boxes. CPMC [18], Randomized Prim's [19], Geodesic [20] and Endres's method [21] generate multiple foreground segments based on foreground seeds. Seeds are placed either randomly [19], uniformly [18], based on boundary cues [21] or at learnt locations [20] in the image. In case of [18] and [21], the proposals are also ranked based on a supervised objectness scoring function. RIGOR [22] reduces the computational load associated with solving multiple foreground segmentation problem by reusing graphs for different seeds.

Another approach to generating segments is to merge fine over-segmentations of an image. Selective search [23] obtains over-segmentations using the graph-based approach of [24] at multiple scales and merges them based on color, textural and geometrical properties of regions. Multiscale Combinatorial Grouping (MCG) [25] segments the image using normalized cuts [26] and hierarchically merges the resulting segments and ranks them. Rantalankila [14] aims to improve the results of [23] by incorporating a global search and using different oversegmentation schemes.

Window-based methods in general, are computationally more efficient compared to region-based ones [27]. However, they lack localization accuracy because the object's exact location is not precise and is only approximated using a bounding box. Region-based methods, albeit computationally more expensive, provide pixel-wise segmentations which make them more suitable for applications that require details about the object's spatial extent e.g. robotics. All region-based proposal methods discussed above are either heavily parameter-dependent or rely on learning some quantitative notion of objectness from ground truth annotations. This introduces a bias towards the categories of objects present in the training dataset, which is not desirable in a category-independent framework.

We propose an unsupervised method for generating proposals that is inspired by the human visual attention where the search for objects starts from the most salient parts of a scene and gradually transitions towards other regions. We employ a state-of-the-art saliency estimation technique, QCUT [28] and modify its formulation to extend the search of objects beyond the salient ones. The rest of the paper is organized as follows. Section 2 describes the proposed method for generating object proposals. Section 3 describes the unsupervised ranking strategy, Section 4 introduces the non-maxima suppression that is used, while Section 5 provides evaluation results in comparison to the state-of-the-art. Section 6 concludes the paper.

2 OBJECT PROPOSAL GENERATION

In QCUT [28], a spectral salient object detection is proposed on a graph formed by image superpixels. This saliency map corresponds to the element-wise square magnitude of the eigenvector ψ_1 corresponding to the minimum eigenvalue E_1 of the Hamiltonian matrix $\boldsymbol{H}$ (see Eq. (1)).

$$\boldsymbol{H}(i,j) = \begin{cases} V(i) + \sum_{k \in N_i} w_{ik} & if\ i = j \\ -w_{ij} & if\ j \in N_i \\ 0 & otherwise \end{cases} \quad (1)$$

In (1), w_{ij} represents the edge weights between two nodes (superpixels i and j) and is proportional to the color similarity between nodes in the Lab color space. $V(i)$ is the unary weight which encodes the background prior. N_i is the set of graph neighbors of node i.

For visual saliency estimation, the superpixels on the boundary of the image are assumed to belong to the background and thus a large background potential is assigned to them. No assumption is made about the foreground/background association of the remaining nodes and their potential is set to zero.

2.1 Decreased Background Potential

As a first step towards extending visual saliency, we decrease the background potential for boundary superpixels in order to increase the likelihood of detecting objects near the image boundary. The effect of this modification is illustrated in Figure 1. We can clearly see that with default potential, the objects of interest in the left side of the image are labelled as background. However, when the potential is decreased, these objects are also highlighted.

Figure 1: (from left to right) Original Image, QCUT output, QCUT output with reduced potential.

2.2 Multiple Eigenstates

To further expand object search, we explore eigenvectors corresponding to progressively higher eigenvalues. We still restrict ourselves to lower energy states as we are interested in foreground only. The saliency maps obtained from these relatively higher energy eigenstates provide locally optimum solutions to the visual saliency estimation problem and can highlight objects that are not salient in the global context [29]. A similar multispectral approach has been employed earlier in [29] to generate salient object segments. We have empirically chosen to explore 50 eigenvectors corresponding to the 50 smallest eigenvalues for our simulations.

