Posudek diplomové práce

Matematicko-fyzikální fakulta Univerzity Karlovy

Autor práce	Luka Ihnatchenko		
Název práce	Deep Neural Networks in Image Processing		
Rok odevzdání	2020		
Studijní program	Informatika	**Studijní obor**	Umělá inteligence

Autor posudku	Mgr. Martin Pilát, Ph.D.	**Role**	oponent
Pracoviště	KTIML MFF UK		

Text posudku:

The goal of the thesis is to study deep learning techniques for classification of mammography images, specifically for breast cancer detection. The student describes a number of deep learning methods, chooses some of them and evaluates them on a dataset with such images.

The whole thesis is divided into six chapters. The first chapter contains the introduction and main goals of the thesis. The second chapter describes the most important techniques used in the rest of the thesis – (deep) neural networks including some specific architectures, semi-supervised learning, and the visualization of internal knowledge representation in neural networks. The chapter contains all the required information, however, the writing could be better. For example, in section 2.4.4 it is not clear what the formulas mean, and how they are connected to the rest of the text.

The third chapter contains description of available mammography datasets and basic visualization of the data. The student here also describes the problems with the datasets and then describes the CBIS-patch dataset (section 3.4.2) that was created to overcome some of these problems. It seem the this dataset was created by the student as part of the thesis, however, it is not explicitly stated.

The fourth chapter describes the software created in the thesis, including how to use it and where the most important parts are implemented. The description contains sufficient detail to be able to work with the software and extend it. My only (subjective) comment here is that this chapter might be moved to an appendix or even attachment as it does not contain any information that would be needed in the following sections and in my opinion it slightly breaks the flow of the thesis (after reading about techniques and datasets, I would like to read directly about the results).

The fifth chapter contains the description of all the experiments performed in the thesis. The experiments are very detailed and the student explains well all the choices made while performing

the experiments. The results are also discussed in detail including statistical evaluation. I also appreciate that the student also discusses the time required to train the models. Finally, the sixth chapter contains the conclusion of the thesis and ideas for future work.

The strongest part of the thesis are the experiments and comparisons performed. As already mentioned, these are performed with great attention to detail, including hyperparameter tuning, and provide a comprehensive comparison of the techniques. The weakest point, on the other hand, is the writing itself. It contains a relatively large number of grammatical mistakes, but in most cases the text is still understandable. The typography could also be improved – in many parts (e.g. end of Section 2.2) the variables are not distinguished from the rest of the text (n is written as n), some lines overflow to the right margin (e.g. on page 39), and there also a few orphans and widows (e.g. the one on page 57, after a page full of figures). Most of these are minor problems which do not affect the quality of work presented in the thesis and therefore I recommend the thesis for defense.

I have only two questions:

1. What are the specificity and sensitivity of the model? Would these be good enough to use the model in the real world? A model with a large number of false positives/negatives may not actually save any time for the radiologists, as they would have to check all the decisions carefully again.

2. Was any preprocessing (e.g. normalization) performed on the data? What about models in the literature, do they perform any preprocessing?

Práci doporučuji k obhajobě.

Práci nenavrhuji na zvláštní ocenění.

V Praze dne 3. září 2020

Podpis: