

Review of Master's Thesis

Student: Lukáč Jakub, Bc.
Title: Tracking People in Video Captured from a Drone (id 23056)
Reviewer: Orság Filip, Ing., Ph.D., DITS FIT BUT

1. **Assignment complexity** **average assignment**
2. **Completeness of assignment requirements** **assignment fulfilled only partially**

The thesis is not fully finished, in my opinion. The reasons for that are two. First of all, I miss more varied treatise on the people tracking and detection algorithms. In the thesis, there are only various types of convolutional neural networks described as the mean of object (person) detection. Even though CNNs are very popular, there are other techniques to detect persons in an image, as well. Besides, it is part of the assignment to study algorithms, which does mean to study more than one. A more serious issue, though, is the lack of better and more exact accuracy estimation of the proposed solution(s). The assignment states the proposed algorithm should determine the accuracy of the obtained trajectories, which is very vaguely mentioned in Chapter 4. I would expect a full report on the measurements, comparison of various methods, and summarization of results obtained from all the provided videos (which algorithm was best, what was the error, influence of the altitude or terrain on the results, proposal of solutions to the issues, etc.). Instead of extensive testing, there is only 4 pages long summarization of some experiments with some general statements. Eg. a sentence "However, the method did not perform that well and is not used in the final comparison" eliminates one method from the comparison because it is not doing well enough. Why did not that method perform well? And, actually, what does men "to perform well"? Where are the numbers representing the error rate? Except for that, I do not fully understand what exactly was the final solution and what was not. The text is rather confusing for me.
3. **Length of technical report** **within minimum requirements**

The number of pages in the thesis is approximately 65, which is lower than the common limit of 80-100 pages.
4. **Presentation level of technical report** **65 p. (D)**

Overall, the thesis is well structured, the chapters do logically divide the text into appropriate parts. However, the size of the individual sections is not proportionate to its content (eg. the subchapter 2.3 or 4.2, both with only two paragraphs, are an example of bad structuring of the text) and the chapter 4, which should be one of the most important chapters, is very sloppy and missing much information I would expect. Other than that, the text would be understandable if the English were not bad in some places.
5. **Formal aspects of technical report** **55 p. (E)**

There are no typographical issues in the text. The equations, figures, and all other elements are well-typed. The biggest downside is the quality of the language, though. I appreciate that the text is written in English. The spelling is mostly correct, but in some places, there are grammatical mistakes (missing words, wrong word order, or inappropriate choice of words), which sometimes makes the content very difficult to understand, and the sentences are sometimes overcomplicated.
6. **Literature usage** **90 p. (A)**

The selection of information sources is split among online sources, conference papers, and some books. The sources are properly cited.
7. **Implementation results** **50 p. (E)**

The software was implemented in Python, which implies the source code is well-structured by default. The code is commented briefly, yet sufficiently. The implementation itself is good enough and its behavior complies with the assignment. The main issue is the extent of the performed experiments, which is not sufficient. The experiments were either not fully performed or not described in the thesis. I would expect a full set of reference solution compared to the results of various algorithms implemented in the thesis. Besides, it seems the proposed solution does perform the detection and recognition on the client-side (on the Raspberry Pi) rather than on the server-side as requested by the assignment.
8. **Utilizability of results**

If the implementation included correction of the height of the terrain, it could be practically used to estimate distance of objects using camera not only on drones but in all similar scenarios.
9. **Questions for defence**
 - Describe the solution you have chosen in your thesis for person detection. Why does it run on the client

rather than on the server-side?

- You should have studied more than one algorithm of object detection. Are there other algorithms appropriate for this task?
- I am missing a more comprehensive report on the accuracy measurements of the algorithm. How was the accuracy estimated? How well does the proposed solution perform? Could you compare it to other algorithms?

10. Total assessment

48 p. failed (F)

In my opinion, the thesis is not fully finished. The assignment has been fulfilled only partially. Chapter 1-3 seem to be good enough, but Chapter 4 is, in contrast, very short and does not provide information I would expect (eg. comparison of various algorithms, or at least experiments with the proposed solution and error estimation on various video records, or even "live" experiment with the drone). Especially a detailed explanation and report of the accuracy estimation are missing. Besides, the thesis does barely fulfill the size requirements, language quality is not high enough and the theoretical part does not include different algorithms of object detection and tracking (only various types of CNN). The quality of the thesis is at the edge and for me, it is not sufficient. I would recommend to add more experiments, describe more algorithms, and rewrite Chapter 4.

In Brno 29 June 2020

Orsg Filip, Ing., Ph.D.
reviewer