

Dear Editor,

please find enclosed the revised manuscript in Word format (file name "83916-Revised Manuscript").

Title: Automated Patellar Height Assessment on High-Resolution Radiographs with a Novel Deep Learning-based Approach.

Authors: Kamil Kwolek, Dariusz Grzelecki, Konrad Kwolek, Dariusz Marczak, Jacek Kowalczewski, Marcin Tyrakowski

Name of Journal: World Journal of Orthopedics

Manuscript NO: 83916

Manuscript Type: ORIGINAL ARTICLE

We would like to thank Reviewers for positive e-mail, as well as for substantive comments, which permitted us to improve. We emended the paper according to the reviewers' comments. We hope this revision will make our manuscript better to be accepted in your journal.

Each comment has been answered accordingly in the manuscript and each text that has been altered was highlighted red in the revised manuscript.

We hope that the revised version will fulfill the requirements for publication in the World Journal of Orthopedics.

Thank you very much.

Reply to ewviewer's comments:

Reviewer #1

Comment #1 and #2: Change the study's title to reflect the goal for which the current study was created. The new title should be twenty words or less and free of abbreviations, as these are the characteristics of a good title in terms of research, according to scientific research methods.

Answer comment #1 and #2: Thank you very much for this valuable comment. Thanks to your remark we reconsidered the new one. In particular new title reflects the goal of the work. The new title is:

"Automated Patellar Height Assessment on High-Resolution Radiographs with a Novel Deep Learning-based Approach"

We hope that you accept it. The new title has thirteen words.

Comment #3: The results section of the study abstract should focus solely on highlighting the most important findings of the current study.

Answer to comment #3: We agree with you. The Results section was shortened from:

“The ICC and SEM values between R#1 and AI for CD and BP indexes were 0.86 (± 0.38) and 0.015, and 0.8 (± 0.33) and 0.013, respectively. For R#2 and AI, the ICC and SEM values were 0.88 (± 0.38) and 0.015 for CD, and for BP were 0.79 (± 0.32) and 0.014. Excellent compliance between manual measurements and automatic measurements was achieved. The Dice score expressing the accuracy of bone segmentation on 92 test images was 95.9% (± 1.26).”

to sentence:

“Excellent agreement between the orthopedic surgeons’ measurements and results of the algorithm has been achieved (ICC > 0.75, SEM < 0.014).”.

Comment #4: The study's abstract is very long and contains fillers in the information. I hope it to shorten further and keep only what is important and serves the research's interests.

Answer to comment #4: Thank you for your valuable comment.

The Background section was shortened from:

“Artificial intelligence and deep learning have shown promising results in medical imaging and interpreting radiographs. On the other hand, the medical community shows a gaining interest in assisted analysis and automatization of routine diagnostics issues and orthopedic measurements.”

To sentence: “Artificial intelligence and deep learning have shown promising results in medical imaging and interpreting radiographs. Moreover, medical community shows a gaining interest in automating routine diagnostics issues and orthopedic measurements.”

The Aim paragraph was shortened from:

“Investigate if automatic measurements based on deep learning algorithms may be a valuable and repeatable tool for calculations of patellar height indexes. The aim of this work was to study the applicability of deep learning algorithms for bone segmentation and estimation of the patellar height on lateral knee radiographs.”

to: “Verification of automated patellar height assessment using deep learning-based bone segmentation and detection on high resolution radiographs.”

In the Methods paragraph of the abstract the sentence:

“The concordance between manual and computer measurements was calculated using the interclass correlation coefficient (ICC) and the standard error for single measurement (SEM). The radiographs have also been evaluated in terms of segmentation accuracy to assess the generalization of the U-Net on the test set. The detection of bones making the patellofemoral joint on high-resolution images was done using a YOLO neural network.”

Was shortened and modified to: “The detection of required bones regions on high-resolution images was utilized using a YOLO (You Only Look Once) neural network. The agreement between manual and automatic measurements was calculated using the interclass correlation coefficient (ICC) and the standard error for single measurement (SEM). To check U-Net's generalization the segmentation accuracy on the test set was also calculated.”

Conclusions section was modified from: “The novelty of this research is an innovative approach to estimating joint lines based on automatic bone segmentation and line fitting to determine patellar height indexes on high resolution radiographs. Our work demonstrates that on a small amount of training data, it is possible to achieve a satisfactory segmentation quality with good generalization on

unknown test radiographs that can be further used to perform estimation of CD and BP indexes. The proposed method exhibited comparable performance to orthopedic surgeons in estimating the patellar height. The method can also be valuable as a pre-operative and, after further development, as a postoperative assessment tool for big volume data analysis in medical practice.”

to “Automatic patellar height assessment can be achieved on high-resolution radiographs with the required accuracy. Determining patellar end-points and the joint line-fitting to the proximal tibia joint surface allows for accurate CD and BP index calculations. The obtained results indicate that this approach can be valuable tool in a medical practice.”

Comment #5: The study's introduction is also very long and should be condensed into two to three paragraphs at most, with the last paragraphs devoted to highlighting the research problem and how to address it within the framework of the goal for which the current study was designed.

Answer to comment #5: This comment permitted us to highlight the research problem. In current manuscript the introduction consists of three paragraphs and has been considerably shortened.

Comment #6: The method section was well-written and does not require modification. I only have one query. Are the images attached to it from the researcher's work or from other sources? I hope author(s) consider property rights if these images come from other sources.

Answer to comment #6: Thank you very much for this comment. In the final version the editorial office will include copyright notice for each image. The images are from the authors’ institutional electronic database. The institutional approval has been submitted to editorial office.

