|  |  |  |
| --- | --- | --- |
| ITU Logo | INTERNATIONAL TELECOMMUNICATION UNION**TELECOMMUNICATIONSTANDARDIZATION SECTOR**STUDY PERIOD 2017-2020 | FG-AI4H-I-042 |
| **ITU-T Focus Group on AI for Health** |
| **Original: English** |
| **WG(s):** | Plenary | E-meeting, 7-8 May 2020 |
| **DOCUMENT** |
| **Source:** | TG-Dental Topic Driver |
| **Title:** | TG-Dental: Artificial intelligence for dental image analysis: A guide for authors and reviewers |
| **Purpose:** | Discussion |
| **Contact:** | Falk SchwendickeCharité – Universitätsmedizin BerlinGermany | Tel: +4930450662556Fax: +40304507562556Email: falk.schwendicke@charite.de  |
| **Contact:** | Joachim KroisCharité – Universitätsmedizin BerlinGermany | Tel: +4930450562328Fax: +40304507562932Email: Joachim.krois@charite.de  |

|  |  |
| --- | --- |
| **Abstract:** | The number of studies employing artificial intelligence (AI), specifically machine and deep learning, for dental image analysis is growing fast. The majority of studies shows weaknesses in planning, implementation and reporting, which in turn results in limited robustness and applicability. We propose to discuss and approve a **process to establish a living, non-authoritative guidance for authors and reviewers** on the conception, implementation and reporting of studies on dental image analysis under the roof of the TG Dentistry in the ITU/WHO Focus Group on AI for Health. |

Applications of artificial AI are fast entering medicine, and one field which is specifically prolific is image analysis (also termed “computer vision”). Deep learning, mainly using convolutional neural networks (CNNs), has found to be able to, in theory, outperform medical experts in detecting diseases (higher accuracy) or assessing structures etc. in seconds (higher speed), thereby possibly making diagnostics and treatment planning safer, more personalized and efficient. Given the importance of imagery in dentistry, computer vision in dental image analysis certainly has potential. However, there are also doubts growing as to the robustness and generalizability, transparency and replicability as well as ethics, effectiveness and, overall, applicability of the results of studies in this field.

Specifically, studies employing CNNs in dental image analysis, where CNNs are used to detect structures (teeth, bone) or pathologies (caries, apical lesions), segment images (cut out the area of interest), and classify them (as an enamel caries lesions, or a cyst), show significant weaknesses: Datasets underlying current studies are rather small, with developed AI solutions possibly lacking robustness and stability. The data generation process is oftentimes unclear and not necessarily fitting the question at hand, with both the data sources and the data characteristics not being fully sufficiently representative or reported, but also the annotation strategy (supervised, unsupervised, number and characteristics of annotators, instance or pixelwise annotation, independent or joint annotation, definition of reference test from annotations) being oftentimes not fully clear or suited. The choice of model, the training and hyperparameter tuning as well as the validation strategy is often unclear, and the metrics chosen to optimize the model against or usually not necessarily clinically relevant. Moreover, the accuracy data reported in many studies is generated in-sample, not on hold-out test datasets or, even better, completely separate independent datasets. It hence often remains unclear if such accuracies can also be reached in other, real-life data, or in prospective sampled groups. Last, it is often not clear if the developed “narrow” application is helpful in clinical practice, and which wider impact it has on health, but also further aspects like costs or ease of treatment provision etc.

In summary, studies on AI in dental image analysis have not been found to adhere to established principles of evidence-based practice, and seldom reach beyond reporting limitedly robust accuracy metrics yielded from a dataset which may be biased. There seems great need to raise awareness of scientists planning and conducting AI studies in dental image analysis, but also for authors, reviewers and editors of journals on which aspects to scrutinize when assessing their own or other researchers’ work. Without rapidly evolving from an experimental field into a solid, scientifically grounded and matured discipline, AI in dental research may disappoint or, worse, lead to harmful decisions in clinical practice.

We plan to establish, discuss and approve a guidance document on how to conceive, conduct and report studies on AI in dental image analysis. Lending from existing reviews in other fields and founded on the principles of evidence-based research practice, a set of guidance items will be defined, assisting future scientists, reviewers and editors in planning, conducting, reporting and evaluating studies on AI in dental image analysis. The items will be derived by topic drivers and members of the TG Dentistry and will be circulated and discussed among the whole focus group. The resulting document will not claim to be authoritative and final, but a living document. However, we strongly believe not taking action and waiting until more definite guidance is available risky; further resources will be wasted in the meantime, futile research findings will be spread and potentially harmful applications of AI in dental medicine translated into clinical practice.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_