2.3 Quantum Superposition

As observed in [28], [30], [31], the ground state eigenvector of $\boldsymbol{H}$ corresponds to the wave function or eigenstate of a quantum mechanical particle at minimum energy. The eigenvectors corresponding to other eigenvalues are thus equivalent to the eigenstates of the particle at other energy levels. The principle of quantum superposition states that any wave function ϕ can be decomposed into a linear weighted combination of eigenstates as shown in (2).

$$\phi = c_1^k \psi_1 + c_2^k \psi_2 + \cdots + c_n^k \psi_n \quad (2)$$

where the coefficients $c_1, c_2, c_3 .. c_n$ obey the normalization criterion. Since the eigenstates are mutually orthogonal, the expression can be used to generate any labelling vector. However, as we are interested only in salient regions and not the background, we restrict ourselves to lower energy eigenstates. Moreover, to produce valid foreground segmentations, we learn the coefficients from pixel-wise ground truth object annotation in PASCAL VOC 2012 dataset [1]. It is worth noticing here that we are not learning the color, textural or spatial features of objects but rather use the ground truth annotations to learn possible ways to combine the basis eigenvectors. This is a very *shallow* form of learning and does not encode any category-specific information about the objects.

Figure 2: (from left to right) Original Image, Saliency Map corresponding to lowest eigenstate, a superposed state.

For each image, we employ 500 superpositions corresponding to 500 clustered sets of coefficients c obtained from ground truth annotations. Figure 2 illustrates an example where a superposed state highlights the object of interest better than the ground state.

2.4 Object Proposal Generation

To extract proposals from the saliency maps obtained as a result of the operations described in Sections 2.2 and 2.3, a

multilevel thresholding operation is applied. Each saliency map is quantized to 256 levels and thresholded at each level. All connected components in the thresholded binary image are added to the pool of proposals. Such an operation ensures that the regions with high saliency are represented more frequently in the generated pool of proposals. A side-effect of this is the introduction of multiple spatially redundant proposals. To rectify this, we employ a non-maximum suppression operation as described in Section 4.

3 UNSUPERVISED RANKING

After the initial pool of proposals has been obtained, segments are ranked based on an unsupervised strategy. The following cues are used in the scoring function.

3.1 Average Local Saliency

Average local saliency ($\Theta_{Saliency}$) refers to the mean saliency value of the proposed region in the saliency map used to generate the proposal. This scoring function penalizes under-segmentations, which produce proposals that contain objects conjoined with part of the background.

3.2 Perimeter Edge Density

The perimeter edge density (Θ_{ED}) cue is similar to the boundary edge density cues utilized in [13]. However, instead of shrinking the original window, we simply average the edge-map values, obtained from [17], for pixels that lie on the segment's perimeter. An example of edge density scores of several proposals is given in Figure 3.

Figure 3: (from left to right) Proposals with edge density scores 0.23,0.27 and 0.31, respectively.

3.3 Eigenvalues

As the eigenvector corresponding to the smallest eigenvalue provides the solution for salient object segmentation, it is more likely to produce a proposal that encloses an object of interest. Although higher eigenstates aid us in expanding the search for objects, they may also produce saliency maps that do not highlight objects of interest accurately. Thus, we use the inverse square root of the eigenvalue as an objectness cue, represented by ($\Theta_{Eigenvalue}$).

3.4 Compactness

The compactness measure ($\Theta_{Compactness}$) aims to penalize regions that have an abnormal shape. This stems from the premise that objects in real life generally have a compact shape. Compactness is defined as the ratio of segment perimeter to its area. Figure 4 illustrates the efficacy of this measure in identifying segments with irregular form-factor.

Figure 4: Compactness scores (left to right): 5.04, 45.44 and 30.1 respectively.