Comment #7: It is critical to review the progress of the results in the current study to avoid any errors that reduce the quality of the research results.

Answer to comment #7: The results were checked. We unified the number of significant digits in the variances in the tables.

Comment #8: The final paragraph of the discussion section should be devoted to highlighting the current study's strengths and weaknesses, as well as explaining the current study's future directions.

Answer to comment #8: We are grateful for this remark, which permitted us to improve the discussion.

According to your suggestion, the last paragraph of the discussion section has been divided into strengths, limitations and future directions. The paragraph is as follows:

“Strengths, Limitations and Future Work. Our algorithm for automatic measurements of patellar indexes permits the evaluation of high volume of data. Taking into account that the proposed algorithm allows measurements on high resolution radiographs, low effort is needed to get patellar indexes, i.e. no manual cropping of area of knee is required as in previous methods. One of the advantage of proposed method is that relatively small amount of manually labelled bones on images is needed to achieve reliable bone segmentation. Although, higher number of manual annotations for training YOLO responsible for knee detection is needed, the labelling of knee regions can be done in relatively short time. This study has some limitations. Firstly, in the current study, the ROIs that were determined on the basis of YOLO detections were resized to size required by the U-Net, i.e., 512x512. This means that the keypoints and lines were determined on images with somewhat smaller resolution than the original radiographs. Secondly, the accuracy of the results depends on: the manual segmentation performed by the researcher during the training phase; the amount and quality of training data, and the architecture of the applied neural network. In current work, the training of neural networks and evaluation of the algorithm was performed on images acquired in a single

institution, i.e. our hospital. Thus, further work is needed to collect radiographs from various hospitals to train networks and assess accuracy of the algorithm on radiographs acquired by different devices. Moreover, the input images may have different levels of intensity and quality. These technical aspects should be emphasized and resolved in medical centers that will implement the radiograph assessment supported by AI. In future work we are planning to combine boundary-aware analysis with landmark-based deep learning measurements. We also plan to extend the U-Net and compare it with recent networks for image segmentation. Additionally images from different hospitals will be used in the research."

Comment #9: The conclusion also has a lot of extravagance and requires a lot of abbreviation, taking into account the answer to the main question that the reader is waiting for. Was the current study successful in solving the research problem?

Answer to comment #9: The rewritten conclusion is as follows:

"The aim of this study was to investigate the reliability of automated patellar height estimation using DL-based bone segmentation and detection on high-resolution images. It showed that reliable automatic patellar height measurements can be achieved on lateral knee radiographs with the accuracy required for the clinical practice. We demonstrated that proximal tibia and patella bones can be segmented precisely (Dice score greater than 95%) by U-Net neural network on knee regions automatically detected by the YOLO network (mean Average Precision mAP greater than 0.96). Determining patellar end-points and the joint line by fitting to points of the proximal tibia joint surface enables calculating the CD and BP indexes with very good reliability. Automated measurements are comparable to measurements performed by orthopedic surgeons (SEM greater than 0.75). Experimental results indicate that our approach can be valuable as a pre-operative and potentially as a postoperative assessment tool for big volume data analysis in medical practice."

Thanks to your comment our work is now concluded better.

Comment #10: Some references need to be updated so that references prior to 2015 are replaced with new ones and what is unnecessary is removed from the current references, as the number of references is 31, which is a large number to use in the current study.

Thank you very much for your comments. We updated references as requested (Phillips et al. 2010, Portner et al. 2011, Kazemi et al. 2011, Kellgren et al. 1957, Caton 2008). Your comments improved references of our manuscript.

Reviewer #2

This manuscript is well and good at innovation and clears the clarity of the reader. It is well structured and well written. The author does a good job of presenting a highly technical and complicated process in an easy-to-understand manner.

Comment #1: Authors need to cross check the reference section by addressing the cited contents in the introduction and related work part.

Answer to comment #1: Thank you for this comment. The references were checked and updated.

Comment #2: The introduction must be an extended version of the abstract. The authors must elaborate on the points highlighted on the abstract and give supportive ideas and references.

Answer to comment #2: This comment permitted us to highlight the research problem. In current manuscript the introduction consists of three paragraphs and has been considerably shortened.

Comment #3: The conclusions in this manuscript are primitive. Rewrite your conclusions.

Answer to comment #3: The rewritten conclusion is as follows:

“The aim of this study was to investigate the reliability of automated patellar height estimation using DL-based bone segmentation and detection on high-resolution images. It showed that reliable automatic patellar height measurements can be achieved on lateral knee radiographs with the accuracy required for the clinical practice. We demonstrated that proximal tibia and patella bones can be segmented precisely (Dice score greater than 95%) by U-Net neural network on knee regions automatically detected by the YOLO network (mean Average Precision mAP greater than 0.96). Determining patellar end-points and the joint line by fitting to points of the proximal tibia joint surface enables calculating the CD and BP indexes with very good reliability. Automated measurements are comparable to measurements performed by orthopedic surgeons (SEM greater than 0.75). Experimental results indicate that our approach can be valuable as a pre-operative and potentially as a postoperative assessment tool for big volume data analysis in medical practice.”

Thanks to your comment our work is now concluded better.

Comment #4: References aren't formatted according to rules. Additional References: The following articles could be useful: • Has the Future Started? The Current Growth of Artificial Intelligence, Machine Learning, and Deep Learning. <https://doi.org/10.52866/ijcsm.2022.01.01.013> • A diagnostic testing for people with appendicitis using machine learning techniques. <https://doi.org/10.1007/s11042-022-11939-8>

Answer to comment #4: We agree with you. The references have been formatted as suggested. We cited the recommended paper (K Aggarwal et al. 2022).