Posudek diplomové práce

Matematicko-fyzikální fakulta Univerzity Karlovy

Autor práce	Hoa Vu Trong
Název práce	Grounding Natural Language Inference on Images
Rok odevzdání	2018
Studijní program	Informatika **Studijní obor** Matematická lingvistika

Autor posudku	doc. RNDr. Pavel Pecina Ph.D. **Role** vedoucí
Pracoviště	ÚFAL MFF UK

Text posudku:

The thesis by Hoa Vu Trong focuses on the area of Natural Language Inference (NLI) where the textual input (a premise and a hypothesis) is enriched by visual information (an image) related to the premise. The task is to predict whether the hypothesis entails, contradicts, or is neutral to the premise. The aim of the presented thesis of is to examine if and how visual information can be used in deep learning models for this task.

The thesis is in English on the total of 67 pages. The text is very well written, with almost no errors, readable, and structured as a typical experimental work. The main textual content spans 56 pages and the author's own contribution is mainly in Chapters 3 (methodology), 4 (experiments), and 5 (conclusions).

The multimodal NLI task is quite new and only a few papers have been published so far – mainly due to unavailability of training and test data. Hoa Vu Trong address this problem by developing a new data set based on an existing data set for NLI (called SNLI) which has been enriched by linking the premises and hypotheses to images from the Flicker30K data set where the premises were originally extracted from. The data set, as discussed in the thesis, has some drawbacks and thus the experiments are also conducted using its "harder" subset.

Hoa Vu Trong worked with two types of models (Bidirectional Long Shor-Term Memory and Bilateral Multi-Perspecive Matching) and explored several ways how to exploit visual information. However, in most configuration the added visual information hurt the prediction accuracy. The best result is reported in Section 4.2.3 where the inclusion of the image in the BiMPM architecture does yield about 0.6% absolute (which is reported as a statistically significant improvement). Though the results are not ground-breaking, the work has been done very well and some of the findings are interesting and potentially important for future work.

The thesis demonstrates that the author well understood the problem, studied the related work, presented his own architecture modification, conducted a large number of experiments and

compared their results with other existing works. I recommend the thesis to be defended.

Práci doporučuji k obhajobě.

Práci nenavrhuji na zvláštní ocenění.

In Prague, 4. 9. 2018

Podpis: