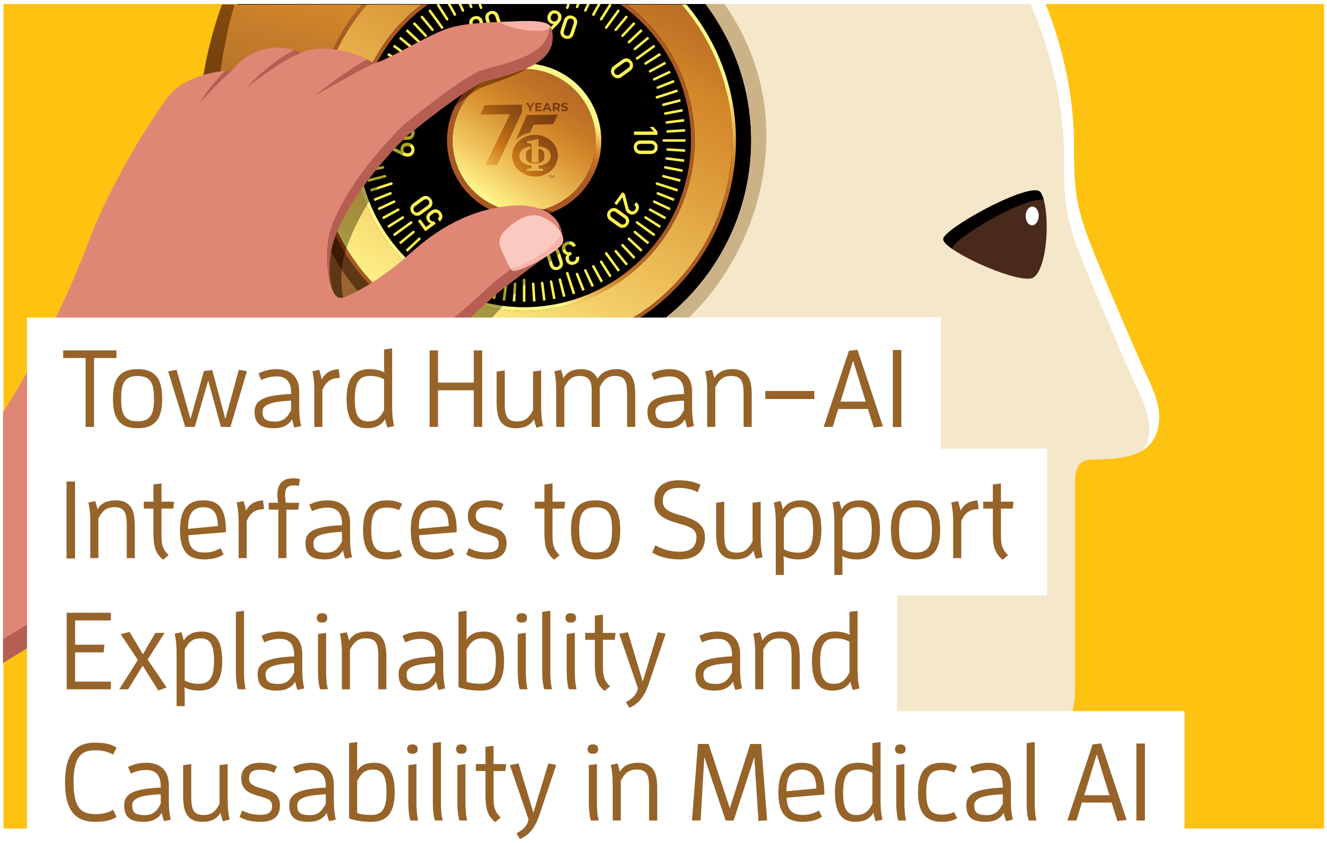

Toward Human–AI Interfaces to Support Explainability and Causability in Medical AI

Andreas Holzinger and Heimo Müller, Medical University Graz

Our concept of causability is a measure of whether and to what extent humans can understand a given machine explanation. We motivate causability with a clinical case from cancer research. We argue for using causability in medical artificial intelligence (AI) to develop and evaluate future human–AI interfaces.

Achieving human-level artificial intelligence (AI) has been the ambition since the emergence of this field. Because of the availability of big data and the necessary computing power, statistical machine learning, especially deep learning, has now made tremendous progress, even in domains as complex as medicine. For example, the work of the Stanford machine learning group on dermatology[1] was popularized in Europe as "AI is better than doctors." The group trained a deep learning model directly from dermatological images by using only pixels and disease labels as the inputs for the classification of skin lesions. For pretraining, they used 1.3 million images from the 2014 ImageNet challenge and then 130,000 clinical images with about 2,000 different diseases. The results, with an average classification performance of 92%, were on par with, or even better than, those of human dermatologists. This is a remarkable achievement, and there is no question that AI will become very important for medicine.

Despite these impressive successes, we must be aware that these previous approaches rely on statistical model-free learning. Relying solely on statistical correlations can be very dangerous, especially in medicine, because correlation must not be confused with causality, which is completely missing in current AI. This is a general

Digital Object Identifier 10.1109/MC.2021.3092610
Date of current version: 24 September 2021

problem. A specific issue is that these methods are so complex, high dimensional, nonlinear, and nonconvex that it is difficult, if not impossible, even for domain experts, to understand how the results were obtained. Consequently, these techniques are called *black-box* models. Both problems motivate us to make use of the expertise of a human in the loop. A human expert can sometimes—but not always—contribute experience, conceptual knowledge, and context understanding.

A very recent work on histopathology[2] showed that machine-generated features correlate with certain morphological constructs and ontological relationships generated by humans. Such overlaps between humans and AI can help to solve the current problems, and human experience can contribute to improve algorithm *robustness* and *explainability*, which are regarded as the grand challenges of current AI.[3] Robustness is closely related to the ability to generalize, and it is no coincidence that we mention robustness and interpretability together: robustness is a ubiquitous feature of biological systems, as we humans are. It ensures that specific functions of the system are maintained—despite external and/or internal perturbations. We all know that our best machine learning algorithms are very fragile and sensitive to even small distortions, arising from the poor data quality in the medical domain.

Consequently, a doctor in the loop[4] will play an important role in medical AI, at least in the foreseeable future. Humans are generally robust, and they can sometimes add expertise, conceptual understanding, and implicit knowledge to machine learning processes. Although humans make mistakes, they can adapt quickly, improvise, and exhibit a plasticity that leads to sensemaking and understanding in the context of an application domain. On the other hand, AI is vulnerable to even small perturbations. We emphasize that the current approaches not only lack robustness and generalization, but more importantly, they are unable to build causal models to support deep understanding in the user.[5] To make the current AI even more successful, we believe that we should try to take advantage of the respective benefits of both statistical machine learning methods and model-based approaches. More precisely, we envision for the future to interactively integrate the existing implicit a priori knowledge, human experience, and conceptual understanding of human experts into statistical learning methods.

This could result in a new hybrid approach that fully exploits the advantages of data-driven machine learning methods and also integrates human conceptual understanding, according to individual needs and depending on the problem. Such an approach will require appropriate human–AI interfaces that enable seamless interaction with machine learning methods. However, let us first summarize the achievements of the explainable AI (XAI) community.

EXPLAINABILITY

The XAI community is very active in developing methods for making black-box approaches explainable.[6] The successful approaches focus on visualizing the elements that have contributed to each decision, for example, through heatmapping, which means highlighting which input parameters contribute most to a certain classification result.[7]

Such "mechanical explanations" can be reached by using various procedures. The simplest method works with gradients as a multivariable generalization of the derivative, where the deep neural network is seen as a function and the explanation relies on the function's gradient, which is available from the backpropagation algorithm.[8]

Another possibility is to use decomposition methods, that is, to break up the more complex larger parts into smaller, more manageable, parts, for example, pixel-wise relevance propagation, layer-wise relevance propagation, or deep Taylor decomposition. These very versatile approaches also work on graph-based data.[9] Other methods include deconvolution, which involves reversing the effects of convolution and generating from two functions a third function that is then the product of both, as well as guided backpropagation.[10]

All of these methods constitute an excellent preprocessing step. Here we want to emphasize two issues: 1) results are human interpretable when they classify objects on the basis of features that a human can perceive and understand and 2) in the medical domain, there is a pressing need for a medical expert to be able to understand the causality of a learned representation and use it for medical decision support. This is called *etiology* in medicine: the science of causes and effects of pathologies.

EXPLAINABILITY AND CAUSABILITY

The terms *interpretation* and *explanation* are often used synonymously. Before we introduce the concept of causability, we distinguish between *interpretation*, which can be defined as a mapping of an abstract statement into a domain or space that the human expert can perceive, comprehend, and understand, and *explanation*, which can be defined as a collection of the features of the interpretable domain or space that have contributed to a given example to produce a statement. In an ideal world, both human and AI statements would be identical and congruent with the *ground truth,* which is defined for AI and humans equally.

However, in the real world we face two problems. 1) The ground truth cannot always be well defined, especially when making a medical diagnosis. 2) Human (scientific) models are often based on causality as an ultimate aim for understanding the underlying explanatory mechanisms, and while correlation is accepted as a basis for decisions, it must be viewed as an intermediate step. This is highly relevant in the medical domain because of the importance of validity and the necessity to build human trust and also the need to build "AI experience."[11] As we have mentioned, the most successful algorithms are based on probabilistic models and provide only a rudimentary basis for establishing causal models. Consequently, when we discuss the explainability of a machine statement, we have to carefully distinguish among the following terms:

1. *Explainability*: In a technical sense, explainability highlights decision-relevant parts of the used machine representations of the algorithms and active parts in the algorithmic model that contribute either to the model accuracy on the training set or to a specific prediction for one particular observation. It does not refer to an explicit human model.
2. *Usability*: This term refers to the measurable extent to which a system achieves a specified level of usability for a user with effectiveness, efficiency, and satisfaction in a specified context of use.
3. *Causability*: Causability is the measurable extent to which an explanation of a statement to a human expert achieves a specified level of causal understanding with effectiveness, efficiency, and satisfaction in a specified context of use. As causability is measured in terms of effectiveness, efficiency, and (human) satisfaction related to causal understanding and its transparency for an expert user, it refers to a human-understandable model.

This is always possible for an explanation of a human statement, as the explanation is, per se, defined related to a human model. However, to measure the causability of an explanation of a machine statement, it must be based on a causal model in the sense of Pearl,[12] that is, to represent causal relationships and to allow inferences about those causal relationships from data. This is not the case for most AI algorithms, so a *mapping* between both must be defined. Here we must distinguish between the explainable model (XAI) and an *explanation interface*, which makes the results gained in the explainable model not only usable but also useful to the expert. When is it useful? It is useful when the system is able to provide *causes of observed phenomena*[13] in a comprehensible and interactive manner. This is done through, for example, a linguistic description or question/answer dialogs of its logical and causal relationships, enabling one to understand the causality of the learned representations relevant not only for sensemaking, but also for judging the quality of explanations. This might be unnecessary for certain areas where we want full automation, but for the medical domain, especially when trying to achieve an explainable medicine, this is indispensable.[14]

HUMAN–AI INTERFACES: EFFECTIVE MAPPING OF EXPLAINABILITY WITH CAUSABILITY

The key to effective human–AI interaction and, consequently, the success of future human–AI interfaces lies in an efficient and consistent *mapping of explainability with causability*. This "mapping" (or "map metaphor") is about establishing connections and relationships between existing areas, not about drawing a new map. Rather, it is about identifying the same, or at least similar, areas in two completely different "maps" of AI explainability and human causability. That is why *mapping* is a very good term. Effective and efficient mapping is necessary, but of course it still will not be sufficient for understanding an explanation. Whether and to what extent an explanation has been understood depends on additional factors, including prior knowledge and expectations on the human side. To fully understand the importance of this mapping, let us look at Figure 1.

In Figure 1 we see that an explanation statement s can either be made by a human s_h or a machine s_m, where s is a function $s = f(r, k, c)$ with the following parameters: r, representations of an unknown (or unobserved) fact u_e related to an entity; k, preexisting knowledge, which is embedded in an algorithm for a machine or made up for a human by explicit, implicit, and/or tacit knowledge; and c, context, which for a machine is the technical runtime environment and for a human is the physical environment in which the decision was made (the pragmatic dimension). An unknown (or unobserved) fact u_e represents a ground truth gt that we try to model with a machine m_m or as a human m_h. Such unobserved variables can be found, for example, in Bayesian models or in hidden Markov models (a special case of a dynamical Bayesian network).

In summary, we can state that causability as a measure of whether and to what extent something is understood consists of two parts. The first part is whether the human understands or can understand the given explanation at all, and the second part is to what extent he/she can understand it, that is, it is expressed in terms of measuring the causability of the explanation e_m of a machine statement s_m under a certain machine model m_m (see Figure 1).

The central goal is that the statement s is identical with the ground truth gt and that a given explanation of this statement is a fit to this ground truth. In an idealistic situation, both the human and the machine statement are *congruent* ($m_h \equiv m_m$) and identical to the ground truth, which is defined for machines and humans within the same framework. The problem in the medical domain is that the ground truth is not well defined and the most successful machine learning models are based on correlation or related concepts of similarity and distance. All of this is probabilistic in nature and therefore must be considered as an intermediate step that can only provide a basis for further causal model building. Explainability in a technical sense highlights the decision-relevant parts of the machine representations r_m and machine models m_m, that is, the parts that contributed to model accuracy in training or to a specific prediction. It is important to emphasize that explainability does not refer to a human model m_h.

Causability is the extent to which an explanation of a statement to a user achieves a specified level of causal understanding with effectiveness, efficiency, and satisfaction in a specified context of use. As causability is measured in terms of effectiveness, efficiency, satisfaction related to causal understanding, and its transparency for a user, it refers to a human-understandable model m_h. This is always possible for an explanation of a human statement as the explanation is intrinsically defined related to m_h. To conclude, to measure the causability of an explanation e_m of a machine statement s_m, either m_h has to be based on a causal model (which is not the case for most machine learning algorithms) or a mapping between m_m and m_h must be defined.

As we can imagine, this is not trivial because future human–AI interfaces

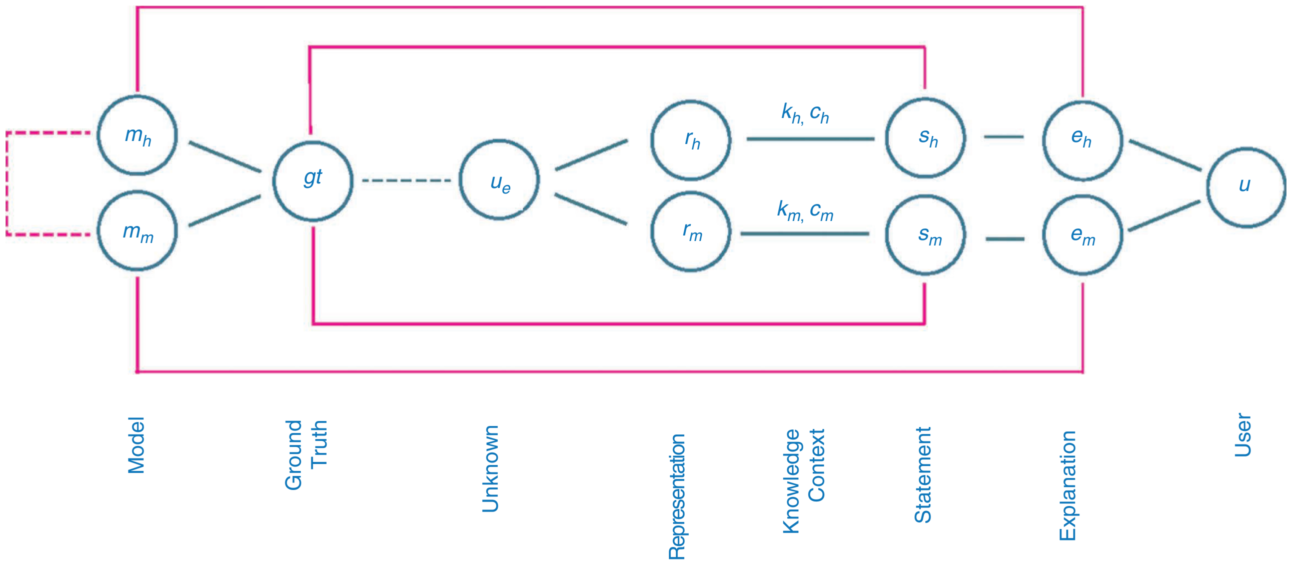

FIGURE 1. The process of explanation. Explanations (e) by humans and machines (subscripts h and m) must be congruent with statements (s) and models (m), which, in turn, are based on the ground truth (gt). Statements are a function of representations (r), knowledge (k), and context (c). (Source: Holzinger et al. 2020,[15] used with permission.)

should allow a constant feedback, indicating whether and to what extent something has been understood, or not. In a human-to-human interaction, this feedback is very much provided by facial expressions. Consequently, concepts of "emotional interfaces"[16] will become an important part of future conversational interfaces for XAI, and dialog systems will become important for these future "explanation interfaces" where affective computing will play an important role. Generally, affective computing[17] describes a system in which the machine interacts with a human by observing his/her actions and varying affordances accordingly. Here it is important to classify, filter, and make use of affective computing methods to

1. measure the effectiveness of explainability (does the user really understand a given explanation), which can be measured via sensors by the success rate to which a given explanation has been understood by humans; see the examples in the section "A Clinical Case"
2. adapt the visual communication according to the mental model (a priori knowledge) of the user.

Using advanced biometric technology including 2D/3D cameras and eye trackers, heart sensors (for electromyography, electrocardiography, photoplethysmography, and galvanic skin response), acoustic microphone arrays, and other sensors (potentially machine vision), an affective computing environment may enhance the effectiveness of navigation. In particular, when an expert user is actually reading text or looking at an image, the system will "know" how much time he/she is spending on what. The longer a user looks at a particular word, sentence, paragraph, or data feature, the more important the system will assume it to be for that user. A very important aspect lies in fusing the information from different sensor signals and also allowing the metrics of importance to be changed in the background, thereby automatically "bringing up" features of the greatest similarity or utility—augmented by a potential explanation of how and why. The user will similarly be able to weight the importance of the affective element via sophisticated interaction elements (face recognition, gestures, and so on). Other tools for navigating the data will include a virtual "important features" pile (possibly on the right of the screen) that can be selected by gaze or mouse click, which can interactively be overridden to show any features of interest. Alternatively, the user data will be "captured" without any user interaction; that is, through a set of rich sensors, the "machine" will get reliable information about the understanding.

There are two main types of sensor input data: 1) facial expression analysis by multiple cameras and 3D sensors (compare the recent developments in augmented reality toolkits) and 2) gaze patterns—which can be experimentally contrasted and fused with other sensor data—and they can be very useful. These data can be used to exploit affective computing methods to massively improve the affordances generated by a system,[18] using the fact that there are computational models that are associated with human behavior.

In Figure 2, we outline a model for the information flow between humans and an AI system. On the interaction surface, which can be seen as a "border" between human intelligence and AI, the information flow is maximal. As one gradually goes "deeper" into the AI system, the information flow decreases; at the same time, the semantic richness (SR) of potential information objects increases. In traditional human–computer interactions, the information flow is extremely asymmetrical; that is, much more information is shown by high-resolution displays compared to mouse and/or textual input—not to mention other input modalities (see the dotted line in Figure 2).