Our final objectness score is combined as in (3). To keep the scoring free from learnt parameters, we simply multiply the objectness cues together.

$$\Theta = \Theta_{ED} * \Theta_{Saliency} * \Theta_{Compactness} * \Theta_{Eigenvalue} \quad (3)$$

4 NON-MAXIMUM SUPPRESSION

After the proposed segments are scored using the objectness measure devised in Section 3, we apply a non-maximum suppression (NMS) to remove spatially redundant proposals. By scoring them first, we increase the chances of retaining higher quality segments while removing comparatively lower quality and redundant proposals for the same object. Given the ranked list of proposals r, our NMS operation proceeds as shown in (4).

$$\boldsymbol{r} \rightarrow \{\boldsymbol{r} - \boldsymbol{s}\} \mid \forall s \in \boldsymbol{s} : o(r_i, s) > \kappa \quad (4)$$

where κ is the NMS threshold which controls the amound of suppression achieved. Setting the threshold at its maximum value of 1 will only suppress the duplicate proposals. As observed in [16], the NMS threshold can be used to optimize the recall at a given Intersection-of-Unions (IoU) threshold with the ground truth.

5 RESULTS

In order to compare our results with the contemporary state-of-the-art methods, we used the toolbox provided by [5], [27] to compare bounding-box proposals on the PASCAL VOC 2007 test set. The segments proposed by region-based methods are transformed to rectangular windows by enclosing the tightest fitting bounding box around them. Although this does not put the region-based methods at a significant disadvantage, it does favor window based methods. However, given a scenario in which a region based method exhibits similar results at par with a window-based method, the region-based method would still be preferred due to its higher localization power.

Table 1: ABO/MABO Measures

	Airplane	Bicycle	Bird	Boat	Bottle	Bus	Car	Cat	Chair	Cow	Table	Dog	Horse	MBike	Person	Plant	Sheep	Sofa	Train	TV	MABO
Objectness [13] (1773)	0.665	0.664	0.615	0.592	0.539	0.707	0.618	0.727	0.602	0.628	0.720	0.707	0.689	0.663	0.621	0.610	0.598	0.736	0.710	0.640	0.653
Rahtu [32] (10000)	0.780	0.762	0.694	0.627	0.517	0.812	0.663	0.879	0.623	0.711	0.828	0.854	0.830	0.783	0.695	0.647	0.696	0.866	0.863	0.766	0.745
Bing [15] (1939)	0.662	0.652	0.643	0.632	0.631	0.659	0.648	0.706	0.631	0.648	0.685	0.677	0.655	0.656	0.652	0.640	0.645	0.693	0.686	0.646	0.657
EdgeBoxes [16] (991)	0.760	0.780	0.728	0.677	0.576	0.804	0.704	0.822	0.698	0.757	0.776	0.826	0.798	0.781	0.699	0.691	0.749	0.812	0.796	0.788	0.751
CPMC [18] (596)	0.743	0.707	0.677	0.577	0.501	0.800	0.685	0.892	0.665	0.734	0.796	0.866	0.783	0.751	0.660	0.641	0.699	0.881	0.819	0.786	0.733
MCG [25] (1939)	0.813	0.822	0.765	0.718	**0.688**	0.870	0.814	0.899	**0.796**	0.820	0.875	0.884	0.841	0.834	0.788	0.765	0.791	0.917	0.872	0.860	0.822
SS [23] (8753)	**0.865**	**0.869**	**0.829**	**0.765**	**0.727**	**0.886**	**0.830**	**0.928**	**0.847**	**0.854**	**0.904**	**0.920**	**0.862**	**0.872**	**0.814**	**0.815**	**0.836**	**0.935**	**0.894**	**0.901**	**0.858**
Rantalankila [14] (2788)	0.763	0.758	0.707	0.582	0.566	0.834	0.716	**0.928**	0.737	0.750	0.874	**0.901**	0.835	0.788	0.713	0.687	0.713	**0.928**	0.865	0.824	0.773
RP [19] (5745)	0.854	**0.834**	0.786	0.733	0.662	0.873	0.794	0.919	**0.800**	0.829	**0.887**	0.899	0.840	**0.844**	0.771	0.757	0.804	**0.932**	0.875	**0.868**	0.828
Rigor [22] (1536)	0.773	0.799	0.746	0.687	0.637	0.860	0.789	0.917	0.762	0.798	0.824	0.887	0.833	0.814	0.739	0.732	0.782	0.904	0.855	0.847	0.799
Geodesic [20] (5529)	0.759	0.830	0.747	0.668	0.638	**0.874**	0.787	0.920	0.784	0.796	0.881	0.897	0.853	0.842	0.767	0.763	0.772	0.915	0.876	0.845	0.811
Endres [21] (660)	0.690	0.786	0.662	0.563	0.536	0.820	0.731	0.880	0.685	0.725	0.835	0.861	0.799	0.809	0.680	0.657	0.680	0.891	0.853	0.753	0.745
QCUT_95 (7682)	**0.879**	**0.843**	**0.845**	**0.776**	0.657	**0.884**	**0.835**	**0.933**	0.796	**0.872**	**0.893**	**0.928**	**0.886**	**0.860**	**0.801**	**0.816**	**0.849**	0.922	**0.899**	**0.866**	**0.852**
QCUT_85 (5289)	**0.858**	0.830	**0.827**	**0.768**	**0.669**	0.862	**0.822**	0.904	0.793	**0.851**	0.870	0.901	**0.864**	0.842	**0.792**	**0.810**	**0.836**	0.902	**0.877**	0.849	**0.836**
QCUT_75 (3200)	0.826	0.796	0.797	0.742	0.649	0.828	0.793	0.863	0.762	0.821	0.834	0.860	0.826	0.808	0.761	0.778	0.803	0.860	0.833	0.816	0.803

