Expert Systems

WILEY

Expert Systems

ORIGINAL ARTICLE

From AI to the Era of Explainable AI in Healthcare 5.0: Current State and Future Outlook

Anichur Rahman[1,2] | Dipanjali Kundu[1] | Tanoy Debnath[3] | Muaz Rahman[4] | Utpol Kanti Das[1] | Abu Saleh Musa Miah[5] | Ghulam Muhammad[6]

[1]Department of CSE, National Institute of Textile Engineering and Research (NITER), Constituent Institute of the University of Dhaka, Dhaka, Bangladesh | [2]Department of CSE, Mawlana Bhashani Science and Technology University, Dhaka, Bangladesh | [3]Department of Computer Science, Stony Brook University, Stony Brook, New York, USA | [4]Department of EEE, National Institute of Textile Engineering and Research (NITER), Constituent Institute of the University of Dhaka, Dhaka, Bangladesh | [5]Department of Computer Science and Engineering, Bangladesh Army University of Science and Technology, Saidpur, Bangladesh | [6]Department of Computer Engineering, College of Computer and Information Sciences, King Saud University, Riyadh, Saudi Arabia

Correspondence: Anichur Rahman (anis_cse@niter.edu.bd) | Ghulam Muhammad (ghulam@ksu.edu.sa)

Received: 6 February 2025 | **Revised:** 28 March 2025 | **Accepted:** 12 April 2025

Funding: The work was funded by the Researchers Supporting Project number (RSP2025R34), King Saud University, Riyadh, Saudi Arabia.

Keywords: artificial intelligence | data analysis | datasets | explainable AI | Healthcare 5.0 applications | healthcare management | machine learning

ABSTRACT

Artificial intelligence (AI) and explainable artificial intelligence (XAI) are advancing rapidly, with the potential to deliver significant benefits to modern society. The healthcare sector, in particular, has experienced transformative changes; overall, these technologies are helping to address numerous challenges, such as cancer cell detection, tumour zone identification in animal bodies, predictions of major and minor diseases, diagnosis, and more. This article provides an in-depth and detailed overview of AI and XAI, focusing on recent trends and their implications for advancing Healthcare 5.0 applications. Initially, the study examines the key concepts and exceptional features of AI, XAI, and Healthcare 5.0. Additional emphasis is placed on state-of-the-art practices currently being implemented in healthcare, particularly those involving AI and XAI. Subsequently, it establishes a coherent link between AI and XAI in Healthcare 5.0, grounded in contemporary advancements. Based on the findings, algorithms are recommended to address initial obstacles to integrating AI into the Healthcare 5.0 framework. Proposals for further enhancing Healthcare 5.0 performance through the integration of XAI and its unique features are discussed in detail. The work also provides in-depth implementation strategies and highlights model-specific trends within AI and XAI frameworks in Healthcare 5.0. Particular attention is given to AI model predictions in healthcare settings, emphasising their contributions to improved patient feedback and the delivery of more sophisticated care. Most importantly, this research highlights the potential for AI and XAI to support sustainable advancements in Healthcare 5.0 applications. Finally, significant issues are analysed, and an open discussion is presented on future guidelines for the blending of AI with XAI, and Healthcare 5.0 applications.

Abbreviations: AI, artificial intelligence; CNN, convolutional neural networks; DL, deep learning; EHR, electronic health records; FHIR, fast healthcare interoperability resources; FL, federated learning; FTL, federated transfer learning; HCU, healthcare control unit; HM, healthcare management; IoMT, internet of medical things; IoT, internet of things; LR, logistic regression; M2M, machine to machine; MCDM, multi-criteria decision-making; ML, machine learning; NLP, natural language processing; P2P, peer to peer; PCA, patient-centric agent; PRISMA, systematic reviews and meta-analyses; SC, smart contact; SDG, sustainable development goals; SH, smart healthcare; XAI, explainable artificial intelligence.

1 | Introduction

Industrial operations have undergone significant changes with the incorporation of artificial intelligence (AI), and the industry is not an exception. Through the assistance of medical data examination, it has become possible to offer better treatment plans, leading to improved and more accurate patient outcomes. Additionally, AI has been integrated into current medical techniques, across various fields, such as surgical robots, virtual assistants, image classification and diagnosis, and predictions (Ueda et al. 2024). There are several areas where the integration of AI has improved the healthcare sector. Machine learning is the first step toward technological advancement in medical science. These algorithms can learn and categorise patterns from large datasets that may surpass human capabilities. Convolutional neural networks (CNNs), among various algorithms, are well known for image-based tasks, including cancer detection from images, tumour detection from radiology scans, COVID-19 detection, and many more (Furriel et al. 2024). Natural language processing (NLP), another subset of AI, has a tremendous impact on emotion recognition and extracting key information from unstructured clinical data, such as patient records. This, in turn, assists healthcare professionals in decision-making (Ahuja 2024). Additionally, AI-powered tools, like IBM Watson (Ofosu-Ampong 2024; Rahman, Debnath, et al. 2024), were developed specifically for oncologists by analysing medical literature, clinical trial data, and patient information. By analysing factors such as disease patterns, survival rates, and patient length of stay, AI-enabled tools assist hospital staff and management in optimising resources and provide patients with clearer insights into treatment plans and survival rates (Kazimierczak et al. 2024).

However, despite the tremendous progress of AI in the medical sector, this integration faces several obstacles. The full-scale implementation of AI remains challenging particularly due to the 'black box' features of AI. Among the key algorithms of AI, Deep Learning (DL) carries out decisions by analysing certain features that are often incomprehensible to people. This unclear pattern of AI-based decisions reduces trust and acceptability of such models in a healthcare facility (Ueda et al. 2024). Medical practitioners often find it difficult to rely on results delivered by AI models since there is a lack of explanations on how the conclusions are reached. Again, the DL method relies heavily on data, which raises the awareness of enhancing data security. There remains a possibility of attacks by external parties towards AI-driven data. The data is generated by sensors that are connected to server clouds. If there is a breach on any single module, the whole structure faces potential collapse resulting in significant losses for both hospitals and patient data. This validates the need for careful ethical considerations to develop a system that is acceptable by all parties (Nozari et al. 2024; Rahman, Khan, et al. 2024).

To minimise the potential data breach and enhance transparency and trust among users of AI-based models, the concept of Explainable AI (XAI) has been brought upon. The model aims to enhance AI models by generating explanations behind decisions made. These explanations are more readily understandable by its users (Williamson and Prybutok 2024). The clear explanations provided in XAI result in higher-accuracy mitigating gap among critical AI-based decisions and medical practices. The advent of Healthcare 5.0 further aims at generating a more robust and interconnected system with its patient-centric approach. The Healthcare 5.0 setting couples AI with XAI to improve the trustworthiness among its users allowing justifications for AI-based results (Patra and Datta 2024). The use of XAI will accelerate the acceptance of AI-driven data and predictions will be more accepted by general patients. The ethical considerations of AI are also dealt with by XAI by permitting an easy-to-monitor platform. Considerable participation from both doctors and patients alike will enhance the acceptance of XAI in healthcare settings as patients will be provided with an elaborate understanding of their treatment procedure (Patel et al. 2024; Rahman, Hossain, et al. 2023).

XAI in Healthcare 5.0 settings offers a unique feature in which a more patient-centric approach is introduced. This concept builds transparency and trust among its users. Hospitals can greatly benefit from the utilisation of XAI models which promote the concepts of understandability, explainability, interpretability, transparency, informativeness, comprehensibility, responsibility, explicitness, and faithfulness (Kök et al. 2023). The aim of this research remains to offer this significant transition in a healthcare setting. The authors detailed an elaborate survey of the current state, focusing on the challenges, and presenting future directions for improving Healthcare 5.0.

1.1 | Paper Rationale

This work is motivated by the urgent need to enhance trust, transparency, and interpretability in AI-based healthcare systems, particularly within the emerging Healthcare 5.0 paradigm. AI has the potential to identify and predict diseases in a healthcare environment. However, many parties often remain critical of AI-based decisions due to their lack of clarity on how conclusions are reached. This opaque nature of AI has created an obstacle on its way to being widely accepted into the healthcare setting. The fundamental concept of trust between medical practitioners and patients using an AI-based platform faces numerous challenges due to transparency issues.

There are notable breakthroughs in medical science with AI leading the race. However, the advancements came with criticism ranging from data privacy to biases in algorithmic decisions. The tendency of AI to lean toward predictive accuracy is not accepted by users. Consequently, this work focuses on the essential requirement to enhance trust, transparency, and interpretability in AI-based healthcare systems. The black box feature of AI is hindering its way toward clinical acceptance. The study concluded in this paper aims to attenuate the existing gap by introducing a novel framework that enhances clarity and transparency concerns in AI-based decisions in Healthcare 5.0.

To analyse the research gaps, the work in this paper presents a detailed review of AI and XAI frameworks in a healthcare setting. The research work presented identifies limitations in current approaches, and offers innovative strategies to enhance the explainability, scalability, and clinical reliability aspects of AI. The focus is on the interpretability of XAI leading towards a sustainable healthcare solution. The unique feature of XAI lies in its ability to deal with a patient-centric approach, offering transparency and trust among its users.

The next parts will go over the technical and practical components of Healthcare 5.0, with a focus on how AI and XAI may handle these difficulties and improve healthcare systems.

1.2 | Related Surveys With Research Gaps

The application of AI to transform medical services has received massive attention in the research community. There is an ongoing tendency among researchers to promote AI and explainable AI (XAI) to enhance patient care within the HC 5.0 framework. However, existing research in this area has its limitations. The work presented by the authors in Al Mamun et al. (2021) delineates an intelligible AI-based system designed to monitor and detect abnormalities in brain functioning. Similarly Pulipeti et al. (2024) significantly examines the involvement of AI in enhancing hospital sector facilities, by observing the challenges faced by AI in Healthcare 5.0 and suggesting the role of explainable AI to help address these concerns. The study further explores key research areas associated with the development of XAI in the sphere of Healthcare 5.0. Chamola et al. (2023) conduct an in-depth analysis of Trustworthy and XAI, focusing on important parameters such as transparency, fairness, and accountability in healthcare and autonomous systems. Their study prioritises practical applications and policies, particularly for AI development in autonomous vehicles. However, the authors identify several limitations, including the lack of established measures for trust, the difficulty of interpreting deep neural networks, and the trade-off between efficiency and explainability. These challenges underscore the need for more standardised frameworks to ensure the reliability and interpretability of AI systems. In their comprehensive review of example-based XAI techniques in medical imaging, the authors of Fontes et al. (2024) highlight how these methods can improve usability, accuracy, and transparency in clinical settings. They examine promising techniques such as CapsNet, GANs, and counterfactual approaches that show promising results in improving ethical decision-making and diagnostic accuracy. However, issues related to consistency, computational efficiency, and ethics remain, and the authors argue that future studies should focus on promoting more flexible and ethically sound XAI techniques for healthcare applications.

The research presented in Giuste et al. (2022) analyses the significance of XAI in increasing the acceptance of AI-based solutions during the COVID-19 pandemic. The authors emphasise how the lack of clarity in AI models has created obstacles in the path of their adoption in clinical practice. XAI is proposed as a method for improving trust, performance, and decision-making by making AI models more interpretable. Despite its potential, the study stresses that clinical adoption continues to face challenges and that more transparent and accountable AI technologies are essential for full integration into healthcare settings. Tjoa and Guan (2021) provide a thorough analysis of XAI methods, emphasising their relevance to the medical industry, where accountability and transparency are essential. They classify various interpretability techniques and discuss the challenges posed by noisy training data, the 'black box' nature of deep learning models, and the risk of explanation manipulation. To address these issues, the authors address the need for specialised training that combines data science and medical expertise. They also argue that human oversight remains crucial in medical AI applications until more reliable interpretability techniques are developed. In Dhiman et al. (2023), the authors evaluate global trends in XAI research within healthcare from 2019 to 2022 using tools such as VOSviewer and Biblioshiny. They note that the USA leads in XAI research, particularly in healthcare classification and diagnostic imaging. However, the study has some limitations, such as its reliance on a single database (Scopus) and the need for more comprehensive keyword searches. The authors emphasise that, particularly in high-stakes medical decision-making, XAI techniques like LIME and SHAP are significant for improving trust and transparency in AI-driven healthcare. Another study (Hassija et al. 2024) provides a comprehensive analysis of XAI methods used to interpret complex AI models in high-stakes industries such as public safety, healthcare, and banking. The authors highlight the ongoing challenges of interpreting black box models, particularly concerning the reduction of false positives and negatives. They stress the need for more structured approaches to improve AI's ethical considerations and transparency and recommend future research directions focused on developing Responsible AI frameworks that prioritise openness, accountability, and human-centred decision-making.

Jiang et al. (2022) explored the development of artificial intelligence, tracing its journey from its early conception to its contemporary applications and potential future paths. The authors highlight significant historical turning points, technical advancements, and the broader societal implications of AI. They divide the evolution of AI into three main phases: the birth and golden era of AI, the 'AI winters' marked by periods of stagnation, and the resurgence driven by breakthroughs in machine learning (ML) and deep learning (DL). Additionally, they emphasise AI's influence on modern technologies such as human-in-the-loop (HITL) systems, autonomous systems, and human-AI symbiosis. In their conclusion, the authors offer a forward-looking assessment of the opportunities and potential challenges facing the AI industry. Zhang and Lu (2021) provided a comprehensive review of artificial intelligence with a focus on industry information integration. The report provides a comprehensive review of the area by looking at background data, motives, technologies, and applications in addition to logical insights on AI's development. The authors argue that their work provides valuable insights for real-world practitioners and contributes to ongoing AI research. Alberto et al. (2023) examined the current landscape of commercial health data vendors, focusing on data sources, challenges related to reproducibility and generalisability, and ethical concerns in data vending. The authors advocated for sustainable strategies to curate open-source health data, aiming to ensure the inclusion of global populations in biomedical research. They emphasised the need for collaboration among key stakeholders to make healthcare datasets more accessible, inclusive, and representative, while also safeguarding individual privacy and rights. Loh et al. (2022) explored the role of explainable artificial intelligence (XAI) in healthcare, with the goal of increasing trust in AI systems by providing transparent explanations for their predictions. Although AI models have demonstrated human-like performance, their adoption in healthcare remains limited due to concerns over their 'black box' nature. The review examined XAI techniques, such as SHAP, LIME, GradCAM, and rule-based systems, based on 99 studies from high-quality Q1 journals. The authors called for further development of AI systems to improve healthcare applications, which could lead to smarter, more efficient healthcare management, particularly in smart cities. Wang, Wu, et al. (2023) investigated the integration of visualisation and AI in data analysis to leverage their complementary strengths. While visualisation aids in intuitive data understanding and interactive exploration, AI excels in learning from data and automating complex tasks. The combination of these approaches is essential for large-scale data scenarios, such as epidemic traceability and city planning. The authors proposed an integration framework, VIS+AI, which allows AI to learn from human interactions and communicate through visual interfaces.

A thorough analysis of XAI methods in healthcare is presented by the authors of Chaddad et al. (2023), with a particular focus on medical imaging applications. They classify various XAI algorithms and emphasise the importance of making AI models more interpretable to ensure transparency in high-stakes medical decisions. The study highlights the need for more naturally interpretable models to build physician trust, even as post hoc techniques like Grad-CAM and LIME remain popular due to their ease of use. To further improve AI performance and usability in clinical contexts, the authors suggest combining XAI with strategies such as domain adaptation and federated learning. The study by Rong et al. (2023) explores the need for human-centred evaluations in XAI. By categorising user research according to trust, comprehension, usability, and human-AI collaboration, it reviews 97 foundational publications. Key strengths of the study include its focus on enhancing user engagement with XAI models and its practical recommendations for conducting effective user studies. However, the study also points out some limitations, including the lack of user assessments in certain areas and the insufficient integration of social and cognitive sciences in XAI research. The authors stress the need for standardised, transparent evaluation methods and call for future research that connects XAI with psychological sciences to better understand and address user needs.

1.3 | Research Questions

Several important research questions arise in light of the revolutionary potential of AI and XAI within the evolving paradigm of Healthcare 5.0. These inquiries aim to explore the potential applications, challenges, and practical integration of these technologies in healthcare settings. The following key research questions guide this study:

- How can AI and XAI be effectively integrated into Healthcare 5.0 to address modern challenges in disease detection and prediction?
- What are the key characteristics of AI and XAI that contribute to advancements in Healthcare 5.0, and how do these technologies improve healthcare outcomes?
- What are the existing technological gaps in Healthcare 5.0, and how can AI and XAI be leveraged to provide long-term, future-ready solutions?
- How do different approaches to implementing AI and XAI in Healthcare 5.0 frameworks influence their scalability and adaptability in real-world applications?
- What critical issues in traditional healthcare systems can AI and XAI address, and what unresolved questions remain for future research?

In addition to advancing our theoretical understanding of AI and XAI in Healthcare 5.0, addressing these research questions will provide valuable insights for developing sustainable, innovative healthcare solutions.

1.4 | Research Contribution

The main contributions of this research are as follows:

- First, the authors of this paper introduce the key concepts of AI, XAI, and Healthcare 5.0, along with their underlying principles and features, and provide a comprehensive, state-of-the-art survey of these technologies.
- Second, we offer conceptual and methodological approaches for integrating AI and XAI with Healthcare 5.0, providing insights for future sustainability and application. While this paper does not implement specific systems, it presents actionable guidelines for addressing practical integration challenges.
- Third, this study identifies key technical challenges in traditional healthcare systems and proposes solution frameworks using modern technologies, including XAI, to address these challenges in Healthcare 5.0 applications.
- Finally, this work analyses open issues and proposes potential future research directions, offering a roadmap for researchers interested in advancing Healthcare 5.0.

Organisation of this Study: The remainder of this survey is structured as follows. Section 2 outlines the methodology used in this manuscript. Section 3 discusses key concepts and features related to AI, XAI, and Healthcare 5.0. Section 4 presents the current state of AI and XAI applications in Healthcare 5.0. Section 5 explores the implementation perspectives of these techniques. Section 6 addresses future sustainability considerations. Section 7 highlights open issues and future opportunities in detail. Finally, Section 8 concludes the paper. To provide additional context, Table 1 provides a chronological summary of key research works categorised by AI, XAI, and their applications in Society 5.0 and Healthcare 5.0. It includes information on the authors, year, key technology, and main findings of each study. In addition, Table 2 provides an arranged collection of research from 2023 to 2024, organised by important technologies such as AI, XAI, and Healthcare 5.0. The table provides a useful comparison by highlighting the authors, methodology, datasets, evaluation criteria, and key focus areas. These studies look at a variety of topics, such as automation in academic publications, the role of AI in scientific writing, the Sustainable Development Goals (SDGs), cloud-based AI decision-making, and explainable AI frameworks for solar energy and healthcare.

TABLE 1 | Study related to Healthcare 5.0, XAI, and AI.

Authors	Year	Key technologies	Key findings
Kousha and Thelwall (2024)	2024	Artificial intelligence	They explored automating publishing tasks, excluding bibliometric data, and identified efficiencies.
Kacena et al. (2024)	2024		The study aimed to assess AI's potential in enhancing scientific writing, finding decreased writing time but requiring human supervision due to errors.
Singh et al. (2024)	2024		The authors analyse global AI research for SDGs, focusing on changes over time, geographical trends, knowledge flows, application areas, and AI approaches for meeting SDG objectives.
Alshahrani et al. (2024)	2024		The study aimed to pinpoint challenges in establishing an IT firm's long-term AI cloud system, utilising a fuzzy integrated hybrid MCDM model.
Aurelia et al. (2024)	2024		The pandemic's effects on education led to a move toward remote proctoring and virtual exams, which have advantages like cost-effectiveness and flexibility.
Dwivedi et al. (2024)	2024		Generative AI integration, exemplified by ChatGPT, promises transformative effects on hospitality.
Rajapaksha et al. (2023)	2023		AI-based intrusion detection systems (IDSs) as effective defenses against cyberattacks, the study exposes the Controller Area Network (CAN) vulnerability.
Alotaibi and Alshehri (2023)	2023		In light of the 2030 Vision, the study highlights the significant role artificial intelligence (AI) plays in Saudi Arabia's higher education institutions.
Nallakaruppan et al. (2024)	2024	Explainable AI	The authors proposed an XAI framework to improve predictability and control of solar energy distribution by explaining machine learning.
Famiglini et al. (2024)	2024		The study results challenge common XAI assumptions and emphasise evidence-based, human-centred design for AI-assisted diagnostics.
Cerekci et al. (2024)	2024		Lower-level feature Class Activation Maps (CAMs) improve diagnostic accuracy in TL fracture detection, especially among experienced physicians, and traditionally coloured CAMs consistently outperform semantic CAMs.
Nazat et al. (2024)	2024		To improve confidence and transparency in the security of vehicular networks, the paper presents an XAI framework for anomaly detection in autonomous vehicle networks.
Magd et al. (2024)	2024		They aimed for long-term growth via Industrial Revolution 4.0, integrating Big Data, IoT, robotics, AI, and society 5.0, focusing on strategic options like health, mobility, and infrastructure.
Sivamohan and Sridhar (2023)	2023		Development of BiLSTM-XAI intrusion detection system in safeguarding Industry 4.0 networks against diverse cyber threats.
Weber et al. (2023)	2023		The paper presents an overview of practical applications of XAI techniques to improve machine learning models, emphasising their potential to improve properties.

