



Supervisor's statement of a final thesis

Student: Jan Rudolf
Supervisor: Ing. Jakub Žitný
Thesis title: Detecting abnormalities in X-Ray images using Neural Networks
Branch of the study: Knowledge Engineering

Date: 10. 6. 2020

<i>Evaluation criterion:</i>	<i>The evaluation scale: 1 to 4.</i>
1. Fulfilment of the assignment	<u>1 = assignment fulfilled,</u> 2 = assignment fulfilled with minor objections, 3 = assignment fulfilled with major objections, 4 = assignment not fulfilled
<i>Criteria description:</i> 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.	
<i>Comments:</i> All parts of the assignment are fulfilled, although the experiments and comparisons could be more complex (described below).	
<i>Evaluation criterion:</i>	<i>The evaluation scale: 0 to 100 points (grade A to F).</i>
2. Main written part	85 (B)
<i>Criteria description:</i> 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.	
<i>Comments:</i> The structure of the work is clear, and the reader can flow smoothly from basic concepts to the core of the thesis and experiments. The theoretical part is good. Most required building blocks are described, and no useless parts are present. It is clear that the student understands the topic. The CLAHE method is mentioned in Experiments, but it is not present in the theoretical part. On the other hand, cross-validation and hyperband are detailed, but these are never used during experimentation, which is a pity. Chapter 5 is very "blogpost-like"; it could use better citations and be more formal and detailed. The hardware and training times are not described at all. Parts of Chapter 4 are unclear as well, especially where assumptions are made ("in my opinion"). Experiments and result descriptions are okay, but there is notable room for improvement. It is not clear why the pre-training on ImageNet was not used. Why are results not cross-validated? The first three models in Table 4.1 have suspiciously similar CK. Figure 4.1 shows a case with a possibly incorrect label — this should have been explored and described in more detail.	
<i>Evaluation criterion:</i>	<i>The evaluation scale: 0 to 100 points (grade A to F).</i>
3. Non-written part, attachments	80 (B)
<i>Criteria description:</i> 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.	
<i>Comments:</i> The tech stack for the experiments is robust and used correctly. Reproducibility is somewhat achievable here; however, the repository is not documented thoroughly, and one would need to spend a while to use it for further experimentation.	
<i>Evaluation criterion:</i>	<i>The evaluation scale: 0 to 100 points (grade A to F).</i>

4. Evaluation of results, publication outputs and awards

80 (B)

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:

Results are impressive; the student compared four high-quality models and used proper pre-processing techniques. Comparison is made only against the "Stanford baseline," which was "defeated" many times, including many submissions from the MURA leaderboard.
Further experimentation here would be required to achieve better 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:

Student was active and self-reliant.

Evaluation criterion:

The evaluation scale: 0 to 100 points (grade A to F).

6. The overall evaluation

83 (B)

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:

Many more things could be tried here, however, the scope and depth of the work are appropriate.

Signature of the supervisor: