FACULTY OF INFORMATION TECHNOLOGY CTU IN PRAGUE

Supervisor's statement of a final thesis

Student:	Ivana Hacajová
Supervisor:	Ing. Jakub Žitný
Thesis title:	Detection of organs in CT images using Neural Networks
Branch of the study:	Knowledge Engineering

Date: 10. 6. 2020

Evaluation criterion:	*The evaluation scale: 1 to 4.*
1. Fulfilment of the assignment	*1 = assignment fulfilled,* ***2 = assignment fulfilled with minor objections,*** *3 = assignment fulfilled with major objections,* *4 = assignment not fulfilled*

Criteria description:
Assess whether the submitted FT defines the objectives sufficiently and in line with the assignment; whether the objectives are formulated correctly and fulfilled sufficiently. In the comment, specify the points of the assignment that have not been met, assess the severity, impact, and, if appropriate, also the cause of the deficiencies. If the assignment differs substantially from the standards for the FT or if the student has developed the FT beyond the assignment, describe the way it got reflected on the quality of the assignment's fulfilment and the way it affected your final evaluation.

Comments:
The main parts of the assignment have been fulfilled; the student analyzed, designed, implemented, and evaluated basic models. However, better experimentation and proper evaluation metrics are missing. Additionally, a comparison with existing results from literature or VerSe leaderboard would be helpful.

Evaluation criterion:	*The evaluation scale: 0 to 100 points (grade A to F).*
2. Main written part	*85 (B)*

Criteria description:
Evaluate whether the extent of the FT is adequate to its content and scope: are all the parts of the FT contentful and necessary? Next, consider whether the submitted FT is actually correct – are there factual errors or inaccuracies? Evaluate the logical structure of the FT, the thematic flow between chapters and whether the text is comprehensible to the reader. Assess whether the formal notations in the FT are used correctly. Assess the typographic and language aspects of the FT, follow the Dean's Directive No. 26/2017, Art. 3. Evaluate whether the relevant sources are properly used, quoted and cited. Verify that all quotes are properly distinguished from the results achieved in the FT, thus, that the citation ethics has not been violated and that the citations are complete and in accordance with citation practices and standards. Finally, evaluate whether the software and other copyrighted works have been used in accordance with their license terms.

Comments:
The text's quality is excellent, the structure of the chapters is reasonable, and the theoretical part is comprehensive and easy to read. The practical part is split into two basic tasks — spine and vertebrae segmentation. Both are properly divided into parts that discuss loss functions, training, and results. Reasonable metrics are used for evaluation, cross-validation is utilized, and all steps are correctly explained (although the loss function comments could use some citations).
On the other hand, chosen metrics are not useful for comparison of trained models with results from VerSe leaderboard, and no such comparisons are made. Many experiments could be added for improvement (pre-processing, hyper-parameter tuning, more complex loss functions, transfer learning, or ensembling). The chosen topic was quite difficult, though, CT data need a lot of custom pre-processing steps, and 3D architectures need more care when tuning.

Evaluation criterion:	*The evaluation scale: 0 to 100 points (grade A to F).*
3. Non-written part, attachments	*90 (A)*

Criteria description:
Depending on the nature of the FT, comment on the non-written part of the thesis. For example: SW work – the overall quality of the program. Is the technology used (from the development to deployment) suitable and adequate? HW – functional sample. Evaluate the technology and tools used. Research and experimental work – repeatability of the experiment.

Comments:
Tech stack and repository are of outstanding quality, published on GitHub. Colab link could be placed into Readme, and notebook cells could be pre-calculated, but these are only small details that could be improved.

Evaluation criterion:	*The evaluation scale: 0 to 100 points (grade A to F).*
4. Evaluation of results, publication outputs and awards	*70 (C)*

Criteria description:
Depending on the nature of the thesis, estimate whether the thesis results could be deployed in practice; alternatively, evaluate whether the results of the FT extend the already published/known results or whether they bring in completely new findings.

Comments:

The student familiarized herself with the tech stack, all parts of the machine-learning flow, and was able to achieve interesting results. The results are not compared (or comparable) with "outside world," and the results from vertebrae segmentation are not good at all. As mentioned already, further experimentation would be needed here in order to achieve more impressive results.

Evaluation criterion:	*The evaluation scale: 1 to 5.*
5. Activity and self-reliance of the student	*5a:* *1 = excellent activity,* *2 = very good activity,* ***3 = average activity,*** *4 = weaker, but still sufficient activity,* *5 = insufficient activity* *5b:* ***1 = excellent self-reliance,*** *2 = very good self-reliance,* *3 = average self-reliance,* *4 = weaker, but still sufficient self-reliance,* *5 = insufficient self-reliance.*

Criteria description:
From your experience with the course of the work on the thesis and its outcome, review the student's activity while working on the thesis, his/her punctuality when meeting the deadlines and whether he/she consulted you as he/she went along and also, whether he/she was well prepared for these consultations (5a). Assess the student's ability to develop independent creative work (5b).

Comments:

The student was totally independent and moderately active.

Evaluation criterion:	*The evaluation scale: 0 to 100 points (grade A to F).*
6. The overall evaluation	*79 (C)*

Criteria description:
Summarize which of the aspects of the FT affected your grading process the most. The overall grade does not need to be an arithmetic mean (or other value) calculated from the evaluation in the previous criteria. Generally, a well-fulfilled assignment is assessed by grade A.

Comments:

To sum up, the thesis is admirable, the theory and text itself is of the highest quality. Due to the difficulty of the task, though, more advanced results have not been achieved.

Signature of the supervisor: