FOURTH - IR
APPLIED INTELLIGENCE

Improving caries detection through Artificial Intelligence

4th-IR collaborated with a Dental Imaging Company to explore how Artificial Intelligence (AI) could be used to identify high-probability caries regions quickly and accurately.

Company Info:

Name: Dental Imaging Company

Location: USA

Industry: Healthcare

Solution

Dental caries is one of the most widespread chronic condition worldwide. Early detection can significantly improve treatment outcomes and reduce the need for invasive procedures. 4th-IR collaborated with a Dental Imaging Company to create a tool that quickly and accurately identifies high-probability caries regions within a bitewing x-ray image. The tool developed by 4th-IR has achieved an impressive level of accuracy, opening up new possibilities in Dental Care.

The 4th-IR team designed a caries AI model that was trained on the set of annotations provided by the Dental Imaging Company and annotated by a team of Oral Maxillofacial Radiologists (OMRs). After multiple rounds of training and testing, the result was that 91% of the predicted boxes from the AI were agreed to be correct by the annotators, indicating that the AI model achieved similar performance as the expert panel.

Methodology

4th-IR measured the internal consistency of the annotators and compared this to the AI in a “golden source” data set. It was found that the AI performed as well as, and often better than, the annotator team in replicating the golden data set. While these initial results are impressive for assessing performance, the greater question is: does the AI perform well enough that a panel of experts cannot tell whether the results are from a human or from the AI.

Data Set

In order to measure performance, the 4th-IR team was provided with thousands of images with a combination of predicted boxes and original annotations that were sent back to the annotators for correction. The data were chosen randomly from a set where the predictions or annotations had at least one inter-proximal caries bounding box.

The validation data set is the best measure of the algorithm's performance. 4th-IR used the annotations data set as a baseline to see if predictions were modified more or less than returned annotations. Since the annotators were not made aware if the boxes were from the AI predictions or the original annotations, this was considered as a fair metric.

Consistency Measurement

The annotators were shown the images with bounding boxes in the annotation platform. They were able to resize, move, add, delete, and change caries type for each box. Similarly to the metric in the golden data set analysis, it was measured if each box in set B overlapped with boxes in set A. If there was an overlap with the correct type, it was considered to be a match. This was done for both reviewer vs sample and sample vs reviewer, where the sample was the "prediction" (or original annotation) and the reviewer set was the "corrections" that the annotators provided upon re-review. While overlapping caries was considered to be a match if they also matched on type, since the inter-proximal caries accounted for around 90% of the data set, this doesn't greatly affect the overall numbers. However, they would likely be even higher if the criteria were type-agnostic.

Analysis and Recommendations

Over 90% of the boxes the AI predicted were considered correct by the reviewers. This is the same rate as for the boxes which were supplied from the annotators, thus the AI achieved expert-level performance. This is related to the false - positive rate for the AI and the annotators. If we had a higher false positive rate, this number would be much lower. Similarly, we see that the reviewers tend to add boxes regardless of whether or not the AI or experts provided the boxes. This shows a bias in the reviewer mentality, which was to be expected. People tend to over-analyze when reviewing, especially when an algorithm is involved. Since the reviewer matching in the AI validation set is actually higher than for the reviewer set, it is showing that the reviewers felt the need to add fewer new caries to the AI predictions than for those from the experts. The absolute value of this number isn't critical, however we see again that the AI and annotators perform at similar levels. It's interesting to note that the reviewer matching percentage is similar to the golden data set consistency.

4th-IR continues to work with the Dental Imaging Company to enhance the performance of the caries detection AI model and pursue new collaborations in computer vision and dental imaging.

FOURTH - IR
APPLIED INTELLIGENCE

Luzern, Switzerland
New York, USA
Vancouver, Canada

USA +1 (630) 345-0912
EU +49 172-8295-895
Email: info @4th-ir.com

www.4th-ir.com