Habilitation Thesis Reviewer's Report

Masaryk University Faculty	Faculty of Informatics
Procedure field	Informatics
Applicant	Ing. Vlad Popovici, , Ph.D.
Applicant's home unit, institution	Faculty of Science, Masaryk University
Habilitation thesis	Computational biomarker discovery: methods and practice
Reviewer	Prof. Daniel Racoceanu, HDR, PhD, MSc, Ing.
Reviewer's home unit, institution	Department of Engineering, Pontifical Catholic University of Peru, Lima Peru *Tenure Profesor, Faculty of Science and Engineering, Sorbonne University, Paris, France*

Dr. Vlad Popovici prepared a habilitation thesis entitled: "Computational biomarker discovery: methods and practice". This document synthetizes the scientific contributions and strategy of the candidate during the last decade of his activity.

The title of the work sounds very pragmatic, showing immediately the wish of the candidate to be effective with his approaches and his synthesis. The effectiveness of the thoughts makes this synthetic document very precise, as easy to read and interpret.

The overall objective of Dr. Popovici's research is dedicated to biomarker discovery and validation. The majority of the results are obtained in the field of oncology, in particular for colon and breast cancers. The approaches developed here are within the area of Integrative Computational Pathology, aiming at refining the existing knowledge about breast and cancer sub-typologies, by bringing a cross expertise issued from the joint study and characterization of the tumor microenvironment and the tumor heterogeneity, together with the understanding of the genomic signature, and the clinical context.

In order to make a full use of the phenotypical features, the research proposed strategy considered a data-driven perspective concerning the whole slide image (WSI) analysis, without integrating – at the first sight - pathologist's expert supervision. The computational approach taken was based on extensions to bag-of-visual features method. These extensions aimed at producing more descriptive dictionaries for the histopathology images and investigated the possibility of structuring the visual dictionaries around some semantical terms, allowing an easier interpretation of the results. The resulting models revealed new features, some of them without a clear correspondence in the present pathology practice. This approach is important since the computer vision is able to extract features which have not been explicitly formalized as critical. Careful correlation approaches shall be set-up in the future to data mine this unsupervised indexing.

Another computational aspect addressed was the optimization of the visual features for the purpose of analyzing pathology images and in-depth analyses were performed using both "classical" image analysis features and convolutional neural networks features. A hierarchical quantification schema was also proposed for building multiresolution visual dictionaries. This approach led to features that were more appealing to an expert pathologist than the convolutional features, but the overall performance of the system was lower. This could lead to an interesting discussion about the role of the pathologist in the future of Integrative Pathology Approaches / Demarches.

In conclusion, the presented studies demonstrated that not only it is possible to combine whole-slide imaging with molecular/gene expression data, but this combination reveals new connections between the "genetic program" and the tissue architecture. Even if the results were obtained on breast and colon cancer, the methodologies developed and employed here are deployable to other cancers, as well. This highlights the genericity of the approaches and encourages to believe that this study will support an important new generation of studies on the way and coming forward.

Reviewer's questions for the habilitation thesis defense

1. You did make use of semantics in some of your approaches. Could you additionally develop your perspective about a more dynamic use of semantics in biomarker discovery/validation and cancer sub-typing / patient stratification refinement?
2. From your studies, some interesting conclusions were extracted, in which sometimes the intuition is not giving the best indexing and the best results ... How do you see the role of the pathologist in the future Integrative Pathology Approaches / Demarches?
3. The data-driven WSI analysis revealed new features, some of them without a clear correspondence in the present pathology practice. Could you detail the critical part of the correlation approaches which could be developed to mine these features?
4. How would you structure a (distributed?) lab dedicated to integrated Computational Pathology? How do you see the structuring characteristics of a sustainable international / national / regional / local collaboration in this area (considering all the parameters involved in such research – i.e. the sensitivity of the data, size of data, patients' involvement, decision organisms involvement etc.)?

Conclusion

The habilitation thesis entitled *"Computational biomarker discovery: methods and practice"* by Vlad Popovici *fulfils* requirements expected of a habilitation thesis in the field of Informatics.

In Lima on 22nd of March 2018

....
signature