# Bachelor Thesis Review

## Faculty of Mathematics and Physics, Charles University

| | |
|---|---|
| **Thesis author** | Oliver Piter |
| **Thesis title** | Unsupervised Cloud Classification from Sky Imagery |
| **Year submitted** | 2024 |
| **Study program** | Computer Science |
| **Specialization** | Artificial Intelligence |

| | | |
|---|---|---|
| **Review author** | Tobias Rittig | Reviewer |
| **Department** | Department of Software and Computer Science Education | |

## Overall

| | good | OK | poor | insufficient |
|---|---|---|---|---|
| Assignment difficulty | | X | | |
| Assignment fulfilled | | X | | |
| Total size                   *. . . text and code, overall workload* | | X | | |

Overall the assignment of unsupervised clustering of cloud images is a challenging topic. However with the availability of code to replicate state of the art methods the implementation workload decreases.

The thesis describes and extensively analyzes various attempts at fulfilling this task with mediocre success. The results show that the perfect solution to this problem has not been found yet. The thesis puts a lot of effort into evaluation but lacks in communicating the connection, purpose and reasoning in the chapters leading up to that.

## Thesis Text

| | good | OK | poor | insufficient |
|---|---|---|---|---|
| Form                          *. . . language, typography, references* | | X | | |
| Structure       *. . . context, goals, analysis, design, evaluation, level of detail* | | | X | |
| Problem analysis | | | X | |
| Developer documentation | X | | | |
| User Documentation | X | | | |

The level of English and text flow leaves room for improvement. Besides the constantly missing definite and indefinite articles, the text reads as a concatenation of factual statements with little connection between consecutive parts and sentences. I would recommend taking some English academic writing courses before the Master thesis.

The structure of the text is also differing quite from academic norm in that there is no method section but instead an overly inflated results section. A basic high level overview is only given as the last section before the conclusion. Information on various experiments is very scattered and most of the times there is no reasoning given on why specific choices have been made.

I am missing an introduction to the classification subject - clouds and atmospheric science. It seems the student has not conducted any background research but instead fell for the fallacy of treating the images as black-box data. Clouds are referred to in colloquial language as "fluffy"and "rainy". A proper background study could have informed a more tailored feature extraction algorithm.

Questions:

- Why have you limited the dataset to only 2/3 locations? With only so little variation one cannot be surprised about the lack of generalization over this axis, i.e. location ending up being one clustering factor.

- The limit of only looking at the first two principle components, which at times was identified as the day/night axis, seems like a missed opportunity. The data points might have been differing in higher-order components which were not analyzed. Did you base this choice on the Eigenvalues of the components or why didn't you consider other components?

- Which dataset augmentations would you consider problematic given the subject is clouds?

## Thesis Code

| | | good | OK | poor | insufficient |
|---|---|---|---|---|---|
| Design ..._architecture, algorithms, data structures, used technologies_ | | | X | | |
| Implementation ..._naming conventions, formatting, comments, testing_ | | | X | | |
| Stability | | | X | | |

The implementation part of this thesis is on the smaller side as it is build on top of an existing open-source clustering code base. I have not run the code myself, but let the student walk me through his implementation in an in-person consultation. Given the documentation in the thesis, the results should be well replicable.

**Overall grade**   Good
**Award level thesis**   No

Date                                     Signature