(Continues)

TABLE 1 | (Continued)

Authors	Year	Key technologies	Key findings
Wazid et al. (2023)	2024	AI and Healthcare 5.0	The authors proposed an incredible model, 'EIDS-HS' for detecting various types of intrusions in a healthcare system driven by 5.0. This model proves security against regular possible attacks.
Alsamhi et al. (2024)	2024		The proposed system, a collaborative style with AI and multiple healthcare institutions ensures patient privacy in medical science.
Ahmad et al. (2024)	2024		This work recaps all dimensions of cutting-edge AI technologies such as Industry 5.0, Healthcare 5.0, and Society 5.0 with its implementation.
Sharma et al. (2023)	2023		This work's main focus is sustainable research for medical multidisciplinary such as quick decision-making, digital integration, predictive analytics with AI, and Industry 4.0.
Date and Thalor (2023)	2023		Highlighted enhancing patient outcomes, robust assurance methods, and optimised resource allocation strategies for healthcare management with AI techniques.
Soni et al. (2023)	2023		This study uses a bibliometric strategy to interpret the interconnection dynamics between digital healthcare systems and AI approaches with regular parameters.
Saraswat et al. (2022)	2022		The authors presented a comprehensive study that supported electrocardiogram (ECG) monitoring with AI and Healthcare 5.0.
Karyamsetty et al. (2024)	2024	XAI and Healthcare 5.0	The author examined how explainable AI advances modern civilisations, integrating Society 5.0 with Industry 5.0 for digital progress.
Viswan et al. (2024)	2024		This study provides a systematic review focused on specific healthcare domains, including detecting Alzheimer's disease using XAI.
Hulsen (2023)	2023		Outline various impediments associated with executing explainable AI in the healthcare sector.
Chaddad et al. (2023)	2023		This work reviews the most delinquent explainable AI techniques utilised in healthcare and related medical imaging applications.
Bharati et al. (2023)	2023		Explored the reasons, strategies, and timing of applying these XAI models, as well as their impact on the healthcare field.
Wang, Chen, et al. (2023)	2023		Rudimentary cross-domain tools and techniques are employed. Four AI-based hospital recommendation procedures are examined in this study.
Saraswat et al. (2022)	2022		The authors proposed an XAI-enabled architecture for COVID-19 diagnosis, integrating CNN-based DL with FTL and an explainable diagnostic module.
Kose et al. (2024)	2024	AI-XAI with Healthcare 5.0	This work emphasises the application of XAI to healthcare issues, aspiring to improve trustworthiness, interpretation, and sustainability in diverse contexts.
Sindiramutty et al. (2024)	2024		Highlighted the complex issues of AI and XAI in medical applications and directed the improvement of deep-learning system elucidations using XAI principles to analyse medical images and text.

(Continues)

TABLE 1 | (Continued)

Authors	Year	Key technologies	Key findings
Parveen and Kannan (2024)	2024		This work investigates rare diseases with AI and XAI techniques in the Gulf region and various noncommunicable diseases common in Asian countries, especially those impacting the elderly.
Gomathi et al. (2023)	2023		The study probes the potentials and opportunities of Healthcare 5.0, such as personalised medicine, advanced diagnostics, telemedicine, and more patient-centric care, enabled by integrating advanced technologies–AI, big data analytics, and robotics.
Blenman et al. (2023)	2023		This study addressed the challenges and ethical concerns associated with emerging technologies—AI, XAI, and healthcare—and analysed possible solutions.
Dhiman et al. (2023)	2023		This work explores various XAI applications in healthcare, including diagnosis, therapy, prevention, and palliation, providing valuable insights for researchers in this area.
Shaikh et al. (2022)	2022		This study leverages AI explainability to enrich trust in the medical field and probes recent advancements in explainable AI, which supports creativity and is periodically essential in practice to increase awareness.

Note: These strategies are used to group the works, which are then reported chronologically.

Our review paper stands out for providing a thorough synthesis of AI and XAI outcomes in Healthcare 5.0. Unlike earlier research, which focuses on individual applications, our study creates an evident connection between AI and XAI in Healthcare 5.0, identifying emerging trends, integration issues, and future research areas. Furthermore, our survey is unique in that it combines technical methods with a strategic discussion about the future sustainability of AI-powered healthcare frameworks. This paper presents an important roadmap for improving AI-driven advances in healthcare by connecting theoretical innovations to practical applications.

2 | Survey Methodology

The research presents a detailed survey to allow methodical discussion of similar studies conducted in this scope. The study includes coverage of the works dealing with XAI in the Healthcare 5.0 framework. The authors also implemented the PRISMA approach to permit a systematic analysis ranging from identifying similar studies to the conclusion of each finding. Following the detailed search of relevant literature, the work presented adapts a Data Mining strategy to sort out the most appropriate studies.

2.1 | Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) Process

The aim of this research is to study a wide range of works that present the advances made in the field of Healthcare 5.0 implementing XAI. Some of the key terms included in this study are 'Explainable AI', 'XAI', 'XAI in healthcare', 'AI in healthcare', 'Interpretability in Healthcare AI', and 'Healthcare 5.0'. The focus remained to sort out the relevant literature from a category of trusted and renowned databases including PubMed, IEEE, ACM Digital Library, Springer, Google Scholar, Nature, and Wiley. The search query was constructed using Boolean operations on these keywords, focusing on topics like XAI in healthcare and AI transparency in Healthcare 5.0. The inclusion criteria for the selected papers are based on relevance, publication date (within the last 4–5 years), peer-reviewed journal articles, and articles published in English to ensure accessibility. Papers were excluded if they did not involve any AI or XAI applications or if they were published more than 6 years ago unless they included core concepts related to AI, XAI, or Healthcare 5.0. Non-peer-reviewed sources were also excluded to maintain the quality of the content. Initially, only titles and abstracts were screened, with papers falling outside the scope of the study being excluded at this stage. Following the initial screening, full-text papers were collected to assess their contribution to the understanding of XAI and its impact in the healthcare domain. The preferred reporting items for systematic reviews and meta-analyses (PRISMA) flowchart is shown in Figure 1.

2.2 | Data Mining Process

From the outset of this survey, the search and selection process was systematically designed by clearly defining both the main objectives and scope. The primary objective is to review and summarise the current status of Explainable AI (XAI) within the framework of Industry Revolution 5.0, specifically in Healthcare 5.0. This involves identifying recent advancements, addressing existing challenges, and providing insights into future research directions. To ensure a structured approach, the classification strategy follows a multi-stage methodology:

TABLE 2 | Literature review based on AI, XAI, and Healthcare 5.0 and reported chronologically.

Author	Year	Key technology	Methods	Dataset	Evaluation	Key focus	Summary
Kousha and Thelwall (2024)	2024	AI	Review	N/A	N/A	Academic peer review and publishing with automation.	The articles analysed the automation of publishing related documents like recommendation system, quality control, grant proposals, and their associated quality control assurance. The article's quality assurance through peer review text scores and their post-publication scores are also mentioned.
Kacena et al. (2024)	2024		Review	N/A	N/A	ChatGPT, peer-reviewed, scientific review articles.	The primary goal of this scientific study was to determine whether AI might be employed in a scientifically suitable way to improve the scientific writing process. Indeed, AI decreased the time required for writing while also introducing considerable errors. The latter means that AI cannot yet be utilised alone, although it might be employed with strict supervision by humans to aid in the drafting of scientific review articles.
Singh et al. (2024)	2024		Analysis	N/A	N/A	Sustainable Development Goals (SDGs), Bibliometric patterns, and concept evolution trajectories.	For SDGs the author analysed global AI research activity and how these have evolved over time focusing on the applications for SDGs for AI. Again, the investigation continues through the knowledge flows in AI research to identify the key application areas with the approaches and AI models to meet the objectives of SDG.
Dar et al. (2022)	2022		Analysis	MIAS, DDSM, TCGA, BreakHis, and INbreast	N/A	Analyse past research on breast cancer detection using ML, DL, and imaging modalities, addressing limitations and future challenges in applying DL to clinical practice.	Nowadays Breast cancer has become the most frequently seen cancer in women. The early detection of breast cancer reduces the risk of spread and improves the outcomes as well. Imaging advancements aid diagnosis but depend on costly, error-prone expert analysis. This study reviews past research on using ML, DL, and reinforcement learning for breast cancer detection explores imaging datasets, and addresses DL's limitations and future challenges in clinical use.

(Continues)

TABLE 1 2 (Continued)

Author	Year	Key technology	Methods	Dataset	Evaluation	Key focus	Summary
Alshahrani et al. (2024)	2024	AI and cloud systems in the IT industry	MCDM model and Delphi, AHP, DEMATEL, ISM, and MICMAC MCDM tools	N/A	N/A	Hybrid multi-criteria decision making, Fuzzy Delphi, Decision-making trial and evaluation, laboratory, Interpretive structural modelling, AI, sustainability.	To identify the key issues for building long-term AI with cloud systems for IT firms. The authors developed a fuzzy-based hybrid MCDM (multi-criteria decision making) model to overcome the challenges.
Nallakaruppan et al. (2024)	2024	XAI	RFR, Local Interpretable LIME, and PDP	Solar Power plant Dataset	the accuracy level of around 99.9% with a variance score of 0.99	XAI, AI, Solar Radiation.	In this study, authors introduced an XAI framework for explaining machine learning models' decision-making processes, which improves the predictability and control of solar energy distribution. The Local Interpretable Model-Agnostic Explainer (LIME) is developed to investigate the impact of crucial factors, like, solar irradiance, the temperature of the module, and ambient on energy generation.
Magd et al. (2024)	2024		Review	N/A	N/A	AI, Big Data, IoT, Smart Industrial IoT, Cloud Computing.	According to the Japanese strategy, the goal is to achieve long-term growth by implementing the IR 4.0 structure, which includes technology like IoT, robotics, and a combination of AI elements. Specifically, five tactical options for Society 5.0 were proposed: healthy living and increased lifespan, the creation of a value chain, smart infrastructure, smart cities, and model towns.
Sethi et al. (2024)	2024		Machine learning, Explainable AI (XAI)	Augmented UCI-1190-11 heart dataset	96.07%	Machine Learning, Explainable AI (XAI), and a user-friendly interface.	They presented a heart disease predictive model combining machine learning, XAI, and an interface to enable early detection, transparent decision-making, and improved management of cardiovascular diseases.

(Continues)

TABLE 1 2 (Continued)

Author	Year	Key technology	Methods	Dataset	Evaluation	Key focus	Summary
Kisten et al. (2024)	2024	XAI and Society 5.0	Review	N/A	N/A	Society 5.0, Industry 5.0, and modern societies.	This study proposed a paradigm that blends XAI with predicted Maintenance (PdM) to provide both predicted insights and explanations across four significant dimensions: data, model, result, and end user. The authors signalled a change in AI for agricultural sectors, and how these technologies are interpreted.
Kisten et al. (2024)	2024	AI and XAI	XAI with PdM in SAF	WoS and Scopus databases	LSTM: 5.81% rise. XGBoost: 7.09% F1 score, raised accuracy about 10.66% and a 4.29% ROC-AUC	SAF, XAI, with PdM, LSTM classifier, and XGBoost classifier.	Authors in this work integrated XAI and PdM based on 4 explainability areas which are data, model, end user, and the outcome. The ROC-AUC curve metric has been used to present the clear train and test results.
Saraswat et al. (2022)	2023	XAI and Health Care 5.0	Review and CNN-based DL techniques and FTL	IEEE Xplore, ACM Digital Library, PubMed	Accuracy 98%	Healthcare 5.0 ecosystem, EXAI-enabled medical image classification and segmentation, COVID-19 detection, explainable diagnostic module (XDM).	The authors suggested an end-to-end EXAI-enabled medical image classification and segmentation architecture that merges CNN-based DL approaches with FTL for COVID-19 identification. This suggested explainability module, the explainable diagnostic module which in short known as XDM, is utilised to analyse the classifier's prediction & offer a decision-making process. A specific use-case taxonomy for healthcare is also offered, combining several AI approaches with EXAI.

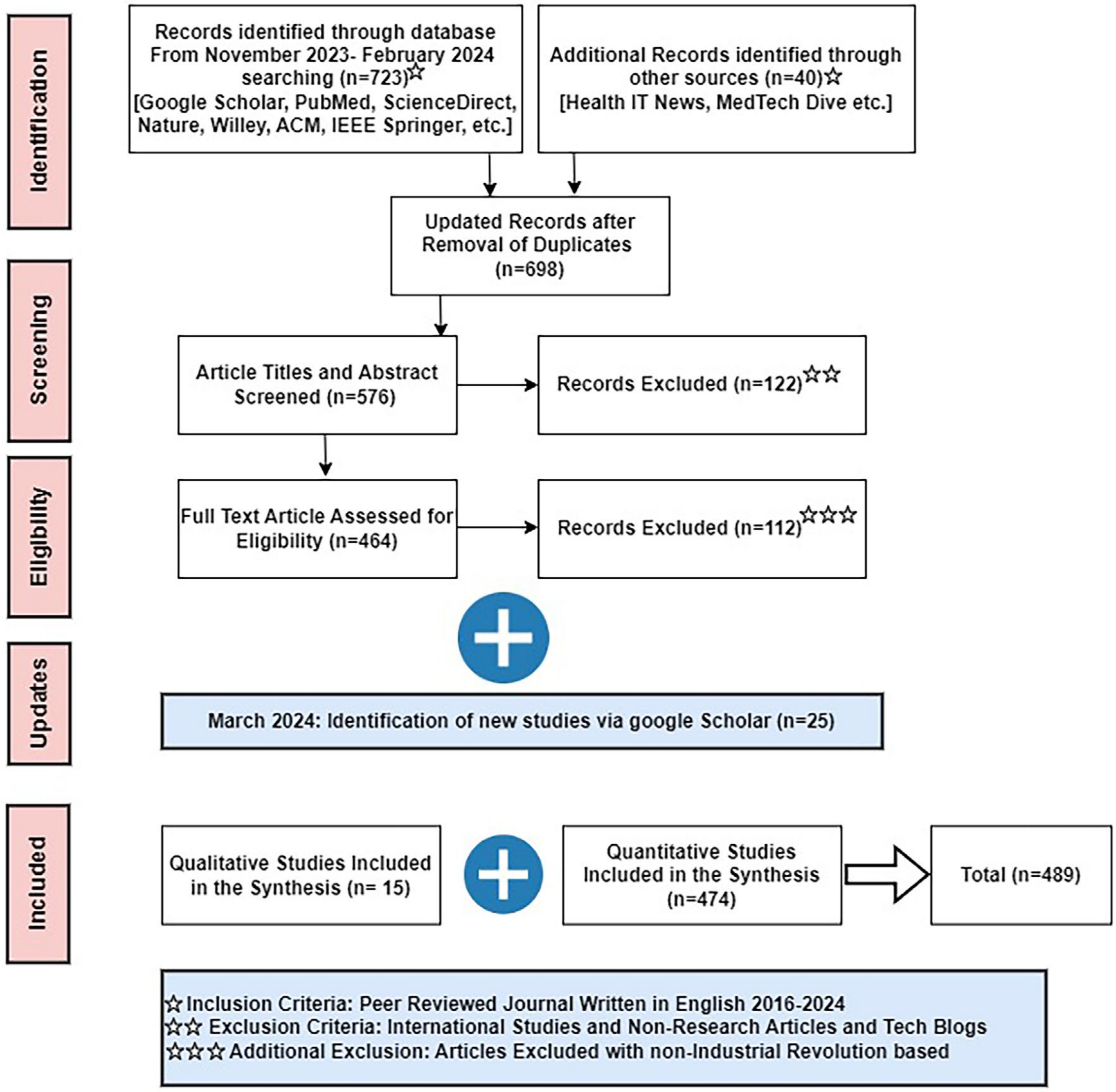

FIGURE 1 | PRISMA flowchart for the systematic review.

1. *Data Collection and Search Strategy*: Relevant literature was collected from reputable databases, including PubMed, IEEE Xplore, Scopus, and also from Google Scholar. Keywords such as 'Explainable AI in Healthcare', 'Healthcare 5.0', 'Industry 5.0 and AI', and 'XAI applications in medicine' were used to filter relevant studies. The search focused on publications from the recent 5 years to make sure the inclusion of the latest trends.
2. *Selection Criteria*: The selection criteria process was strictly confined to studies focusing on XAI and Healthcare 5.0. The study only considered high-impact factor journals, and research work presented in well-established journals. However, studies that did not present a clear application aspect of XAI or lacked sufficient experimental data were omitted.
3. *Classification and Categorisation*: The research papers surveyed were grouped as per the key attributes of XAI in a healthcare setting. During the research work, categories were subsequently refined when a new theme emerged.
4. *Data Synthesis and Trend Analysis*: The literature works were thoroughly studied to figure out critical details, patterns, and existing knowledge gaps. During the survey, the key findings were constantly updated. This allowed the incorporation of the most recent advances made in the field of Healthcare 5.0 using XAI.

This technique permitted the survey to carry out a detailed landscape of existing research works concerning XAI in Healthcare 5.0. Figure 2 illustrates the technological progression of the healthcare industry from 1.0 to 5.0 and beyond.

3 | Artificial Intelligence, Explainable Artificial Intelligence, and Healthcare 5.0 System: Approaches of Key Concepts and Features

The concept of AI has fundamentally changed the dimensions of various industries, and its integration into healthcare has the potential to improve efficient healthcare systems significantly.

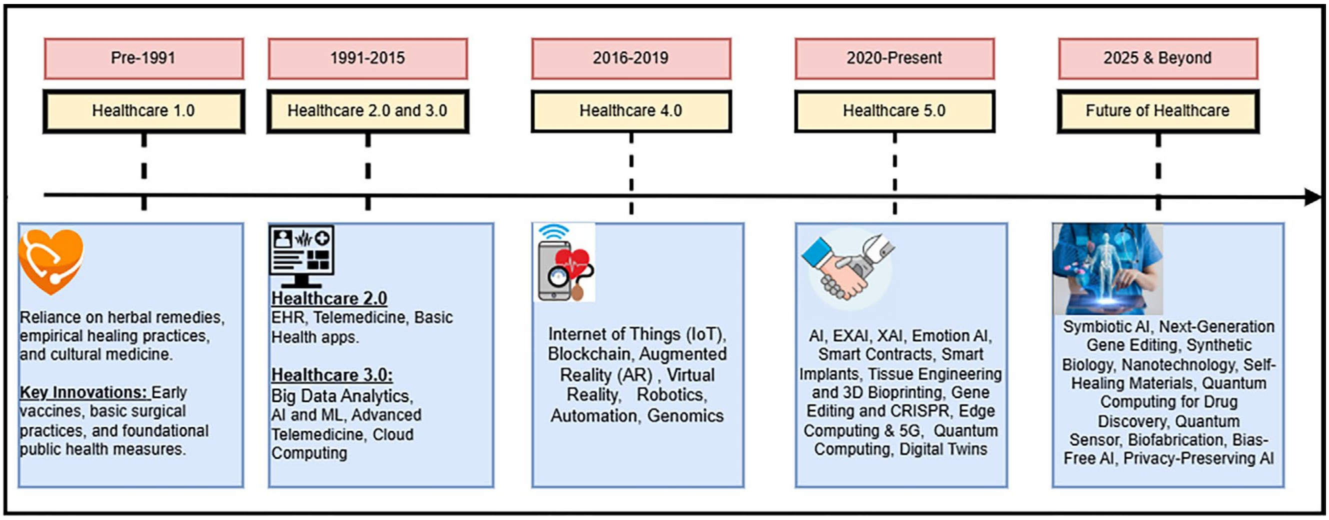

FIGURE 2 | The evolution of Healthcare from 1.0 to 5.0 and beyond.

AI systems are designed to mimic human cognitive functions, enabling them to perform tasks that would typically require human intelligence, which has led to the development of intelligent systems that can process vast amounts of medical data, assist in diagnostic processes, and predict patient outcomes.

3.1 | Artificial Intelligence (AI)

The Impact of AI in healthcare lies in its ability to process and analyse large datasets, which has proven invaluable for addressing the complexities of modern medicine. From diagnostic imaging to personalised treatment plans, AI's application has shown immense promise by encompassing various methodologies and technologies.

3.1.1 | Key Concepts of AI

The application of artificial intelligence in healthcare offers promising solutions to improve patient services. The healthcare sector is poised for significant transformation, with its budget expected to exceed USD 158 billion in the coming years (Baker and Xiang 2023). A major challenge for medical professionals using IoT-connected devices is processing the vast amounts of data these devices generate in order to extract valuable information. This challenge has led to the adoption of IoT systems powered by AI, contributing to the shift toward Healthcare 5.0 (Mbunge et al. 2021). AI-driven healthcare systems enable remote patient monitoring, intelligent management, personalised medicine, and other key features, overcoming the limitations of Healthcare 4.0 Despite the promise of AI-IoT integration, it remains underexplored. A study in Qadri et al. (2020) surveyed AI applications in healthcare but did not address the explainability of AI systems. Another study (Markus et al. 2021) focused specifically on the explainability issue in AI healthcare applications. AI techniques, such as convolutional neural networks (CNNs), have been applied to medical image processing (Shin et al. 2016) and neurodevelopmental predictions (Saha et al. 2020), as well as detecting COVID-19 (Jia et al. 2021; Rahman, Chakraborty, et al. 2022). Support vector machines (SVMs) have been used to diagnose heart conditions based on variability parameters (Goumopoulos and Potha 2023) and detect COVID-19 from lab results (Abdulkareem et al. 2021), as well as identifying Parkinson's disease through voice patterns (Lahmiri and Shmuel 2019). Long short-term memory (LSTM) networks have gained popularity in healthcare research, being used for tasks like identifying cardiac arrhythmias (Yildirim et al. 2019) and detecting epilepsy (Xu et al. 2021). Autoencoder models have also shown promise in medical imaging applications, such as creating noiseless retinal tomography images (Tajmirriahi et al. 2021) and enhancing MRI resolution (Park et al. 2021). Additionally, autoencoders have been used to identify COVID-19 from chest x-rays (Shome et al. 2021). Random forest (RF) models have been applied in healthcare to improve diagnostic accuracy, including breast cancer detection (Wang et al. 2020) and forecasting mortality in kidney disease patients (Lin et al. 2019). Attention models, which complement traditional CNN and LSTM models, are also gaining traction in healthcare research. They have been used to detect dementia (Kherchouche et al. 2022), classify abnormalities in ECG recordings (Wang et al. 2021), and assess heart rate from video data (Guo et al. 2021). However, attention models have faced criticism for inconsistencies in performance, such as mismatches in predicting diabetes based on health features compared to general statistical data (Kim et al. 2020).

3.1.2 | Key Features of AI

The rise of IoT has led to an increasing number of smart measuring devices (Helbing et al. 2019). To extract meaningful information from this vast amount of data, it is essential to implement intelligent systems. This is where AI comes into play, performing tasks such as knowledge gathering, forecasting, learning, and executing actions with the ability to reason and manipulate results. AI can identify errors and evaluate situations autonomously, without user intervention. AI dramatically boosts performance by advancing conventional data-driven technologies, producing results quickly and without delays. AI encompasses several subfields, including ML, CNN, and DL. Supervised Learning is a type of ML that uses labelled data and applies past learnings to new datasets. It relates input to output, and when errors occur, it learns from feedback to improve its performance (Jiang et al. 2017; Sikder, Das, and Chakma 2021). In contrast, Unsupervised Learning uses unlabelled data, with no direct correlation between input and output. As a result of its complexity, unsupervised learning is less practiced than supervised learning. This system includes techniques like clustering and anomaly detection (Sikder, Sarek, et al. 2021; Sikder, Das, and Anwar 2021). Semi-supervised Learning integrates elements of both supervised and unsupervised learning, making it preferable for scenarios where certain data is labelled and some not. This ML algorithm trains on labelled data and applies its findings to unlabelled data (Jiang et al. 2017). Finally, Reinforcement Learning is an ML algorithm that presents high-accuracy output and penalises errors. This approach involves generating efficient output while reducing negative consequences, requiring significant interaction with the environment to achieve optimal results.

3.2 | XAI

AI has demonstrated significant capabilities in healthcare, it often operates as a black box, making it challenging to understand the rationale behind its decision-making process for the medicator. This is where XAI comes in aiming to make AI systems more transparent and interpretable, enabling healthcare professionals to trust and verify AI-driven decisions.

3.2.1 | Key Concepts of XAI

In recent years, the application of XAI in clinical settings has received great interest among researchers. Authors in Born et al. (2021) improved the detection of infected lungs through explainable ultrasound image analysis. Again, the analysis in Jia et al. (2021) showcased a modified algorithm for classifying chest x-ray images and promoted a new ResNet architecture for computerised tomography (CT) image classification. The results demonstrated improved accuracy in detecting COVID-19, tuberculosis, and both viral and bacterial pneumonia. In Shen et al. (2021), the authors introduced a novel neural network algorithm to identify unique features of medical images. This model successfully detected benign and malignant lesions in a breast cancer screening dataset. A deep learning-based model developed in Song et al. (2021) aimed to improve COVID-19 detection among patients. The study analysed CT scans from patients with COVID-19, bacterial pneumonia, and healthy individuals for comparative purposes. The model effectively extracted significant features to assist medical personnel in making accurate diagnoses. Another study in Lu et al. (2022) explored the use of AI for accurate COVID-19 diagnosis using CT images. The authors proposed a novel neighbouring-aware graph neural network algorithm to classify COVID-19 cases. In Alsinglawi et al. (2022), an explainable AI approach was presented to predict patients' hospital stay durations. Researchers in Sutton et al. (2022) developed an AI-assisted diagnosis mechanism for ulcerative colitis in endoscopy images. The study used XAI techniques combined with gradient-weighted class activation maps to improve the interpretability of medical images. The XAI method used in Hu et al. (2022) helped medical personnel better understand chest x-ray images. This study focused on the chest x-ray dataset and the ISIC 2017 skin lesion dataset to identify features indicating the need for urgent intervention for COVID-19 patients requiring emergency treatment. In Du et al. (2022), XAI was used to predict gestational diabetes mellitus in patients. Another study in Moncada-Torres et al. (2021) employed XAI to improve the predictability and understanding of breast cancer patients' conditions.

3.2.2 | Key Features of XAI

XAI is a research domain focused on making the results of AI systems easier to understand. This field has gained renewed interest from researchers due to the growing application of AI and ML in industrial contexts. XAI provides a means to interpret detailed information, which is essential for important decision-making. From social, ethical, and legal perspectives, it has become increasingly important to develop models that make decisions in an explainable, trustworthy, and understandable way. From a technical standpoint, XAI lacks a standard definition. It generally refers to efforts aimed at making AI systems more transparent, thereby addressing trust issues. According to Gunning and Aha (2019), XAI ensures that AI systems remain highly effective while also enabling users to understand, trust, and manage the next generation of AI tools. XAI encompasses several key features, including interpretability, accountability, responsibility, transparency, explicitness, and faithfulness. Interpretability is closely related to explainability, both referring to systems that users can comprehend. XAI must also consider societal norms, as well as moral and ethical values, to be deemed responsible. Accountability involves the need to explain and justify decision-making processes to users. Transparency refers to the necessity of clarifying, inspecting, and replicating the processes through which XAI systems make decisions and adapt to their environments. Responsibility highlights the obligation of both the system and its creators to address preferences, errors, or unforeseen outcomes. Faithfulness is another crucial feature of XAI, emphasising the importance of developing models that are both credible and true to their underlying algorithms. Unfaithful explanations fail to provide a clear and accurate description of the implemented algorithm, leading to ineffective or misleading information (Rudin 2019; Rahman, Islam, et al. 2020).

3.3 | Healthcare 5.0 System

Healthcare 5.0 represents a shift toward more efficient and patient-centric healthcare by integrating AI, the IoT, and smart wearables into healthcare systems. This section focuses on Healthcare 5.0, which aims to enhance the quality of healthcare services by improving patient outcomes, enabling remote patient monitoring, and fostering better disease prevention strategies.

3.3.1 | Key Concepts of the Healthcare 5.0 System

Healthcare 5.0 was introduced to transform traditional medical services comprehensively, making healthcare systems more efficient, convenient, and secure while ensuring privacy. Existing research has explored various aspects of Healthcare 5.0 and its benefits. In Wazid et al. (2022), the authors discuss the security framework for Healthcare 5.0, detailing implementations, key concerns, and future directions for security requirements. The functionalities and capabilities of sensors, 5G technology, blockchain-based algorithms, IoT, and AI, as well as their potential to transform the healthcare industry, are examined by researchers in Mbunge et al. (2021) and Rahman, Wadud, et al. (2024). This gap in existing research is addressed by the authors in Saraswat et al. (2022), who emphasise the necessity of Explainable AI (EXAI) in Healthcare 5.0, focusing on its challenges and opportunities. Another study in Chi et al. (2023) explores Healthcare 5.0 in the context of fog/cloud computing, offering a novel forecast for its innovative application in Industry 5.0. Security concerns in Healthcare 5.0 are critically addressed in Rehman et al. (2022), where the authors present a secure Healthcare 5.0 system that implements blockchain and federated learning to detect malware attacks, enabling medical professionals to treat patients securely. Further research in Wu et al. (2024) and Rahman, Debnath, et al. (2024) proposes a Blockchain-based secure Healthcare 5.0 system for medical surgeries. This methodology improves patient data security, improves efficiency, and ensures better access to healthcare services. A patient-centric approach to Healthcare 5.0 is explored in Gomathi et al. (2023), focusing on the challenges and opportunities of applying Industry 5.0 in healthcare. Researchers in Natarajan et al. (2023) propose an energy-efficient Healthcare 5.0 system that addresses security concerns while enhancing energy efficiency using a novel elliptic curve cryptography protocol. In Khan et al. (2023), IoMT coupled with transfer learning is used to detect cancer in a Healthcare 5.0 setting. Healthcare 5.0 also supports the efficient operation of medical institutions, as demonstrated in Bhavin et al. (2021), and helps identify the risk of critical illnesses, reducing the likelihood of disease development. This process is carried out error-free, as noted in Garg et al. (2022).

3.3.2 | Key Features of the Healthcare 5.0 System

Healthcare 5.0 represents a transformation from traditional healthcare systems, emphasising a patient-centric approach, the integration of advanced technologies and sensors to improve patient well-being, and a focus on disease prevention. A key component of Healthcare 5.0 is the effective remote monitoring of patients, achieved through the use of sensors, health-based applications, and smart wearables, as shown in Figure 3. For example, the use of smart wearables to monitor health parameters and detect COVID-19 is detailed by the authors in Seshadri et al. (2020). Other important aspects of remote patient monitoring include tracking oxygen saturation, respiratory issues, sleep patterns, heart rate, and more. These systems notify patients about their health status and help prevent the progression to critical stages. Virtual clinics also play a role in reducing the spread of diseases by minimising direct patient contact. Additionally, they improve patient care by reducing waiting times, allowing patients to receive treatment from the comfort of their homes (Gordon et al. 2020). Another core feature of Healthcare 5.0 is precision medicine. Leveraging AI and big data analytics, Healthcare 5.0 has the potential to improve diagnoses and provide tailored treatments for patients. With access to comprehensive data, healthcare providers can deliver personalised diagnoses based on individual patient needs. Healthcare 5.0 also emphasises wellness monitoring, offering real-time feedback on health conditions and enabling prompt intervention when needed. This feature is crucial for both disease treatment and overall individual fitness. Improved interoperability is another significant advantage of Healthcare 5.0. In the future, healthcare systems will integrate advanced technologies to better meet patient needs by seamlessly connecting healthcare providers, medical data, and patient information. This is a stark contrast to traditional healthcare models, making Healthcare 5.0 a more precise and forward-thinking framework (Rahman, Nasir, et al. 2020). In summary, Figure 4 presents the number of research publications on AI, XAI, AI in Healthcare 5.0, and XAI in Healthcare 5.0 across various Scopus-indexed journals.

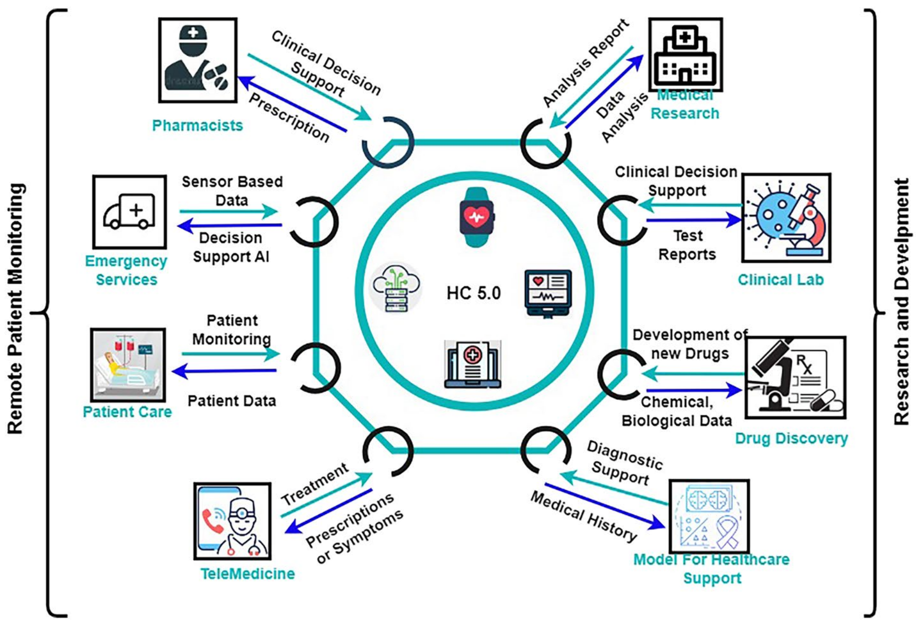

FIGURE 3 | Healthcare 5.0 features and scopes under remote patient monitoring, research and development field.

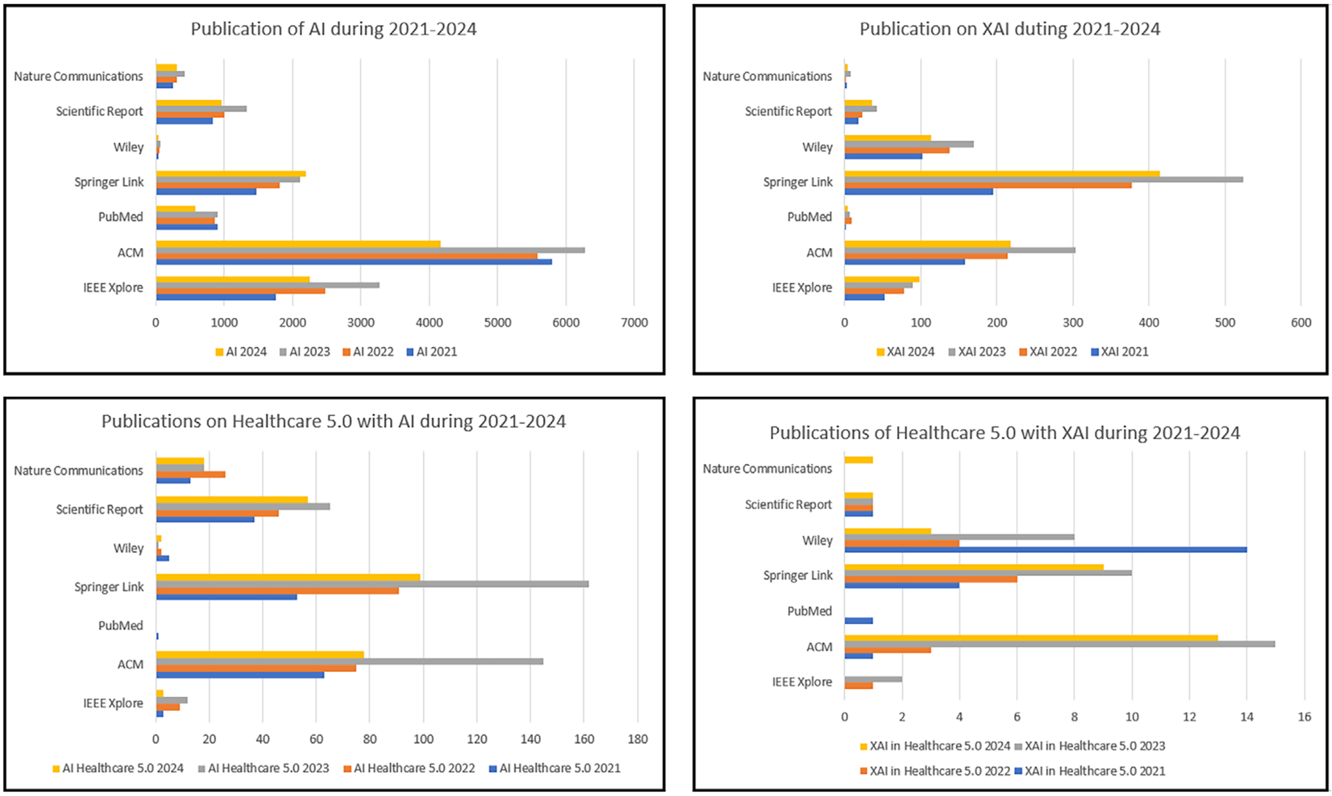

FIGURE 4 | Number of research published in AI, XAI, AI in Healthcare 5.0 and XAI in Healthcare 5.0 in different Scopus indexed journals.

4 | AI and XAI in the Healthcare 5.0 System: Current State-of-the-Art

The integration of AI and XAI within the context of Healthcare 5.0 is paving the way for a new era of patient care due to the healthcare landscape continues to evolve. This section explores the current advancements and state-of-the-art applications of AI and XAI in Healthcare 5.0, highlighting the transformative potential these technologies hold for the future of healthcare.

4.1 | AI in Healthcare 5.0

The next phase of artificial intelligence's evolution in the healthcare industry is represented by AI in Healthcare 5.0. This paradigm leverages cutting-edge technologies, including biotechnology, augmented reality, and quantum computing to unlock extraordinary opportunities and address unique challenges. A unique opportunity for AI remains to explore accelerated drug detection and a patient-centric approach that permits customised treatment for each patient. The implementation of AI however remains challenging with concerns regarding data security, the need for a legal framework, and the result of biased decisions made by AI. The full-phased implementation of AI further requires investments in a trained workforce and infrastructural changes in medical facilities. The research presented in Date and Thalor (2023) and Rahman, Islam, Montieri, et al. (2023): presents some answers to the obstacles faced through AI integration in Healthcare 5.0.

1. Differential Privacy Algorithm for Data Security and Privacy: The following solution addresses a critical security concern where no patient data is identified. How it assists: This technique allows patient data to be secured during data analysis and training AI models.
2. Transfer Learning Algorithm for Data Availability and Quality: This feature includes training an AI algorithm through a single dataset and further refining it through a different dataset. How it assists: This allows AI models to gain knowledge from a wide range of datasets to develop predictions with limited data.
3. XAI Algorithm Regulatory Compliance: Different models of XAI including LIME and SHAP enhance the clarity and transparency aspects of AI model decisions. How it assists: The regulatory compliance issue is dealt with by XAI through a transparent explanation which further fosters trust between patients and caregivers.
4. Fast Healthcare Interoperability Resources (FHIR) Algorithm for Interoperability: This feature improves the treatment plans by providing patients with real-time data. This further permits reduced errors and redundancy issues in data or medical records of patients. The end result is that clinicians are able to make more informed decisions leading to better outcomes. On top of this, FHIR permits AI-driven applications such as predictive analytics and personalised medicine, enhancing diagnostic accuracy and patient monitoring. FHIR also permits for automated data transfer between EHR systems, test facilities, and insurance parties. This permits a seamless operation that reduces manual data input and mitigates operational costs.
5. Fairness-aware Algorithms for Ethical AI: A principal objective of AI should be to reduce biases and discrimination

in AI predictions. This feature allows AI algorithms to focus on fairness issues such as adversarial debiasing and reweighted loss functions. How It assists: The algorithms equipped with this feature address critical AI concerns such as ensuring fairness and equity, dealing with ethical concerns, and generating fair healthcare outcomes. These recommended algorithms provide solutions to key challenges in integrating AI into healthcare systems, supporting more responsible and efficient AI applications in this important field. However, it is essential to adapt the chosen algorithms to the unique challenges and use cases specific to healthcare environments.

4.2 | XAI With Healthcare 5.0

Healthcare 5.0 is the next big advancement in the field of medical facilities for patients. It ensures a patient-centric approach generating better results, predictive behaviour, tailor-made patient care, and improved trust and transparency among its stakeholders. Previously, AI-based decisions lacked clarity and its wide-scale adaptation faced hindrance due to their black box feature. However, Healthcare 5.0 powered by XAI presents patients with a detailed insight into how conclusions are reached in AI-driven technology. The improved feature offered by XAI is detailed in the following section:

1. Improved Decision Making: Medical practitioners provides care to their patients. In an advanced healthcare setting, where AI-based decisions are utilised to prepare treatment plans for patients, it is important for the doctors to understand the reasoning behind the information provided by AI. XAI assists the medical personnel with proper tools and guidance while also delivering the rationale behind its proposed treatment process.

2. Improved Trust and Adoption: XAI provides clarity in its decision making process. This fosters trust among the stakeholders using the XAI algorithm in a healthcare setting. As a result of this enhanced trust between medical personnel and patients, the acceptance of AI in the healthcare environment becomes more feasible.

3. Patient Empowerment: The diagnoses of a patient in a Healthcare 5.0 setting utilising XAI is more insightful. This clarity in how their treatment plan is going to proceed enables a more fruitful collaboration among all parties. The patient as such actively participates in how the medical facility plans to proceed with their treatment plan.

4. Regulatory Compliance: It is important for compliance to ensure that there is room for clarity and accountability when it comes to AI-based decisions, particularly in a healthcare setting. The importance of such compliance is simplified with XAI, as it is able to provide a transparent rationale behind its decision-making. Thus, XAI and its transparent feature allow for a healthier environment while also ensuring precautionary standards are adequately met. That healthcare systems meet necessary safety and efficacy standards.

5. Continuous Learning and Improvement: The clarity pattern of XAI allows medical personnel to accurately categorise probable causes for error within the decision-making process. This transparency over a certain period of time enables medical caregivers to train the model with more accurate data so that it can generate fitting treatment plans for diverse patient groups.

6. Ethical Considerations: Healthcare 5.0 places a strong emphasis on ethical principles such as accountability, transparency, and fairness. XAI promotes these values by empowering stakeholders to recognise and address ethical concerns related to AI in healthcare. In order to guarantee that AI technologies meet the standards and ideals of patients, healthcare professionals, and society at large, XAI promotes ethical AI practices.

7. Research and Innovation: XAI supports collaboration between researchers, data scientists, and healthcare professionals, driving innovation in the healthcare sector. Through providing interpretable insights into AI models, XAI accelerates research on diseases, treatment outcomes, and healthcare interventions, ultimately enhancing patient care (Saraswat et al. 2022).

Example (Diabetes Management with Healthcare 5.0 and Explainable AI): In the Healthcare 5.0 paradigm, the integration of Explainable AI improves personalised treatments and supports greater trust among patients in the decision-making process. For instance, in diabetes management, continuous glucose monitoring (CGM) devices are paired with AI algorithms to predict blood sugar levels and provide real-time, personalised treatment recommendations.

- Real-Time Monitoring and Prediction: Continuous data streams from CGM devices provide real-time blood glucose levels. An AI algorithm processes these data to predict short-term glucose patterns, such as the likelihood of hypoglycaemia or hyperglycaemia within the next hour.

- Explainable Recommendations: Using XAI techniques, the AI model identifies the key factors contributing to glucose fluctuations, such as recent meals, exercise, and insulin doses. For instance, the system might provide a recommendation like: 'Your glucose level is expected to drop in the next 30 min due to recent insulin administration', the system would suggest. To avoid hypoglycaemia, think about eating 15 g of carbs.

- Integration with IoT Devices: IoT-enabled insulin pumps automatically adjust insulin delivery based on AI predictions, ensuring optimal glucose management.

- Integration with Wearable and IoT Devices: Wearable devices and IoT-enabled insulin pumps work together to adjust insulin delivery based on AI predictions, ensuring effective glucose control while minimising the need for manual intervention.

- Patient and Clinician Transparency: XAI interfaces allow doctors to verify recommendations and help build trust with patients by providing visual explanations of predictions, such as the impact of meals or physical activity on glucose levels.

