International Journal of Engineering Sciences Paradigms and Researches (IJESPR)
Volume 54, Issue 01 and Publication Date: 11th March, 2025
An Indexed, Referred and Peer Reviewed Journal with ISSN (Online): 2319-6564
www.ijesonline.com

Explainable AI for Cancer Care: Opportunities and Challenges

Ritika Gupta[1], Dr Bijender Bansal[2], Dr Vinit Kumar[3] and Dr Deepak Goyal[4]

[1]M. Tech. CSE Student, Vaish College of Engineering, Rohtak

[2]Professor CSE, Vaish College of Engineering, Rohtak
bijender.vce@gmail.com

[3]Assistant Professor CSE, Vaish College of Engineering, Rohtak
lohanvinit@gmail.com

[4]Professor CSE, Vaish College of Engineering, Rohtak
deepakgoyal.vce@gmail.com

Abstract

Despite recurrent issues with trust which limits clinical adoption due to its “black box” nature, artificial intelligence has the ability to transform aspects of cancer treatment and improve the quality of life for patients. AI, or Explainable Artificial Intelligence (XAI), enhances reason for trust, confidence, and transparency in regard to the reliability of the outcome, therefore solves these concerns. This trust ensures that widespread use is adopted. The paper aims to address the most significant concerns and advantages of XAI in oncology. To analyze cancer care through XAI, dependable and transparent insights into the prognosis need to be provided. This covers executing interpretable image analysis and issuing personalized treatment recommendations based on patient data and XAI driven decisions. Some of the methods we cover include SHAP Values, LIME, and counterfactual explanations pertinent to cancer tasks. There are also numerous challenges that need to be worked on such as developing reliable and scalable XAI frameworks for complex data in cancer care, creating human-centered explanation interfaces which can be easily interpreted by both clinicians and patients, focused on the ethical aspects of XAI.

Keywords: *Artificial Intelligence, Cancer Care, Black Box, DNN, Grad CAM.*

1. Introduction

More recent advancements in artificial intelligence (AI) have enabled Deep Neural Networks (DNNs) to achieve remarkable strides in multiple domains like image classification and medical diagnosis [1,2], analysis of chemical systems, and forecasting the spread of infectious diseases [3,4,5,6]. Although progress in these areas is noteworthy, one key hurdle still stands: there exists no universal acceptance of DNNs, even when they deliver high accuracy.

One of the main factors contributing toward this is AI bias. Just like we do, AI models develop biases as part of their learning which affect their predictions. Unlike human biases which at some level can be flagged and rectified, the biases present in AI models tend to be sophisticated and buried. A number of works [7] has shown that bias can impact AI forecasts, but only a handful of biases have actually been brought to light. Due to the so-called “black box” nature of these models, users are entirely unable to understand how decisions are made, which is crucial for discovering and resolving biases. Among the biases bias is included the race, gender, and socio economic background of a person or a group of people [7].

AI models' absence of clarity raises moral and real-world concerns. Understanding the internal workings of algorithms that function as ‘black boxes’, AI systems which cannot be directly observed, is difficult. Such systems can lead to ethical breaches, sensitive data manipulations, and damaging outcomes, with modern AI integration pacing across industries like finance, healthcare, and government as top priorities.

In response, researchers introduced concepts of AI without obscurity or methods that assist the AI in providing explanations of its own processes. These arguments underpin Explainable AI (XAI), which makes sure that its AI systems offer transparent reasoning and pertinent explanations for their outputs.

XAI is not only an AI decision-interpretation tool but also a key to fairness and trust in AI systems [9]. It empowers humans to audit AI algorithms, ensuring transparency and fairness. By explaining the AI decision-making processes, XAI enables researchers and developers to identify defects, weaknesses, and biases, ultimately improving the credibility of AI models [9]. Additionally, XAI is valuable in knowledge discovery, as it will uncover new patterns and facts learned by the AI models [10].

Cancer is a complex and heterogeneous disease involving uncontrolled growth of cells and possible spread of cells into remote organs within the body. Early detection is very important to save lives because it will enable clinicians to provide the appropriate drug at the appropriate time.

Artificial Intelligence (AI) and Machine Learning (ML) are now part of the backbone of contemporary healthcare [11]. Their incorporation into cancer diagnosis has transformed medical imaging by enhancing early detection using CT scans, MRIs, mammograms, and other diagnostic modalities. Moreover, genomic profiling, which entails analyzing a person's entire collection of DNAs, has improved personalized treatment plans [13-15].

Deep learning, specifically Convolutional Neural Networks (CNNs), has been particularly critical in the detection of cancer through identifying anomalies that would otherwise not be noticed by the naked eye [16, 17]. Beyond CNNs, other AI methodologies including Ensemble Learning, Natural Language Processing (NLP), Transfer Learning, and Graph-Based Methods have been driving AI's impact in oncology [18].

With AI increasingly revolutionizing cancer diagnosis and treatment, ethical issues have come to the forefront. Bias, justice, and accountability all need to be addressed to ensure that AI is utilized ethically in medicine. XAI becomes a central solution, offering the transparency, trust, and reliability needed to ensure that AI is utilized responsibly in cancer treatment. Through improving explainability, XAI gives clinicians, patients, and researchers a better grasp of AI-recommended decisions, reduces bias, and guarantees that AI continues to be an effective tool in saving lives through early and precise detection of cancer.

2. Methodology Used:

2.1. Literature Search and Selection Criteria

To ensure a comprehensive and unbiased review, research articles were collected from multiple academic databases, including:

- PubMed (for medical and clinical studies)
- IEEE Xplore (for AI applications in healthcare)
- Google Scholar (for a broad range of AI and machine learning studies)
- ArXiv (for recent advancements in Explainable AI—XAI)

2.2. Selection and Exclusion Criteria

To maintain the relevance and reliability of the study, the following criteria were applied:

- Only studies published within the last decade were considered to ensure up-to-date research.
- Priority was given to research papers focusing on XAI methodologies applied to cancer diagnosis.
- Preference was given to studies comparing different XAI techniques, particularly those evaluating performance metrics.
- Papers unrelated to experimental validation of XAI methods, medical imaging, pathology, or oncology AI models were excluded.
- Non-English publications were excluded to ensure accessibility and consistency.

2.3. Analysis and Validation

- AI models and their approaches were examined with a primary focus on improving model interpretability and understanding how XAI is perceived in medical care.
- Special attention was given to identifying research gaps and real-world deployment challenges of XAI in clinical settings.
- Findings were cross-referenced with systematic reviews and meta-analyses to ensure consistency**,** reliability**,** and clinical relevance.

This study follows a systematic methodology to analyze XAI applications in cancer care and treatment**,** ensuring that the findings are evidence-based, reliable, and clinically applicable**.**

3. Explainable AI (XAI): Transparency and Trust in cancer care

Explainable AI (XAI) deals with and pushes forward artificial intelligence (AI) by making decision-making procedures observable and understandable to humans. Transparency, here, means intelligibility and understandability of judgments made by AI so that customers can perceive and have faith in AI predictions to a greater extent [19]. XAI reduces the mysticism behind AI's "black box" by showing hidden principles and rationale that are otherwise hidden from users [19]. This decrease in opacity is particularly significant in industries that require high levels of rationality and moral responsibility, such as the health sector, banks, and justice systems.

Emerging studies have recently concentrated on constructing new approaches and refining current practices to enhance explainability in AI. For instance, a paper [20] proposed a framework for transparency that encourages reflective decision-making, whereby AI-generated insights are aligned with ethical concerns. The framework uses active inference to offer decision-making processes that are interpretable as well as unopaque, and hence more tolerable to society and regulatory institutions.

In cancer diagnosis, transparency is particularly valued by clinicians as it helps build trust in AI-driven recommendations, facilitating unhindered and effective utilization of AI in decision-making processes. Studies have demonstrated that SHAP (SHapley Additive exPlanations) is one of the widely used model-agnostic XAI techniques in breast cancer research [21]. SHAP has been instrumental in: Prediction analysis, Biomarker diagnostics, and Prognosis and survival analysis.

