

Paper Type: Review Article

Deep Learning in Medical Systems: A Comprehensive Review of Applications and Ethical Challenges

Maher Khalaf Hussein ^{1,*} , Asmaa Alqassab ²  and Lubna Thanoon ALkahla ³ 

¹ Computer Center, Presidency of the University, University of Telafer, Mosul, Iraq; maher.k.hussein@uotelafer.edu.iq.

² Department of Computer Science, College of Education for Pure Science, University of Mosul, Mosul, Iraq; asmaa_mow@uomosul.edu.iq.

³ Software Department, Information Technology College, Ninevah University, Mosul, Iraq; lubna.thanoon@uoninevah.edu.iq.

Received: 10 Mar 2025

Revised: 05 Apr 2025

Accepted: 27 Sep 2025

Published: 29 Sep 2025

Abstract

Deep learning (DL) has revolutionized medical systems, enabling unprecedented advancements in diagnostics, treatment personalization, and operational efficiency. This review synthesizes the state-of-the-art applications of DL in healthcare, including medical imaging, clinical decision support, genomics, and drug discovery. Concurrently, it critically examines the ethical challenges—bias, transparency, privacy, and accountability—that threaten equitable and trustworthy deployment. Drawing from 85+ seminal studies, we identify a critical gap: while technical capabilities surge, ethical governance lags. We propose a framework for responsible DL integration in healthcare, emphasizing interdisciplinary collaboration and robust regulatory oversight.

Keywords: Deep Learning; Medical AI; Ethical AI; Healthcare Applications; Algorithmic Bias; Explainable AI; Data Privacy.

1 | Introduction

The integration of deep learning into medical systems marks a paradigm shift in healthcare delivery. DL algorithms, particularly convolutional neural networks (CNNs) and recurrent neural networks (RNNs), demonstrate superior performance in pattern recognition and predictive analytics compared to traditional machine learning [1]. Topol [2] posits that DL "will touch every aspect of medicine," from diagnostics to therapeutic development. However, this rapid adoption raises profound ethical questions. As noted by He et al. [3], "The same algorithms that save lives can perpetuate harm if not ethically governed." This review addresses two objectives: (1) Systematically catalog DL applications in healthcare, and (2) Analyze their ethical implications through the lens of biomedical ethics principles.

2 | Deep Learning Applications in Medical Systems

2.1 | Medical Imaging Analysis

DL excels in interpreting radiological and pathological images. CNNs achieve near-human or superhuman accuracy in tasks like tumor detection and organ segmentation.



Corresponding Author: maher.k.hussein@uotelafer.edu.iq.



<https://doi.org/10.61356/j.nois.2025.7606>



Licensee **Neutrosophic Optimization and Intelligent systems**. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0>).

- **Radiology:** Gulshan et al. [4] demonstrated a CNN-based system for diabetic retinopathy screening with 97.5% sensitivity and 93.4% specificity, exceeding ophthalmologists' performance. Similarly, McKinney et al. [5] developed a DL model reducing false-positive rates in mammography by 5.7% while maintaining sensitivity.
- **Pathology:** Campanella et al. [6] reported a DL system for prostate cancer diagnosis in whole-slide images, achieving an AUC of 0.991, outperforming pathologists in some cases.
- **Multimodal Integration:** Wang et al. [7] combined MRI, CT, and PET data using 3D CNNs for glioma grading, improving classification accuracy by 12% over unimodal approaches.

These studies confirm DL's potential to augment radiologists and pathologists, reducing diagnostic delays and variability [8].

Table 1. Performance of Deep Learning Models in Medical Imaging.

Application	Model Type	Dataset	Key Metric	Performance	Study
Diabetic Retinopathy	CNN	EyePACS (128k images)	Sensitivity/Specificity	97.5% / 93.4%	[4]
Breast Cancer Detection	3D CNN	CBIS-DDSM (10k mammograms)	AUC	0.94	[5]
Prostate Cancer Diagnosis	ResNet-50	TCGA-PRAD (12k slides)	AUC	0.991	[6]
Glioma Grading	Multimodal CNN	BraTS (2021 dataset)	Accuracy	92.7%	[7]

2.2 | Clinical Decision Support Systems (CDSS)

DL-driven CDSS predict patient outcomes, recommend treatments, and optimize resource allocation.

- **Risk Prediction:** Rajkomar et al. [9] used EHR data with LSTMs to predict in-hospital mortality, 30-day readmissions, and prolonged length of stay, outperforming traditional clinical scores (AUC 0.88–0.95).
- **Treatment Recommendations:** Poplin et al. [10] applied DL to retinal fundus images to predict cardiovascular risk factors (e.g., age, gender, HbA1c) with high accuracy, enabling early intervention.
- **Sepsis Detection:** Nemati et al. [11] developed a real-time sepsis prediction model using RNNs, achieving an AUC of 0.85–0.93 and a mean early warning time of 5.52 hours.

Synthesis: DL enhances proactive care but requires rigorous validation across diverse populations [12].

