Developing Interactive Systems for Enhanced Communication and Feedback Between Medical Experts and Explainable AI Models

Jeroen Geerdink, Innovation Manager, Health & Information Technology, Hospital Group Twente
Xun Zhao, MS; Oswald Kessels, MD; Esmée Dijkstra, MD, PhD; Lars Bosboom, MD; Jeroen Veltman MD, PhD; Johannes Hegeman MD, PhD; Maurice van Keulen, PhD; Shenghui Wang, PhD

Background/Problem Being Solved

Despite the promise of explainable AI models like PIP-Net in medical imaging, no user-facing interface has been developed to bridge these models with real-world clinical workflows. While PIP-Net offers intrinsic interpretability by classifying images based on key visual concepts (prototypical parts) and addressing undesirable behaviors like biases and shortcut learning, its practical effectiveness and alignment with medical workflows remain unexplored. To fully realize the potential of such models in clinical decision-making, an interactive interface is needed to facilitate user engagement, feedback, and integration into real-world medical practices.

Intervention(s)

The application framework contains three main components: (1) integration of PIP-Net and YOLO models for hip fracture inference, enabling users to compare outputs from interpretable prototype networks and segmentation-based models; (2) feedback tools allowing users to validate or reject model explanations and provide corrective annotations to refine the model's performances; and (3) an intuitive interface equipped with a message center for guidance and explanations, as well as a "find similar images" feature to contextualize model outputs by referencing relevant training data. These components aim to align AI functionality with clinical workflows, enhancing usability, trust, and model accuracy.

Barriers/Challenges

One of the key challenges is guiding users on how to make accurate annotations, as different users may have varying standards. Some annotations may be too large and imprecise for further model training, while others may be too small, leading to a lower perceived performance of the model compared to human evaluation.

Outcome

The outcome of the study showed that most users were satisfied with the interface, particularly with the message center for the guidance and the "find similar images" feature for enhancing context. Additionally, all users found the YOLO model to perform better in both classification accuracy and explainability compared to other models.

Conclusion/Statement of Impact/Lessons Learned

In conclusion, the integration of PIP-Net and YOLO models into a user-friendly interface successfully facilitated user interaction and feedback, enhancing the model's alignment with clinical needs. The positive response to the interface's features, such as the message center and "find similar images" tool, demonstrates its potential to support clinicians in real-world applications. The preference for the YOLO model highlights its effectiveness in providing accurate classifications and transparent explanations, underscoring the importance of model performance and interpretability in medical imaging tasks. Further refinement of the annotation process and interface will improve usability and model accuracy.

Figure(s)