By making AI models more interpretable, SHAP increases trust and uptake in healthcare. Furthermore, the combination of Convolutional Neural Networks (CNNs) with XAI frameworks has also demonstrated great potential in mammography. Studies indicate that the combination of CNNs with XAI enhances: Diagnostic accuracy, Understandability of AI-provided explanations, and Clinician trust in AI usage [22][23].

Another investigation revealed that model-agnostic XAI methods are used mostly in medical decision-support systems, with deep learning models being used more than conventional machine learning models [24]. Nevertheless, even with all these developments, XAI user interfaces are still limited, having few interactive and visual elements that have the potential to enhance clinician engagement and comprehension. This point of limited usability hinders effective collaboration between medical professionals and AI specialists, restricting AI solution development tailored to clinical environments [24].

Through several XAI methodologies, including SHAP, LIME, and Grad-CAM, it has been identified their potential for bridging the difference between complex AI models and actual healthcare uses. Through incorporation with machine learning and deep learning algorithms, XAI promotes: medical professionals' trust and transparency, improved prediction performance, and more accurate and trustworthy AI-based diagnoses [25].

Even with these breakthroughs, several obstacles to broader deployment of XAI in cancer diagnosis continue: the intricacies of medical information, the necessity of in-depth verification of AI models, and the need for smooth incorporation into clinical workflows. Future studies must focus on developing more understandable, interactive XAI tools which are aligned with actual medical work flows while maintaining high accuracy, equity, and reliability.

4. Literary Survey

XAI has becomes a pivotal point of research in enhancing transparency and Trust in AI driven models in medical diagnostics and treatments.

Table 4.1: Literary Survey

Paper	Focus Area	Contributions
[1]	Deep Learning	Introduced AlexNet, pioneering CNN-based image classification.
[2]	Medical Imaging	Survey of deep learning techniques in radiology and pathology.
[16]	Skin Cancer Detection	Developed CNNs for dermatologist-level accuracy in skin cancer.
[8]	XAI	Comprehensive survey on XAI techniques for interpreting black box AI.
[21]	XAI in cancer care	Identified SHAP as a dominant technique in breast cancer risk prediction, improving interpretability.
[22]	XAI and CNNs	Integrated CNNs with Grad-CAM &SHAP for breast

		cancer mammography.
[23]	Trust in AI	Explored trustworthy deep learning models for cancer diagnosis, focusing on clinical adoption.
[24]	XAI in medicine	Systematic Review of XAI methods used in medical decision support for oncology and beyond.
[29]	Ethical AI in healthcare	Proposed an ethical framework for AI in healthcare.
[25]	XAI in breast cancer	Comparative Study of different XAI techniques in Breast Cancer Diagnostics.

- In [26], the author surveys explainable deep learning approaches for medical picture categorization and discusses Grad-Cam's uses in several investigations. They also examine the use of TCAV in medical picture categorization, emphasizing the possibility for context-based explanations. Applications of LRP in medical picture classification, with an emphasis on its utility in offering hierarchical explanations, were also discussed.
- In [27], the focus is on LIME approximating the complex models with interpretable ones locally to explain individual predictions and highlights the use rule-based methods in providing explanations for AI models in biomedical imaging.
- In [21], The discussion of using counterfactual explanations to better understand AI models in breast cancer research. Indications of how little changes to the input can affect the model's forecast, aiding in the comprehension of decision boundaries.
- In [22], the use of CNNs with their evaluation and Integration with Explainable AI is discussed, especially in breast cancer domain.
- In [24], models like decisions trees and rule-based systems were thought of as more straightforward decision paths, making them inherently interpretable. This research highlights the application of such models in medical decision support, emphasizing their transparency and ease of understanding.
- According to [28], there are still obstacles in integrating XAI in cancer care. Approximately 83 percent of research did not engage clinicians in developing and evaluating XAI tools, which may limit their practical application.
- In [29], the necessity of human centred AI design to make the AI driven models transparent and trustable to clinicians were discussed as well as the role of XAI in reducing algorithmic biases that may affect the under present patient populations was one of the key points. The demands for AI standards to ensure clinical safety and ethical deployment were stated.

5. Future Research Directions: Advancing XAI for Equitable and Effective Oncology

The evolution of Explainable Artificial Intelligence (XAI) in cancer needs to direct fairness, effectiveness, and accessibility so that AI-based diagnoses and treatment are ethical, clear, and beneficial across patient groups. AI models, despite their vast potential, may inadvertently perpetuate biases, resulting in inequalities in cancer diagnosis and therapy. Mitigating these biases requires the development of accurate and broad AI tools and the establishment of expert panels for reviewing AI-driven predictions, detecting biases, and suggesting modifications. This approach preserves the justice and fairness of AI-based applications in medicine while fostering greater confidence among patients and clinicians [30].

To be effectively used in oncology, XAI systems must not only be reliable but also nicely integrated into clinical workflows. Interpretability of such systems is critical in assisting clinicians with diagnostic and biomarker classification tasks, thereby generating confidence in AI-aided recommendations. The fusion of multimodal data, such as medical imaging and genomic information, has proven to have vast potential in enhancing accuracy and interpretability. However, validity and generalizability of such AI algorithms for clinical use entail extensive cross-validation across different patient groups and environments. Extensive rigorous testing, therefore, must be carried out for real-world oncology deployment possibilities as well as for managing bias in predictive practice [31].

In short, moving XAI ahead in oncology necessitates a collaborative and interdisciplinary process of trust establishment, mitigation of biases, and enhancement of interpretability strategies compatible with the complexity of cancer treatment. These issues should be given priority so that subsequent research will be able to contribute towards the development of AI tools not only more efficient and equitable, but also confidence-inducing in healthcare professionals. Ultimately, such progress will lead the way towards improved cancer diagnosis, customized treatment strategies, and improved patient outcomes, continuing to advance the role of AI as an innovative solution for oncology [30,31].

References

[1] Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional neural networks." Advances in neural information processing systems 25 (2012).

[2] Litjens, Geert, et al. "A survey on deep learning in medical image analysis." Medical image analysis 42 (2017): 60-88.

[3] Mohammad Ali, and Sulaiman Al Riyaee. "Skin cancer detection: Applying a deep learning based model driven architecture in the cloud for classifying dermal cell images." Informatics in Medicine Unlocked 18 (2020): 100282.

[4] Wang, Yalin, et al. "A novel deep learning based fault diagnosis approach for chemical process with extended deep belief network." ISA transactions 96 (2020): 457-467.

[5] Shahid, Farah, Aneela Zameer, and Muhammad Muneeb. "Predictions for COVID-19 with deep learning models of LSTM, GRU and Bi-LSTM." Chaos, Solitons & Fractals 140 (2020): 110212.

[6] Weber, Leander, et al. "Beyond explaining: Opportunities and challenges of XAI-based model improvement." Information Fusion 92 (2023): 154-176.

[7] Stock, Pierre, and Moustapha Cisse. "Convnets and imagenet beyond accuracy: Understanding mistakes and uncovering biases." Proceedings of the European conference on computer vision (ECCV). 2018.

[8] Guidotti, Riccardo, et al. "A survey of methods for explaining black box models." ACM computing surveys (CSUR) 51.5 (2018): 1-42.

[9] Adadi, Amina, and Mohammed Berrada. "Peeking inside the black-box: a survey on explainable artificial intelligence (XAI)." IEEE access 6 (2018): 52138-52160.

[10] Gerlings, Julie, Arisa Shollo, and Ioanna Constantiou. "Reviewing the need for explainable artificial intelligence (xAI)." arXiv preprint arXiv:2012.01007 (2020).

[11] Smith, J., and A. Johnson. "The role of artificial intelligence in healthcare." Med J 12.3 (2020): 45-52.