A CLINICAL CASE

A clinical case of lung cancer will serve as a practical example. The survival of patients with lung cancer can be improved by the use of immunotherapies directed against programmed cell death 1-ligand 1 (PD-L1) and its receptor PD-1. PD-L1 is a surface protein involved in the inhibition of the immune response. PD-L1 protein expression has emerged as an effective biomarker that can predict which patients are more likely to respond to immunotherapy.[19] Whether a tumor can spread in the body depends on whether one's immune system recognizes the degenerate cells leading to the tumor as a potential danger and attacks them. The immune system scans all cells to distinguish the body's own cells from foreign ones. T cells are an important component of the immune system in this process. Cancer cells can be recognized by T cells, but they are not attacked by the immune system because they are able to camouflage themselves by producing the protein PD-L1. PD-L1 is like a disguise that helps cancer cells to conceal themselves and remain undetected, and one way to detect PD-L1 is by using immunohistochemistry. In Figure 3 we see a typical part of a whole slide image that is used in an ongoing validation study (see the ethics declaration in the

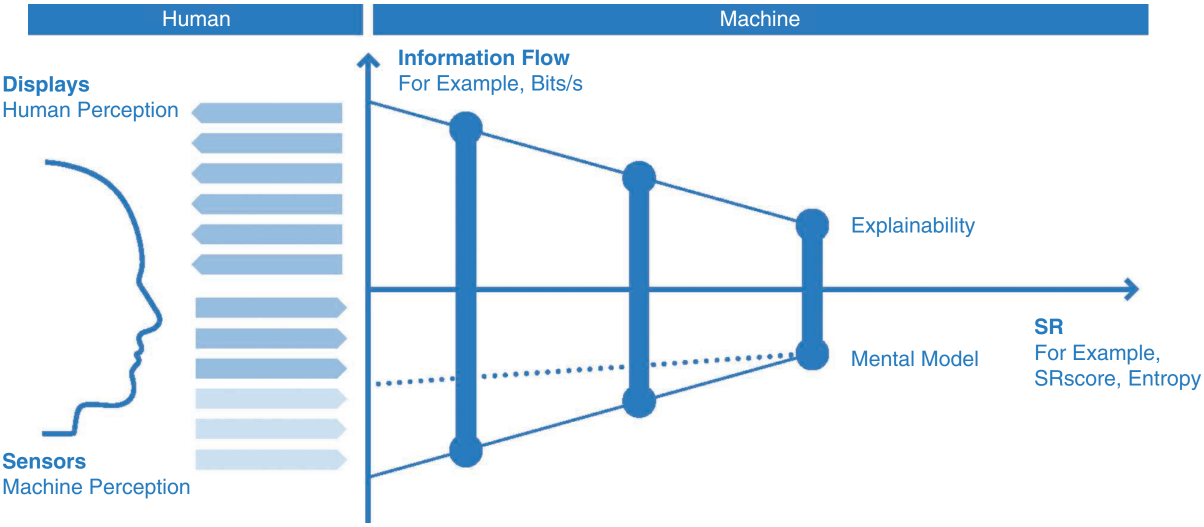

FIGURE 2. On the upper part of vertical line, explanations are visualized and displayed to the user. Parallel to the visualization, sensor data are captured and analyzed (see the lower part of vertical line) with less information density. Only when a rich set of sensor inputs, for example, cameras, microphones, and movement sensors, is used can the explanation process adapt to the prior knowledge of the user. Semantic richness (SR) can be measured via scores and entropy measures and used for appropriate feedback.

"Acknowledgments"), focusing both on the clinical performance and also on the human–AI interface.

To this end, the pathologist surveys the tissue by encircling PD-L1 stained immune cell (IC) scored aggregates (the IC score indicates the percentage of the area of PD-L1 positive immune cells from the area of vital tumor cells). The pathologist then places them in relation to the tumor area and can estimate the percentage relevant for making a diagnosis.

In Figures 4 and 5 we can see the process assisted by an AI algorithm. With the results in Figure 4, the pathologist can see from the corner of his/her eye that the explanation of the algorithm is correct; thus, there is congruence, as previously explained in the model, and the pathologist will be satisfied and in agreement.