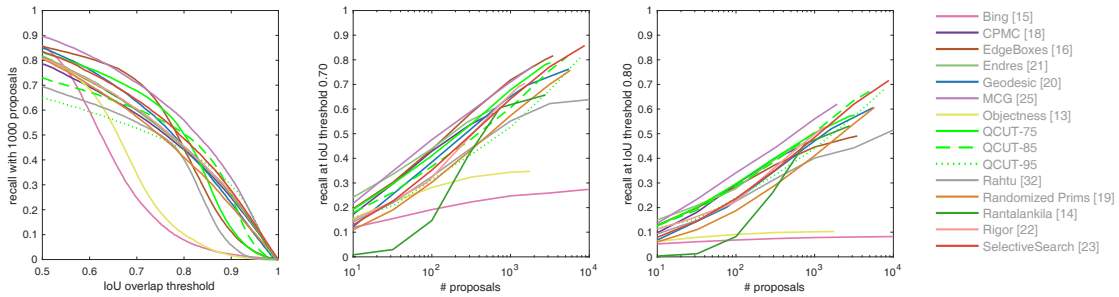

Figure 5: (From left to right) Recall vs IoU overlap with ground truth for top 1000 proposals, Recall at IOU 0.7 and 0.8.

Table 2: Effect of NMS threshold on number of proposals and average recall.

NMS Threshold (κ)	Number of Proposals	Average Recall
1	22920	0.6986
0.95	7682	0.6652
0.85	5289	0.6364
0.75	3200	0.5728

For the evaluation, we compare three variants of our method corresponding to three different NMS thresholds; 0.95, 0.85 and 0.75. Table 2 shows the threshold used, the corresponding number of proposals and the achieved recall, as calculated in [5], [27]. As can be seen, NMS can achieve almost a three-fold reduction in the number of proposals (at threshold 0.95) where the loss of recall is only 3%.

Figure 5 shows the recall vs number of proposals at three fixed IoU thresholds with the ground truth. At IoU threshold of 0.5, the proposed method with NMS threshold 0.75 achieves a recall of 0.8316 in the top 1000 proposals.

