Reviews For Paper

Paper ID	2283
Title	Timely Object Detection

Masked Reviewer ID: Assigned_Reviewer_1
Review:

Question	
Please briefly describe the paper's contributions, and list its positive and negative points.	The paper considers the problem of anytime multi class object detection. The authors propose to study the Time-vs-AP curve (in particular the area under the curve) as performance metric. They then propose to model the sequential decision problem of choosing which object detector to apply as a POMDP. They present experiments on the PASCAL VOC 2007 data set, demonstrating improvement over random ordering. The problem considered is interesting, and the proposed performance metric makes sense. However, I have several serious concerns, both about the technical approach and about its presentation.
Overall Rating	Definitely Reject
Please explain your rating. If the paper is so unclear that it should be rejected, please explain that. If the paper is not novel please explain, citing the work that makes it so. You should take into account whether this paper is of wide interest, or addresses only a narrow community. You should mention whether the paper cites prior work fairly or not. You should comment	The authors motivate the use of POMDP planning to find an optimal policy. However, in the end, they apply a simple greedy algorithm (p5:

on the paper's correctness and experimental evaluation. A paper that is definitely correct has conclusions supported by flawless arguments (including correct proofs, formulas, and so on) or by well- designed and executed experiments, or perhaps by both. A convincing paper has strong arguments or limited but compelling experiments. Minor technical errors or experimental design and execution count against a paper, but it might have redeeming features (which you would explain). If a paper has major problems, it should be rejected. Your rating should not consider whether or not the paper might become an oral or poster.

L506), which greedily chooses the detector for the object class that maximizes the probability $P(C)$ of the object being present, and the average improvement in AP (ΔAP_{avg}) that this detector yielded in a training set. Through the many heuristic approximations, it is unclear how this simple greedy algorithm relates to the overall POMDP approach that the authors describe. Perhaps even more crucial is that both $P(C)$ and ΔAP_{avg} are themselves estimated in a heuristic manner. $P(C)$ is estimated inconsistently according to (P4: L351-L356), rather than, e.g., by computing marginals in some graphical model. ΔAP_{avg} does not take into account interaction among the performance of different object detectors (e.g., that multiple detectors may improve the chances of detection, but with diminishing returns). The attempt in equation (3) to account for this effect is not well motivated and evaluated. For more complex detector actions (e.g., not just choosing the detector, but also the region where it should be applied), I would expect this issue to become even more prominent. Figure 4 actually demonstrates the potential concerns about these shortcomings. For example, the fact that the "Fixed-Ordering" approach (which seems to disregard $P(C)$) outperforms the "Back-off" policy indicates potential issues with the estimation of the conditional object presence probabilities $P(C)$.

Additional comments to author(s)	The presentation of the paper is problematic. While the introduction is quite well written, the technical description is vague and leaves many details open, likely making it difficult to reproduce the results. There's also a fairly large number of typos (see below for some examples). The results (Figures 4 & 5, Table 2) are also not well described. Typos: 4: yields the highest improve 4: This is best modeled with a CRF for which there are well known algorithm for exact inference.

4: This model will not satisfy our requirements too good
6: promising relaxes approaches

Masked Reviewer ID: Assigned_Reviewer_2
Review:

Question	
Please briefly describe the paper's contributions, and list its positive and negative points.	This paper investigates a new problem of object detection given a certain time constrain. The assumption in this paper is that performing multi-object detections on the image is expensive. This paper aims to achieve higher detection average precision given any time constrains by actively choosing which detector to use . Therefore, this paper proposes a reward formulation for the object detector, and actively selects the next detector to perform at each step given the current observation. The major contribution of this paper would be the novel research problem of selecting the best object detector to run given time constrains. The negative points would be the algorithm and the experiment parts of the paper. The terms used in the algorithm are not well motivated (for example the prob. term). The experiments are not well organized and explained. These diminish the contribution of the paper. Besides, many parts in the paper are confusing and unexplained, which makes the paper hard to understand and fully judge the improvement given by the proposed algorithm.
Overall Rating	Weakly Reject
Please explain your rating. If the paper is so unclear that it should be rejected, please explain that. If the paper is not novel please explain, citing the work that makes it so. You should take into account whether this paper is of wide interest, or addresses only a narrow community. You should mention whether the paper cites prior work fairly or not. You should comment on the paper's correctness and experimental	For the algorithm part, basically this paper uses the average precision (AP) and time (delta t) to model the rewards for each object detector, and based on this to perform the selection. This algorithm part is intuitive and simple. However, there are problems with the notations. In the section 4.1, first this paper proposes to use a Bayes model to calculate the object observation at each stage, then tries to smooth the equation in line 352. However this equation is not a proper-defined probability. This equation adds different probabilities instead of multiplying, and it is not correct if assuming independence as naïve Bayes model. This interpretation of the conditional needs to be explained more. Equation in line 365 seems incorrect. What are P_1 to P_k? Are they C_1 to C_k? For the paragraph 407 to 428, the paper tries to argue about using SVM instead of just thresholding. However line 421 reports the results in Table 2. Is it Table 1? However in table 1, the author did not show the result of thresdholding, but two different SVMs.

