View Reviews

Paper ID
126

Paper Title
Few-shot Semantic Segmentation with Support-induced Graph Convolutional Network

Reviewer #1

Questions

1. [Paper Summary] What are the key ideas, what is the significance and how are the ideas validated? Please be concise (3-5 sentences).
This paper addresses the problem of appearance difference between support and query images in few-shot semantic segmentation. The authors propose a GCN variant that uses a novel support-induced graph reasoning module to update the graph with the guidance of support foreground prototypes. It improves the query context features by including distinct parts between query and support objects. In addition, they introduce an instance association module to simultaneously capture the context of support and query instances.

2. [Paper Strengths] Please summarise the strengths of the paper. (Eg: novelty, insight, theoretical strength, state of the art performance, thorough evaluation). Please provide a clear explanation of why they are valuable.
This paper proposes a new method to update the graph using information from the support in the few-shot segmentation.

The authors show the effectiveness of each of the proposed methods through ablation and achieve state-of-the-art performance in most experimental cases.

3. [Paper Weaknesses] Please summarize the weaknesses of the paper. (E.g., Lack of novelty, technical errors, insufficient evaluation, etc). You should clearly justify your criticisms with precise and factual comments (E.g., with an explanation of technical errors, citation to prior work if novelty is an issue). Please note: It is not appropriate to ask for comparison with unpublished arXiv papers, and papers published after the BMVC deadline. Please be polite and constructive.
There is no comparison with other methods in the qualitative results, so it is not known how effective the proposed network is.

For ablation study, analysis such as meaning or cause is needed, not just a list of results.

4. [Questions for Authors] Questions that reviewers expect to be addressed in rebuttal? Please avoid asking for more experiments here (if an experiment is needed, it should be in Weakness). Please note: It is not appropriate to ask for comparison with unpublished arXiv papers, and papers published after the BMVC deadline.
What do the superscripts r and p mean in the notation of the activation map?

The paper says "These updated query features, which can be seen as different
object instances" and " the updated query graphs are further transformed into query instance features". How do you get an instance feature (v) from the graph? I wonder how updated query features can be considered as different instances.

In most cases, mIoU is higher compared to HSNet, while FB-IoU is lower. What is the cause of the difference between these two results?

I think the larger the k, the more reliable the prototype of support can be obtained. However, the result of k=7 is similar to that of k=1,3. Is there any analysis of this?

5. [Overall Rating]
Borderline Accept

6. [Justification of Rating] Please explain how the different strengths and weakness were weighed together to produce the overall rating. Please provide details of questions or ambiguities to be addressed in the rebuttal that might change your rating.
This paper proposes novel ideas and achieves state-of-the-art performance in most experimental cases. Although there are some grammatical errors and explanations of some symbols are insufficient, there is no big problem with the flow of the contents of the paper. However, there is a slight lack of explanation for the methodology.

Reviewer #2

Questions

1. [Paper Summary] What are the key ideas, what is the significance and how are the ideas validated? Please be concise (3-5 sentences).
This work presents a support-induced GCN model (SiGR) for few-shot semantic segmentation. By "support-induced", it refers to the use of support prototypes as convolution kernels in the constructed query graph for feature propagation.
They also propose an instance association (IA) module to learn instance-level contextual information, which essentially uses a cross-attention technique.
According to the experiments in Table 3, SiGR brings a clear performance gain.

2. [Paper Strengths] Please summarise the strengths of the paper. (Eg: novelty, insight, theoretical strength, state of the art performance, thorough evaluation). Please provide a clear explanation of why they are valuable.
The paper is mostly well written and the method well explained, though with some minor details missing.
The experiments and ablation studies are comprehensive.

3. [Paper Weaknesses] Please summarize the weaknesses of the paper. (E.g., Lack of novelty, technical errors, insufficient evaluation, etc). You should clearly justify your criticisms with precise and factual comments (E.g., with an explanation of technical errors, citation to prior work if novelty is an issue). Please note: It is not appropriate to ask for comparison with unpublished arXiv papers, and papers published after the BMVC deadline. Please be polite and constructive.
The motivation of using support prototypes as convolution kernels (support induced) is not clear. More analysis and insights are preferred.