Moreover, the drop in recall for higher IoU thresholds is not as significant as in the case of other methods. At 0.75 IoU threshold, the recall (for the same proposed method) is 0.6117 which is almost equal to Edge Boxes [16] score of 0.6155 and only slightly behind that of MCG [25] score of 0.6394. At an even more challenging IoU threshold of 0.8, our proposed methods with NMS thresholds 0.75 and 0.85 achieve a recall of 0.5031 and 0.5053, respectively which is bested only by MCG [25].

Figure 5 (b-c) shows the recall vs number of proposals at two different IoU thresholds with the ground truth. Across all the thresholds, our results are competitive and at par with the state-of-the art methods. At 0.8 IoU, the performance of our unsupervised methods is better than some supervised ones for all number of proposals. Furthermore, our maximum recall is consistently in the top three methods as shown in Figure 5.

The ABO (Average Best Overlap), as defined in [32], measures a method's performance for a given class of objects in the dataset.

Table 1 shows the ABO for all the methods for the classes present in PASCAL VOC 2007 test set. The top three methods are colored red, green and blue respectively. We can see that our methods are among the top three for all the classes in the dataset. Two out of our three variants are among the top three methods with best MABO (Mean Average Best Overlap) [32]. Furthermore, our proposed method outperform those which learn objectness from the same classes such as [13], [15] among others, which is evidence of the success of the category independent nature of the proposed methods.

6 CONCLUSION

In this study, a novel category independent object proposal method is presented based on visual saliency estimation. Inspired by the human visual system, the search for objects was initialized from the salient objects and extended towards others via exploiting the principle of quantum superposition. The object proposals generated in this way are category independent. Evaluation results showed that the proposed unsupervised method achieve competitive results with the state-of-the-art methods that employ supervised approaches to object proposals.

7. REFERENCES

[1] M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman, "The pascal visual object classes (VOC) challenge," *IJCV*, vol. 88, no. 2, pp. 303–338, 2010.

[2] P. Viola and M. Jones, "Robust Real-Time Face Detection," *IJCV*, vol. 57, no. 2, pp. 137–154, 2004.

[3] N. Dalai, B. Triggs, I. Rhone-Alps, and F. Montbonnot, "Histograms of oriented gradients for human detection," in *CVPR*, 2005, vol. 1, p. 0.

[4] P. F. Felzenszwalb, R. B. Girshick, D. Mcallester, and D. Ramanan, "Object Detection with Discriminatively Trained Part Based Models," *IEEE Trans. PAMI*, vol. 32, no. 9, pp. 1–20, 2009.

[5] J. Hosang, R. Benenson, P. Dollár, and B. Schiele, "What Makes for Effective Detection Proposals?," *IEEE Trans. PAMI*, vol. 38, no. 4, pp. 814–830, Apr. 2016.

[6] X. Wang, M. Yang, S. Zhu, and Y. Lin, "Regionlets for Generic Object Detection," *IEEE Trans. PAMI*, vol. 37, no. 10, pp. 2071–2084, 2015.

[7] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proceedings of the IEEE CVPR*, 2014, pp. 580–587.

[8] C. Szegedy, S. Reed, D. Erhan, D. Anguelov, and S. Ioffe, "Scalable, High-Quality Object Detection," *ArXiv e-prints*, 2014. [Online]. Available: http://arxiv.org/abs/1412.1441.

[9] K. He, X. Zhang, S. Ren, and J. Sun, "Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition," *IEEE Trans. PAMI*, vol. 37, no. 9, pp. 1904–1916, 2015.

[10] R. G. Cinbis, J. Verbeek, and C. Schmid, "Segmentation Driven Object Detection with Fisher Vectors," in *IEEE International Conference on Computer Vision*, 2013, pp. 2968–2975.

[11] R. Girshick, "Fast R-CNN," in *2015 IEEE International Conference on Computer Vision (ICCV)*, 2015, pp. 1440–1448.

[12] J. D. J. Deng, W. D. W. Dong, R. Socher, L.-J. L. L.-J. Li, K. L. K. Li, and L. F.-F. L. Fei-Fei, "ImageNet: A large-scale hierarchical image database," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2009, pp. 2–9.