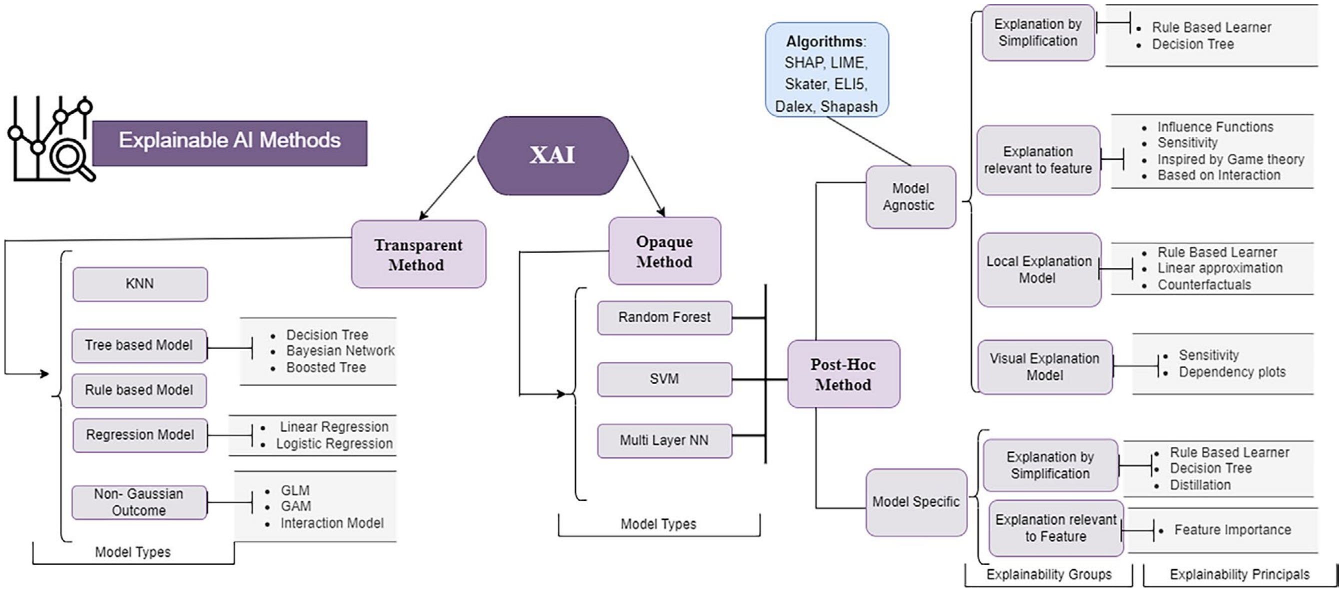

FIGURE 5 | 'Explainable AI' approaches with algorithms, and corresponding methods.

- Real-World Implementation: In practice, AI-driven tools like these are already being used in clinics to support endocrinologists in adapting treatment plans, adjusting insulin regimens, and educating patients about self-care. A randomised trial at XYZ Hospital found that using this AI-based approach reduced hypoglycaemia incidents by 25%.

Clinical Impact: The patient-centric approach of XAI in a Healthcare 5.0 setting provides a more accurate decision-making model. This, in turn, improves patient health conditions as each individual is treated with custom treatment plans. The transparency, and clarity aspects of XAI are critical for applications in areas such as healthcare, finance, and similar autonomous systems. XAI has the potential to perform a crucial role in many applications, including healthcare facilities. Figure 5 illustrates the various approaches to XAI.

5 | Implementation Perspectives of AI and XAI in the Healthcare 5.0 System

The implementation of AI and XAI in Healthcare 5.0 is driving a significant transformation in the way healthcare systems operate. As we move from Healthcare 4.0 to Healthcare 5.0, integrating advanced technologies such as AI and robotics is key in enhancing patient care and healthcare efficiency. This section explores various implementation perspectives of AI and XAI in the Healthcare 5.0 system, focusing on how these technologies are applied to healthcare delivery.

5.1 | Approaches to Implementing AI in the Healthcare 5.0 Field

Healthcare 4.0 is built on key technologies such as AI, big data, and the IoT, which are used to perform various significant tasks. With technological advancements, Healthcare 5.0 extends this framework by integrating cutting-edge techniques such as robotics, nanotechnology, and model explainability. These modern technologies assist healthcare professionals in providing improved services to both patients and clinical staff. Among these smart technologies, AI plays a significant role in advancing Healthcare 5.0. AI algorithms can analyse large datasets, significantly impacting patient monitoring, and treatment, while helping to reduce operational challenges. Leveraging AI-based methods, future disease prediction and potential solutions are also becoming increasingly feasible. Furthermore, AI can aid in image analysis, such as detecting abnormalities in CT scans and other medical images. This allows healthcare professionals to make informed primary decisions and plan further procedures. AI-based systems also enable remote patient monitoring by analysing real-time health data collected through various sensors and AI-powered platforms. Additionally, AI models can assist in the cost analysis of hospitals and predict patient monitoring needs and length of stay, helping hospital staff optimise resources and forecast future utility requirements. Figure 6 shows the implementation perspectives of AI and XAI in the Healthcare 5.0 System.

5.2 | Implementation Approaches of XAI in the Healthcare 5.0 Field

Healthcare is undergoing a significant transformation, and we are now entering the era of Healthcare 5.0. The integration of modern, sophisticated technologies is essential for ensuring patient satisfaction, as they provide comprehensive insights into the results generated by various models. In healthcare, transparency and reliability are paramount. XAI plays a crucial role in offering clear explanations of model predictions by identifying which input factors contribute to positive outcomes and which lead to negative ones. Table 3 illustrates the application of XAI in Healthcare 5.0 across various domains. In this context, XAI ensures transparency and builds trust in the system's predictions. Researchers in the field of Healthcare 5.0 have highlighted the

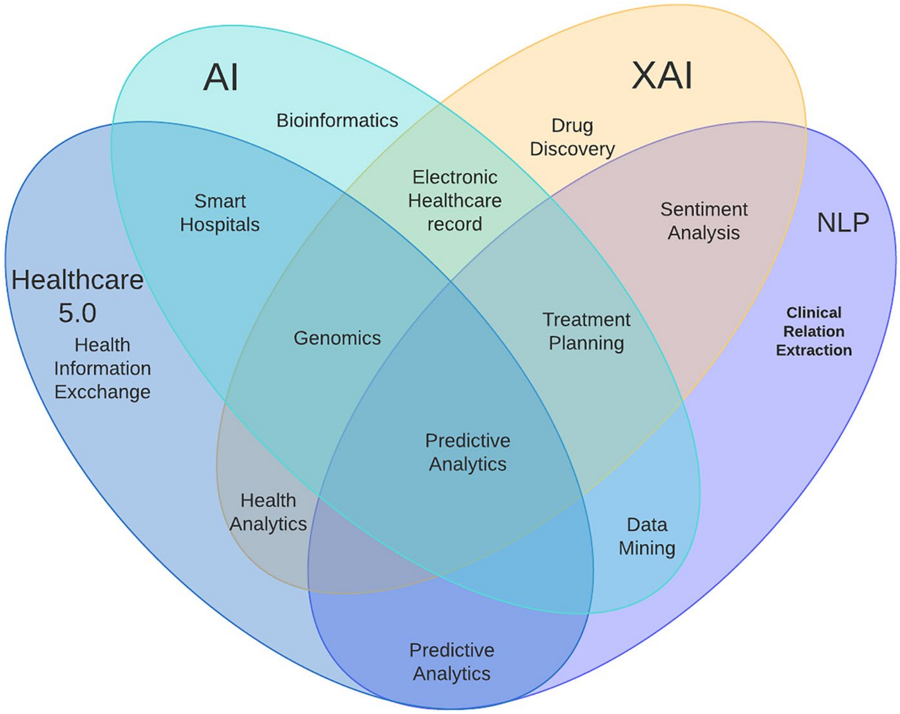

FIGURE 6 | Implementation perspectives of AI and XAI in Healthcare 5.0 System.

following strategies for deploying XAI (Kiruthika et al. 2024; Arslanoğlu and Karaköse 2024; Khan and Naim 2024; Jain et al. 2024):

- *Model Interpretability*: Ensuring the interpretability of models without sacrificing performance is essential in healthcare operations (Hassija et al. 2024; Dablain et al. 2024). The healthcare industry often relies on ML and DL models for prediction tasks; however, interpreting these models can be challenging due to their non-linearity and complex network structures. XAI addresses this challenge by providing insights into a DL model's decisions. Activations across different neural network (NN) layers help the model understand the features learned during training. One way XAI aids in this process is through feature visualisation and statistical significance analysis. NN models have several layers, and layer-wise relevance testing is significant for identifying which parts of the input contribute most to the output. Through propagating from the output to the input, this process helps reveal the relevance of each feature. Additionally, feature variation impacts the model's predictions. A Partial Dependency Plot (PDP) illustrates the relationship between specific features and the model's forecast (Molnar et al. 2023; Parr et al. 2024), shedding light on how changes in a given attribute affect the prediction. For models utilising attention mechanisms, visualising attention weights helps identify the most relevant portions of the input sequence. Combining XAI approaches, DL and ML models can be made more understandable, thereby increasing trust in their decisions (Arreche et al. 2024).
- *Model Explainability*: Understanding how a model reaches its conclusions is crucial, especially when using complex neural networks like deep learning models. XAI can improve model explainability by showing how input characteristics influence predictions. Techniques such as SHAP or LIME can be employed (Alabi et al. 2023; Aldughayfiq et al. 2023). LIME, for example, creates surrogate models that help users better understand the predictions based on individual instances, providing a local approximation of the model's decision boundary around those instances (Cakiroglu et al. 2024).
- *Visualisation and Dashboards*: XAI is important for visualising model predictions in the medical domain. Activation maps, for example, visualise the activations of various CNN layers, helping healthcare professionals understand which features are being learned and processed by the model. Plots that display feature significance highlight the most influential factors in the model's predictions (Ha et al. 2024). Gradient-based approaches can also be employed to improve model performance by incorporating the effects of input data on model predictions. Additionally, XAI is valuable for developing dashboards that display performance metrics such as confusion matrices, accuracy, specificity, and ROC curves, aiding in the evaluation of the model's overall performance. These dashboards allow healthcare professionals to adjust parameters and monitor how changes impact predictions in real-time, ensuring that model outputs remain relevant and up to date (Maddigan et al. 2024). Moreover, XAI generates comprehensive reports with justifications for the model's outputs, offering transparency and enhancing trust among medical professionals (Bhattacharya et al. 2024).
- *Quantifying Uncertainty*: In healthcare, quantifying uncertainty is crucial, as decisions can have significant consequences. XAI techniques, combined with uncertainty quantification methods, can make machine learning models

TABLE 3 | Application of AI and XAI in Healthcare 5.0.

Objectives	Work	Dataset	Model	Hyperparameters used
Patient length of Stay prediction	Ma et al. (2020)	The Data collected from the PhysioNet website containing 4000 records of ICU patient data.	One-class JITL-ELM	AUC, G-Mean accuracy, specificity, and sensitivity
	Chrusciel et al. (2021)	Centre Hospitalier de Troyes hospitals' admitted patients between 01/01/2023 and 24/09/2023	Random Forest	Recall, Specificity, Precision, Accuracy, *F*1 Score
	Rahman, Kundu, et al. (2022)	New York State Department health dataset of 2015 of 10 hospitals	Federated Learning with Lasso, Ridge and Linear Regressor	MAE, RMSE, R-2
	Saadatmand et al. (2023)	963 ICU admitted patients of Iranian hospital between 7/09/2024 and 07/03/2021	XGB, KNN, Random Forest, bagged-CART, and LogitBoost	AUC, Sensitivity, Accuracy
	Kadri et al. (2023)	Patient data from Lille hospital from 01/01/2011 to 31/12/2012	GAN, AdaBoost	RMSE, MAE, MAPE
Cancer detection	Jopek et al. (2024)	Liquid Biopsy Data	Deep NN with XAI	Accuracy, Sensitivity, Confusion Matrix
	Demir et al. (2024)	130 samples of whole blood cell and urine	ResNet and Deep NN	AUC, Accuracy, Sensitivity, Specificity
	Keskenler et al. (2024)	International Skin Image Collaboration	Multilayer machine learning Model with Decision tree, Neural network, SVM	Accuracy, Precision, *F*1 Score, Recall
	Sivakumar et al. (2024)	(ISIC) Dataset	CNN with ResNet50	Accuracy, *F*1 score
	Mohanty and Das (2024)	HAM10000 dataset from Kaggle	Rule based neuro Fuzzy classification	Accuracy, Confusion Matrix.
	Hernandez et al. (2023)	(ISIC) Dataset	ResNet50	Accuracy, Confusion Matrix.

(Continues)

TABLE 3 | (Continued)

Objectives	Work	Dataset	Model	Hyperparameters used
ICU care	Moulaei et al. (2024)	Patients with methanol poisoning from Loghman Hakim Hospital in Iran with a total of 897 samples and among them 202 cases required intubation	CNN, LSTM, DNN, FNN, SVM, XGB	Accuracy, specificity, and sensitivity ROC, SHAP & Lime value
	Alsinglawi et al. (2024)	Real time Hospital Data and Al-Ain hospital EHR	XGBoost, RF, LR, MLP, GB	Recall, Specificity, Precision, Accuracy, AUC, PR-AUC, K-Fold 'Cross-Validation' Score
	Moreno-Sánchez et al. (2024)	Medical Information Mart for Intensive Care database (Total 425,000 samples)	XGBoost, XAI, RF, LR	AU-ROC, Accuracy, SHAP
	Joshi et al. (2024)	ICU Patient dataset	ANN	Accuracy, Sensitivity
	Kadri et al. (2023)	ICU Patient data of tertiary university	LSTM, MP	MSE, Accuracy, Confusion Matrix
Sepsis management	Zhang et al. (2024)	MIMIC-III Dataset	LR, RF, LSTM, Gradient Boosting	AUC accuracy
	Bukkapatnam et al. (2024)	Open source dataset of sepsis and not sepsis images	CNN	Accuracy, Loss
	Iqbal et al. (2024)	388 sepsis patient data	Naive Bayes, RF, Bagging, LR, and J48 models	ROC, Accuracy
	Liu et al. (2024)	MIMIC-IV	RL (Q-learning, Monte Carlo, and n-step TD)	SOFA scores
Pox virus detection	Kundu et al. (2024)	MPox Dataset with 1300 samples	FL, GAN, ViT, MobileNetV2, ResNet50	Accuracy, Precision, Recall, *F*1 Score
	Bukkapatnam et al. (2024)	MSID, DermNet	attention-based MobileNetV2, LIME, Grad-CAM	Cohen's Kappa Score, Youden's *J* Index, Accuracy, Sensitivity, *F*1 Score
	Raha et al. (2024)	MCSI Dataset	MobileNetV3Large, LIME, Grad-CAM	*F*1 Score, Recall, Sensitivity, Precision, Accuracy
	Dihan et al. (2024)	Monkeypox Skin Lesion Dataset	VGG16, DenseNet121, and ResNet50	*F*1 Score, Recall, Sensitivity, Precision, Accuracy, Loss
	Kundu et al. (2022)	MPox Dataset with 1300 samples	ViT, KNN, SVM	Accuracy, Sensitivity, *F*1 Score

more understandable and reliable (Thomas et al. 2024). XAI helps ensure that predictive models in healthcare adhere to appropriate uncertainty measurement standards and can assist in explaining the process of uncertainty quantification to regulatory bodies and other stakeholders (Seoni et al. 2024). It is also important to develop systems that track model performance and update uncertainty assessments over time. XAI can detect when model performance or uncertainty estimates change, signalling the need for retraining or recalibration (He et al. 2024). Models can provide probabilistic predictions, such as predictive intervals or probability distributions, rather than deterministic outputs. XAI strategies can help healthcare professionals interpret these probabilistic forecasts, allowing them to better understand the model's uncertainty and make more informed decisions (Lammert et al. 2024).

5.3 | Implementation Approaches of AI-XAI in the Healthcare 5.0 Field

LIME and SHAP are methods used to make the results of AI models more interpretable by providing local-level approximations of complex classifiers. These techniques are employed after the training and prediction tasks are completed, helping to clarify the rationale behind the predictions. AI models, including CNNs, logistic regression, and vision transformers, utilise patient data as input to generate predictions. LIME and SHAP then provide local explanations for these predictions, highlighting the features that significantly contribute to the results. SHAP, in particular, ensures transparency in AI model decisions by providing a clear record of the factors that influence predictions. This transparency helps organisations comply with legal standards such as GDPR and HIPAA, reducing legal liabilities and promoting accountability in healthcare settings (Rahman, Islam, Kundu, et al. 2023). SHAP integration facilitates regulatory compliance and improves trust in AI-driven decision-making by guaranteeing that forecasts are comprehensible and can be connected to specific inputs. Another useful method is ELI5 (Explain Like I am 5), which simplifies the process of understanding a model's inner workings by analysing feature weights or coefficients. With this method, it becomes easier to examine the dataset and identify key features. Additionally, the DALEX method is a model-agnostic tool that can be used with any machine learning model, regardless of its underlying structure. This versatility allows users to evaluate predictions from various models, such as linear models, tree-based models, and neural networks. DALEX offers both local and global explanations: local interpretation provides specific insights into predictions for individual instances, while global interpretation offers a broader understanding of how the model behaves across the entire dataset. DALEX also assesses model stability by checking for overfitting or underfitting, thus enhancing the reliability of the model.

Figure 7 illustrates the integration of AI and XAI in Healthcare 5.0, highlighting key features and challenges associated with data collection. This data includes text, image, and audio data, all of which can be processed by AI systems and analysed using XAI techniques. Moreover, Figure 8 shows the primary drivers of AI and XAI in Healthcare 5.0, emphasising areas such as data management, IoT and wearable integration, predictive analytics, and AI-personalised interventions. XAI focuses on ensuring openness, interpretability, accountability, and a user-centred design, which helps make AI systems more understandable, fair, and safe. These characteristics improve the precision, effectiveness, and trustworthiness of AI applications, ultimately leading to better patient outcomes and improved healthcare delivery.

5.4 | Mathematical Foundations of Explainable AI in Healthcare 5.0

In the era of Healthcare 5.0, the integration of XAI requires the use of mathematical frameworks that not only improve model interpretability but also ensure transparency and preserve model performance. This section explores several key mathematical concepts and their application in improving the understanding and reliability of AI systems within the healthcare domain.

One of the primary methods for achieving interpretability in AI models is by determining the importance of various model features. A widely-used approach for this purpose is the Shapley value, rooted in cooperative game theory (Lundberg 2017). The Shapley value provides a reliable method to assess how each feature x_i contributes to the model's prediction $f(x)$:

$$\phi_i = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(|N|-|S|-1)!}{|N|!} \left[f(S \cup \{i\}) - f(S) \right] \quad (1)$$

This formula offers a fair assessment of a feature's contribution by summing the contributions of all possible subsets S of the feature set N that do not include x_i.

Additionally, to strike a balance between explainability and accuracy, we can develop a regularised loss function that penalises model complexity (Caruana et al. 2015):

$$\mathcal{L}_{\text{explain}}(f) = \mathcal{L}(f(x), y) + \lambda \cdot \text{Complexity}(f) \quad (2)$$

In this case, the conventional loss is represented by $\mathcal{L}(f(x), y)$, while the model's interpretability is preserved by $\text{Complexity}(f)$, which controls extraneous complexity through the regularisation parameter λ.

Additionally, decision trees, which are widely regarded as interpretable models, use Gini impurity to determine the best splits (Breiman 2017):

$$I_G(p) = 1 - \sum_{i=1}^{n} p_i^2 \quad (3)$$

The decision-making process in decision trees is guided by the Gini impurity $I_G(p)$, which measures the likelihood of incorrect labelling at a given node.

In the medical field, managing uncertainty is crucial. Bayesian neural networks (BNNs) address this by treating the output as a probability distribution, allowing them to model the uncertainty in predictions (Blundell et al. 2015):

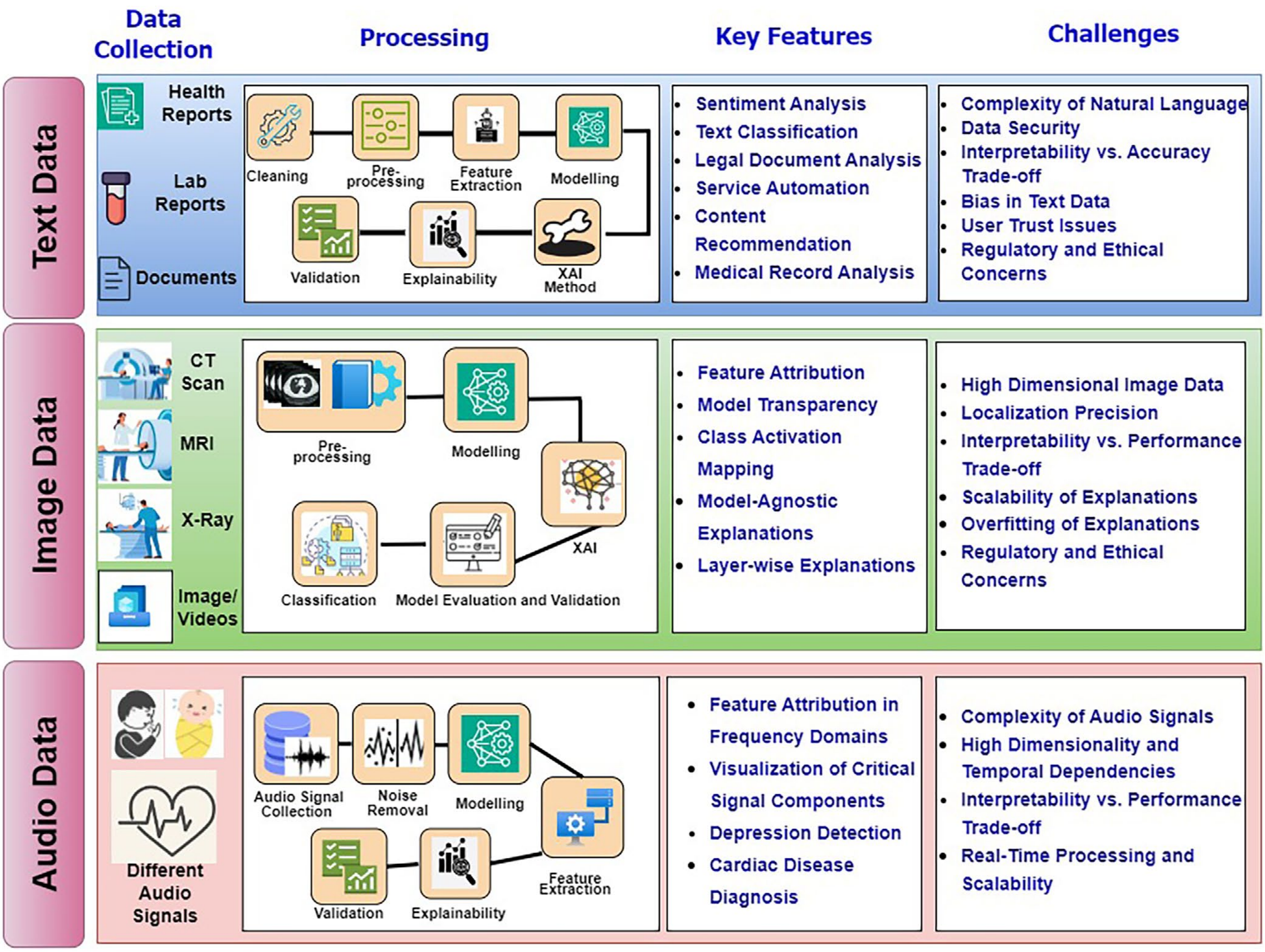

FIGURE 7 | AI and XAI integrated Healthcare 5.0 key features and challenges.

$$P(y|x, \mathcal{D}) = \int P(y|x, w) P(w|\mathcal{D})\, dw \tag{4}$$

This integral reflects the uncertainty in the predictions by considering all possible model parameters w.