[12] Lee, S. J. "Machine learning applications in cancer prognosis and prediction." Computers in Biology and Medicine 95 (2018): 93-103.

[13] LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. "Deep learning." nature 521.7553 (2015): 436-444.

[14] Obermeyer, Ziad, and Ezekiel J. Emanuel. "Predicting the future—big data, machine learning, and clinical medicine." New England Journal of Medicine 375.13 (2016): 1216-1219.

[15] Hosny, Ahmed, et al. "Artificial intelligence in radiology." Nature Reviews Cancer 18.8 (2018): 500-510.

[16] Esteva, Andre, et al. "Dermatologist-level classification of skin cancer with deep neural networks." nature 542.7639 (2017): 115-118.

[17] LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. "Deep learning." nature 521.7553 (2015): 436-444.

[18] Ravì, Daniele, et al. "Deep learning for health informatics." IEEE journal of biomedical and health informatics 21.1 (2016): 4-21.

[19] Meske, Christian, and Enrico Bunde. "Transparency and trust in human-AI-interaction: The role of model-agnostic explanations in computer vision-based decision support." Artificial Intelligence in HCI: First International Conference, AI-HCI 2020, Held as Part of the 22nd HCI International Conference, HCII 2020, Copenhagen, Denmark, July 19–24, 2020, Proceedings 22. Springer International Publishing, 2020.

[20] Albarracin, Mahault, et al. "Designing explainable artificial intelligence with active inference: A framework for transparent introspection and decision-making." International Workshop on Active Inference. Cham: Springer Nature Switzerland, 2023.

[21] Ghasemi, Amirehsan, et al. "Explainable artificial intelligence in breast cancer detection and risk prediction: A systematic scoping review." Cancer Innovation 3.5 (2024): e136.

[22] Ahmed, Maryam, et al. "Enhancing Breast Cancer Diagnosis in Mammography: Evaluation and Integration of Convolutional Neural Networks and Explainable AI." arXiv preprint arXiv:2404.03892 (2024).

[23] Olumuyiwa, Badaru I., The Anh Han, and Zia U. Shamszaman. "Enhancing Cancer Diagnosis with Explainable & Trustworthy Deep Learning Models." arXiv preprint arXiv: 2412.17527 (2024).

[24] Prentzas, Nicoletta, Antonis Kakas, and Constantinos S. Pattichis. "Explainable AI applications in the Medical Domain: a systematic review." arXiv preprint arXiv:2308.05411 (2023).
[25] Bai, Samita, et al. "Breast Cancer Diagnosis: A Comprehensive Exploration of Explainable Artificial Intelligence (XAI) Techniques." arXiv preprint arXiv: 2406.00532 (2024).
[26] Patrício, Cristiano, João C. Neves, and Luís F. Teixeira. "Explainable deep learning methods in medical image classification: A survey." ACM Computing Surveys 56.4 (2023): 1-41.
[27] Nazir, Sajid, Diane M. Dickson, and Muhammad Usman Akram. "Survey of explainable artificial intelligence techniques for biomedical imaging with deep neural networks." Computers in Biology and Medicine 156 (2023): 106668.
[28] Mohamed, Yusuf Abas, et al. "Decoding the black box: Explainable AI (XAI) for cancer diagnosis, prognosis, and treatment planning-A state-of-the art systematic review." International Journal of Medical Informatics (2024): 105689.
[29] Nasir, Sidra, Rizwan Ahmed Khan, and Samita Bai. "Ethical framework for harnessing the power of ai in healthcare and beyond." IEEE Access 12 (2024): 31014-31035.
[30] Bennett, Chloe R., et al. "Interdisciplinary Expertise to Advance Equitable Explainable AI." arXiv preprint arXiv:2406.18563 (2024).
[31] Viswanathan, Vidya Sankar, Vani Parmar, and Anant Madabhushi. "Towards equitable AI in oncology." Nature Reviews Clinical Oncology (2024): 1-10.