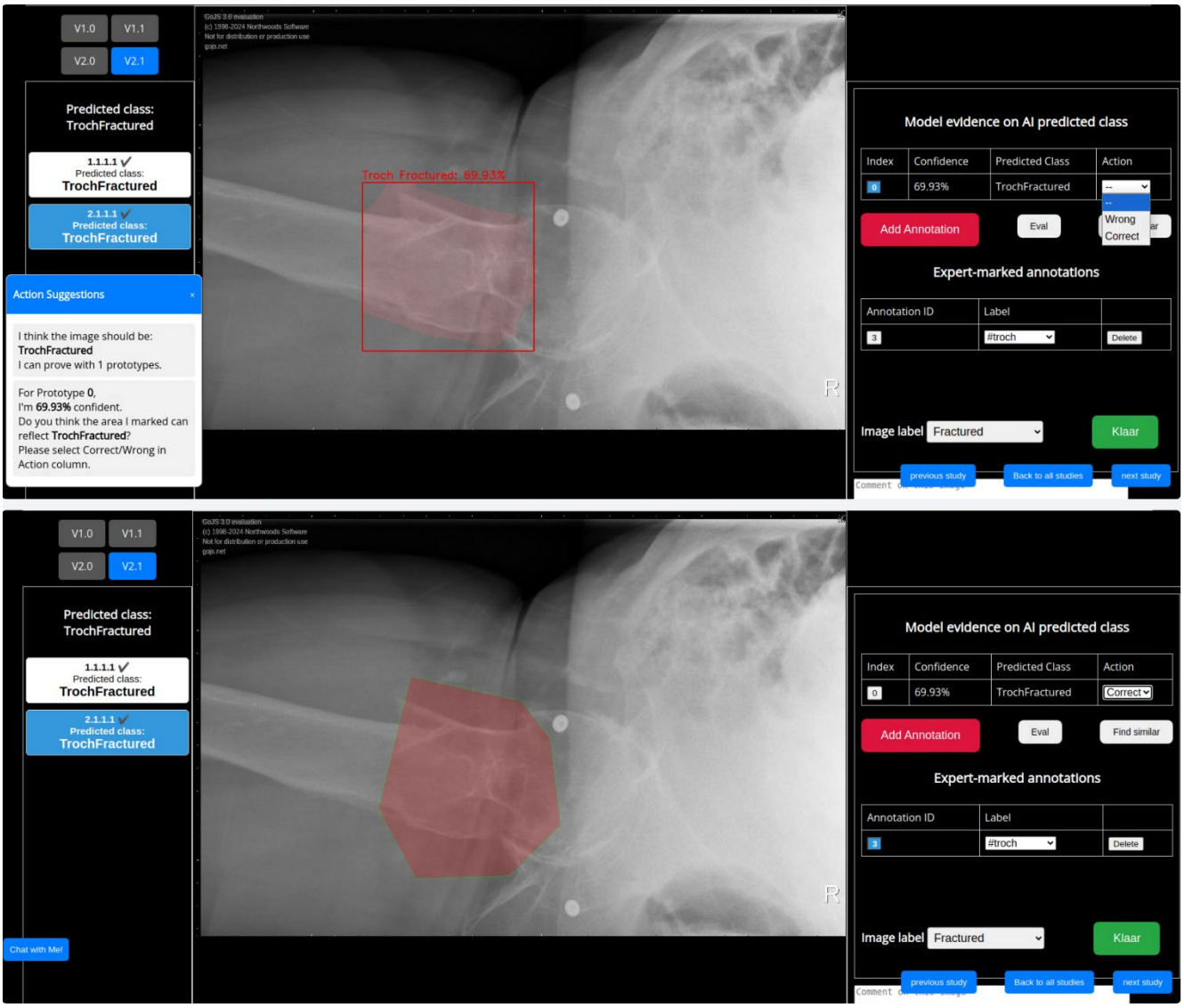

Figure 1. Single Study View for interpreting a image and providing feedbacks

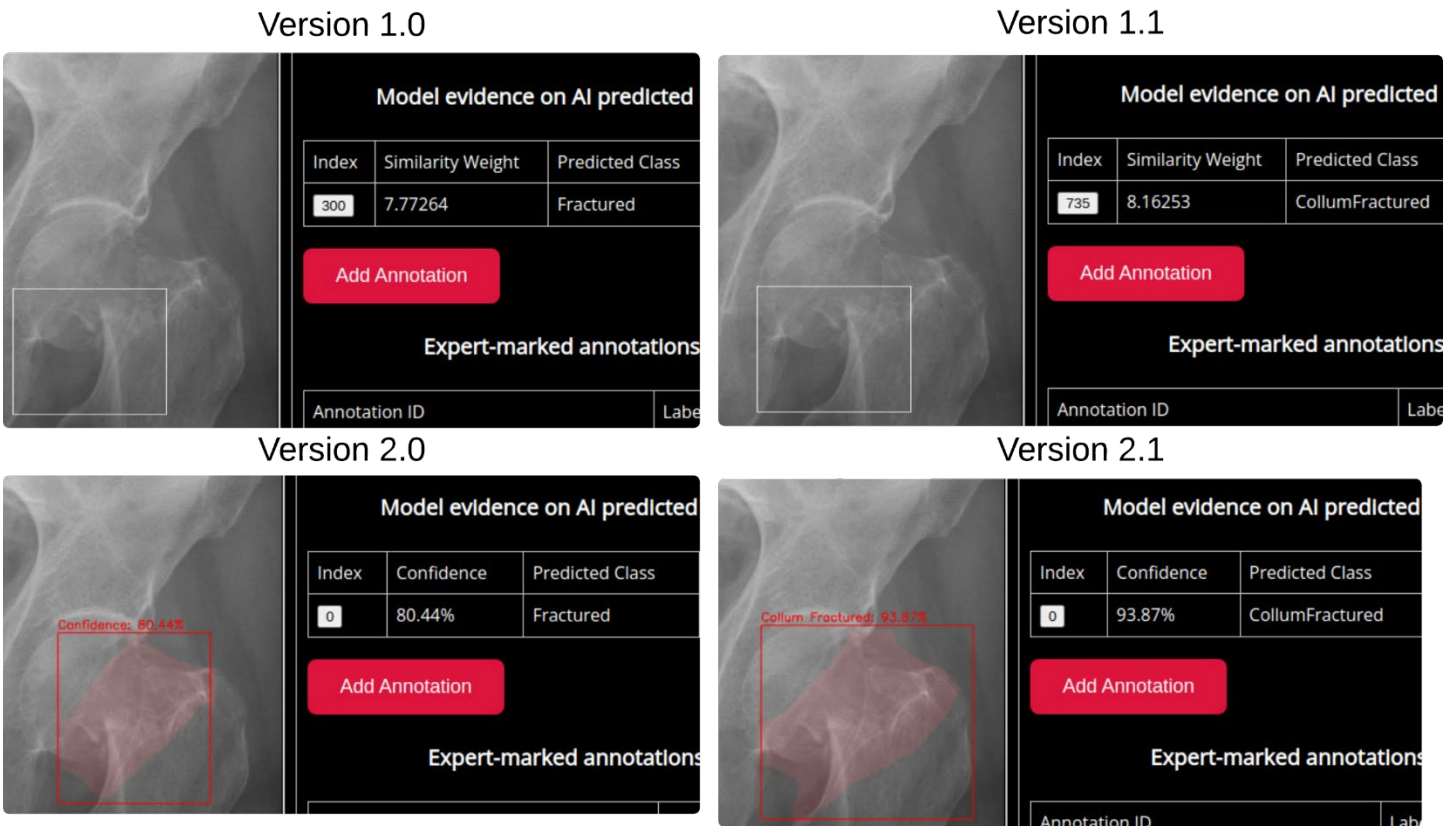

Figure 2. Four model's performance in classification and explanation: (1) Version 1.0: PIP-Net binary-class model (2) Version 1.1: PIP-Net three-class model (3) YOLO binary-class model (4) YOLO three-class model

Keywords

Applications; Clinical Workflow & Productivity; Imaging Research