In the context of XAI optimisation, there is often a trade-off between predictability, interpretability, and fairness (Ross et al. 2017):

$$\min_f \mathcal{L}(f(x), y) + \alpha \cdot \text{Interpretability}(f) + \beta \cdot \text{Fairness}(f) \tag{5}$$

In this equation, α and β serve to balance the need for interpretability and fairness in the model, respectively, while the primary objective remains minimising prediction error.

One of the fundamental strategies for model interpretability is LIME, which approximates a complex model with a simpler, more interpretable model in the vicinity of a specific prediction. This approximation provides insights into individual predictions and helps in understanding the model's behaviour on a localised scale.

The LIME method can be expressed mathematically as follows (Ribeiro et al. 2016):

$$\min_{g \in G} \mathcal{L}\left(f, g, \pi_x\right) + \Omega(g) \tag{6}$$

In this case, $\mathcal{L}\left(f, g, \pi_x\right)$ represents the loss function that measures the similarity between the interpretable model g and the original model f, while $\Omega(g)$ is a complexity penalty that ensures g remains interpretable and simple. The local neighbourhood surrounding the instance x is defined by the word π_x.

Neural networks often use attention mechanisms to improve model performance by focusing on relevant parts of the input data. In the context of XAI, attention weights can be leveraged to identify which features or input components are most significant in the decision-making process (Bahdanau 2014).

One way to mathematically represent the attention mechanism is as follows (Vaswani 2017):

$$\alpha_t = \frac{\exp\left(e_t\right)}{\sum_{t'} \exp\left(e_{t'}\right)}, \quad e_t = a\left(s_{t-1}, h_t\right) \tag{7}$$

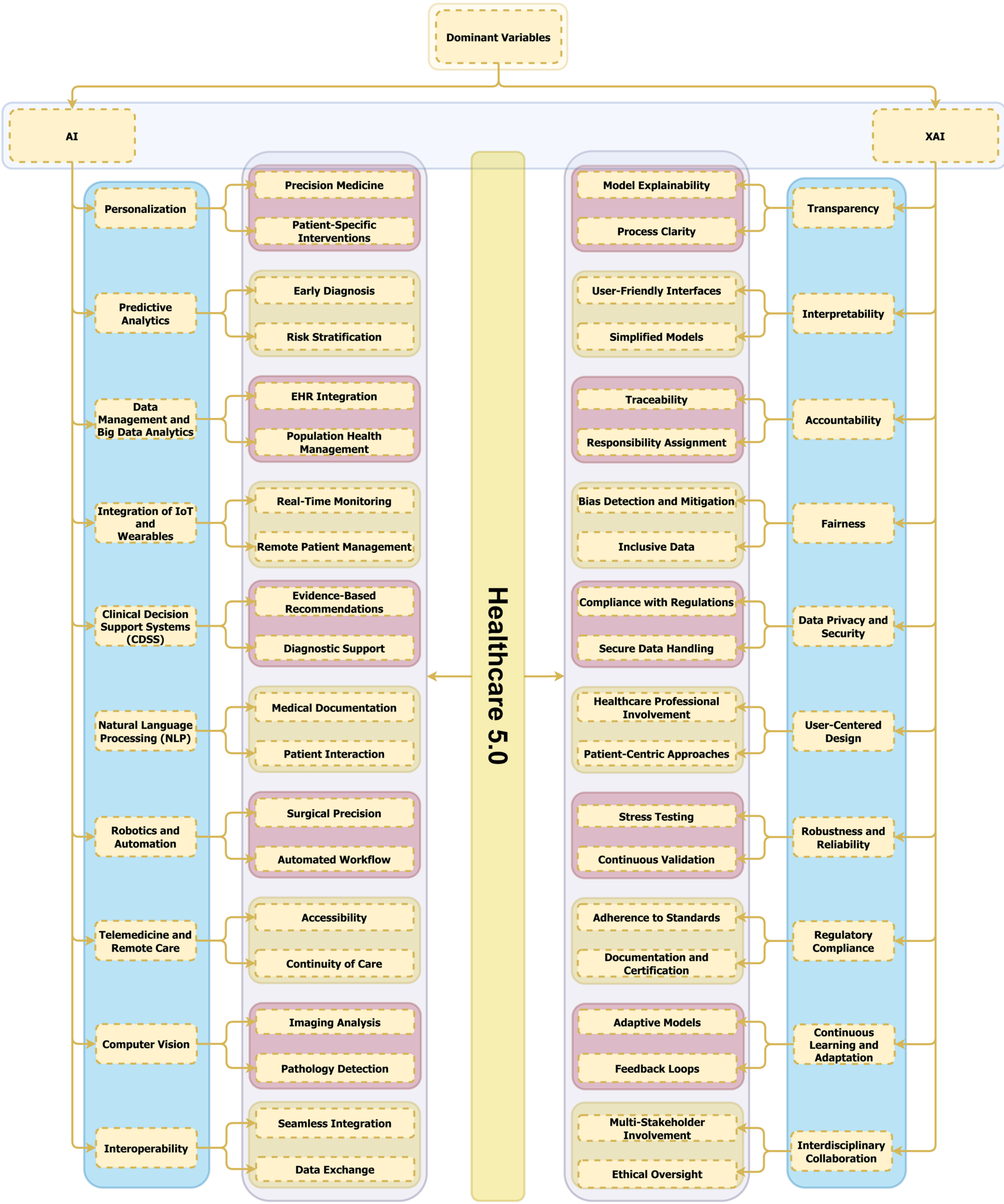

FIGURE 8 | Dominant variables in AI and XAI for Healthcare 5.0.

The alignment score is denoted by e_t, the previous hidden state by s_{t-1}, the current hidden state by h_t, and the attention weight assigned to the input at time step t is represented by α_t. This approach makes the incoming data interpretable by highlighting important areas.

Kullback–Leibler KL divergence, an information-theoretic metric, quantifies the difference between two probability distributions. In the context of XAI, KL divergence can be used to compare the output distribution of a complex model with that of a simpler, interpretable model, helping to understand the degree to which the simplified model diverges from the original.

The KL divergence is computed as follows (Kullback and Leibler 1951):

$$D_{KL}(P \parallel Q) = \sum_i P(i) \log \frac{P(i)}{Q(i)} \tag{8}$$

In this instance, the probability distributions being compared are P and Q. This metric is particularly useful when evaluating the interpretability trade-offs in model simplification. When the value of P and Q are both equal then it tends to zero meaning no difference. The method is not symmetic meaning $D_{KL}(P \parallel Q)$ is not equal to $D_{KL}(Q \parallel P)$. The larger value from this method denotes greated dissimilarity between the distributions.

Counterfactual explanations provide a natural way to understand model predictions by illustrating how the input features must change in order to influence the prediction. This approach is especially valuable in the healthcare industry, where decision-making often relies on understanding what could alter a model's output (Wachter et al. 2017).

Counterfactual explanations can be mathematically expressed as follows (Ustun et al. 2019):

$$\text{argmin}_{x'} \left\| x - x' \right\|_p \quad \text{s. t.} \quad f\left(x'\right) = y', y' \neq f(x) \tag{9}$$

With the disclaimer that the prediction of the model shifts to a different result y', this equation minimises the distance between the original instance x and the counterfactual instance x' under a norm p. This is basically the fundamental concept in the field of adversarial ML. With very little changes to the input function can turn the model towards wrong predictions.

Saliency maps and other gradient-based techniques are frequently used to explain deep neural network predictions. These techniques visualise the features that contribute most to the prediction by displaying the gradient of the output with respect to the input (Simonyan et al. 2013).

One computes the saliency map S_{ij} as follows (Zeiler and Fergus 2014):

$$S_{ij} = \frac{\partial y}{\partial x_{ij}} \tag{10}$$

This partial derivative highlights the most important characteristics and illustrates how sensitive the output y is to the input feature x_{ij}, providing a clear framework for interpretability. From this equation it is depicted that if we get a large value of S_{ij} then with small changes in x_{ij} indicates high sensitivity. For the ML model this concept is used for backpropagation to compute gradients.

Surrogate models are commonly used to approximate complex models with simpler, more interpretable alternatives. The trustworthiness of a surrogate model g in relation to the original model f can be quantified as follows (Ribeiro et al. 2016):

$$\text{Fidelity}(g, f) = \frac{1}{N} \sum_{i=1}^{N} \mathbb{I}\left[g\left(x_i\right) = f\left(x_i\right)\right] \tag{11}$$

In this case, the indicator function $\mathbb{I}$ represents the alignment over N samples between the surrogate model g and the original model f.

The total variation distance (TVD) is a reliable tool for comparing two models or assessing a model's complexity. It measures the greatest difference between two probability distributions, providing a precise indication of how much a simpler model deviates from a more sophisticated one (Gibbs and Su 2002).

The definition of the total variation distance is (Sriperumbudur et al. 2011):

$$\delta(P, Q) = \frac{1}{2} \sum_x |P(x) - Q(x)| \tag{12}$$

When evaluating how closely a surrogate model mimics the behaviour of the original model, this statistic is crucial for XAI.

One way to link a model's prediction to its input attributes is through the use of integrated gradients. This technique satisfies several desirable features, such as completeness and sensitivity, and ensures that the attribution is consistent with the model's prediction.

For a feature x_i, the integrated gradients are computed as follows (Sundararajan et al. 2017):

$$\text{IntegratedGradient}_i = \left(x_i - x_i'\right) \times \int_{\alpha=0}^{1} \frac{\partial F\left(x' + \alpha \times \left(x - x'\right)\right)}{\partial x_i} d\alpha \tag{13}$$

As the input changes from a baseline x' to the actual input x, this integral calculates the average gradient of the model's prediction F. The result provides an understandable attribution of the prediction to the input features.

Finally, robust metrics like the area under the receiver operating characteristic curve (AUROC) are essential for assessing model performance in the healthcare industry (Bradley 1997):

$$\text{AUROC} = \int_0^1 TPR\left(FPR^{-1}(\alpha)\right) d\alpha \tag{14}$$

The AUROC evaluation analyses the model's ability to differentiate between positive and negative classes across various threshold applications, which is crucial in healthcare, where misinterpretation can have serious consequences.

The mathematical formulas discussed in this section emphasise the importance of transparency and comprehensibility in AI models, particularly in the sensitive domain of healthcare. Employing these strategies, we can move toward a future where AI systems are trustworthy, understandable, and powerful—all aligning with the objectives of Healthcare 5.0.

Figure 8 illustrates the key variables in AI and XAI for Healthcare 5.0.

6 | Future Sustainability of AI and XAI in Healthcare 5.0 Applications

While the future of artificial intelligence (AI) and explainable AI (XAI) in healthcare, particularly within the context of Healthcare 5.0 applications, depends on several important factors, it holds great promise. This study categorises these factors into two main areas: the sustainability of XAI and the sustainability of AI in Healthcare 5.0.

6.1 | Sustainability of AI in Healthcare 5.0

The sustainability of AI in Healthcare 5.0 involves ensuring that AI technologies are effectively implemented and continuously developed to improve healthcare delivery while addressing long-term challenges. Below are key factors that contribute to the sustainability of AI in Healthcare 5.0 (Mennella et al. 2024; Suha and Sanam 2023; Moyano-Fernández et al. 2024; Rahman, Hasan, et al. 2023):

- To ensure sustainability, AI systems must be developed and deployed ethically, ensuring patient privacy, autonomy, and fairness. Adhering to regulations such as GDPR and HIPAA supports trust and confidence in AI applications (Mennella et al. 2024).
- AI systems in Healthcare 5.0 must be transparent and explainable, clearly outlining their decisions and recommendations. XAI techniques help healthcare providers and patients understand how AI algorithms arrive at their conclusions, promoting trust and collaboration between humans and technology (Alam et al. 2023).
- Sustainable AI in healthcare depends on high-quality data that is diverse, accessible, and representative of different patient populations. Improving data quality, correcting biases, and ensuring data interoperability enable AI systems to provide equitable healthcare solutions (Albahri et al. 2023).
- AI technologies must undergo rigorous clinical validation to ensure their efficacy, safety, and clinical value before being integrated into healthcare workflows. Seamless integration with existing clinical systems and electronic health records (EHRs) is essential to improve AI's impact on patient care while minimising disruptions to healthcare operations (Olawade et al. 2024).
- AI systems in Healthcare 5.0 must be designed for continuous learning and improvement. Feedback from healthcare professionals, patients, and real-world data can be used to fine-tune algorithms and adapt to evolving healthcare needs. Regular model retraining, validation, and monitoring are necessary to ensure AI systems remain effective and relevant over time (Hamood Alsamhi et al. 2023).
- Sustainable AI solutions in healthcare must be both cost-effective and scalable. These solutions should deliver value to healthcare organisations while remaining accessible and affordable. Strategies to optimise resource usage, reduce implementation costs, and establish an ROI are crucial to ensuring the long-term success of AI in healthcare settings (Alowais et al. 2023).
- Healthcare practitioners must be adequately trained and supported in using AI technologies in clinical practice. Ongoing education and skill development initiatives help healthcare personnel maximise the potential of AI while addressing concerns about job displacement and workflow disruptions (Metta et al. 2024).
- Sustainable AI in Healthcare 5.0 empowers patients by providing personalised, data-driven healthcare solutions and encouraging active participation in their care. AI technologies can be used to deliver personalised interventions, facilitate collaborative decision-making, and improve health outcomes (Bohr and Memarzadeh 2020).

Through addressing these critical factors, stakeholders can ensure the long-term viability of AI in Healthcare 5.0, which will support innovation, improve patient outcomes, and drive the future of healthcare delivery.

6.2 | Sustainability of XAI in Healthcare 5.0

For AI technologies to be responsibly and effectively integrated into healthcare systems, XAI in Healthcare 5.0 must be sustainable. Here's how XAI contributes to making healthcare more sustainable:

- XAI approaches provide both patients and healthcare professionals with a clear understanding of how AI algorithms make decisions. The clarity aspects of AI foster trust among its stakeholders. This aids in the wide-scale adaptation of AI in a sustainable manner within a healthcare framework (Ali et al. 2023).
- Another feature of XAI is its ability to deal with the ethical concerns of AI algorithms. XAI promotes accountability and transparency factors. XAI in a healthcare setting promotes morals and with a higher degree of responsibility (Vainio-Pekka et al. 2023).
- XAI helps identify biases, errors, and limitations within AI models, reducing the risk of harm to patients and healthcare systems. Enabling stakeholders to recognise and address potential risks, XAI improves the safety and reliability of AI applications, contributing to their sustainable deployment in healthcare (Dalvi-Esfahani et al. 2023).
- XAI provides interpretable insights into AI-generated predictions and recommendations, facilitating collaboration between AI systems and human specialists. This leads to more confident and well-informed decision-making by healthcare professionals, ultimately resulting in better patient outcomes and more sustainable healthcare delivery (Olaoye et al. 2024).
- XAI methodologies allow stakeholders to monitor AI performance, identify areas for improvement, and incrementally improve algorithms. XAI supports the long-term sustainability of AI solutions by encouraging continuous learning

and adaptation, which keeps them applicable and efficient in changing clinical settings (Mahbooba et al. 2021).

- XAI assists healthcare institutions in meeting legal and industry standards regarding data protection, accountability, and transparency. Through offering interpretable explanations for AI decisions, XAI supports compliance with regulations such as GDPR and HIPAA, helping ensure the sustainable deployment of AI in healthcare (Kose et al. 2024).
- XAI enables patients to actively participate in their healthcare decisions by explaining AI-generated insights and recommendations. Promoting patient empowerment and engagement, XAI supports sustainable healthcare models that prioritise the needs and preferences of patients (Grover and Dogra 2024).
- XAI methods contribute to educational programs that increase awareness about AI among patients, healthcare providers, and other stakeholders. XAI helps stakeholders comprehend how AI algorithms work and how to interpret their results by promoting a culture of constant learning. The long-term incorporation of AI into healthcare practice is supported by this continuing education (Khosravi et al. 2022).

In conclusion, the sustainability of XAI in Healthcare 5.0 is essential to ensure the responsible, ethical, and effective application of AI technologies in healthcare. Through transparency, accountability, and stakeholder empowerment, XAI plays an important role in securing the long-term success and viability of AI applications, driving advancements in healthcare delivery, and improving patient outcomes.

7 | Open Issues and Future Guidelines

As the healthcare industry continues to adopt and integrate AI and XAI technologies, several open issues and challenges must be addressed to ensure these systems are effective and trustworthy. This section discusses these challenges and outlines the open issues that need attention for the successful integration of XAI into healthcare systems.

7.1 | Discussion on Open Issues

There are both benefits and challenges associated with the transition from traditional AI to XAI. The integration of AI is becoming increasingly widespread due to the advancements in automation and intelligence. As automated systems continue to grow in prevalence, it is crucial to maintain transparency and build trust, especially when these systems interact with critical healthcare stakeholders and patients.

7.1.1 | Transparency

AI systems, particularly those based on NLP, ML, and DL models, often operate as 'black box' models. These systems are capable of performing tasks such as data analysis, classification, and prediction, but the reasoning behind their decisions remains obscure. As Industry 5.0 progresses, the need for model transparency becomes more pressing (Awosika et al. 2024). When AI models only produce outputs without explaining the reasoning behind their predictions, it raises significant concerns about transparency. This lack of clarity becomes even more significant in healthcare, where AI-based decisions can directly impact the lives of patients. In such a context, model interoperability and trustworthiness are important, ensuring that users can rely on the system to make informed decisions on complex issues. To address this issue, XAI approaches, such as LIME or SHAP-based models, can provide valuable insights. These methods help to illuminate how a model reaches its predictions, offering a clearer understanding of its decision-making process (Metta et al. 2024).

7.1.2 | Observance of Regulations

The use of AI software models in the healthcare and pharmaceutical industries has led to increasing calls for transparency in predictive healthcare models. Regulatory bodies such as the European Medicines Agency (EMA) and the United States Food and Drug Administration (FDA) have emphasised the importance of these models being able to justify their decisions or findings explicitly (Chettri and Ravi 2024). This transparency is essential for gaining the trust of medical practitioners. However, implementing these regulatory requirements presents challenges, particularly when it comes to the legal deployment of AI technologies in healthcare settings. Failing to meet these regulations could delay the adoption of AI, potentially resulting in legal consequences, such as lawsuits. To ensure the legal application of AI in healthcare, it is crucial to develop XAI models that comply with regulatory standards. This involves creating processes that clearly disclose and validate the decision-making grounds of the AI, thus fulfilling legal requirements related to accountability and traceability.

7.1.3 | Ethical Considerations

In some cases, the data used to train AI models may be incomplete or inaccurate, leading to biased predictions that could have serious ethical implications (Longo et al. 2024). Ethical use of AI in healthcare is essential to minimise harm and ensure that AI systems do not exacerbate existing inequalities. Additionally, maintaining privacy and trust in these systems requires ethical handling of the data and decision-making processes. XAI has the potential to identify and address biases in AI systems, providing stakeholders with insight into why certain decisions are made. This transparency enables stakeholders to identify and correct biases, ensuring that AI systems are used fairly and responsibly. Incorporating fairness-conscious machine learning techniques and conducting regular audits of AI models are key measures that should be standard practice in the industry (Cacciamani et al. 2024).

7.1.4 | Clinical Integration

Integrating AI into healthcare workflows can be valuable but also challenging. In some cases, healthcare workers may struggle to understand and trust AI-driven systems, leading to reluctance in their adoption. While AI-integrated systems can assist with tasks that are difficult for humans to perform, their success in clinical settings depends on their seamless integration into the decision-making process. For AI systems to function effectively in hospitals or clinics, all system users must understand how the technology fits into their workflows. In this context, XAI plays a crucial role by providing the necessary insights for practitioners to trust and effectively use AI recommendations. It helps to build intuitive user interfaces that present information clearly, making it easier for healthcare professionals to rely on AI-driven advice (Hulsen 2024).

