Digital Health and Informatics Innovations for Sustainable Health Care Systems
J. Mantas et al. (Eds.)

doi:10.3233/SHTI240579

Explaining Deep Learning Models Applied in Histopathology: Current Developments and the Path to Sustainability

Cătălin-Mihai PESECAN[a,1] and Lăcrămioara STOICU-TIVADAR[a]
[a] *Politehnica University of Timișoara, Timișoara, Romania*
ORCiD ID: Cătălin-Mihai PESECAN https://orcid.org/0009-0008-4031-9303
Lăcrămioara STOICU-TIVADAR https://orcid.org/0000-0001-9700-7732

Abstract. The digital pathology landscape is in continuous expansion. The digitalization of slides using WSIs (Whole Slide Images) fueled the capacity of automatic support for diagnostics. The paper presents an overview of the current state of the art methods used in histopathological practice for explaining CNN classification useful for histopathological experts. Following the study we observed that histopathological deep learning models are still underused and that the pathologists do not trust them. Also we need to point out that in order to get a sustainable use of deep learning we need to get the experts to trust the models. In order to do that, they need to understand how the results are generated and how this information correlates with their prior knowledge and for obtaining this they can use the methods highlighted in this study.

Keywords. Artificial Intelligence, Deep Learning, Diagnostic Imaging, Sustainability

1. Introduction

The medical field is benefiting more from Information Technology integration. Deep learning platforms for "clinical data extraction, aggregation, management and analysis can support clinicians by effectively stratifying subjects to understand specific scenarios and optimize decision making". The shortage of histopathological experts is significant in multiple parts of the world [1] leading to a lower standard of care and prolonged patient waiting time for a histopathological diagnosis. The evolution of artificial intelligence connected with the health informatics data could lead to the improved patient care especially in the areas where the experts are in short supply. A big problem for this is the lack of trust in the models, even when they perform at the level of an M.D (this is valid for both the patients and the M.D.). To investigate this path the authors decided to try to answer to the following questions:

1. What are the latest types of Explainable Artificial Intelligence methods available for Convolutional Neural Networks?

[1] Corresponding Author: Cătălin-Mihai PESECAN, Politehnica University of Timișoara, Timișoara, Romania; E-Mail: catalin.pesecan@student.upt.ro.

2. How using XAI could increase the sustainability of artificial intelligence in the medical field?

Answering the first question we will support the idea that trust can be increased based on XAI and raising the interpretability rate, while answering the second one we can understand how to use explainability in the future.

In order to grasp the basic information to support the answers we conducted a systematic review using the **Clarivate Web of Science** database. Other reviews do not take into account the sustainability issue. We consider that this part is of utmost importance because if we want that the use of deep learning to become part of the daily clinical practice, we need to use it in a sustainable manner. A review for prostate cancer detection on histopathological images [2] concluded that AI models applied on histopathology images could benefit for reducing intra- and inter-pathologist variability in grading tumors.

2. Materials and Methods

This paper analyzes the main developments in explainable artificial intelligence (XAI) for image classification tasks that can be used in the medical domain. A systematic literature review has been conducted on Web of Science on explainable artificial intelligence approaches that can be applied for the histopathological image data using the search terms “explain CNNs”, “explain Convolutional Neural Networks”, “interpret CNNs”. For the current study we considered the results published between 01.2019-02.2024 in order to make sure that we capture the newest and up to date innovations. The identified papers were abstract screened and after that we selected the papers to be included in the review.

We decided to include papers that correspond to the following ***inclusion criteria***:

- The paper must be related to images.
- The model in discussion must be black-box (Convolutional Neural Network).
- The paper must be relatable to the domain of histopathology.
- The paper must be connected to image classification and offer an explainability solution.

2.1. Study selection and data extraction

Table 1. Initial number of search results

Search term	Initial number of results
explain CNNs	233
explain Convolutional Neural Networks	1036
interpret CNNs	256

2.2. Main XAI methods that can be applied on images

In this section we will present the main methods that were selected during this study. The survey in [3] presents multiple setups where CNNs can be useful for histopathological practice. The authors of [4] presents an example of a disease that benefitted from the explanations.

Several researchers used concept-based explanation [5] to offer a better perspective for predicting survival in patients with lung cancer. They compare that with the more common Grad-CAM used by [4]. An interesting development in this area is the presence of approximations for explanations that offers a decrease in computational times for the explaining part (important for the emergency situations in medical area). Another type of method that can be used for better explanation is using attention mechanisms [6], which are an improvement for CAM-based methods.

3. Results

We will present the main advantages of methods that resulted after the systematic review and how the methods presented in this paper can be used to get better sustainability for the final AI user. In order to get the AI to be used long term and in critical areas (like the medical domain) all the parties involved (MDs, regulators, nurses, patients etc.) must be able to trust and understand the models. No matter where in the medical profession the models are used, they must be trustworthy and help the MDs perform a better and faster job.

