

## OPPONENT'S REVIEW OF A DOCTORAL THESIS

Author: RNDr. Aleš Zita

Title: Analysing Videokymograms Using Classical and Deep Learning Methods

Opponent: Radovan Jiřík, Ph.D.

The main topic of the dissertation is automatic analysis of images acquired using 1D line-scanning of vocal folds and recording the time evolution, so called videokymograms. The topic is very actual and presents a challenging research area of image processing with very practical applications being computer-aided medical diagnosis of laryngeal diseases.

The proposed methods might improve the reliability and time efficiency of routine vocal fold diagnosis.

The text of the dissertation is very well structured. It starts with a comprehensive introduction to vocal-fold examinations and focuses on the particular examination method, videokymography. The hardware and software aspects of the method are described. The dissertation is a part of long-term unique research and development activities of the author's institute in the field of videokymography, performed in close connection with the producer company.

The main goals and structure of the thesis are clearly stated. The given literature review shows that the dissertation follows up on several conventional image processing approaches applied in this field and extends it by estimating several new features. Furthermore, the dissertation introduces deep-learning approaches to the field (together with one previous publication of another team).

The text is written in very good English and on a very good formal level. The thesis is written as a collection of 7 conference and journal papers, namely 2 journal papers (one 1<sup>st</sup>-authored) and 5 conference papers (three 1<sup>st</sup>-authored). Together with the additional author's publications listed in the bibliography of the thesis, the authors's publication activity is clearly sufficient for a successful defense of the dissertation. A structured list of all author's publications is missing and would be appropriate.

The dissertation presents state-of-the-art image processing methods which are not new on their own, but the novelty lies in their new application in the field of videokymography and their appropriate and optimized combination. I also see a great value of the dissertation in the fact that it presents two compact well evaluated software tools for analysis of videokymograms, one based on conventional methods (paper 1) and one based on deep-learning methods (paper 7). This may potentially have high impact on clinical diagnosis of larynx.

The quality of the used methodology and the achieved results is indicated by being published in high-level journals and conferences and also by winning the ImageCLEF2020 competitions in two categories (papers 4, 5), which shows that the author(s) have mastered the limited-datasets deep-learning methodology on a very good level.

In conclusion, the author has clearly proved his ability to work systematically on the given scientific topic and showed a good potential for further work in research. The results presented in this thesis are of high impact for the image processing and medical imaging communities, as well as for routine clinical diagnostics. In my opinion, the author has qualified to be awarded the title Ph.D.

Questions:

1. In the last paragraph of Section 6.3.3, results achieved in two studies are presented. Can you elucidate what evaluation results are required in order to employ your software in routine clinical practice?
2. Can you explain the issue of certification of the software mentioned in Section 3.2? Is a certification intended also for the software of Paper 1 and/or Paper 7? What is the relationship of the mentioned certification to FDA and CE approvals?
3. Could you please specify more clearly your contribution to Paper 2? The formulation of the first item in Section 6.2.4 is not clear. Is it design of the methodology, implementation, taking part in processing of the data...?

In Brno, Feb 17, 2023

Radovan Jiřík, Ph.D., opponent