[13] B. Alexe, T. Deselaers, and V. Ferrari, "What is an object ?," in *ICCV*, 2010.

[14] P. Rantalankila, J. Kannala, and E. Rahtu, "Generating object segmentation proposals using global and local search," in *Cvpr*, 2014, pp. 2417–2424.

[15] M.-M. Cheng, Z. Zhang, W.-Y. Lin, and P. Torr, "BING: Binarized Normed Gradients for Objectness Estimation at 300fps," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2014, pp. 3286–3293.

[16] C. L. Zitnick and P. Doll, "Edge Boxes:Locating Object Proposals from Edges," in *European Conference on Computer Vision*, 2014, pp. 1–15.

[17] P. Dollár and C. L. Zitnick, "Fast Edge Detection Using Structured Forests," *IEEE Trans. PAMI*, vol. 37, no. 8, pp. 1558–1570, 2014.

[18] J. Carreira and C. Sminchisescu, "CPMC: Automatic object segmentation using constrained parametric min-cuts," *IEEE Trans. PAMI*, vol. 34, no. 7, pp. 1312–1328, 2012.

[19] S. Manen, M. Guillaumin, and L. Van Gool, "Prime Object Proposals with Randomized Prim's Algorithm," in *Proceedings of the 2013 IEEE ICCV*, 2013, pp. 2536–2543.

[20] P. Krähenbühl and V. Koltun, "Geodesic Object Proposals," in *Computer Vision -- ECCV 2014: September 6-12, 2014, Proceedings, Part V*, Cham, 2014, pp. 725–739.

[21] I. Endres and D. Hoiem, "Category-Independent Object Proposals with Diverse Ranking," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 2, pp. 222–234, Feb. 2014.

[22] A. Humayun, F. Li, and J. M. Rehg, "RIGOR: Reusing inference in graph cuts for generating object regions," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 336–343, 2014.

[23] K. E. A. Van De Sande, J. R. R. Uijlings, T. Gevers, and A. W. M. Smeulders, "Segmentation as Selective Search for Object Recognition," *Int. Conf. Comput. Vis.*, no. November, pp. 1879–1886, 2011.

[24] P. F. Felzenszwalb and D. P. Huttenlocher, "Efficient graph-based image segmentation," *IJCV*, vol. 59, no. 2, pp. 167–181, 2004.

[25] J. Pont-Tuset, P. Arbelaez, J. T. Barron, F. Marques, and J. Malik, "Multiscale Combinatorial Grouping for Image Segmentation and Object Proposal Generation," pp. 1–14, 2015.

[26] J. Shi and J. Malik, "Normalized Cuts and Image Segmentation," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 22, no. June, pp. 731–737, 1997.

[27] J. Hosang, R. Benenson, and B. Schiele, "How good are detection proposals, really?," *Bmvc*, pp. 1–25, 2014.

[28] C. Aytekin, E. C. Ozan, S. Kiranyaz, and M. Gabbouj, "Visual saliency by extended quantum cuts," *Proc. - Int. Conf. Image Process. ICIP*, vol. 2015–Decem, no. November 2016, pp. 1692–1696, 2015.

[29] Ç. Aytekin, S. Kiranyaz, and M. Gabbouj, "Learning to rank salient segments extracted by multispectral Quantum Cuts," *Pattern Recognit. Lett.*, vol. 72, pp. 91–99, 2015.

[30] Ç. Aytekin, E. C. Ozan, S. Kiranyaz, and M. Gabbouj, "Extended quantum cuts for unsupervised salient object extraction," *Multimed. Tools Appl.*, 2016.

[31] C. Aytekin, S. Kiranyaz, and M. Gabbouj, "Automatic object segmentation by quantum cuts," *Proc. - Int. Conf. Pattern Recognit.*, pp. 112–117, 2014.

[32] J. R. R. Uijlings, K. E. A. Van De Sande, T. Gevers, and A. W. M. Smeulders, "Selective search for object recognition," *Int. J. Comput. Vis.*, vol. 104, no. 2, pp. 154–171, 2013.