

Recommendation for the defense of Ph.D. thesis

Ph.D. candidate: Kateřina Žmolíková
advisor: Jan "Honza" Černocký

I am happy that Kateřina (Katka) Žmolíková has finished her thesis under my supervision, and that her PhD years have converged to the submission of an excellent doctoral dissertation at FIT BUT. The term "supervision" must be precised - at the beginning of her research career, John Hershey and Shinji Watanabe contributed to making Katka a specialist in the use of neural network embeddings together with speech signal processing, and in the continuation, she received a great support and many ideas from Marc Delcroix and his colleagues from NTT labs in Japan. In Brno, she naturally benefited from the know-how accumulated in the group and personified by its research director Lukáš Burget.

Her thesis is in the important area of robust speech processing aiming at real-life scenarios. The target speaker extraction (TSE) she has investigated, has significant importance both as a stand-alone block (producing the signal with enhanced target speaker and suppressed other speakers and noise) and in combination with other "target" systems (for example speech recognition or speaker diarization) that can include TSE as a front-end trainable in end-to-end fashion.

I know Katka since the fall semester 2012/13 when she took my course "Introduction to Signals and Systems" (ISS). It is impossible to remember everyone taking ISS as it is compulsory to all Bc.-level students at FIT (several hundreds each year), but she was among the few to attend the Friday morning lecture group (taught in English), and, with four other students, obtained 98/100 points - the best result in ISS that graduation year.

She started to work with BUT Speech@FIT group and gradually progressed to true scientific excellence, making her one of the best PhD students in my group, and, in my opinion, at the whole BUT. Let me select a few important points:

- In summer 2014, she visited Toshiba (Cambridge, UK) research group led by Prof. Yannis Stylianou to work on KALDI-based speech recognition. Her UK supervisor was not expecting a bachelor student, but at the end, he was very enthusiastic about her performance.
- In summer 2015, she attended the prestigious Second Frederick Jelinek Memorial Summer Workshop at Johns Hopkins University. For six weeks, she worked in research team "Far-Field Enhancement and Recognition in Mismatched Settings" under the direction of John Hershey and Shinji Watanabe. This work led to several conference papers and to her deep interest in neural network embeddings used for speech signal processing and recognition.
- Based on the results from this workshop, she started a collaboration with NTT Communication Science Laboratories in Japan (Marc Delcroix and Keisuke Kinoshita). Her work aims at embeddings for speaker-aware speech separation and speech recognition. This work still continues, and has led to several contractual research agreements between NTT and BUT, to several conference papers, two journal publications, and three patent applications.

At BUT, Katka has coordinated a research project funded by the Technology Agency of the Czech Republic, participated in several Czech and European research projects, and she has been leading the cooperation with NTT. She was the main person coordinating the BUT

automatic speech recognition efforts in 2020 CHiME-6 challenge concentrating on distant multi-microphone conversational speech diarization and recognition.

In addition, Katka she been active in teaching speech- and signal-oriented courses at the Faculty of Information technology of BUT, led several successful bachelor and diploma thesis and is highly appreciated by her students for excellent technical supervision and very helpful attitude.

Katka has a publication track surpassing many senior faculty members, she is author or co-author of several papers accepted to top conferences in our field (IEEE ICASSP, ISCA Interspeech, and IEEE ASRU), of one book chapter, and of two papers in journals of which the IEEE Journal of Selected Topics in Signal Processing (**impact factor 6.688**) is clearly outstanding. Despite her young age, she has **h-index 11** at Google Scholar and **10** at SCOPUS.

To conclude, **I fully recommend Kateřina Žmolíková's Ph.D. thesis for the defense.** I am wishing her all the best in her future professional and personal life, and I am looking forward to continuing to cooperate with her.

In Brno, March 21st 2022

Dr. Jan "Honza" Černocký, Professor
Head of Department of Computer Graphics and Multimedia
Responsible of BUT Speech@FIT group
Faculty of Information Technology, Brno University of Technology
Božetěchova 2, 612 66 Brno, Czechia, Tel: +420 5 41141284
<mailto:cernocky@fit.vutbr.cz>, <http://www.fit.vutbr.cz/~cernocky>
<http://www.fit.vutbr.cz/> <https://cs-cz.facebook.com/FIT.VUT>
<http://speech.fit.vutbr.cz> <https://www.facebook.com/BUT-Speech/>