However, in Figure 5, we see an image where the AI solution found the wrong

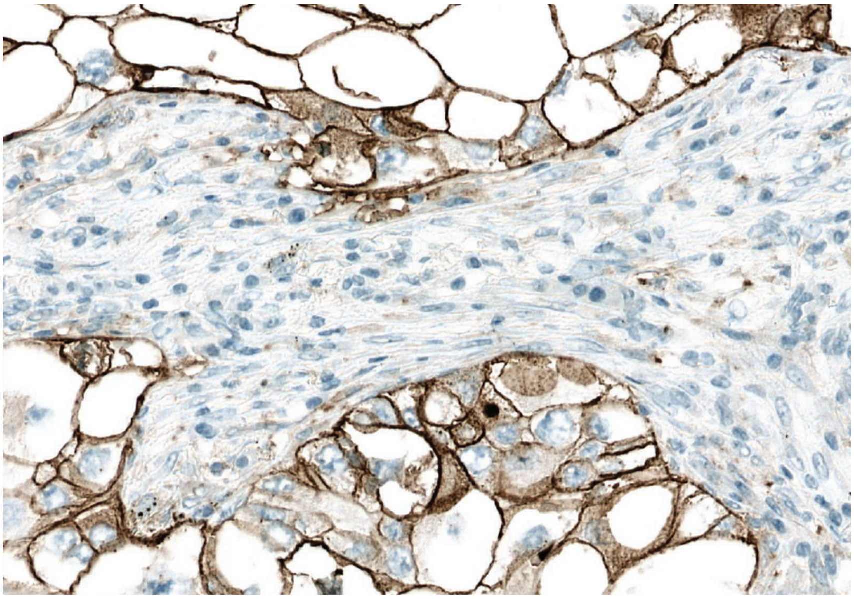

FIGURE 3. The human pathologist sees the tumor above and below in the image. In the middle the stroma can be seen (the tumor identification is 100% positive).

tumor areas. In this case, the emotion on the pathologist's face will be immediately visible, and the pupil widths will increase, which can be easily detected by the sensors of modern eye-tracking systems. These indicators, which can be measured very precisely by sensors and are therefore also comparable, will show that there is no congruence between the explanation provided by the machine and the explanation of the human pathologist.

Statistical machine learning is extremely successful and has made AI very popular again, even in the complex application area of medicine. Whether AI will be used or not in this field is not up for debate; in

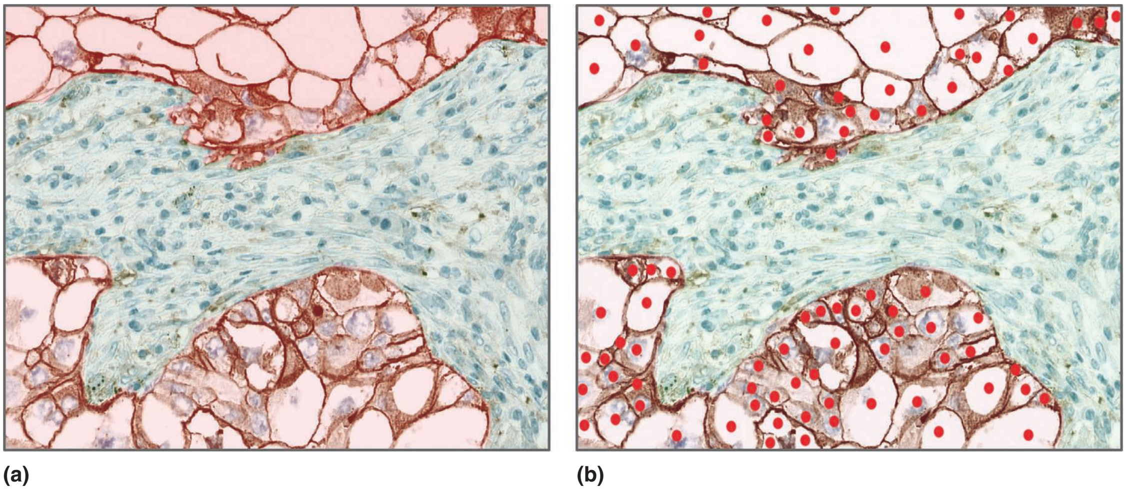

FIGURE 4. (a) A visualization of the benign (nontumorous area) in the middle and the malignant (tumorous areas) above and below. (b) The detected positive cells are marked. If a pathologist evaluates this, he/she intuitively sees with "one look" that the explanation of the AI solution is correct and sufficient. In this case, there is *congruence* between machine and human ($m_h \equiv m_m$).