7.1.5 | Patient-Centric Care

The primary goal of any healthcare system is to provide comprehensive patient care. In today's healthcare landscape, numerous AI-driven solutions facilitate remote patient care, especially for elderly patients living at home who require long-term care but may need immediate hospital assistance when necessary. These solutions have become increasingly prevalent with the integration of AI technologies. However, one major challenge in this sector is trust. Many patients hesitate to rely on AI systems for making important decisions or judgments about their care. To address this, it is crucial to develop systems that can explain their conclusions, decisions, or assessments in a clear and accessible manner, avoiding technical jargon. While AI models are created to solve specific problems, they often introduce new challenges, such as rigid models that may not adapt well to patients' varying needs. To overcome this, the system must be flexible enough to allow patients to customise it according to their personal preferences. AI-based healthcare tools often require specific skills from users. Training patients to understand and use these systems effectively can be a difficult task. Moreover, the development process does not end once the system is deployed—ongoing user feedback is essential for refining and improving the solution. Additionally, safeguarding personal health data is a critical issue. Ensuring that patient information is protected from unauthorised access or leaks is a top priority. Ultimately, the most crucial function of patient-centric tools is to secure personal health information and ensure its confidentiality (Biswas and Talukdar 2024).

7.1.6 | Education and Training

Successfully implementing a fully automated healthcare system requires extensive knowledge and a robust training program to transition from traditional manual systems to fully automated systems. Achieving the goals of Healthcare 5.0 requires that all individuals involved in the healthcare automation process—such as healthcare professionals, data scientists, and engineers—receive sufficient education (Patel et al. 2024). However, there are significant limitations in the educational systems of some countries, particularly in less developed or developing regions, where training programs may lack the necessary infrastructure and qualified instructors. Moreover, there are instances where AI engineers and data scientists fail to fully grasp the unique constraints and specifics of healthcare environments, which can lead to challenges when deploying AI solutions. It is also crucial to consider the ethical and legal implications of using AI in healthcare, which are often overlooked. To ensure the successful and sustainable operation of automated healthcare systems, both AI developers and healthcare personnel must receive ongoing training and resources. One key realisation is that effective education must start at the foundational level, which may require updating outdated curricula to ensure that new generations are adequately prepared (Nazar et al. 2021).

7.1.7 | Interdisciplinary Collaboration

One of the key considerations when integrating XAI with AI in healthcare is supporting interdisciplinary collaboration among diverse stakeholders such as regulatory agencies, physicians, ethicists, and AI technologists. These groups often have differing perspectives, yet they must work together to address the challenges outlined earlier, all of which require solutions that can be provided by effective XAI systems. Collaboration ensures that the integration of XAI into healthcare workflows takes into account the viewpoints and needs of all stakeholders. To achieve this, it is essential to form multidisciplinary teams and establish forums for ongoing communication and cooperation. This approach facilitates a more holistic solution to complex healthcare problems (Rubegni et al. 2024). However, achieving such collaboration in the healthcare industry can be challenging due to the diverse backgrounds and perspectives of the personnel involved. Each stakeholder brings their own ethical considerations and expertise, which can sometimes lead to disagreements or misunderstandings. For instance, healthcare professionals may prioritise fair distribution and patient outcomes, while AI developers may focus on optimising model accuracy. To address these differing priorities, it is essential that system requirements are clearly specified and balanced throughout the development process (Sadeghi et al. 2024).

Adopting an XAI paradigm for Healthcare 5.0 requires a multidimensional approach that emphasises the principles of transparency, personalisation, education, engagement, ethical use, and inclusivity. Addressing these core issues with practical solutions, healthcare systems can gain the trust of patients and improve health outcomes through AI technologies that are truly beneficial to patients. This approach not only improves AI adoption within healthcare but also supports the concurrent evolution of technology and patient-centred care.

7.2 | Future Opportunities

In this section, the convergence of AI, XAI, and Healthcare 5.0 presents significant opportunities to advance medical care, improve patient outcomes, and optimise healthcare systems. Below are some potential avenues for future research in these fields.

7.2.1 | Integration of AI With Emerging Technologies

The goal is to assemble seamless, integrated systems that improve data accuracy, patient monitoring, and secure, compliant data usage across various platforms (Rahman et al. 2021; Ahmed et al. 2022). AI applications span across multiple sectors, from smart healthcare to smart agriculture. A prominent trend in modern urban development is the creation of smart cities, which involve both technological integration and advanced data collection. However, simply collecting data is insufficient to solve the design challenges of a smart city. Decision-making systems are essential for creating an intelligent network of connected devices. These systems can improve various aspects of urban life, including environmental control, horticulture, water and energy management, and more. AI is also transforming smart parking and traffic management, becoming increasingly important as urban populations and vehicle numbers rise. Leveraging AI, solutions like drone-based surveillance, door-to-door services for the elderly, and more efficient monitoring systems can be achieved in the context of smart cities. Similarly, in the manufacturing sector, AI-driven approaches have revolutionised factory production, from raw material collection to efficient processes. Intelligent factory management systems rely on smart devices, sensors, and advanced AI algorithms. AI also allows for tasks traditionally requiring human involvement to be delegated to robots (Islam et al. 2021). Overall, AI must be integrated with emerging technologies such as Federated Learning, IoT, 5G, large language models, and Blockchain for real-time health monitoring and secure data management. The goal is to assemble seamless, integrated systems that improve data accuracy, patient monitoring, and secure, compliant data usage across various platforms (Rahman et al. 2021; Ahmed et al. 2022). AI applications span across multiple sectors, from smart healthcare to smart agriculture. A prominent trend in modern urban development is the creation of smart cities, which involve both technological integration and advanced data collection. However, simply collecting data is insufficient to solve the design challenges of a smart city. Decision-making systems are essential for creating an intelligent network of connected devices. These systems can improve various aspects of urban life, including environmental control, horticulture, water and energy management, and more. AI is also transforming smart parking and traffic management, becoming increasingly important as urban populations and vehicle numbers rise. Leveraging AI, solutions like drone-based surveillance, door-to-door services for the elderly, and more efficient monitoring systems can be achieved in the context of smart cities. Similarly, in the manufacturing sector, AI-driven approaches have revolutionised factory production, from raw material collection to efficient processes. Intelligent factory management systems rely on smart devices, sensors, and advanced AI algorithms. AI also allows for tasks traditionally requiring human involvement to be delegated to robots (Islam et al. 2021). Overall, AI must be integrated with emerging technologies such as Federated Learning, IoT, 5G, large language models, and Blockchain for real-time health monitoring and secure data management. The goal is to assemble seamless, integrated systems that improve data accuracy, patient monitoring, and secure, compliant data usage across various platforms (Rahman et al. 2021; Ahmed et al. 2022). AI applications span across multiple sectors, from smart healthcare to smart agriculture. A prominent trend in modern urban development is the creation of smart cities, which involve both technological integration and advanced data collection. However, simply collecting data is insufficient to solve the design challenges of a smart city. Decision-making systems are essential for creating an intelligent network of connected devices. These systems can improve various aspects of urban life, including environmental control, horticulture, water and energy management, and more. AI is also transforming smart parking and traffic management, becoming increasingly important as urban populations and vehicle numbers rise. Leveraging AI, solutions like drone-based surveillance, door-to-door services for the elderly, and more efficient monitoring systems can be achieved in the context of smart cities. Similarly, in the manufacturing sector, AI-driven approaches have revolutionised factory production, from raw material collection to efficient processes. Intelligent factory management systems rely on smart devices, sensors, and advanced AI algorithms. AI also allows for tasks traditionally requiring human involvement to be delegated to robots (Islam et al. 2021). Overall, AI must be integrated with emerging technologies such as Federated Learning, IoT, 5G, large language models, and Blockchain for real-time health monitoring and secure data management.

7.2.2 | Augmentation of Human Capabilities With XAI

The goal is to assemble seamless, integrated systems that improve data accuracy, patient monitoring, and secure, compliant data usage across various platforms (Rahman et al. 2021; Ahmed et al. 2022). AI applications span across multiple sectors, from smart healthcare to smart agriculture. A prominent trend in modern urban development is the creation of smart cities, which involve both technological integration and advanced data collection. However, simply collecting data is insufficient to solve the design challenges of a smart city. Decision-making systems are essential for creating an intelligent network of connected devices. These systems can improve various aspects of urban life, including environmental control, horticulture, water and energy management, and more. AI is also transforming smart parking and traffic management, becoming increasingly important as urban populations and vehicle numbers rise. Leveraging AI, solutions like drone-based surveillance, door-to-door services for the elderly, and more efficient monitoring systems can be achieved in the context of smart cities. Similarly, in the manufacturing sector, AI-driven approaches have revolutionised factory production, from raw material collection to efficient processes. Intelligent factory management systems rely on smart devices, sensors, and advanced AI algorithms. AI also allows for tasks traditionally requiring human involvement to be delegated to robots (Islam et al. 2021). Overall, AI must be integrated with emerging technologies such as Federated Learning, IoT, 5G, large language models, and Blockchain for real-time health monitoring and secure data management. The potential of augmenting human capabilities through AI and robotics is rapidly transforming fields such as surgery, rehabilitation, and patient care. AI can improve precision in surgical procedures, enhance rehabilitation strategies, and increase the efficiency of care delivery (Patidar et al. 2024). XAI plays an important role in the redesign of healthcare by enabling significant advancements in both cognitive and physical capacities. XAI frameworks enhance clinical decision-making by offering clear and intelligible justifications for AI-generated recommendations, which fosters confidence and boosts medical professionals' adoption (Arrieta et al. 2020). Furthermore, AI and XAI extend human potential beyond traditional limitations, offering new possibilities for improving the quality of life and functional abilities across diverse populations. As this landscape evolves, research into the ethical, social, and technical implications of augmenting human capabilities becomes increasingly important. Ensuring that these technologies are developed and deployed responsibly will promote well-being and accessibility for all (Gupta 2024). The integration of these advancements also necessitates ongoing investigation into the ethical concerns and societal impacts of augmented human capabilities, ensuring their fair and inclusive progression.

7.2.3 | Ethical and Social Implications of AI in Healthcare

The ethical considerations surrounding AI and XAI in healthcare are vast and complex, with significant concerns about bias, fairness, and accountability. As AI-powered systems become more embedded in healthcare processes, they introduce profound ethical challenges related to privacy, autonomy, and potential biases in decision-making. Implementing XAI approaches is crucial to detecting and addressing biases in AI models in a transparent and understandable way for all stakeholders (Taddeo et al. 2023). To ensure the ethical implementation of AI, it is essential to prioritise patient privacy and create regulatory frameworks that evolve alongside technological advancements. Legally, integrating AI into healthcare raises important questions about accountability and liability, particularly in cases of diagnostic errors or treatment failures. Ensuring clear guidelines and responsible practices will be essential for addressing these concerns and ensuring that AI technologies are used in a way that is both ethical and beneficial to patients (Rigby 2019).

7.2.4 | Smart IIoT Applications in Manufacturing Areas

With the advent of the IR 4.0 revolution, the field of smart Industrial Internet of Things (IIoT) applications has expanded significantly. While electrical technologies were introduced in IR 3.0, their integration has become more widespread with the introduction of IR 4.0. The success of IIoT applications in the smart industry is evident when a central factory is connected to multiple other factories within a network, allowing for the rapid transfer of data. This results in a more efficient and reliable machine-to-machine (M2M) communication setup, which empowers users to make better decisions and meet critical factory requirements. Leveraging IIoT enables the efficient and secure transfer of diverse data from one location to another. The application of IIoT, however, is not limited to factories alone. It extends to agriculture, where data on factors like moisture levels, water levels, and other relevant metrics can be collected and transmitted to stakeholders, facilitating informed decision-making and better utilisation of available resources (Tabaa et al. 2020). In the manufacturing sector, IIoT plays a crucial role in managing customer data, which is essential for business growth and product management systems. Through enabling the fast and secure collection and transmission of data, IIoT improves management efficiency in these scenarios, supporting effective decision-making and operations (Zhan et al. 2021).

7.2.5 | AI-Empowered Drug Discovery and Expansion

A promising future for the pharmaceutical industry lies in the improvement of new therapeutics, the identification of potential drug candidates, and the prediction of pharmacokinetic and pharmacodynamic efficacy through AI models. These advancements aim to reduce the need for expensive and time-consuming laboratory experiments. Additionally, the synergy between AI and emerging biotechnologies, such as CRISPR and bioprinting, holds great potential in personalising drug treatments. Considering an individual's genetic makeup, lifestyle, and environment, AI can help optimise treatments to improve efficacy and minimise adverse effects, particularly for complex diseases such as cancer and autoimmune disorders. Furthermore, research could explore the use of AI in drug repurposing, which involves identifying new uses for existing drugs. AI can also assist in analysing vast biomedical datasets, electronic health records, and genomic data to uncover novel treatments, ultimately expediting the development of new therapies and improving access to cutting-edge treatments. This approach has the potential to rapidly address emerging health crises, such as pandemics, and improve healthcare outcomes on a global scale.

7.2.6 | Training and Education in AI and XAI for Healthcare Professionals

To effectively integrate AI and XAI into healthcare practice, it is crucial to equip healthcare professionals with the knowledge and skills to use these tools confidently for clinical applications. Research should focus on developing comprehensive training programs, including interactive workshops, online courses, and hands-on sessions, to teach fundamental AI concepts and XAI systems to medical staff. Another important area is the evaluation of these training programs to assess healthcare professionals' proficiency and confidence in utilising AI tools. Collaboration between AI experts, educators, and healthcare providers will be essential in creating standardised educational resources that address the specific needs and challenges faced by the healthcare sector. Through ensuring that these technologies are used safely, effectively, and ethically, we can improve patient outcomes and help advance the future of healthcare delivery.

7.2.7 | XAI in Data-Driven System

Enhancing the transparency of XAI within data-driven systems involves several key areas of focus. One significant area is the development of advanced XAI algorithms that can explain the decision-making mechanisms behind AI processes while offering a user-friendly interface. This includes creating visualisation tools that translate complex algorithmic decisions into interpretable data, helping clinicians understand and trust AI outputs more effectively. Additionally, integrating XAI into electronic healthcare systems ensures that AI-driven insights are both explainable and supportive of clinical decision-making. Ethical and regulatory considerations are also essential, particularly in maintaining privacy standards and ensuring that AI solutions are unbiased and provide equitable healthcare recommendations. Lastly, longitudinal studies are necessary to assess the impact of XAI on patient outcomes and clinician adoption. These studies will be crucial in validating the benefits of XAI technologies and guiding their effective integration into clinical practice.

7.2.8 | Quantum Computing Could Further Transform AI in Healthcare

This work provides a wide ranging overview of AI and XAI in HC 5.0 applications, highlighting the potentiality of changing healthcare through quantum computing. The capacity of quantum computing to handle enormous volumes of data at rates that were previously unprecedented can greatly improve the capabilities of AI and machine learning models used in diagnosing diseases, predicting patient outcomes, drug discovery, and personalised treatment planning. A promising area of research involves developing quantum-improved algorithms to reduce the time required to train complex models on large-scale genomic datasets. Furthermore, the integration of quantum computing with existing AI healthcare systems could lead to more accurate models for detecting diseases such as cancer, neurological disorders, Alzheimer's, and cardiovascular diseases by efficiently processing large volumes of imaging and genetic data. Future studies should focus on making these technologies accessible, ensuring their ethical application, and securely integrating them into existing healthcare infrastructures to align with patient safety and data integrity standards.

8 | Conclusion

In this research, we provide a complete discussion of the integration of AI and XAI within the framework of HC 5.0, emphasising their transformative potential for modern healthcare systems. In contrast to the previous studies that typically concentrate on AI or XAI independently, this work bridges the gap by analysing their combined application within Healthcare 5.0. The survey not only explores the foundational concepts but also delves into state-of-the-art tools, techniques, and sustainable applications, highlighting their significant roles in enhancing transparency, interpretability, and efficiency in patient care. The originality of this study lies in its dual focus on AI's computational power and XAI's interpretability, offering actionable insights for the formation of patient-centred, morally sound, and globally aligned healthcare solutions. In contrast to prior works, this survey uniquely integrates a forward-looking perspective, providing recommendations for optimising XAI frameworks to address current challenges in transparency, legal considerations, and interdisciplinary collaboration. Despite the comprehensive scope and strengths of this study, the evolving nature of global standards, ethical compliance, and data privacy regulations calls for further investigation to establish clear guidelines for real-world applications. In future, we aim to refine these frameworks to better address practical challenges, enhancing model interpretability and security while ensuring alignment with patient-centred principles. We hope that our study has explored innovative and sustainable applications, opening the door for the broader integration of AI and XAI in diverse medical domains to meet the growing demands of modern healthcare systems. Ultimately, this endeavour aims to make a meaningful and lasting impact on the advancement of HC 5.0.

Acknowledgements

We acknowledge the usage of AI-based tools, specifically ChatGPT, for grammar and language editing, while Grammarly is employed for style and consistency checks. The intellectual contributions and conclusions presented in this manuscript are solely the result of the authors' efforts.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

Abdulkareem, K. H., M. A. Mohammed, A. Salim, et al. 2021. "Realizing an Effective COVID-19 Diagnosis System Based on Machine Learning and IOT in Smart Hospital Environment." *IEEE Internet of Things Journal* 8, no. 21: 15919–15928.

Ahmad, S. S., W. Almobaideen, M. N. Abdelhakim, and R. Rani. 2024. *Artificial Intelligence-Enabled Cutting-Edge Technologies: Innovation for Industry 5.0, Healthcare 5.0, and Society 5.0*, 140–158. IGI Global.

Ahmed, I., G. Jeon, and F. Piccialli. 2022. "From Artificial Intelligence to eXplainable Artificial Intelligence in Industry 4.0: A Survey on What, How, and Where." *IEEE Transactions on Industrial Informatics* 18: 5031–5042.

Ahuja, K. 2024. *Emotion AI in Healthcare: Application, Challenges, and Future Directions*, 131–146. Elsevier.

Al Mamun, S., M. S. Kaiser, and M. Mahmud. 2021. *An Artificial Intelligence Based Approach Towards Inclusive Healthcare Provisioning in Society 5.0: A Perspective on Brain Disorder*, 157–169. Springer.

Alabi, R. O., M. Elmusrati, I. Leivo, A. Almangush, and A. A. Mäkitie. 2023. "Machine Learning Explainability in Nasopharyngeal Cancer Survival Using LIME and SHAP." *Scientific Reports* 13, no. 1: 8984.

Alam, M. N., M. Kaur, and M. S. Kabir. 2023. "Explainable AI in Healthcare: Enhancing Transparency and Trust Upon Legal and Ethical Consideration." *International Research Journal of Engineering and Technology* 10, no. 6: 1–9.

Albahri, A. S., A. M. Duhaim, M. A. Fadhel, et al. 2023. "A Systematic Review of Trustworthy and Explainable Artificial Intelligence in Healthcare: Assessment of Quality, Bias Risk, and Data Fusion." *Information Fusion* 96: 156–191.

Alberto, I. R. I., N. R. I. Alberto, A. K. Ghosh, et al. 2023. "The Impact of Commercial Health Datasets on Medical Research and Health-Care Algorithms." *Lancet Digital Health* 5, no. 5: e288–e294.

Aldughayfiq, B., F. Ashfaq, N. Jhanjhi, and M. Humayun. 2023. "Explainable AI for Retinoblastoma Diagnosis: Interpreting Deep Learning Models With LIME and SHAP." *Diagnostics* 13, no. 11: 1932.

Ali, S., T. Abuhmed, S. El-Sappagh, et al. 2023. "Explainable Artificial Intelligence (XAI): What We Know and What Is Left to Attain Trustworthy Artificial Intelligence." *Information Fusion* 99: 101805.

Alotaibi, N. S., and A. H. Alshehri. 2023. "Prospers and Obstacles in Using Artificial Intelligence in Saudi Arabia Higher Education Institutions—The Potential of AI-Based Learning Outcomes." *Sustainability* 15, no. 13: 10723.

Alowais, S. A., S. S. Alghamdi, N. Alsuhebany, et al. 2023. "Revolutionizing Healthcare: The Role of Artificial Intelligence in Clinical Practice." *BMC Medical Education* 23, no. 1: 689.

Alsamhi, S. H., R. Myrzashova, A. Hawbani, et al. 2024. "Federated Learning Meets Blockchain in Decentralized Data-Sharing: Healthcare Use Case." *IEEE Internet of Things Journal* 11: 19602–19615.

Alshahrani, R., M. Yenugula, H. Algethami, et al. 2024. "Establishing the Fuzzy Integrated Hybrid MCDM Framework to Identify the Key Barriers to Implementing Artificial Intelligence-Enabled Sustainable Cloud System in an IT Industry." *Expert Systems With Applications* 238: 121732.

Alsinglawi, B., O. Alshari, M. Alorjani, et al. 2022. "An Explainable Machine Learning Framework for Lung Cancer Hospital Length of Stay Prediction." *Scientific Reports* 12, no. 1: 607.

Alsinglawi, B. S., F. Alnajjar, M. S. Alorjani, O. Alshari, M. Novoa, and O. Mubin. 2024. "Predicting Hospital Stay Length Using Explainable Machine Learning." *IEEE Access* 12: 90571–90585.

Arreche, O., T. R. Guntur, J. W. Roberts, and M. Abdallah. 2024. "E-XAI: Evaluating Black-Box Explainable AI Frameworks for Network Intrusion Detection." *IEEE Access* 12: 23954–23988. https://doi.org/10.1109/ACCESS.2024.3365140.

Arrieta, A. B., N. Díaz-Rodríguez, J. Del Ser, et al. 2020. "Explainable Artificial Intelligence (XAI): Concepts, Taxonomies, Opportunities and Challenges Toward Responsible AI." *Information Fusion* 58: 82–115. https://doi.org/10.1016/j.inffus.2019.12.012.

Arslanoğlu, K., and M. Karaköse. 2024. *An Overview for Trustworthy and Explainable Artificial Intelligence in Healthcare*, 1–5. IEEE.

Aurelia, S., R. Thanuja, S. Chowdhury, and Y. C. Hu. 2024. "AI-Based Online Proctoring: A Review of the State-of-the-Art Techniques and Open Challenges." *Multimedia Tools and Applications* 83, no. 11: 31805–31827.

Awosika, T., R. M. Shukla, and B. Pranggono. 2024. "Transparency and Privacy: The Role of Explainable Ai and Federated Learning in Financial Fraud Detection." *IEEE Access* 12: 64551–64560. https://doi.org/10.1109/ACCESS.2024.3394528.

Bahdanau, D. 2014. "Neural Machine Translation by Jointly Learning to Align and Translate." arXiv Preprint arXiv:1409.0473.

