Peer review

Name: Sem Kluiver

Paper to review: "Impact of Cloud Removal on Urban Classification Tasks"

Please fill in your findings on the paper to review according to the following criteria. We grade the peer review according to:

- Depth: by showing that you have read the draft pointing to both strong and weak aspects.
- Clarity: by mentioning the weaknesses along with actionable suggestions for improvement.

Use the remainder of this document as an answer sheet.

Readability: Does the text make sense? Are all parts (abstract, intro, related work, methods, results, conclusion) included? Are all sentences complete? Are there typos? Etc.

Make sure your grammar/spelling is correct and consistent. (F.e., cloud remove/removal?)

Some sentences are quite long and may benefit from being split into two sentences.
Check some word order in some sentences as well.

Some misuse of semicolons: use a regular colon before an enumeration or list.

Make sure to not have empty sections. For example, Section 1 (Instroduction) has no text before Section 1.1 (Motivation). You should put some general introductory text before.

Make sure to clearly explain what Sentinel is about. At first, I thought you made a typo when you wrote Sentinel-1. It was when I looked online that I understood that Sentinel-2 is not the mission, but is one of the two satellites, the other being Sentinel-1.

I think you should not use reference words in a section when talking about something in a previous section (section 2.2 refers to something in section 2.1 with "these").

Abstract: Is the abstract short? Is it intriguing? Does it summarize all important elements of the general problem, solution, and the results achieved?

Abstract is short, but maybe a bit too short. I would add that you compare "the resulting images" to the cloudless images to determine whether images where clouds have been removed can still be predicted correctly.

There is, as of yet, no tease of the results you have found. I suppose you will add this later on.

Introduction: Is it clear that: (i) We have a relevant research question, (ii) This is not an easy research question and there are theoretical/methodological challenges to finding an answer to this question, (iii) Your contribution is addressing (and overcoming) those challenges, (iv) Positioning of the work next to earlier related work.

i) Research question seems relevant to me.
ii) It is also not easily answered and requires research, training models and testing.
iii) Work seems to be relevant in addressing the issue.
iv) Earlier work is mentioned and is relevant.

I would suggest splitting your relevant work section. You can keep the missions and previous papers/methods there, but I think you should put the datasets you will be using in a separate (sub)section about datasets. At least 'sort' the methods and the datasets: currently, you interweave used datasets and previous methods.

"Previous research" in the Eurosat subsection is a bit ambiguous. What model was used?

Methods: Are the methods clearly described? Could another researcher reproduce the results? Are there procedures that can better be represented in form of pseudocode?

Methods are clear. However, I am unsure about using a model's prediction as a ground truth. If there is no alternative, I guess this is the best available. Is there no other available dataset?

Also, motivate why you use ResNet-50 for the land cover type classification task. You currently have a paragraph consisting of one sentence, the statement that you will use ResNet-50: expand this with a motivation.

Considering you expect pix2pix to perform worse, is there maybe another model that you think could measure itself better with DSen2-CR? It may be more interesting to look at that model instead. Also keep in mind that, as far as my knowledge goes, pix2pix only uses RGB images, and is unable to use any/all of the other bands offered by Sentinel imagery. It would be good to mention the limitations of pix2pix in this regard.

Don't forget to add in parameters you have used, and, if relevant, the train/test split.

Other than that, the method seems solid: comparing the prediction on 'ground truth' images when compared to images where clouds have been removed seems like a good metric.
Do keep in mind that some images may have a much higher cloud coverage than others, and that the performance may vary depending on how much cloud coverage there is. This may also be interesting to investigate in your results.

Results: Are the results presented and interpreted in a meaningful manner? Is the link and reference to data provided? Are the necessary metadata about the dataset provided? Are the baselines introduced and explained? Is the motivation behind the data, baselines, performance metrics clear?

No results as of yet.

Conclusion: Do the authors link back to the Introduction? Did they deliver the promised contribution? Do they include disadvantages/ discussion points/ future research?

No conclusion as of yet.

Plots: Can you understand the plots? Does it use a good color scheme? Are the plots black/white proof? Is the text font readable? Do the plots have a correct format (labels on axes, title, legend, etc). Are all figures explained and referred to in text? Are the captions undertandable?

No plots as of yet.

References: Are there any claims that require a reference? Are there any errors in the bibliography?

Create a reference for the earth being covered by 67% of clouds.

I do not think you are required to write the entire original paper name (Motivation section). You can just write: “The authors of [2] propose an extension to the DSen-2 [5] …”

Use a consistent reference style. Sometimes, work is referenced with [1], and at other times it is referenced with [Research Group et al.].

I think you should also refer to ResNet when talking about ResNet-50 in your methods section.