4. [Questions for Authors] Questions that reviewers expect to be addressed in rebuttal? Please avoid asking for more experiments here (if an experiment is needed, it should be in Weakness). Please note: It is not appropriate to ask for comparison with unpublished arXiv papers, and papers published after the BMVC deadline.
1. In Eq.7, the definitions of A_m^p and A_h^p are not given any where in the paper.

2. In Table 1, 2, the bold-highlighted results are not the best (the best FB-IoU are achieved by HSNet). As long as the proposed idea is interesting and has some contributions to the comminity, the ratings of the paper will not be affected even if the proposed method can not beat state-of-the-art methods.

3. The performance of using different support instance size s shown in Fig. 4(a) looks a bit strange, i.e., setting s=1 yields almost same results as setting s=10. Any explanations on why is that?

5. [Overall Rating]
Borderline Accept

6. [Justification of Rating] Please explain how the different strengths and weakness were weighed together to produce the overall rating. Please provide details of questions or ambiguities to be addressed in the rebuttal that might change your rating.
See above

Reviewer #6

Questions

1. [Paper Summary] What are the key ideas, what is the significance and how are the ideas validated? Please be concise (3-5 sentences).
The paper tackles the problem of few-shot semantic segmentation, where only a few annotated samples are used. The authors proposed a novel support-induced graph convolutional network, which consists of a support-induced graph reasoning module and an instance association module. The proposed method has been evaluated on two large public datasets.

2. [Paper Strengths] Please summarise the strengths of the paper. (Eg: novelty, insight, theoretical strength, state of the art performance, thorough evaluation). Please provide a clear explanation of why they are valuable.
1. The proposed network is novel.
In the paper, the authors proposed a support-induced graph reasoning. The construction of the graph is very interesting, by cleverly utilizing a salience matrix. In this way, the graph can be constructed dynamically on the fl y and does not suffer from "fixed structure".
2. The motivation of the method is reasonable.
I agree with the author's statement "The appearance variations between objects from the same category could be extremely large, leading to unreliable feature matching and query mask prediction."
3. The performance is good.
The proposed method achieves state-of-the-art performance in the setting of 1-shot on both PASCAL and COCO datasets.
4. Ablation studies are well performed.
Several ablation studies have been conducted to validate the design of the network and the choice of hyper-parameters.

3. [Paper Weaknesses] Please summarize the weaknesses of the paper. (E.g., Lack of novelty, technical errors, insufficient evaluation, etc). You should clearly justify your criticisms with precise and factual comments (E.g., with an explanation of technical errors, citation to prior work if novelty is an issue). Please note: It is not appropriate to ask for comparison with unpublished arXiv papers, and papers published after the BMVC deadline. Please be polite and constructive.
1. Fixed 1D kernel used for the proposed GCN.
In line 230, it is said that fixed kernel is used for the convolution. Is there any justification on this? How about using a trainable kernel (maybe initialized by the above mentioned fixed kernel)? Would that yield worse performance?
2. Boldness of numerical results in Table 1.
Usually, it is the best performance that is bolded in the table. For "5-shot" results in Table 1, the "HSNet (ICCV'21)" achieves the best performance. Instead of bolding the results of this submitted manuscript, the results of the HSNet should be bolded. Otherwise, people would be potentially misled by the table, if they don't read carefully.
3. Impact of different number of prototypes.
In Table 5, it is shown that k=5 gives the best result, while k=7 decreases the performance to that of k=1. Is there any hypothesis that could possibly explain why using more prototypes would surprisingly incur worse performance?
4. Information gathering process in section 3.4 is very simple yet interesting. Just wondering is there any reason why the authors did not choose a traditional attention mechanism?
5. Related work.

The authors proposed to use 1D convolution for the implementation of the GCN. There is a work [1] where the authors proposed to use 2D convolutions as the implementation of the GCN. Maybe this work deserves a mention in the related work section.
[1] "SIA-GCN: A Spatial Information Aware Graph Neural Network with 2D Convolutions for Hand Pose Estimation." (BMVC 2020)

5. [Overall Rating]
Accept

6. [Justification of Rating] Please explain how the different strengths and weakness were weighed together to produce the overall rating. Please provide details of questions or ambiguities to be addressed in the rebuttal that might change your rating.
Overall the paper is in a good shape. I rate to accept, after considering of the pros and cons of the paper.