ICCV 2017
International Conference on Computer Vision 2017
Oct 22, 2017 - Oct 29, 2017, Venice, Italy

Reviews For Paper

Paper ID 1992

Title Dynamic Computational Time for Visual Attention

Masked Reviewer ID: Assigned_Reviewer_1

Review:

Question	
Paper Summary. Please summarize in your own words what the paper is about.	This paper aims to accelerate the processing of recurrent visual attention (RAM) by introducing a dynamic stopping mechanism. The average computational time can be saved while keeping the original recognition accuracy.
Paper Strengths. Positive aspects of the paper. Be sure to comment on the paper's novelty, technical correctness, clarity and experimental evaluation. Notice that different papers may need different levels of evaluation: e.g., a theoretical paper vs. an application paper	1. The major strength of this paper is to accelerate computational time while keeping the same accuracy. 2. Most parts of this paper is easy to understand. 3. The experiments setting is clear and sufficient.
Paper Weaknesses. Discuss the negative aspects: lack of novelty or clarity, technical errors, insufficient experimental evaluation, etc. Please justify your comments in great detail. If you think the paper is not novel, explain why and provide evidences	1. The contribution of this paper is limited. The major strength of the proposed method is its computational acceleration. But does speed really constitute an important bottleneck for previous Recurrent Attention Model (RAM)? The experiments only compared the number of steps of RAM and DT-RAM. So, what are their actual inference speed? If the speed is not a bottleneck for RAM, then why we need a dynamic extension? The authors need to explain their contribution in more details. 2. Unfair comparison with previous published methods In Table 3 and Table 4. Different CNN models are adopted in different methods. RAM and DT-RAM have the highest score, since they use ResNet-50. However, there lacks necessary explanations, which may mislead reader's judgement.
Preliminary Rating	Borderline
Preliminary Evaluation. Please indicate to the AC, your fellow reviewers, and the authors your current opinion on the paper. Please	The proposed method mainly to accelerate the inference speed of RAM. It would be better if the authors could explain some issues.

summarize the key things you would like the authors to include in their rebuttals to facilitate your decision making.	
Confidence. Write "Very Confident" to stress you are absolutely sure about your conclusions (e.g., you are an expert working in the area), "Confident" to stress you are mostly sure about your conclusions (e.g., you are not an expert but are knowledgeable). "Not Confident" in all the other cases.	Not Confident

Masked Reviewer ID: Assigned_Reviewer_2

Review:

Question	
Paper Summary. Please summarize in your own words what the paper is about.	The authors augment a recurrent visual attention network with a "decide whether to continue" decision, made at each timestep. They train the system using REINFORCE, and find that curriculum learning and intermediate supervision improve performance. Their system attains competitive fine-grained classification performance, and is more efficient than the baseline visual attention network.
Paper Strengths. Positive aspects of the paper. Be sure to comment on the paper's novelty, technical correctness, clarity and experimental evaluation. Notice that different papers may need different levels of evaluation: e.g., a theoretical paper vs. an application paper	The proposed system decides when to stop in a principled way, saving computation. The authors perform systematic experiments to justify every component of their training pipeline. The technique is applicable to a wide range of vision applications. The paper is clear and well-written.
Paper Weaknesses. Discuss the negative aspects: lack of novelty or clarity, technical errors, insufficient	The intermediate supervision loss feels like it's not an honest reflection of the problem. The empirical improvement over a fixed-policy approach is relatively small.

experimental evaluation, etc. Please justify your comments in great detail. If you think the paper is not novel, explain why and provide evidences	
Preliminary Rating	Strong Accept
Preliminary Evaluation. Please indicate to the AC, your fellow reviewers, and the authors your current opinion on the paper. Please summarize the key things you would like the authors to include in their rebuttals to facilitate your decision making.	For some applications, the relatively small accuracy gain (compared to fixed-policy stopping) may not be worth replicating the authors' relatively complex training pipeline. However, for some it may be. Additionally, the systematic analysis of reinforcement learning in this context will be useful for the community as a whole.
Confidence. Write "Very Confident" to stress you are absolutely sure about your conclusions (e.g., you are an expert working in the area), "Confident" to stress you are mostly sure about your conclusions (e.g., you are not an expert but are knowledgeable). "Not Confident" in all the other cases.	Confident

Masked Reviewer ID: Assigned_Reviewer_3

Review:

Question	
Paper Summary. Please summarize in your own words what the paper is about.	The paper proposes a framework for utilizing spatial attention for improving accuracy and efficiency of visual recognition. It builds off of the Recurrent Visual Attention Model (RAM) [6]. The difference from RAM is (1) the addition of a binary output variable that determines whether to stop or continue looking (Sec 3.3) and introduces intermediate supervision for the classification network at every step (Sec 3.4). Experimental results on MNIST and the fine-grained CUB-200 and Stanford Cars datasets demonstrate some improvement in both accuracy and efficiency over RAM.
Paper Strengths. Positive aspects of the paper. Be sure	Overall the paper is well-written and well-executed. The proposed method adds reasonable improvements to RAM to enable dynamic test-time computation. The writing is clear, and the overview of related literature is thorough (to the best of

to comment on the paper's novelty, technical correctness, clarity and experimental evaluation. Notice that different papers may need different levels of evaluation: e.g., a theoretical paper vs. an application paper	my knowledge). The importance of intermediate supervision is evidenced by the 7% improvement in Table 9.
Paper Weaknesses. Discuss the negative aspects: lack of novelty or clarity, technical errors, insufficient experimental evaluation, etc. Please justify your comments in great detail. If you think the paper is not novel, explain why and provide evidences	The key concern is that the work is incremental. The additional innovations over RAM are only (1) an extra binary variable for deciding whether to stop taking further attention and (2) adding intermediate supervision after every action. The binary stopping variable is very similar to e.g., the binary variable used in [30] to determine whether to output a prediction after every action. Intermediate supervision (and/or reward shaping) is well-known to be critical in reinforcement learning. Similarly, the results demonstrate marginal improvement: (1) with the same number of computational steps, the accuracy improvement over RAM is 0.1-0.2% on MNIST (Table 2), 0.3% on CUB-200 (Table 6), (2) the reduction in computational cost is < 50%: 1.9 steps vs 3 steps on Stanford Cars (Table 4), 3.6 vs 6 steps on CUB-200 (Table 6).
Preliminary Rating	Weak Reject
Preliminary Evaluation. Please indicate to the AC, your fellow reviewers, and the authors your current opinion on the paper. Please summarize the key things you would like the authors to include in their rebuttals to facilitate your decision making.	The work is well-executed; however, the paper is too incremental to warrant publication in ICCV in my opinion. There are no novel insights and only marginal improvements in accuracy and efficiency.
Confidence. Write "Very Confident" to stress you are absolutely sure about your conclusions (e.g., you are an expert working in the area), "Confident" to stress you are mostly sure about your conclusions (e.g., you are not an expert but are knowledgeable).	Confident

"Not Confident" in all the other cases.	