

## I. IDENTIFICATION DATA

<b>Thesis title:</b>	<b>Learning Segmentation from Multiple Datasets with Different Label Sets</b>
<b>Author's name:</b>	<b>Elnaz Babayeva</b>
<b>Type of thesis :</b>	master
<b>Faculty/Institute:</b>	Faculty of Electrical Engineering (FEE)
<b>Department:</b>	Computer Science
<b>Thesis reviewer:</b>	Radim Šára
<b>Reviewer's department:</b>	Cybernetics

## II. EVALUATION OF INDIVIDUAL CRITERIA

<b>Assignment</b>	<b>challenging</b>
<i>How difficult was the project assignment?</i>	
The student had faced a novel, previously unsolved problem. Some ideas were possible to take from a related literature but the problem required an original solution. The topic required a good understanding of multiple topics in machine learning, pattern recognition, and semantic segmentation.	

<b>Fulfilment of assignment</b>	<b>fulfilled with minor objections</b>
<i>How well does the submitted thesis fulfil the assignment? Were the primary goals achieved? Which points of the assignment fell short or were extended? Justify your answer.</i>	
All points of the assignment were fulfilled, except the state-of-the-art comparison which was not possible because no sufficiently similar method has been found in the literature. Nevertheless, there is an extensive chapter with experimental results that clearly demonstrates performance of the proposed methods using a reproducible evaluation methodology. If this work is published, other authors will be able to make a comparison. Still, besides stating the experimental results, the thesis would benefit from a discussion on the importance of the results with respect to the state of the art in the field.	

<b>Methodology</b>	<b>correct</b>
<i>Comment on correctness of approach and/or solution methods.</i>	
I have no comments on the methodology used in this work.	

<b>Technical level</b>	<b>A - excellent.</b>
<i>Is the thesis technically sound? How well did the student employ expertise in the field of his/her field of study? Does the student explain clearly what he/she have done?</i>	
The student used her expertise well. There are just a number of design choices on the low algorithmic level whose elegance and/or efficiency is difficult to assess for this reviewer, who is not doing active research in semantic or instance segmentation.	

<b>Formal and language level, scope of thesis</b>	<b>B - very good.</b>
<i>Are formalisms and notations used properly? Is the thesis organized in a logical way? Is the thesis extent adequate? Does the thesis meet typographical standards? Is the language clear and understandable? Is the English satisfactory?</i>	
The thesis is written in good English, it is well organized and easy to follow for a specialist but harder to follow for a non-specialist due to large technical complexity. I have some comments on the typographical style standards, see Additional Comments below.	

<b>Selection of sources, citation correctness</b>	<b>A - excellent.</b>
<i>Does the thesis make adequate reference to earlier contributions? Was the source selection adequate? Is his/her original work clearly distinguishable from prior art? Do the bibliographic citations meet the standards?</i>	
The sources are adequate. Citation standards are fulfilled.	

<b>Additional comments and evaluation (optional)</b>
--

*Comment on the overall quality of the thesis, its novelty and impact on the field, its strengths and weaknesses, utility of the developed solution, theoretical/formal level, student's skillfulness, etc.*

**Detailed comments:**

- It seems that the matrix table in Fig. 2.11 does not correspond to the graph. Is it a mistake or a misunderstanding on the reviewer's side?
- The connection between  $L_x$ ,  $L_c$  in (2.7) and  $V_x$ ,  $V_c$  in (2.9) is not clear. One of the proposed methods is inspired by this work so I would expect that there is a good understanding of the work [28,29], so that it is not difficult for the student to summarize the method well.
- The term 'sub-vector' in the explanation of (2.9) is a confusing term, it is perhaps better to avoid it.
- I do not understand the criticisms in the first two sentences on the beginning of p.27. Why exactly is the approach inefficient?
- Sec. 4.1 (which is important for understanding the method) is written a bit too chaotically. Discussions on several different concepts are crammed in a single paragraph (pg. 4). Crucially, it is not clear if metric learning has been implemented or not (It wasn't, as I understood later).
- It is not clear on p. 30 that MLML refers to [28]. Overall, the thesis would benefit from a table of abbreviations.

**Typographical and stylistic errors:**

- If all numbers in a sentence are zero to ten then they should all be written in words. See eg. the caption of Fig 2.8 but also elsewhere.
- There should be no forward references, see eg. (2.8) on p. 21 that is referenced before it is shown.
- A sentence cannot start with a formal (mathematical) symbol or otherwise formatted expression, unless defined by the typographical style template, see e.g. p. 21, where the symbol  $L_x$  starts even a paragraph. More examples are found elsewhere, e.g. on pp. 22, 26, etc.
- Figures, tables, algorithms, etc should be formatted as floats and appear at the beginning of a page or at its bottom, they should be inlined only rarely. If this rule was followed, we would see no orphan lines of text, as, for instance on pp. 30-31, 31-32, 37-38.
- The font used in all graphical elements should be similar in size to the font of the main body text. An extreme example of this is in Fig. 2.12 on p. 23. Also the bounding boxes and their labels in Fig. 5.8 are very hard to read.

### III. OVERALL EVALUATION, QUESTIONS FOR DEFENSE, CLASSIFICATION SUGGESTION

*Summarize your opinion on the thesis and explain your final grading. Pose questions that should be answered during the defense.*

The topic of the thesis is highly relevant and interesting. The work shows a good level of originality and creative work. The thesis is organized well. Experimental method and results are presented well. Results are encouraging. On the negative side, the thesis would benefit from a discussion on the broader importance of the results with respect to the state of the art in the field. Another weakness is that elementary typographical standards are sometimes violated.

There is an interesting attempt to probe into the relative inefficiency of the kNN rule for Faster R-CNN + kNN method in Sec. 5.3.2, which I appreciate. My **question** is if the results discussed in that section could mean that the "sample-level smoothness assumption" is in fact invalid in practice. Or it is perhaps valid, but it manifests itself only in an exponentially larger number of training samples.

I evaluate the submitted thesis with classification grade **B - very good**.

Date: **13.6.2019**

Signature: