Pawel Herman, Associate Professor

School of Computer Science and Engineering
KTH Royal Institute of Technology
Lindstedtsvägen 5, 11428 Stockholm, Sweden

Review of doctoral thesis

This review is concerned with a doctoral thesis titled *"Computer-Aided Image-Based Diagnosis: Extensions of Computational Models with Domain Knowledge"*, authored by Mateusz Buda under the main supervision of Prof. Artur Jerzy Przelaskowski and co-supervision of Dr. Maciej Andrzej Mazurowski.

1. General characteristics

The PhD thesis under review is in the scientific discipline of Information and Communications Technology and the field of Engineering and Technology at Warsaw University of Technology, Poland. The research reported in the thesis is centered on machine learning approaches to computer-aided diagnosis in radiology. The author identifies and examines key challenges in developing and deploying these computational techniques in clinical contexts to support radiological diagnosis, and proposes concrete measures to effectively address them. The subject is relevant and up to date, particularly given fast evolving digitalisation trends and unprecedented advances in so-called artificial intelligence (AI). The treatment of the subject offers scientifically valuable insights and has strong utilitarian value for potential applications in the clinical domain.

The 179-page-long thesis report is written in English and contains a summary (alongside keywords) in Polish. The body of the thesis is composed of seven chapters (pages 9-44) along with two appendices - the first one containing key author's publications and the second one includes an extensive collection of the authorship statements corresponding to the attached publications. The chapters are preceded by the acknowledgments, abstract in both English and Polish, and the table of contents.

2. Evaluation of the main scientific problem formulation, aims and context of the research reported in the thesis and its scope

The research reported in the thesis is positioned within a fast-developing area of computer-aided diagnosis, which has undergone a major paradigm shift with the widespread use of deep learning methods. The author's work not only leverages these advances but also significantly contributes to the applicability of machine learning approaches to medical imaging. In the era of digital transformation, accelerating at an

unprecedented pace, healthcare and medical diagnostics have been considered as one of the most promising target application domains for machine learning especially given remarkably growing availability of data. Yet, the deployment of machine learning, and particularly deep learning methods, in clinical practice is still somewhat limited due to the intrinsic complexity of medical diagnostics problems and of legal regulations among others. The author identifies some key underlying technical challenges and tackles them in his work. He makes an important attempt to provide a wider and more generalising perspective on the applicability of the machine learning based approach to computer-aided diagnosis in radiology despite studying different imaging modalities and diagnosis problems. This perspective is valuable and timely given the urgent need to scale up clinical radiological diagnostics and, in the longer run - treatments. As far as the realm of data science and machine learning is concerned, the thesis also offers valuable insights since the technical challenges associated with processing medical images can be generalised to other domains. In this regard, the problems identified by the author as well as novel solutions proposed and evaluated in the presented body of work represent a meaningful contribution to the field. It should be noted that some of the findings reported by the author constitute a body of benchmark results for future algorithms to compare. This is important given the pace of technical advancements in deep learning technology with newer and computationally more potent architectures already being in use. From this perspective the author's work can be considered as the relevant reference point in the fast-moving trajectory of technological progress in a wider domain of today's artificial intelligence.

The aim formulated in the thesis work is to "improve machine learning methods to support diagnostic imaging in the context of selected (...) challenges" by proposing new solutions and refining the existing ones so that those challenges and requirements identified within the realm of image-based computer-aided diagnosis are met. This overarching aim is rather broad so the scope and level of ambition is more tangibly determined by the range and calibre of the aforementioned challenges. They can be deduced from what the author calls "a complimentary list of more specific research hypotheses". In particular, the urgent problem of class imbalance, insufficient and biased data, the relevance of transfer learning to improve generalisation capabilities of deep learning techniques, effective way of exploiting existing "approximate annotations" for segmentation purposes as well as utilising auxiliary information to boost generalisation of deep learning models for ultrasound image analysis are clearly listed. Given the spectrum of different medical imaging modalities the author has worked with, the overall scope and desirable balance of the research breadth and width, outlined in the introductory chapter, matches the ambition of doctoral studies and provides a good basis for the positive assessment of the problem formulation.

What could have been improved however in the PhD thesis is the construction of the main hypotheses. In my view, they are excessively broad and insufficiently specific as they imply merely a feasibility study with rather vague criteria ("more effectively"). This PhD work definitely represents a more impactful and well targeted scientific enquiry. In other words, I do not feel that the "two-fold research theses" give justice to the wealth of the author's research agenda. In addition, the second research thesis concerned with an assessment or radiology workflow (diagnostic process) is hard to follow up on since the term is rather broadly defined in the literature. In fact, it may account for all processes leading up to the diagnostic decision or radiological assessment in the clinical context. Integrating machine learning into that workflow should therefore be also concerned with the broader problematics of human-in-the-loop AI systems. While the incorporation of deep learning in a clinical decision making pipeline for radiological management of thyroid nodules in ultrasound images is mentioned in two publications, it is somewhat less visible, less exposed in the overall research contribution account in the thesis. From this perspective, I would advise a

more conservative approach to the formulation of this particular research hypothesis. All in all, I would like to strongly emphasise that the issue raised here is just related to the formulation of the hypotheses since the main goals presented and contextualised clearly by the author, as discussed above, offer sufficient insights to understand the scientific agenda of the doctoral research. In fact, the formulation of the grand research hypotheses or the leading problem statement could have been limited to what the authors call "a complimentary list of more specific research hypotheses" together with the tangible overarching aim referred to above. They sufficiently define the scientific agenda on their own, without the existing two-fold feasibility hypothesis.

The body of the thesis is composed of seven chapters along with two appendices - the first one containing key author's publications and the second one includes an extensive collection of the authorship statements corresponding to the attached publications. The chapters are preceded by the acknowledgments, abstract in both English and Polish, and the table of contents. Below a short summary of thesis chapters' scope is given without reviewing details of the author's scientific publications referred to in the subsequent chapters.

In **Chapter 1 Introduction** (7 pages) the author contextualises the presented research starting from a broad introduction to machine learning based approach to computer-aided diagnosis support without really reviewing the state of the art. Furthermore, the author identifies key limitations and challenges for this approach when deployed to support diagnosis in radiology across different imaging modalities. Importantly, the first chapter specifies main objectives and defines research hypotheses to provide a scientific framework for this work. It also lists key author's publications accounted for in the thesis and outlines thesis organisation. Overall, the introductory chapter sets the adequate tone and rather clearly presents a research theme by formulating and motivating the core problem undertaken in this work. Here, the complementary list of objectives is particularly appreciated. Perhaps the problematics could be more intrinsically anchored in the existing body of scientific work with a more thorough literature review or a more representative use of literature references.

Chapter 2 (3 pages) summarises the first two publications from the list of the author's publications. In particular, it stresses the relevance and impact of deep learning technology on diagnostic support in radiology and discusses the problem of class imbalance inherent to convolutional neural networks.

Chapter 3 (4 pages) summarises the third and fourth of the author's publications, where the author reported on the segmentation task performed on publicly available magnetic resonance imaging data with the use of a so-called U-Net architecture of a convolutional neural network and investigated the effect of transfer learning from natural images to magnetic resonance imaging data in a classification task (discrimination between tumour genomic subtypes) with a convolutional neural network, respectively.

Chapter 4 (5 pages) briefly reviews the fifth, sixth and seventh publications in the author's publication list. It demonstrates how approximate or auxiliary annotations can be leveraged by deep learning in a segmentation task in ultrasound images. In a similar spirit it also points out how auxiliary label annotations (extra imaging feature prediction tasks) boost the diagnostic performance of convolutional neural networks in estimating the likelihood of malignancy of thyroid nodules by exploiting a multi-task learning approach. Finally, the author reports how risk stratification in the radiological assessment of thyroid nodules with the American College of Radiology Thyroid Image Reporting and Data System can be optimised with the support of genetic algorithms broadly referred to as an AI method. Importantly, this contribution is unique in the thesis as it touches upon clinical workflows strongly exposed in the thesis introduction and reveals

the potential of machine learning to support guidelines for thyroid nodule interpretation within the aforementioned radiological thyroid image reporting and data system.

Chapter 5 (4 pages) outlines the main findings reported in the eighth and the ninth of the author's publications in the context of digital breast tomosynthesis. One of the key contributions elaborated in this chapter is the curation of what constitutes now a publicly available dataset of three-dimensional digital breast tomosynthesis volumes with annotations, helping evaluate this new technique in breast cancer screening. In a similar vein, the author also contributed a baseline deep learning model for benchmarking purposes. The chapter also summarises the author's and his colleagues' efforts towards more effective use of imaging data from healthy individuals, as they are more easily and broadly available, for developing deep learning models that can identify actual lesions in clinical patients. Despite promising results the author explains key shortcomings of this work.

Chapter 6 (6 pages) serves as a summary for the main body of the thesis. At first the author briefly reviews research problems and challenges addressed in respective studies reported in the publications, emphasising his original contributions. Then a valuable discussion of computer-aided diagnosis systems is included alongside reflections about multi-aspect challenges present at different stages of designing, developing, maintaining, evaluating and deploying such systems in clinical contexts.

In **Chapter 7** (2 pages) the author provides an overview of his academic achievements supported by bibliographic statistics and a record of his contributions to the academic community. Importantly, the author stresses his dedication to the open science practices.

As far as the scope of the work reported in chapters 2-5 is concerned, it accounts for 9 publications addressing a spectrum of different clinical and technical challenges across diverse imaging modalities. It reflects the breadth of the author's research though the focus on selected machine learning problems clearly represents a common thread for most of the author's contributions. To better appreciate this unifying perspective and, at the same time, to gain deeper insights into the complex nature of research problems and technicalities concerned with devising machine learning based solutions in radiological diagnostics, it is necessary to study the author's publications in more detail. It feels that chapters 2-5 could better facilitate this process, helping the reader to focus on all details important to appreciate the depth of synergies across different contributions.

From a linguistic perspective, the thesis report represents high quality and clearly, though almost excessively concisely at times, communicates the content.

3. Original contributions and major highlights

The work reported in the thesis is concerned with the effective use of machine learning for computer aided diagnosis with a heavy focus on deep learning approaches to image based radiology diagnosis support. In particular, the author identifies a set of relevant methodological challenges in radiology workflows (discussed in section 2 of this review) and in a series of computer studies tackles them relying on novel advancements of machine/deep learning techniques or their novel use, including their hybridisation, in this emerging domain. The underlying publications, co-authored with multiple researchers, demonstrate these original data science pipelines and report on their systematic evaluation as well as insightful comparative analyses. As highlighted in section 2, this research line is well motivated and timely with a considerable

potential for high impact in the clinical context. At the same time, the nature of challenges in radiology workflows to be addressed with machine learning methodology has offered a tangible opportunity for the author's contributions to the field of machine/deep learning itself. Consequently, the author's research is highly interdisciplinary and in a broad sense caters for both radiology and machine/deep learning audiences with different flavours of contributions to radiology assessment/diagnosis and machine learning/data science, respectively. One first-author and one second-author publications feature in *Radiology* journal and another second-author publications in *Journal of Magnetic Resonance Imaging*, where the clinical and radiologist perspectives of effectively leveraging machine learning based diagnosis support are prioritised. Stronger focus on deep learning methodology in image based diagnosis is reflected in two first-author publications in more technical journals such as *Computers in Biology and Medicine* and *Radiology: Artificial Intelligence*, and in *Scientific Reports* (third author). Mateusz Buda has also authored an impactful publication on the class imbalance problem in convolutional neural networks in *Neural Networks* outside any radiology (even clinical) context. This particular publication is concerned with fundamental aspects of deep neural network architectures and has thus received considerable attention in the field. All these publications were written in collaboration or under the supervision (deduced from the last position in the author's lists) of Dr. M.A. Mazurowski, Mateusz Buda's co-supervisor at Duke University. Interestingly, the main supervisor, Prof. A.K. Przelaskowski is not in the author's list of those publications.

I consider the following advances and findings as the most valuable original contributions and research highlights documented in the author's publications:

i) a systematic analysis of the effects of class imbalance on the performance of convolutional neural networks along with a review and evaluation of techniques to counteract negative implications with focus on sampling methods - published in *Neural Networks* (more than 2800 citations since 2018);

ii) an identification and quantification of the predictive power of tumor shape characteristics about tumour genomic subtypes; to the best of my knowledge, this relation has not been established before;

iii) a systematic evaluation of the effect of transfer learning on the performance of convolutional neural networks in radiology; the application domain and specific configurations of transfer learning with respect to available domain specific data are of special importance here;

iv) novel deployments of deep learning methods (convolutional neural networks) for thyroid nodule assessment based on ultrasound images and, in particular, a systematic evaluation and insightful comparison with the radiologists' assessment; this study has set the baseline for future computer-aided thyroid nodule assessment based on ultrasound input especially in relation to the human expert performance;

v) contribution to the advancement of clinically relevant Thyroid Imaging Reporting and Data System by optimising guidelines for interpretation of thyroid nodules in ultrasound images.

vi) a significant contribution to the community with the mindset to promote the advancement of machine learning support for breast cancer screening by curating and publicly sharing real-world clinical data obtained with a recently emerging and promising digital breast tomosynthesis - this contribution reflects the author's commitment to open science.

All in all, the author's publications, with the support of authorship statements, reflect the scope of his original contributions and testify to his strong competence, deeper understanding of main problems and potential to propose innovative solutions. It is worth noting that bibliometric data also reveals a considerable

interest of the research community in these publications. Above all, the author has demonstrated a critical scientific mindset and a mature problem oriented research approach.

4. Critical remarks and review of key limitations

Despite the appreciation for the scope, research depth and impact of the work reported in the thesis, there are some limitations worth mentioning. As some of the key limitations and shortcomings of technical nature, along with suggestions for addressing them in future work, are discussed by the author and his colleagues in individual research publications, here I mainly list overarching, still rather minor, concerns and share critical remarks about the entire thesis report.

i) Despite a systematic evaluation strategy adopted in the studies discussed within the thesis framework, it is felt that the decision about the architecture of the base model is not fully motivated. In fact, given different algorithmic manipulations, e.g. transfer learning configurations or multi-task-learning setup among others, introduced to examine their effect on the model performance it would be insightful to study this question more broadly across alternative base architectures. There is some evidence in the literature that suggests varying implications of transfer learning or multi-task learning depending on the deep learning model. This is a minor remark since the base network architecture that the author often exploits is well established in this application domain (though not explicitly motivated in the author's publications) so the conclusions may be most representative for the field.

ii) The author builds a narrative about technical challenges of deploying deep learning methods in image based computer-aided diagnosis partly exploiting his high-impact study on the effect of class imbalance on the robustness of convolutional neural networks in non-medical images. Although extending the argumentation from ordinary images to radiological modalities is intuitive, the negative implications of class imbalance may actually manifest themselves in different ways. In my limited experience working with deep learning method to support ovarian tumor diagnostics based on ultrasound images, the class imbalance had somewhat different effects on the performance. Also, more recent publications on the problem of class imbalance in the radiology domain demonstrate class specific effects. In any case, this is a minor remark since, as mentioned, the line of reasoning to generalise problem setting from normal images to clinically relevant radiological modalities is reasonable. It could be just highlighted and critically discussed at more depth.

iii) As briefly mentioned in section 2, a rather imprecise formulation of the key hypotheses in the thesis represents a certain weakness. In fact, the notion itself and the scope for consideration given to clinical workflows could be more precise. The nature of the author's contribution in this regard could be better contextualised in the larger deployment framework for computer-aided diagnosis systems in clinical practice. Although section 6.1 shows the ambition for a valuable discussion along this line, it could address the topic at more depth and offer novel insights. The discussion should also be strongly anchored in the existing literature and produce clearer directions for future work. It is also felt that ethical issues and different sustainability aspects could be given more consideration particularly in this domain of research and innovation.

iv) It is felt that the summary of the research content in chapters corresponding to subsequent publications is very concise. This in itself is not concerning. However, what could help convey the scientific

perspective is the actual account of the literature context. Challenges and problems the author identifies and tackles could be briefly discussed in relation to the state-of-the-art developments in the field accounting also for vast applications of deep learning methods in medical imaging.

5. Conclusion and final evaluation

In conclusion, I state that the author has addressed all the research objectives and more specific research hypotheses listed in the introduction. Overall, Mateusz Buda has made a meaningful and significant contribution to the field of computer-aided image-based diagnosis as well as machine learning in clinical applications. His research work therefore represents a valuable contribution towards the advancement of the scientific discipline of information and communications technology. The thesis meets in my view all the subject related requirements for PhD defence.

Paweł Herman, 13.01.2025
Stockholm