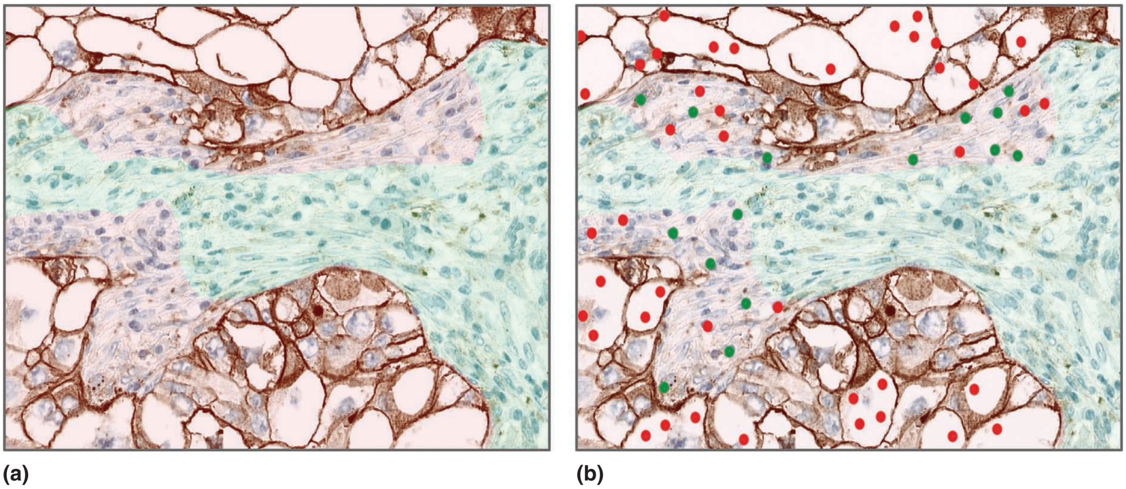

FIGURE 5. (a) The same visualization method is applied, but now (b) the AI solution found the wrong tumor areas (in the middle); also the cell detection, for both positive and negative cells, did not work. In this case we have *incongruence* between machine and human. ($m_h \not\equiv m_m$).

the future, the use of medical AI will be practically indispensable.

However, the most successful AI methods are so-called *black boxes*, which means that they are so complicated that it is difficult or even impossible for a human expert to understand how a result was obtained. Because of increasing legal requirements,[20] in the future it will be imperative for results to be made retraceable, comprehensible, and understandable to a human expert. The fast-growing XAI research community has already developed a number of very successful methods of explainability. Explainability in the technical sense of the XAI community means highlighting decision-relevant parts of a result. However, another important aspect is that these explanations do not refer to an explicit human model. This motivated us to introduce our concept of causability. Causability is intended as a measure of whether and to what extent something is understood, and it basically consists of two parts. The first part is whether a human understands or can understand a given explanation at all, and the second part is to what extent, that is, in terms of measuring, it is understandable by a human. Consequently, causability will become important for the design, development, testing, and evaluation of future human–AI interfaces. Such interfaces are needed not only for enabling a human to understand an explanation, but also to capacitate an interaction of the human with the AI. This is needed because a human in the loop can (sometimes) provide conceptual knowledge, experience, and context understanding—which to date no AI can do.

In our newly established causability laboratory, we currently study how pathologists make judgments and decisions. Central questions include the following: How do physicians make causal judgments? What role, if any, do counterfactuals play in this process? From theory, we know that counterfactual theories of causal judgments predict that people compare what actually happened with what would have happened if the possible cause had not been present. Common theories also state that people focus only on what actually happened to judge the mechanism linking the cause and the outcome. To test this, for example, we propose in the future not only to record the expert's eye movements, but also to compare them with other circumstantial evidence relevant to the decision. Here it is important to analyze in real time, and we plan to do this according to our causability model, using various methods, including analysis of eye movements, facial expressions, and micromovements such as head nods. In this future work, we plan to combine and analyze these metrics. The results obtained will be useful for the development of novel human–AI interfaces that will benefit medical experts and also lead to further contributions to the international XAI research community. C

ACKNOWLEDGMENTS

We are grateful for the very useful comments of the anonymous reviewers and the valuable substantial comments of the editor. We declare that real-world data were used for this work, but only the principles of the human–AI interface were investigated and used for the illustrative example in the section "A Clinical Case." We are grateful for comments from Luca Brcic and Markus Plass for the positive ethic vote of the Roche Ventana Study. Parts of this work have received funding from the European Union's Horizon 2020 research and innovation program under grant agreement 857122 (CY-Biobank), 824087 (European Open Science Cloud-Life), 874662 (Human Exposome Assessment Platform), and 826078 (Feature Cloud). Parts of this work have received funding from the Austrian Research Promotion Agency under grant agreement 879881 (Ecosystem for Pathology Diagnostics With AI Assistance) and by the Austrian Science Fund, project P-32554, Explainable Artificial Intelligence.

REFERENCES

1. A. Esteva et al., "Dermatologist-level classification of skin cancer with deep neural networks," *Nature*, vol. 542, no. 7639, pp. 115–118, 2017. doi: 10.1038/nature21056.
2. K. Faust et al., "Intelligent feature engineering and ontological mapping of brain tumour histomorphologies by deep learning," *Nature Mach. Intell.*, vol. 1, no. 7, pp. 316–321, 2019. doi: 10.1038/s42256-019-0068-6.
3. R. Hamon, H. Junklewitz, and I. Sanche, *Robustness and Explainability of Artificial Intelligence—From Technical to Policy Solutions*. Luxembourg: Publications Office of the European Union, 2020.
4. A. Holzinger, "Interactive machine learning for health informatics: When do we need the human-in-the-loop?" *Brain Inf.*, vol. 3, no. 2, pp. 119–131, 2016. doi: 10.1007/s40708-016-0042-6.
5. B. M. Lake, T. D. Ullman, J. B. Tenenbaum, and S. J. Gershman,

DISCLAIMER

This publication reflects only the authors' views. The European Commission is not responsible for any use that may be made of the information it contains.

