* [Submissions](https://cmt3.research.microsoft.com/ACCV2018/Submission/Index)
* **Select Your Role :**
* [**Author**](https://cmt3.research.microsoft.com/ACCV2018/Submission/ReviewsSnapshot/1183)
* [ACCV2018](https://cmt3.research.microsoft.com/ACCV2018/Submission/ReviewsSnapshot/1183)
* [Oliver Schulte](https://cmt3.research.microsoft.com/ACCV2018/Submission/ReviewsSnapshot/1183)

 Print

**View Reviews**

Paper ID

1183

Paper Title

Dynamic Gated Graph Neural Networks for Scene Graph Generation

**Reviewer #2**

**Questions**

* **1. [Summary] Please provide a short summary of the paper and its contributions.** 
  + This work presented a dynamic version of Gated Graph Neural Networks to generate a graph of relations between objects appeared on the target image.
* **2. [Paper Strengths] Please discuss the 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: an algorithm paper may need fewer experiments, while an application paper may require thorough comparisons to existing methods. Also, please be sure to justify your comments in detail. For example, if you think the work is novel, not only say so, but also explain in detail why you think this is the case.** 
  + The main novelty of this work is the extension of GGNN by adding two Q-Learning networks to search for optimal link weights and connections.
* **3. [Paper Weaknesses] Please discuss the negative aspects of the paper: lack of novelty or clarity, technical errors, insufficient experimental evaluation, etc. Justify your comments in detail; don't provide just generic critique. It is not reasonable to ask for comparisons with unpublished, non-peer-reviewed papers (eg. arXiv), or papers published after the ACCV 2018 deadline.** 
  + \* The work claimed that DGGNN outperformed the study in citation [18], however that was not the case in Table 1. Only [11] and [12] were included. Is that possible to directly use GCNN for the task?

Our bad! We did compare with reference [18] Xu et al, but mislabelled the row in Table 1 as [12]. So the results in the second row of Table 1 are for [18]. The results of the method in Reference [12] is not state-of-the-art (worse than reference [18]).   
  
\* For the result in Table 2, it appeared better than VRL in [19]. Was there a significant test involved?

We didn’t carry out a formal test, but with a sample size of over 30,000 test cases, a 2% improvement must be significant.

\* The training time of DGGNN seems quite long. How does that compare to other methods? Was that a case of trading computational cost for a slightly better performance?

The training time given is for the challenging dataset VG1.4-b with 1,750 object categories. For the less challenging dataset VG1.4-a, training takes less than 2 days. The only reference method that has been applied to VG1.4-b dataset is [19], which takes a comparable amount of training time to ours (10 days for [19] vs. 2 weeks for our method).

\* … at the end of the abstract  
  
\* Q-value needs to be explained when first introduced  
  
\* Section 4.1 is a bit odd, looks like more of experiments rather than methodology. In addition, how much does DGGNN rely on the pre-training using COCO and ImageNet?

The pre-training is only for the object detector. The baseline methods [18] and [19] use the same pre-training for object detection.   
  
\* why the coefficient in the formula in between Formulae (3) and (4) was set to 0.6?

Set experimentally by trying a few different values. This coefficient is not very important.  
  
\* Formula (4) is not explained.

We will add an explanation. It’s the standard formula for the output of a 2-layer neural net with hyperbolic tangent.

* **4. [Recommendation Justification / Detailed Comments ] Please explain to the AC, your fellow reviewers, and the authors your current recommendation for the paper. This explanation may include how you weigh the importance of the various strengths and weaknesses you described above.**
  + This paper needs quite a bit further development to reach a publishable standard. See the detailed comments above. Therefore I recommend a weak rejection.
* **5. [Overall Rating]**
  + Weak Accept

**Reviewer #3**

**Questions**

* **1. [Summary] Please provide a short summary of the paper and its contributions.** 
  + The authors present a dynamic version of the gated graph neural network for extracting scene graph given object bounding boxes. The DGGNN uses reinforcement learning framework to determine the labels for edges and nodes with the reward being ground truth, even in the absence of information. They evaluate their proposed method on the visual genome dataset.
* **2. [Paper Strengths] Please discuss the 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: an algorithm paper may need fewer experiments, while an application paper may require thorough comparisons to existing methods. Also, please be sure to justify your comments in detail. For example, if you think the work is novel, not only say so, but also explain in detail why you think this is the case.** 
  + Method is well explained. Its adding a dynamic element to an existing graph neural network which makes it interesting.
* **3. [Paper Weaknesses] Please discuss the negative aspects of the paper: lack of novelty or clarity, technical errors, insufficient experimental evaluation, etc. Justify your comments in detail; don't provide just generic critique. It is not reasonable to ask for comparisons with unpublished, non-peer-reviewed papers (eg. arXiv), or papers published after the ACCV 2018 deadline.** 
  + The text seems incomplete, for example the abstract. The evaluation seems preliminary and lacking.
* **4. [Recommendation Justification / Detailed Comments ] Please explain to the AC, your fellow reviewers, and the authors your current recommendation for the paper. This explanation may include how you weigh the importance of the various strengths and weaknesses you described above.**
  + While this paper presents a RL framework on a GNN and applies it to an interesting problem, the experiments are quite minimal and preliminary.
* **5. [Overall Rating]**
  + Borderline