Baker, S., and W. Xiang. 2023. "Artificial Intelligence of Things for Smarter Healthcare: A Survey of Advancements, Challenges, and Opportunities." *IEEE Communications Surveys & Tutorials* 25: 1261–1293.

Bharati, S., M. R. H. Mondal, and P. Podder. 2023. "A Review on Explainable Artificial Intelligence for Healthcare: Why, How, and When?" *IEEE Transactions on Artificial Intelligence* 5: 1429–1442.

Bhattacharya, A., S. Stumpf, L. Gosak, G. Stiglic, and K. Verbert. 2024. "EXMOS: Explanatory Model Steering Through Multifaceted Explanations and Data Configurations." arXiv Preprint arXiv:2402.00491.

Bhavin, M., S. Tanwar, N. Sharma, S. Tyagi, and N. Kumar. 2021. "Blockchain and Quantum Blind Signature-Based Hybrid Scheme for Healthcare 5.0 Applications." *Journal of Information Security and Applications* 56: 102673.

Biswas, A., and W. Talukdar. 2024. "Intelligent Clinical Documentation: Harnessing Generative AI for Patient-Centric Clinical Note Generation." arXiv Preprint arXiv:2405.18346.

Blenman, K. R., R. Hull, C. Maimela, T. Molefi, R. Khanyile, and Z. Dlamini. 2023. *Society 5.0 Healthcare: Ethics, Legal Rights, Human Rights, Safety and Security*, 267–291. Springer.

Blundell, C., J. Cornebise, K. Kavukcuoglu, and D. Wierstra. 2015. *Weight Uncertainty in Neural Network*, 1613–1622. PMLR.

Bohr, A., and K. Memarzadeh. 2020. *The Rise of Artificial Intelligence in Healthcare Applications*, 25–60. Elsevier.

Born, J., N. Wiedemann, M. Cossio, et al. 2021. "Accelerating Detection of Lung Pathologies With Explainable Ultrasound Image Analysis." *Applied Sciences* 11, no. 2: 672.

Bradley, A. P. 1997. "The Use of the Area Under the ROC Curve in the Evaluation of Machine Learning Algorithms." *Pattern Recognition* 30, no. 7: 1145–1159.

Breiman, L. 2017. *Classification and Regression Trees*. Routledge.

Bukkapatnam, A., S. Manjaly, and R. Senthil. 2024. "A Novel Approach to Sepsis Diagnosis: Using Artificial Intelligence to Assist Clinicians and Innovate." *Procedia Computer Science* 231: 237–242.

Cacciamani, G. E., A. Chen, I. S. Gill, and A. J. Hung. 2024. "Artificial Intelligence and Urology: Ethical Considerations for Urologists and Patients." *Nature Reviews Urology* 21, no. 1: 50–59.

Cakiroglu, C., S. Demir, M. H. Ozdemir, B. L. Aylak, G. Sariisik, and L. Abualigah. 2024. "Data-Driven Interpretable Ensemble Learning Methods for the Prediction of Wind Turbine Power Incorporating SHAP Analysis." *Expert Systems With Applications* 237: 121464.

Caruana, R., Y. Lou, J. Gehrke, P. Koch, M. Sturm, and N. Elhadad. 2015. *Intelligible Models for Healthcare: Predicting Pneumonia Risk and Hospital 30-Day Readmission*, 1721–1730. ACM.

Cerekci, E., D. Alis, N. Denizoglu, et al. 2024. "Quantitative Evaluation of Saliency-Based Explainable Artificial Intelligence (XAI) Methods in Deep Learning-Based Mammogram Analysis." *European Journal of Radiology* 173: 111356.

Chaddad, A., J. Peng, J. Xu, and A. Bouridane. 2023. "Survey of Explainable AI Techniques in Healthcare." *Sensors* 23, no. 2: 634.

Chamola, V., V. Hassija, A. R. Sulthana, D. Ghosh, D. Dhingra, and B. Sikdar. 2023. "A Review of Trustworthy and Explainable Artificial Intelligence (XAI)." *IEEE Access* 11: 78994–79015. https://doi.org/10.1109/ACCESS.2023.3294569.

Chettri, B., and R. Ravi. 2024. "A Comparative Study of Medical Device Regulation Between Countries Based on Their Economies." *Expert Review of Medical Devices* 21: 467–478.

Chi, H. R., F. Domingues, H. Zhu, C. Li, K. Kojima, and A. Radwan. 2023. "Healthcare 5.0: In the Perspective of Consumer Internet-of-Things-Based Fog/Cloud Computing." *IEEE Transactions on Consumer Electronics* 69: 745–755.

Chrusciel, J., F. Girardon, L. Roquette, D. Laplanche, A. Duclos, and S. Sanchez. 2021. "The Prediction of Hospital Length of Stay Using Unstructured Data." *BMC Medical Informatics and Decision Making* 21, no. 1: 351.

Dablain, D., C. Bellinger, B. Krawczyk, D. W. Aha, and N. Chawla. 2024. "Understanding Imbalanced Data: XAI & Interpretable ML Framework." *Machine Learning* 113: 3751–3769.

Dalvi-Esfahani, M., M. Mosharaf-Dehkordi, L. W. Leong, T. Ramayah, and A. M. J. Kanaan-Jebna. 2023. "Exploring the Drivers of XAI-Enhanced Clinical Decision Support Systems Adoption: Insights From a Stimulus-Organism-Response Perspective." *Technological Forecasting and Social Change* 195: 122768.

Dar, R. A., M. Rasool, A. Assad, et al. 2022. "Breast Cancer Detection Using Deep Learning: Datasets, Methods, and Challenges Ahead." *Computers in Biology and Medicine* 149: 106073.

Date, S., and M. Thalor. 2023. "AI in Healthcare 5.0: Opportunities and Challenges." *International Journal of Educational and Psychological Sciences* 1, no. 4: 189–191.

Demir, R., S. Koc, D. G. Ozturk, et al. 2024. "Artificial Intelligence Assisted Patient Blood and Urine Droplet Pattern Analysis for Non-Invasive and Accurate Diagnosis of Bladder Cancer." *Scientific Reports* 14, no. 1: 2488.

Dhiman, P., A. Bonkra, A. Kaur, et al. 2023. "Healthcare Trust Evolution With Explainable Artificial Intelligence: Bibliometric Analysis." *Information* 14, no. 10: 541.

Dihan, F. J., S. A. Murad, A. J. M. Muzahid, et al. 2024. "MpoxSLDNet: A Novel CNN Model for Detecting Monkeypox Lesions and

Performance Comparison With Pre-Trained Models." arXiv Preprint arXiv:2405.21016.

Du, Y., A. R. Rafferty, F. M. McAuliffe, L. Wei, and C. Mooney. 2022. "An Explainable Machine Learning-Based Clinical Decision Support System for Prediction of Gestational Diabetes Mellitus." *Scientific Reports* 12, no. 1: 1170.

Dwivedi, Y. K., N. Pandey, W. Currie, and A. Micu. 2024. "Leveraging ChatGPT and Other Generative Artificial Intelligence (AI)-Based Applications in the Hospitality and Tourism Industry: Practices, Challenges and Research Agenda." *International Journal of Contemporary Hospitality Management* 36, no. 1: 1–12.

Famiglini, L., A. Campagner, M. Barandas, G. A. La Maida, E. Gallazzi, and F. Cabitza. 2024. "Evidence-Based XAI: An Empirical Approach to Design More Effective and Explainable Decision Support Systems." *Computers in Biology and Medicine* 170: 108042.

Fontes, M., J. D. S. De Almeida, and A. Cunha. 2024. "Application of Example-Based Explainable Artificial Intelligence (XAI) for Analysis and Interpretation of Medical Imaging: A Systematic Review." *IEEE Access* 12: 26419–26427.

Furriel, B. C., B. D. Oliveira, R. Prôa, et al. 2024. "Artificial Intelligence for Skin Cancer Detection and Classification for Clinical Environment: A Systematic Review." *Frontiers in Medicine* 10: 1305954.

Garg, N., M. Wazid, J. Singh, D. P. Singh, and A. K. Das. 2022. "Security in IoMT-Driven Smart Healthcare: A Comprehensive Review and Open Challenges." *Security and Privacy* 5, no. 5: e235. https://doi.org/10.1002/spy2.235.

Gibbs, A. L., and F. E. Su. 2002. "On Choosing and Bounding Probability Metrics." *International Statistical Review* 70, no. 3: 419–435.

Giuste, F., W. Shi, Y. Zhu, et al. 2022. "Explainable Artificial Intelligence Methods in Combating Pandemics: A Systematic Review." *IEEE Reviews in Biomedical Engineering* 16: 5–21.

Gomathi, L., A. K. Mishra, and A. K. Tyagi. 2023. *Industry 5.0 for Healthcare 5.0: Opportunities, Challenges and Future Research Possibilities*, 204–213. IEEE.

Gordon, W. J., D. Henderson, A. DeSharone, et al. 2020. "Remote Patient Monitoring Program for Hospital Discharged COVID-19 Patients." *Applied Clinical Informatics* 11, no. 5: 792–801.

Goumopoulos, C., and N. Potha. 2023. "Mental Fatigue Detection Using a Wearable Commodity Device and Machine Learning." *Journal of Ambient Intelligence and Humanized Computing* 14, no. 8: 10103–10121.

Grover, V., and M. Dogra. 2024. *Challenges and Limitations of Explainable AI in Healthcare*, 72–85. IGI Global.

Gunning, D., and D. Aha. 2019. "DARPA'S Explainable Artificial Intelligence (XAI) Program." *AI Magazine* 40, no. 2: 44–58.

Guo, H., Y. Zhao, and Y. Lian. 2021. *Video Based Heart Rate Extraction Using Skin ROI Segmentation and Attention CNN*, 1–4. IEEE.

Gupta, R. 2024. *AI-Based Technologies, Challenges, and Solutions for Neurorehabilitation: A Systematic Mapping*, 1–25. Elsevier.

Ha, C., T. Go, and W. Choi. 2024. "Intelligent Healthcare Platform for Diagnosis of Scalp and Hair Disorders." *Applied Sciences* 14, no. 5: 1734.

Hamood Alsamhi, S., A. Hawbani, A. V. Shvetsov, and S. Kumar. 2023. "Advancing Pandemic Preparedness in Healthcare 5.0: A Survey of Federated Learning Applications." *Advances in Human-Computer Interaction* 2023, no. 1: 9992393.

Hassija, V., V. Chamola, A. Mahapatra, et al. 2024. "Interpreting Black-Box Models: A Review on Explainable Artificial Intelligence." *Cognitive Computation* 16, no. 1: 45–74.

He, C., V. Raj, H. Moen, et al. 2024. *VMS: Interactive Visualization to Support the Sensemaking and Selection of Predictive Models*, 229–244. ACM.

Helbing, D., B. S. Frey, G. Gigerenzer, et al. 2019. "Will Democracy Survive Big Data and Artificial Intelligence?" In *Towards Digital Enlightenment: Essays on the Dark and Light Sides of the Digital Revolution*, 73–98. Springer.

Hernandez, R. M., A. A. Ragasa, L. Macatangay, L. Palad, C. Mendez, and C. Mendez. 2023. *Enhanced Super-Resolution Generative Adversarial Networks (ESRGAN) on Melanoma Skin Cancer Detection*, 137–142. ACM.

Hu, B., B. Vasu, and A. Hoogs. 2022. *X-Mir: Explainable Medical Image Retrieval*, 440–450. IEEE.

Hulsen, T. 2023. "Explainable Artificial Intelligence (XAI): Concepts and Challenges in Healthcare." *Ai* 4, no. 3: 652–666.

Hulsen, T. 2024. "Artificial Intelligence in Healthcare: ChatGPT and Beyond."

Iqbal, F., P. Chandra, L. E. S. Lewis, et al. 2024. "Application of Artificial Intelligence to Predict the Sepsis in Neonates Admitted in Neonatal Intensive Care Unit." *Journal of Neonatal Nursing* 30, no. 2: 141–147.

Islam, M. J., A. Rahman, S. Kabir, et al. 2021. "Blockchain-SDN-Based Energy-Aware and Distributed Secure Architecture for IoT in Smart Cities." *IEEE Internet of Things Journal* 9, no. 5: 3850–3864.

Jain, P., R. K. Mishra, A. Deep, and N. Jain. 2024. *Xplainable AI for Deep Learning Model on PCOD Analysis*, 131–152. Elsevier.

Jia, G., H. K. Lam, and Y. Xu. 2021. "Classification of COVID-19 Chest X-Ray and CT Images Using a Type of Dynamic CNN Modification Method." *Computers in Biology and Medicine* 134: 104425.

Jiang, F., Y. Jiang, H. Zhi, et al. 2017. "Artificial Intelligence in Healthcare: Past, Present and Future." *Stroke and Vascular Neurology* 2, no. 4: 230–243.

Jiang, Y., X. Li, H. Luo, S. Yin, and O. Kaynak. 2022. "Quo Vadis Artificial Intelligence?" *Discover Artificial Intelligence* 2, no. 1: 4.

Jopek, M. A., K. Pastuszak, S. Cygert, et al. 2024. "Deep Learning-Based, Multiclass Approach to Cancer Classification on Liquid Biopsy Data." *IEEE Journal of Translational Engineering in Health and Medicine* 12: 306–313. https://doi.org/10.1109/JTEHM.2024.3360865.

Joshi, S., R. Kumar, V. Rai, A. Dwivedi, and V. C. Tripathi. 2024. *A Detailed Structured Monitoting System for ICU Patient Mortality Rate by Using AI Algorithm*. Vol. 5, 476–480. IEEE.

Kacena, M. A., L. I. Plotkin, and J. C. Fehrenbacher. 2024. "The Use of Artificial Intelligence in Writing Scientific Review Articles." In *Current Osteoporosis Reports*, 1–7. IEEE.

Kadri, F., A. Dairi, F. Harrou, and Y. Sun. 2023. "Towards Accurate Prediction of Patient Length of Stay at Emergency Department: A GAN-Driven Deep Learning Framework." *Journal of Ambient Intelligence and Humanized Computing* 14, no. 9: 11481–11495.

Karyamsetty, H. J., S. A. Khan, and A. Nayyar. 2024. "Envisioning Toward Modernization of Society 5.0—A Prospective Glimpse on Status, Opportunities, and Challenges With XAI." In *XAI Based Intelligent Systems for Society 5.0*, 223–267. Springer.

Kazimierczak, N., W. Kazimierczak, Z. Serafin, P. Nowicki, J. Nożewski, and J. Janiszewska-Olszowska. 2024. "AI in Orthodontics: Revolutionizing Diagnostics and Treatment Planning—A Comprehensive Review." *Journal of Clinical Medicine* 13, no. 2: 344.

Keskenler, M. F., E. Çelik, and D. Dal. 2024. "A New Multi-Layer Machine Learning (MLML) Architecture for Non-Invasive Skin Cancer Diagnosis on Dermoscopic Images." *Journal of Electrical Engineering & Technology* 19: 2739–2755.

Khan, S. A., and A. Naim. 2024. *XAI in Society 5.0 Through the Lens of Marketing and HRM*, 327–363. Elsevier.

Khan, T. A., A. Fatima, T. Shahzad, et al. 2023. "Secure IoMT for Disease Prediction Empowered With Transfer Learning in Healthcare 5.0, the Concept and Case Study." *IEEE Access* 11: 39418–39430.

Kherchouche, A., O. Ben-Ahmed, C. Guillevin, et al. 2022. "Attention-Guided Neural Network for Early Dementia Detection Using MRS Data." *Computerized Medical Imaging and Graphics* 99: 102074.

Khosravi, H., S. B. Shum, G. Chen, et al. 2022. "Explainable Artificial Intelligence in Education." *Computers and Education: Artificial Intelligence* 3: 100074.

Kim, J., S. Lee, E. Hwang, et al. 2020. "Limitations of Deep Learning Attention Mechanisms in Clinical Research: Empirical Case Study Based on the Korean Diabetic Disease Setting." *Journal of Medical Internet Research* 22, no. 12: e18418.

Kiruthika, M., K. Moorthi, M. A. Devi, and S. A. Roseline. 2024. *Role of XAI in Building a Super Smart Society 5.0*, 295–326. Elsevier.

Kisten, M., A. E. Ezugwu, and M. S. Olusanya. 2024. "Explainable Artificial Intelligence Model for Predictive Maintenance in Smart Agricultural Facilities." *IEEE Access* 12: 24348–24367.

Kök, I., F. Y. Okay, Ö. Muyanl, and S. Özdemir. 2023. "Explainable Artificial Intelligence (XAI) for Internet of Things: A Survey." *IEEE Internet of Things Journal* 10: 14764–14779.

Kose, U., N. Sengoz, X. Chen, and J. A. M. Saucedo. 2024. *Explainable Artificial Intelligence (XAI) in Healthcare*. CRC Press.

Kousha, K., and M. Thelwall. 2024. "Artificial Intelligence to Support Publishing and Peer Review: A Summary and Review." *Learned Publishing* 37, no. 1: 4–12.

Kullback, S., and R. A. Leibler. 1951. "On Information and Sufficiency." *Annals of Mathematical Statistics* 22, no. 1: 79–86.

Kundu, D., M. M. Rahman, A. Rahman, et al. 2024. "Federated Deep Learning for Monkeypox Disease Detection on GAN-Augmented Dataset." *IEEE Access* 12: 32819–32829. https://doi.org/10.1109/ACCESS.2024.3370838.

Kundu, D., U. R. Siddiqi, and M. M. Rahman. 2022. *Vision Transformer Based Deep Learning Model for Monkeypox Detection*, 1021–1026. IEEE.

Lahmiri, S., and A. Shmuel. 2019. "Detection of Parkinson's Disease Based on Voice Patterns Ranking and Optimized Support Vector Machine." *Biomedical Signal Processing and Control* 49: 427–433.

Lammert, O., B. Richter, C. Schütze, K. Thommes, and B. Wrede. 2024. "Humans in XAI: Increased Reliance in Decision-Making Under Uncertainty by Using Explanation Strategies." *Frontiers in Behavioral Economics* 3: 1377075.

Lin, Y., M. Bariya, H. Y. Y. Nyein, et al. 2019. "Porous Enzymatic Membrane for Nanotextured Glucose Sweat Sensors With High Stability Toward Reliable Noninvasive Health Monitoring." *Advanced Functional Materials* 29, no. 33: 1902521.

Liu, J., Y. Xie, X. Shu, et al. 2024. "Value Function Assessment to Different RL Algorithms for Heparin Treatment Policy of Patients With Sepsis in ICU." *Artificial Intelligence in Medicine* 147: 102726.

Loh, H. W., C. P. Ooi, S. Seoni, P. D. Barua, F. Molinari, and U. R. Acharya. 2022. "Application of Explainable Artificial Intelligence for Healthcare: A Systematic Review of the Last Decade (2011–2022)." *Computer Methods and Programs in Biomedicine* 226: 107161.

Longo, L., M. Brcic, F. Cabitza, et al. 2024. "Explainable Artificial Intelligence (XAI) 2.0: A Manifesto of Open Challenges and Interdisciplinary Research Directions." *Information Fusion* 106: 102301.

Lu, S., Z. Zhu, J. M. Gorriz, S. H. Wang, and Y. D. Zhang. 2022. "NAGNN: Classification of COVID-19 Based on Neighboring Aware Representation From Deep Graph Neural Network." *International Journal of Intelligent Systems* 37, no. 2: 1572–1598.

Lundberg, S. 2017. "A Unified Approach to Interpreting Model Predictions." arXiv Preprint arXiv:1705.07874.

Ma, X., Y. Si, Z. Wang, and Y. Wang. 2020. "Length of Stay Prediction for ICU Patients Using Individualized Single Classification Algorithm." *Computer Methods and Programs in Biomedicine* 186: 105224.

Maddigan, P., A. Lensen, and B. Xue. 2024. "Explaining Genetic Programming Trees Using Large Language Models." arXiv Preprint arXiv:2403.03397.

Magd, H., M. S. A. Ansari, and S. A. Khan. 2024. *Need for Explainable Artificial Intelligence Ethnic Decision-Making in Society 5.0*, 103–127. Elsevier.

Mahbooba, B., M. Timilsina, R. Sahal, and M. Serrano. 2021. "Explainable Artificial Intelligence (XAI) to Enhance Trust Management in Intrusion Detection Systems Using Decision Tree Model." *Complexity* 2021, no. 1: 6634811.

Markus, A. F., J. A. Kors, and P. R. Rijnbeek. 2021. "The Role of Explainability in Creating Trustworthy Artificial Intelligence for Health Care: A Comprehensive Survey of the Terminology, Design Choices, and Evaluation Strategies." *Journal of Biomedical Informatics* 113: 103655.

Mbunge, E., B. Muchemwa, J. Batani, et al. 2021. "Sensors and Healthcare 5.0: Transformative Shift in Virtual Care Through Emerging Digital Health Technologies." *Global Health Journal* 5, no. 4: 169–177. https://doi.org/10.1016/j.glohj.2021.11.008.

Mennella, C., U. Maniscalco, G. De Pietro, and M. Esposito. 2024. "Ethical and Regulatory Challenges of AI Technologies in Healthcare: A Narrative Review." *Heliyon* 10, no. 4: e26297. https://doi.org/10.1016/j.heliyon.2024.e26297.

Metta, C., A. Beretta, R. Pellungrini, S. Rinzivillo, and F. Giannotti. 2024. "Towards Transparent Healthcare: Advancing Local Explanation Methods in Explainable Artificial Intelligence." *Bioengineering* 11, no. 4: 369.

Mohanty, M. N., and A. Das. 2024. "Skin Cancer Detection From Dermatoscopic Images Using Hybrid Fuzzy Ensemble Learning Model." *International Journal of Fuzzy Systems* 26, no. 1: 260–273.

Molnar, C., T. Freiesleben, G. König, et al. 2023. *Relating the Partial Dependence Plot and Permutation Feature Importance to the Data Generating Process*, 456–479. Springer.