The first advantage is the increased access to high quality care [7] for the patients where the time is of the most importance.

Another advantage could be the usage of an attribution method, like the one described in [8] to explain elements of the objects that could be used to better understand subparts of the histopathological images that are clinically relevant.

Finally, the third most relevant benefit is the possibility to use multiple methods could offer different perspectives on the same image, each opening another dimension of the model in question, similar to the model in [9]. Even using methods with increasing explanations (like different types of CAM) could offer better explainability.

We must also take into account that the training and preparation of models takes time and using explainability will make the models more useful and will reduce the energy needed for their training [10] in this way decreasing the energy need and ultimately the impact on the environment.

4. Discussion

The adoption of artificial intelligence is raising globally. It can enable multiple benefits for the histopathological experts. It is a continuous discussion if we can explain or at least interpret the results of applying a neural network to a medical process or activity. As neural networks are using algorithms and clear steps, one can explain the steps and the fact the result is obtained using matrix multiplication, calls, functions, the usage of optimization algorithms.

Explainability to this point, technically exists. But when one wants to explain how the NN gave the result that a patient has disease X from symptoms Y it becomes more complicated. The parameters of the network are in huge number (millions, billions, etc.) and reasoning what patterns or features its detecting results in a reduced grade of explanation. A physician can explain how he selected diagnosis X, the shapes in the image, the colors, the positions, etc. For future work we will investigate how we can connect the technical basis with the cognitive human explanations.

Further investigations in this area could consider input gathering from multiple individuals in the medical community and adapting the explanations to their specific needs. Another improvement is changing the way the methods deliver the results to the user according to his feedback, in order to help him achieve his particular goal.

5. Conclusions

The need for support in histopathology will continue to increase as the time passes and the economics medical care will put more pressure on all the actors involved. A robust deep learning system that could explain histopathology could be then translated into a standard “modus operandi” for this medical science and lots of patients, especially in places where histopathology expertise is scarce could benefit from it.

References

[1] USCAP 2022 Abstracts: Quality and Patient Safety (1249-1299). Modern Pathology. 2022 Mar;35:1473–522.
[2] Bhattacharya I, Khandwala YS, Vesal S, Shao W, Yang Q, Soerensen SJC, et al. A review of artificial intelligence in prostate cancer detection on imaging. Therapeutic Advances in Urology. 2022 Jan;14:175628722211287.
[3] Abdelsamea MM, Zidan U, Senousy Z, Gaber MM, Rakha E, Ilyas M. A survey on artificial intelligence in histopathology image analysis. WIREs Data Min & Knowl [Internet]. 2022 Nov [cited 2023 Feb 6];12(6). Available from: https://onlinelibrary.wiley.com/doi/10.1002/widm.1474
[4] Yiğit T, Şengöz N, Özmen Ö, Hemanth J, Işık AH. Diagnosis of Paratuberculosis in Histopathological Images Based on Explainable Artificial Intelligence and Deep Learning. TS. 2022 Jun 30;39(3):863–9.
[5] Sauter D, Lodde G, Nensa F, Schadendorf D, Livingstone E, Kukuk M. Validating Automatic Concept-Based Explanations for AI-Based Digital Histopathology [Internet]. Vol. 22, Sensors. 2022. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85135129275&doi=10.3390%2fs22145346&partnerID=40&md5=53f1a9cccfde72138aa95d3c0b093e4b

[6] Song B, Zhang C, Sunny S, Kc DR, Li S, Gurushanth K, et al. Interpretable and Reliable Oral Cancer Classifier with Attention Mechanism and Expert Knowledge Embedding via Attention Map. Cancers. 2023 Feb 23;15(5):1421.
[7] Kourounis G, Elmahmudi AA, Thomson B, Hunter J, Ugail H, Wilson C. Computer image analysis with artificial intelligence: a practical introduction to convolutional neural networks for medical professionals. Postgraduate Medical Journal. 2023 Nov 20;99(1178):1287–94.
[8] Horii K, Maeda K, Ogawa T, Haseyama M. [Papers] Interpretable Convolutional Neural Network Including Attribute Estimation for Image Classification. MTA. 2020;8(2):111–24.
[9] Rguibi Z, Hajami A, Zitouni D, Elqaraoui A, Bedraoui A. CXAI: Explaining Convolutional Neural Networks for Medical Imaging Diagnostic. Electronics. 2022 Jun 2;11(11):1775.
[10] Jean-Quartier C, Bein K, Hejny L, Hofer E, Holzinger A, Jeanquartier F. The Cost of Understanding—XAI Algorithms towards Sustainable ML in the View of Computational Cost. Computation. 2023 May 4;11(5):92.