ABOUT THE AUTHORS

ANDREAS HOLZINGER is the head of the Human-Centered Artificial Intelligence (AI) Laboratory at the Institute of Medical Informatics and Statistics at the Medical University Graz, Graz, 8010, Austria, and a visiting professor of explainable AI at the Alberta Machine Intelligence Institute of the University of Alberta, Canada. His research interests include human-centered AI to put the human-in-control of AI and align AI with human values, ensuring privacy, security and safety. Holzinger received a Ph.D. in cognitive science from Graz University and a second Ph.D. in computer science from the Graz University of Technology. He is a Senior Member of IEEE. Contact him at andreas.holzinger@medunigraz.at.

HEIMO MÜLLER is head of the Information Science and Machine Learning Laboratory at the Diagnostics and Research Institute of Pathology of the Diagnostics and Research Center for Molecular Biomedicine of the Medical University Graz, Graz, 8010, Austria. His research interests include human–artificial intelligence (AI) interfaces for explainable AI and machine learning, particularly for medical AI in digital pathology. Müller received a Ph.D. in mathematics from the Vienna University of Technology with a work on data semantics spaces. Contact him at heimo.mueller@medunigraz.at.

"Building machines that learn and think like people," *Behav. Brain Sci.*, vol. 40, no. e253, 2017. doi: 10.1017/S0140525X16001837.

6. A. B. Arrieta et al., "Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI," *Inform. Fusion*, vol. 58, pp. 82–115, 2020. doi: 10.1016/j.inffus.2019.12.012.
7. S. Bach, A. Binder, K.-R. Müller, and W. Samek, "Controlling explanatory heatmap resolution and semantics via decomposition depth," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, 2016, pp. 2271–2275. doi: 10.1109/ICIP.2016.7532763.
8. G. Montavon, "Gradient-based vs. propagation-based explanations: An axiomatic comparison," in *Explainable AI: Interpreting, Explaining and Visualizing Deep Learning*, W. Samek, G. Montavon, A. Vedaldi, L. K. Hansen, and K.-R. Müller, Eds. Cham: Springer International Publishing, pp. 253–265.
9. A. Holzinger, B. Malle, A. Saranti, and B. Pfeifer, "Towards multi-modal causability with graph neural networks enabling information fusion for explainable AI," *Inform. Fusion*, vol. 71, no. 7, pp. 28–37, 2021. doi: 10.1016/j.inffus.2021.01.008.
10. M. D. Zeiler and R. Fergus, "Visualizing and understanding convolutional networks," in *Proc. European Conf. Comput. Vision*, 2014, pp. 818–833. doi: 10.1007/978-3-319-10590-1-53.
11. F. Cabitza, A. Campagner, and C. Balsano, "Bridging the 'last mile' gap between AI implementation and operation: 'data awareness' that matters," *Ann. Transl. Med.*, vol. 8, no. 7, p. 501, 2020. doi: 10.21037/atm.2020.03.63.
12. J. Pearl, "The seven tools of causal inference, with reflections on machine learning," *Commun. ACM*, vol. 62, no. 3, pp. 54–60, 2019. doi:[10.1145/3241036.
13. J. Peters, D. Janzing, and B. Schölkopf, *Elements of Causal Inference: Foundations and Learning Algorithms*, Cambridge, MA, 2017.
14. A. Holzinger, G. Langs, H. Denk, K. Zatloukal, and H. Müller, "Causability and explainability of artificial intelligence in medicine," *Wiley Interdiscipl. Rev., Data Mining Knowl. Discovery*, vol. 9, no. 4, pp. 1–13, 2019. doi: 10.1002/widm.1312.
15. A. Holzinger, A. Carrington, and H. Müller, "Measuring the quality of explanations: The system causability scale (SCS). Comparing human and machine explanations," *KI – Künstliche Intelligenz (German J. Artif. Intell.)*, vol. 34, no. 2, pp. 193–198, 2020. doi: 10.1007/s13218-020-00636-z.
16. R. W. Picard, A. Wexelblat, and C. I. Nass, "Future interfaces: Social and emotional," in *Proc. CHI'02 Extended Abstracts Human Factors Comput. Syst.*, 2002, pp. 698–699.
17. R. W. Picard, "Perceptual user interfaces: Affective perception," *Commun. ACM*, vol. 43, no. 3, pp. 50–51, 2000. doi: 10.1145/330534.330539.
18. S. Poria, E. Cambria, R. Bajpai, and A. Hussain, "A review of affective computing: From unimodal analysis to multimodal fusion," *Inform. Fusion*, vol. 37, pp. 98–125, 2017. doi: 10.1016/j.inffus.2017.02.003.
19. H. Yu, T. A. Boyle, C. Zhou, D. L. Rimm, and F. R. Hirsch, "PD-L1 expression in lung cancer," *J. Thoracic Oncol.*, vol. 11, no. 7, pp. 964–975, 2016. doi: 10.1016/j.jtho.2016.04.014.
20. D. Schneeberger, K. Stoeger, and A. Holzinger, "The European legal framework for medical AI," in *Proc. Int. Cross-Domain Conf. Machine Learning Knowl. Extraction*, Cham: Springer-Verlag, pp. 209–226. doi: 10.1007/978-3-030-57321-8-12.