Moncada-Torres, A., v M. Maaren C, M. P. Hendriks, S. Siesling, and G. Geleijnse. 2021. "Explainable Machine Learning Can Outperform Cox Regression Predictions and Provide Insights in Breast Cancer Survival." *Scientific Reports* 11, no. 1: 6968.

Moreno-Sánchez, P. A., M. Aalto, and v M. Gils. 2024. "Prediction of Patient Flow in the Emergency Department Using Explainable Artificial Intelligence." *Digital Health* 10: 20552076241264194.

Moulaei, K., M. R. Afrash, M. Parvin, et al. 2024. "Explainable Artificial Intelligence (XAI) for Predicting the Need for Intubation in Methanol-Poisoned Patients: A Study Comparing Deep and Machine Learning Models." *Scientific Reports* 14, no. 1: 15751.

Moyano-Fernández, C., J. Rueda, J. Delgado, and T. Ausín. 2024. "May Artificial Intelligence Take Health and Sustainability on a Honeymoon? Towards Green Technologies for Multidimensional Health and Environmental Justice." *Global Bioethics* 35, no. 1: 2322208.

Nallakaruppan, M., N. Shankar, P. B. Bhuvanagiri, S. Padmanaban, and S. B. Khan. 2024. "Advancing Solar Energy Integration: Unveiling XAI Insights for Enhanced Power System Management and Sustainable Future." *Ain Shams Engineering Journal* 15: 102740.

Natarajan, R., G. H. Lokesh, F. Flammini, A. Premkumar, V. K. Venkatesan, and S. K. Gupta. 2023. "A Novel Framework on Security and Energy Enhancement Based on Internet of Medical Things for Healthcare 5.0." *Infrastructures* 8, no. 2: 22.

Nazar, M., M. M. Alam, E. Yafi, and M. M. Su'ud. 2021. "A Systematic Review of Human–Computer Interaction and Explainable Artificial Intelligence in Healthcare With Artificial Intelligence Techniques." *IEEE Access* 9: 153316–153348.

Nazat, S., L. Li, and M. Abdallah. 2024. "XAI-ADS: An Explainable Artificial Intelligence Framework for Enhancing Anomaly Detection in

Autonomous Driving Systems." *IEEE Access* 12: 48583–48607. https://doi.org/10.1109/ACCESS.2024.3383431.

Nozari, H., J. Ghahremani-Nahr, and A. Szmelter-Jarosz. 2024. *AI and Machine Learning for Real-World Problems*, 1–12. Elsevier.

Ofosu-Ampong, K. 2024. "Artificial Intelligence Research: A Review on Dominant Themes, Methods, Frameworks and Future Research Directions." *Telematics and Informatics Reports* 14: 100127.

Olaoye, F., K. Potter, and L. Doris. 2024. "Explainable AI: Interpreting and Understanding Machine Learning Models." Tech. Rep., EasyChair.

Olawade, D. B., A. C. David-Olawade, O. Z. Wada, A. J. Asaolu, T. Adereni, and J. Ling. 2024. "Artificial Intelligence in Healthcare Delivery: Prospects and Pitfalls." *Journal of Medicine, Surgery, and Public Health* 3: 100108. https://doi.org/10.1016/j.glmedi.2024.100108.

Park, S., H. M. Gach, S. Kim, S. J. Lee, and Y. Motai. 2021. "Autoencoder-Inspired Convolutional Network-Based Super-Resolution Method in MRI." *IEEE Journal of Translational Engineering in Health and Medicine* 9: 1800113.

Parr, T., J. Hamrick, and J. D. Wilson. 2024. "Nonparametric Feature Impact and Importance." *Information Sciences* 653: 119563.

Parveen, A., and G. Kannan. 2024. "Healthcare Transformed: A Comprehensive Survey of Artificial Intelligence Trends in Healthcare Industries." In *Digital Healthcare in Asia and Gulf Region for Healthy Aging and More Inclusive Societies*, 395–424. Elsevier.

Patel, A. U., Q. Gu, R. Esper, D. Maeser, and N. Maeser. 2024. "The Crucial Role of Interdisciplinary Conferences in Advancing Explainable AI in Healthcare." *BioMedInformatics* 4, no. 2: 1363–1383.

Patidar, N., S. Mishra, R. Jain, et al. 2024. "Transparency in AI Decision Making: A Survey of Explainable AI Methods and Applications." *Advances of Robotic Technology* 2, no. 1: 000110.

Patra, G., and S. Datta. 2024. "XAI for Society 5.0: Requirements, Opportunities, and Challenges in the Current Context." In *XAI Based Intelligent Systems for Society 5.0*, 269–293. Springer.

Pulipeti, S., P. Chithaluru, M. Kumar, P. Narsimhulu, and V. UM. 2024. *Explainable AI: Methods, Frameworks, and Tools for Healthcare 5.0*, 71–86. Springer.

Qadri, Y. A., A. Nauman, Y. B. Zikria, A. V. Vasilakos, and S. W. Kim. 2020. "The Future of Healthcare Internet of Things: A Survey of Emerging Technologies." *IEEE Communications Surveys & Tutorials* 22, no. 2: 1121–1167.

Raha, A. D., M. Gain, R. Debnath, et al. 2024. "Attention to Monkeypox: An Interpretable Monkeypox Detection Technique Using Attention Mechanism." *IEEE Access* 12: 51942–51965. https://doi.org/10.1109/ACCESS.2024.3385099.

Rahman, A., C. Chakraborty, A. Anwar, et al. 2022. "SDN–IoT Empowered Intelligent Framework for Industry 4.0 Applications During COVID-19 Pandemic." *Cluster Computing* 25: 2351–2368.

Rahman, A., T. Debnath, D. Kundu, et al. 2024. "Machine Learning and Deep Learning-Based Approach in Smart Healthcare: Recent Advances, Applications, Challenges and Opportunities." *AIMS Public Health* 11, no. 1: 58–109.

Rahman, A., T. Debnath, D. Kundu, et al. 2024. *Unlocking the Potential of IoT, AI, and Blockchain in Transforming Public and Private Industries*. Cambridge Scholars Publishing.

Rahman, A., K. Hasan, D. Kundu, et al. 2023. "On the ICN-IoT With Federated Learning Integration of Communication: Concepts, Security-Privacy Issues, Applications, and Future Perspectives." *Future Generation Computer Systems* 138: 61–88.

Rahman, A., M. S. Hossain, G. Muhammad, et al. 2023. "Federated Learning-Based AI Approaches in Smart Healthcare: Concepts, Taxonomies, Challenges and Open Issues." *Cluster Computing* 26, no. 4: 2271–2311.

Rahman, A., J. Islam, D. Kundu, et al. 2023. "Impacts of Blockchain in Software-Defined Internet of Things Ecosystem With Network Function Virtualization for Smart Applications: Present Perspectives and Future Directions." *International Journal of Communication Systems* 38, no. 1: e5429. https://doi.org/10.1002/dac.5429.

Rahman, A., M. J. Islam, S. S. Band, G. Muhammad, K. Hasan, and P. Tiwari. 2023. "Towards a Blockchain-SDN-Based Secure Architecture for Cloud Computing in Smart Industrial IoT." *Digital Communications and Networks* 9, no. 2: 411–421.

Rahman, A., M. J. Islam, A. Montieri, et al. 2021. "Smartblock-Sdn: An Optimized Blockchain-Sdn Framework for Resource Management in Iot." *IEEE Access* 9: 28361–28376.

Rahman, A., M. J. Islam, Z. Rahman, et al. 2020. "Distb-Condo: Distributed Blockchain-Based Iot-Sdn Model for Smart Condominium." *IEEE Access* 8: 209594–209609.

Rahman, A., M. S. I. Khan, A. Montieri, et al. 2024. "BlockSD-5GNet: Enhancing Security of 5G Network Through Blockchain-SDN With ML-Based Bandwidth Prediction." *Transactions on Emerging Telecommunications Technologies* 35, no. 4: e4965.

Rahman, A., M. K. Nasir, Z. Rahman, A. Mosavi, S. Shahab, and B. Minaei-Bidgoli. 2020. "Distblockbuilding: A Distributed Blockchain-Based Sdn-Iot Network for Smart Building Management." *IEEE Access* 8: 140008–140018.

Rahman, A., M. A. H. Wadud, M. J. Islam, et al. 2024. "Internet of Medical Things and Blockchain-Enabled Patient-Centric Agent Through Sdn for Remote Patient Monitoring in 5g Network." *Scientific Reports* 14, no. 1: 5297.

Rahman, M. M., D. Kundu, S. A. Suha, U. R. Siddiqi, and S. K. Dey. 2022. "Hospital Patients' Length of Stay Prediction: A Federated Learning Approach." *Journal of King Saud University, Computer and Information Sciences* 34, no. 10: 7874–7884.

Rajapaksha, S., H. Kalutarage, M. O. Al-Kadri, A. Petrovski, G. Madzudzo, and M. Cheah. 2023. "Ai-Based Intrusion Detection Systems for In-Vehicle Networks: A Survey." *ACM Computing Surveys* 55, no. 11: 1–40.

Rehman, A., S. Abbas, M. Khan, T. M. Ghazal, K. M. Adnan, and A. Mosavi. 2022. "A Secure Healthcare 5.0 System Based on Blockchain Technology Entangled With Federated Learning Technique." *Computers in Biology and Medicine* 150: 106019. https://doi.org/10.1016/j.compbiomed.2022.106019.

Ribeiro, M. T., S. Singh, and C. Guestrin. 2016. *"Why Should I Trust You?" Explaining the Predictions of Any Classifier*, 1135–1144. ACM.

Rigby, M. J. 2019. "Ethical Dimensions of Using Artificial Intelligence in Health Care." *AMA Journal of Ethics* 21, no. 2: 121–124.

Rong, Y., T. Leemann, T. T. Nguyen, et al. 2023. "Towards Human-Centered Explainable AI: A Survey of User Studies for Model Explanations." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 46: 2104–2122.

Ross, A. S., M. C. Hughes, and F. Doshi-Velez. 2017. "Right for the Right Reasons: Training Differentiable Models by Constraining Their Explanations." arXiv Preprint arXiv:1703.03717.

Rubegni, E., O. Ayoub, and S. M. R. Rizzo. 2024. *Designing for Complementarity: A Conceptual Framework to Go Beyond the Current Paradigm of Using XAI in Healthcare*, 277–296. Springer.

Rudin, C. 2019. "Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead." *Nature Machine Intelligence* 1, no. 5: 206–215.

Saadatmand, S., K. Salimifard, R. Mohammadi, A. Kuiper, M. Marzban, and A. Farhadi. 2023. "Using Machine Learning in Prediction of ICU Admission, Mortality, and Length of Stay in the Early Stage of Admission of COVID-19 Patients." *Annals of Operations Research* 328, no. 1: 1043–1071.

Sadeghi, Z., R. Alizadehsani, M. A. Cifci, et al. 2024. "A Review of Explainable Artificial Intelligence in Healthcare." *Computers and Electrical Engineering* 118: 109370. https://doi.org/10.1016/j.compeleceng.2024.109370.

Saha, S., A. Pagnozzi, P. Bourgeat, et al. 2020. "Predicting Motor Outcome in Preterm Infants From Very Early Brain Diffusion MRI Using a Deep Learning Convolutional Neural Network (CNN) Model." *NeuroImage* 215: 116807.

Saraswat, D., P. Bhattacharya, A. Verma, et al. 2022. "Explainable AI for Healthcare 5.0: Opportunities and Challenges." *IEEE Access* 10: 84486–84517. https://doi.org/10.1109/ACCESS.2022.3197671.

Seoni, S., F. Molinari, U. R. Acharya, et al. 2024. "Application of Spatial Uncertainty Predictor in CNN-BiLSTM Model Using Coronary Artery Disease ECG Signals." *Information Sciences* 665: 120383. https://doi.org/10.1016/j.ins.2024.120383.

Seshadri, D. R., E. V. Davies, E. R. Harlow, et al. 2020. "Wearable Sensors for COVID-19: A Call to Action to Harness Our Digital Infrastructure for Remote Patient Monitoring and Virtual Assessments." *Frontiers in Digital Health* 2: 8.

Sethi, A., S. Dharmavaram, and S. Somasundaram. 2024. *Explainable Artificial Intelligence (XAI) Approach to Heart Disease Prediction*, 1–6. IEEE.

Shaikh, T. A., W. A. Mir, and S. Sofi. 2022. *Decrypting the Black Boxing of Artificial Intelligence Using Explainable Artificial Intelligence in Smart Healthcare*, 53–82. Springer.

Sharma, B., D. Kaushal, M. Sharma, S. Joshi, S. Gopal, and P. Gupta. 2023. *Integration of AI, Digital Twin and Internet of Medical Things (IoMT) for Healthcare 5.0: A Bibliometric Analysis*, 1313–1318. IEEE.

Shen, Y., N. Wu, J. Phang, et al. 2021. "An Interpretable Classifier for High-Resolution Breast Cancer Screening Images Utilizing Weakly Supervised Localization." *Medical Image Analysis* 68: 101908.

Shin, H. C., H. R. Roth, M. Gao, et al. 2016. "Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning." *IEEE Transactions on Medical Imaging* 35, no. 5: 1285–1298.

Shome, D., T. Kar, S. N. Mohanty, et al. 2021. "Covid-Transformer: Interpretable Covid-19 Detection Using Vision Transformer for Healthcare." *International Journal of Environmental Research and Public Health* 18, no. 21: 11086.

Sikder, J., U. K. Das, and A. S. Anwar. 2021. *Cancer Cell Segmentation Based on Unsupervised Clustering and Deep Learning*, 607–620. Springer.

Sikder, J., U. K. Das, and R. J. Chakma. 2021. "Supervised Learning-Based Cancer Detection." *International Journal of Advanced Computer Science and Applications* 12, no. 5: 863–869.

Sikder, J., K. I. Sarek, and U. K. Das. 2021. "Fish Disease Detection System: A Case Study of Freshwater Fishes of Bangladesh."

Simonyan, K., A. Vedaldi, and A. Zisserman. 2013. "Deep Inside Convolutional Networks: Visualising Image Classification Models and Saliency Maps." arXiv Preprint arXiv:1312.6034.

Sindiramutty, S. R., W. J. Tee, S. Balakrishnan, et al. 2024. *Explainable AI in Healthcare Application*, 123–176. IGI Global.

Singh, A., A. Kanaujia, V. K. Singh, and R. Vinuesa. 2024. "Artificial Intelligence for Sustainable Development Goals: Bibliometric Patterns and Concept Evolution Trajectories." *Sustainable Development* 32, no. 1: 724–754. https://doi.org/10.1002/sd.2706.

Sivakumar, M. S., L. M. Leo, T. Gurumekala, V. Sindhu, and A. S. Priyadharshini. 2024. "Deep Learning in Skin Lesion Analysis for Malignant Melanoma Cancer Identification." *Multimedia Tools and Applications* 83, no. 6: 17833–17853.

Sivamohan, S., and S. Sridhar. 2023. "An Optimized Model for Network Intrusion Detection Systems in Industry 4.0 Using XAI Based bi-LSTM Framework." *Neural Computing and Applications* 35, no. 15: 11459–11475.

Song, Y., S. Zheng, L. Li, et al. 2021. "Deep Learning Enables Accurate Diagnosis of Novel Coronavirus (COVID-19) With CT Images." *IEEE/ACM Transactions on Computational Biology and Bioinformatics* 18, no. 6: 2775–2780.

Soni, T., D. Gupta, M. Uppal, and S. Juneja. 2023. *Explicability of Artificial Intelligence in Healthcare 5.0*, 1256–1261. IEEE.

Sriperumbudur, B. K., K. Fukumizu, and G. R. Lanckriet. 2011. "Universality, Characteristic Kernels and RKHS Embedding of Measures." *Journal of Machine Learning Research* 12, no. 7: 2389–2410.

Suha, S. A., and T. F. Sanam. 2023. "Exploring Dominant Factors for Ensuring the Sustainability of Utilizing Artificial Intelligence in Healthcare Decision Making: An Emerging Country Context." *International Journal of Information Management Data Insights* 3, no. 1: 100170.

Sundararajan, M., A. Taly, and Q. Yan. 2017. *Axiomatic Attribution for Deep Networks*, 3319–3328. PMLR.

Sutton, R. T., O. R. Zaiane, R. Goebel, and D. C. Baumgart. 2022. "Artificial Intelligence Enabled Automated Diagnosis and Grading of Ulcerative Colitis Endoscopy Images." *Scientific Reports* 12, no. 1: 2748.

Tabaa, M., F. Monteiro, H. Bensag, and A. Dandache. 2020. "Green Industrial Internet of Things From a Smart Industry Perspectives." *Energy Reports* 6: 430–446.

Taddeo, M., P. Jones, R. Abbas, K. Vogel, and K. Michael. 2023. "Socio-Technical Ecosystem Considerations: An Emergent Research Agenda for AI in Cybersecurity." *IEEE Transactions on Technology and Society* 4, no. 2: 112–118.

Tajmirriahi, M., R. Kafieh, Z. Amini, and H. Rabbani. 2021. "A Lightweight Mimic Convolutional Auto-Encoder for Denoising Retinal Optical Coherence Tomography Images." *IEEE Transactions on Instrumentation and Measurement* 70: 1–8.

Thomas, I., S. Y. Oh, and D. A. Szafir. 2024. *Assessing User Trust in Active Learning Systems: Insights From Query Policy and Uncertainty Visualization*, 772–786. ACM.

Tjoa, E., and C. Guan. 2021. "A Survey on Explainable Artificial Intelligence (XAI): Toward Medical XAI." *IEEE Transactions on Neural Networks and Learning Systems* 32, no. 11: 4793–4813. https://doi.org/10.1109/TNNLS.2020.3027314.

Ueda, D., T. Kakinuma, S. Fujita, et al. 2024. "Fairness of Artificial Intelligence in Healthcare: Review and Recommendations." *Japanese Journal of Radiology* 42, no. 1: 3–15.

Ustun, B., A. Spangher, and Y. Liu. 2019. *Actionable Recourse in Linear Classification*, 10–19. ACM.

Vainio-Pekka, H., M. O. O. Agbese, M. Jantunen, et al. 2023. "The Role of Explainable Ai in the Research Field of Ai Ethics." *ACM Transactions on Interactive Intelligent Systems* 13, no. 4: 1–39.

Vaswani, A. 2017. "Attention is All You Need." *Advances in Neural Information Processing Systems* 30: 2389–2410.

Viswan, V., N. Shaffi, M. Mahmud, K. Subramanian, and F. Hajamohideen. 2024. "Explainable Artificial Intelligence in Alzheimer's Disease Classification: A Systematic Review." *Cognitive Computation* 16, no. 1: 1–44.

Wachter, S., B. Mittelstadt, and C. Russell. 2017. "Counterfactual Explanations Without Opening the Black Box: Automated Decisions and the GDPR." *Harvard Journal of Law & Technology* 31: 841.

Wang, J., X. Qiao, C. Liu, et al. 2021. "Automated ECG Classification Using a Non-Local Convolutional Block Attention Module." *Computer Methods and Programs in Biomedicine* 203: 106006.

Wang, S., Y. Wang, D. Wang, Y. Yin, Y. Wang, and Y. Jin. 2020. "An Improved Random Forest-Based Rule Extraction Method for Breast Cancer Diagnosis." *Applied Soft Computing* 86: 105941.

Wang, X., Z. Wu, W. Huang, et al. 2023. "VIS+ AI: Integrating Visualization With Artificial Intelligence for Efficient Data Analysis." *Frontiers of Computer Science* 17, no. 6: 176709.

Wang, Y. C., T. C. T. Chen, and M. C. Chiu. 2023. "An Improved Explainable Artificial Intelligence Tool in Healthcare for Hospital Recommendation." *Healthcare Analytics* 3: 100147.

Wazid, M., A. K. Das, N. Mohd, and Y. Park. 2022. "Healthcare 5.0 Security Framework: Applications, Issues and Future Research Directions." *IEEE Access* 10: 129429–129442.

Wazid, M., J. Singh, A. K. Das, and J. J. Rodrigues. 2023. "An Ensemble-Based Machine Learning-Envisioned Intrusion Detection in Industry 5.0-Driven Healthcare Applications." *IEEE Transactions on Consumer Electronics* 139: 109560.

Weber, L., S. Lapuschkin, A. Binder, and W. Samek. 2023. "Beyond Explaining: Opportunities and Challenges of XAI-Based Model Improvement." *Information Fusion* 92: 154–176.

Williamson, S. M., and V. Prybutok. 2024. "Balancing Privacy and Progress: A Review of Privacy Challenges, Systemic Oversight, and Patient Perceptions in AI-Driven Healthcare." *Applied Sciences* 14, no. 2: 675.

Wu, C., Y. M. Tang, W. T. Kuo, H. T. Yip, and K. Y. Chau. 2024. "Healthcare 5.0: A Secure and Distributed Network for System Informatics in Medical Surgery." *International Journal of Medical Informatics* 186: 105415.

Xu, Z., T. Wang, J. Cao, Z. Bao, T. Jiang, and F. Gao. 2021. "BECT Spike Detection Based on Novel EEG Sequence Features and LSTM Algorithms." *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 29: 1734–1743.

Yildirim, O., U. B. Baloglu, R. S. Tan, E. J. Ciaccio, and U. R. Acharya. 2019. "A New Approach for Arrhythmia Classification Using Deep Coded Features and LSTM Networks." *Computer Methods and Programs in Biomedicine* 176: 121–133.

Zeiler, M. D., and R. Fergus. 2014. *Visualizing and Understanding Convolutional Networks*, 818–833. Springer.

Zhan, P., S. Wang, J. Wang, et al. 2021. "Temporal Anomaly Detection on IIoT-Enabled Manufacturing." *Journal of Intelligent Manufacturing* 32, no. 6: 1669–1678.

Zhang, C., and Y. Lu. 2021. "Study on Artificial Intelligence: The State of the Art and Future Prospects." *Journal of Industrial Information Integration* 23: 100224.

Zhang, S., J. Yu, X. Xu, et al. 2024. *Rethinking Human-AI Collaboration in Complex Medical Decision Making: A Case Study in Sepsis Diagnosis*, 1–18. ACM.