For the relatively new task of scene graph generation, VG1 is the only available large dataset; it is the one used in the baseline method papers [19] and [18].

**Reviewer #4**

**Questions**

* **1. [Summary] Please provide a short summary of the paper and its contributions.** 
  + The paper tackles several scene graph prediction tasks (i.e. prediction of objects' location and their pairwise relationships). The method infers the graph structure simultaneously with node and edge embeddings. The method does this by iteratively updating the graph structure and computing embeddings of graph nodes and edges using at each iteration. The graph growing procedure is modeled as a deep reinforcement learning problem. The main contribution of the paper is using a graph neural network for computing the node and edge embeddings.
* **2. [Paper Strengths] Please discuss the 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: an algorithm paper may need fewer experiments, while an application paper may require thorough comparisons to existing methods. Also, please be sure to justify your comments in detail. For example, if you think the work is novel, not only say so, but also explain in detail why you think this is the case.** 
  + The paper presents the background (sec. 3) and the proposed idea (sec. 4) quite clearly, the full algorithm description in Page 8 is very helpful. The paper extends [19] by replacing the graph node and edge embeddings to be computed by a Graph Neural Network, which makes sense and reduces dependency of embeddings on action order.
* **3. [Paper Weaknesses] Please discuss the negative aspects of the paper: lack of novelty or clarity, technical errors, insufficient experimental evaluation, etc. Justify your comments in detail; don't provide just generic critique. It is not reasonable to ask for comparisons with unpublished, non-peer-reviewed papers (eg. arXiv), or papers published after the ACCV 2018 deadline.** 
  + Concerning experimental evaluation, the paper lacks an analysis of errors. Given that accuracies of the hardest SG-GEN are low, it would be very useful to have a systematic study of error sources (objects mislocalized? relationship types wrongly guessed? inexistent relationships hallucinated?). The paper would also benefit from a supplementary material pages with more examples of generated graphs for positive and negative cases.  
      
    Typos:  
    023 - ... at the end of abstract  
    176 - As s result

Are we allowed supplementary materials? That would also address reviewer #3

* **4. [ Rebuttal Requests] Please note specific points that you would like the authors to address in their rebuttal.**
  + 1. Some systematic error analysis in SG-GEN task would be very good to see. It is not clear how the alleged improvement due to better graph embedding would show itself in a detailed error analysis.  
      
    2. Discussion of training difficulty of the reinforcement learning for graph growing. Training is reported to take 2 weeks. How fast does the training converge to a SG-GEN result comparable to the state-of-the-art [19]? Did the authors attempt to speed up the training? What may be the steps to improve convergence? ICLR Workshop paper "Learning Deep Generative Models of Graphs" discusses such difficulties (sec. 5), it would be interesting to read authors' reflection on the subject. Is there any particular motivation for the breadth-first order used for graph growing?

Hm, good questions

* **5. [Recommendation Justification / Detailed Comments ] Please explain to the AC, your fellow reviewers, and the authors your current recommendation for the paper. This explanation may include how you weigh the importance of the various strengths and weaknesses you described above.**
  + I think the proposed improvement of [19] makes sense and is confirmed by the experiments. However, I would like to see more discussion of the Q-learning training procedure and a more thourough error analysis.
* **6. [Overall Rating]**
  + Weak Accept

[Go Back](https://cmt3.research.microsoft.com/ACCV2018/Submission/Index)

© 2018 Microsoft Corporation [About CMT](https://cmt3.research.microsoft.com/Content/CMT.html" \t "_blank) | [Terms of Use](https://www.microsoft.com/en-us/servicesagreement/" \t "_blank) | [Privacy & Cookies](https://go.microsoft.com/fwlink/?LinkId=521839" \t "_blank) | [Request CMT Site](https://cmt3.research.microsoft.com/CMTSRM/" \t "_blank) [CMT Support](mailto:support@msr-cmt.org)