evaluation. A paper that is definitely correct has conclusions supported by flawless arguments (including correct proofs, formulas, and so on) or by well- designed and executed experiments, or perhaps by both. A convincing paper has strong arguments or limited but compelling experiments. Minor technical errors or experimental design and execution count against a paper, but it might have redeeming features (which you would explain). If a paper has major problems, it should be rejected. Your rating should not consider whether or not the paper might become an oral or poster.	CSC starts to appear from line 426, but is not defined throughout the paper. Line 427 says "detector-fast is CSC with parameters settings to speed it up to about half the time". However these settings are not revealed. For equation (1) in line 474, AP_1 is defined in line 476, "AP_1 is the difference between 1 and multi-class …" 1 is not defined here and confusing. The experiments are also not very well organized. We are unable to identify what are the performances of the proposed algorithm from the baselines. There is little analysis and discussion about the experiment. For evaluation, the explanation of the "oracle" between line 549 an d 553 is confusing. It seems that the oracle did not use any extra information but achieves the best result. More explanation is necessary for the simulation of the less robust classifier on line 563 - what is this and what is it used for. Equation (3) is not a correct formulation for expectation. This may result in a negative value for expectation. In table 2, the method "advanced", "(double)", "(pair)" are not defined and explained, as well as the numbers. The table is not refered to from the text. Fig.4 and Fig.5 show their final experiment, however there is very little explanations. Also the numbers in the legends are not defined and explained. This makes us fail to evaluate the improvement of their proposed algorithm over the baseline. If fixed does better than backoff and makes more sense since it uses the real conditional probability, then why use the backoff probability.
Additional comments to author(s)	Fig. 2 move caption underneath. Table 2, not explained and not referred to in text. Line 408, "tresholding" to "thresholding" line 567, "described below in Section 1.1", and "evaluated in section 1.1" in line 431, 510. Do you mean "evaluation metrics are defined in Section 1.1"?

Masked Reviewer ID: Assigned_Reviewer_3
Review:

Question	
	The paper proposes a technique for deciding the order in which a pool of object detectors for different classes should be run on an image. The goal is to maximize the total AP performance if the system would be stopped at

	any point in time (as opposed to always running all detectors and checking AP only at the end).
Please briefly describe the paper's contributions, and list its positive and negative points.	Positive points: + 'anytime performance' is a novel goal in object detection. The idea that the co-occurrence statistics between different classes can drive the selection of what detector to run next is interesting. Negative points: - the proposed technique only establishes an order between detectors, without modifying the order inside a detector (e.g. the order in which windows are evaluated). The atomic unit of processing is running a detector for a class fully on the whole image. This is only marginally interesting, as it means that the speedup obtained by stopping the system early comes at the cost of not even trying to detect certain classes at all. I am not sure this has any plausible application scenario (as opposed to choosing which image regions to evaluate, as the introduction suggests). I personally would not consider getting a speedup in return for discarding entire classes. Unfortunately there are a number of other negative points in this paper, which I detail in the next review box.
Overall Rating	Weakly Reject
Please explain your rating. If the paper is so unclear that it should be rejected, please explain that. If the paper is not novel please explain, citing the work that makes it so. You should take into account whether this paper is of wide interest, or addresses only a narrow community. You should mention whether the paper cites prior work fairly or not. You should comment on the paper's correctness and experimental evaluation. A paper that is definitely correct has conclusions	Clarity and quality of writing ----------------------------- This is the weakest point of the paper. The writing is dispersive, often too informal, repetitive and unclear. Instead of presenting a clear, straight line of thought, it circles around or pages, repeatedly hit-and-go around the ideas it wants to present. It took me quite some effort to understand, even if not completely, what the system is doing. In addition to the above style issue, the paper also contains numerous English writing mistakes, along with twisted and overly long sentences. A few important concepts are mentioned long before they are defined, e.g. 'value function'. Overall, the quality of writing and clarity are insufficient for CVPR. Experimental evaluation ---------------------- The experiment in sec. 4.1 (table 1) makes little sense. It shows the performance of the proposed object-presence classifier, but it does not compare it to the baseline the authors talk about (i.e. only counting highly scored detections). The 'detector-fast' results are reported without any comment on how having a 2x speedup in evaluating this classifier would be of any practical advantage. It feels like the major cost is running the detector over the image. However, no timings are given.

supported by flawless arguments (including correct proofs, formulas, and so on) or by well- designed and executed experiments, or perhaps by both. A convincing paper has strong arguments or limited but compelling experiments. Minor technical errors or experimental design and execution count against a paper, but it might have redeeming features (which you would explain). If a paper has major problems, it should be rejected. Your rating should not consider whether or not the paper might become an oral or poster.	The main evaluation in section 5 lacks some important explanations of the experimental protocol: - did you test on the full PASCAL VOC 2007 test set? Or on val? Or on a subset? - most importantly: the AP measure in fig. 4 is not well defined. Is it the average AP over all classes? If so, AP should not decrease for the oracle as more detectors are run, as the AP of classes for which the detector did not run yet should have AP = 0. This might suggest the AP in fig. 4 is some other thing, e.g. the mean AP over only the classes for which detectors were run. But in this case it is not clear why the other curves would grow? The experiment in section 6 is a contrived simulation that does not add anything significant to the paper. A more interesting experiment would have been to add a real secondary detector, e.g. a bag-of-words on dense SIFT features. Overall both sec. 5 and 6 lack a proper discussion of the results. In fact the only statement positioning the results of the proposed method over the random baseline and the oracle is given in the figure caption! From that caption it seems that, by following the proposed policy, one can get about a 2x speedup at the price of not detecting _any_ instance of several classes (as their detectors are not run at all). This does not seem an attractive trade-off to me. Overall evaluation ------------------ Overall this paper discusses a potentially interesting general direction, but proposes a scheme of limited interest. Moreover, the paper is poorly written and the results unclear.
Additional comments to author(s)	none