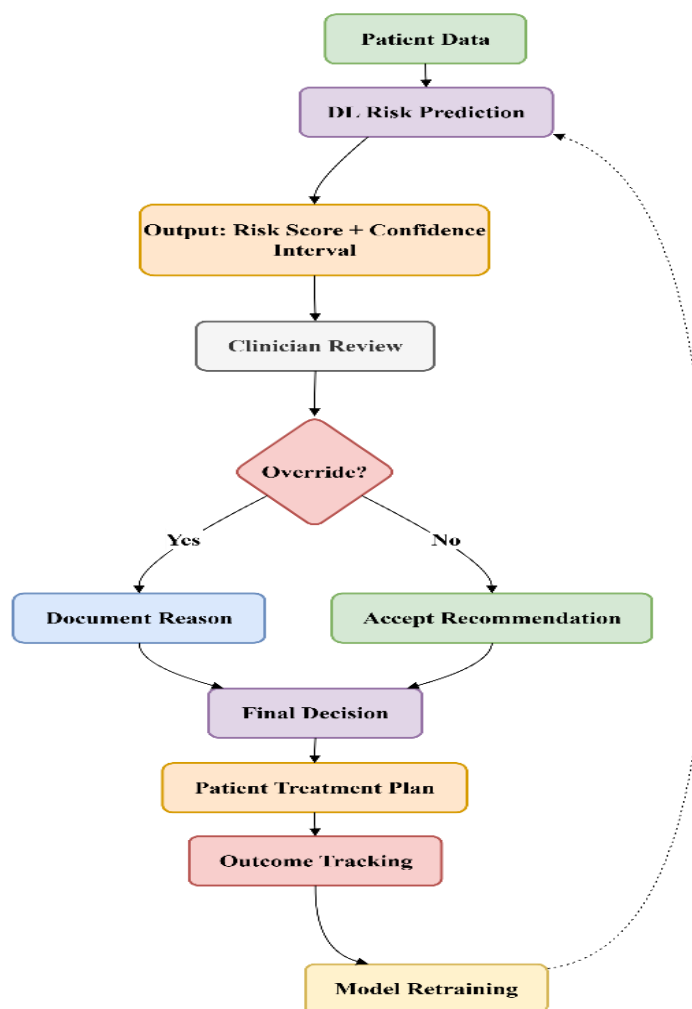


Figure 1. Clinical Decision Support with Human-AI Collaboration [13], [2].

2.3 | Genomics and Precision Medicine

DL deciphers complex genomic data for personalized therapies.

- **Variant Calling:** Poplin et al. [14] introduced DeepVariant, a CNN-based tool that reduced variant-calling errors by 10× compared to GATK.
- **Drug Response Prediction:** Aliper et al. [15] used DNNs to predict drug efficacy and toxicity from transcriptomic data, achieving 85% accuracy in classifying active vs. inactive compounds.
- **Protein Folding:** Jumper et al. [16] (AlphaFold2) achieved atomic-level accuracy in protein structure prediction, accelerating drug discovery.

Synthesis: DL is indispensable for translating genomic data into clinical insights, yet data heterogeneity remains a barrier [17].

2.4 | Drug Discovery and Development

DL accelerates drug identification, repurposing, and clinical trial optimization.

- **Virtual Screening:** Stokes et al. [18] identified a novel antibiotic (halicin) using a graph neural network, demonstrating DL's ability to discover structurally unique compounds.
- **Clinical Trial Optimization:** Kadur et al. [19] employed NLP models to analyze clinical trial protocols, reducing patient recruitment time by 30%.

- **Drug Repurposing:** Zhou et al. [20] applied DL to molecular and clinical data, identifying baricitinib as a COVID-19 treatment later validated in trials.

DL reduces drug development timelines from decades to years, but validation in biological systems is critical [21].

3 | Ethical Challenges in Medical Deep Learning

3.1 | Algorithmic Bias and Health Disparities

DL models trained on non-representative data propagate biases, exacerbating health inequities.

- **Demographic Bias:** Obermeyer et al. [22] found a widely used healthcare algorithm systematically underestimated risk for Black patients by 28.5% due to biased training data.
- **Geographic Disparities:** Larson et al. [23] showed skin cancer detection DL models performed poorly on darker skin tones due to underrepresentation in datasets.
- **Socioeconomic Factors:** Chen et al. [24] demonstrated that DL models for predicting heart failure readmissions were less accurate for low-income patients (AUC 0.71 vs. 0.89 for high-income).

Bias mitigation requires diverse datasets, fairness-aware algorithms, and continuous monitoring [25].

3.2 | Transparency and Explainability

The "black-box" nature of DL undermines trust and accountability in clinical settings.

- **Clinical Trust:** Topol [2] argues that "without explainability, clinicians cannot responsibly act on AI outputs."
- **Explainability Techniques:** Ribeiro et al. [26] introduced LIME, and Lundberg and Lee [27] developed SHAP to interpret DL predictions. However, these methods often provide post-hoc approximations rather than true causality [28].
- **Regulatory Demands:** The EU AI Act mandates "high-risk" systems (including medical AI) to be "sufficiently transparent" [29].

Explainable AI (XAI) is essential for clinical adoption but remains technically and conceptually immature [30].

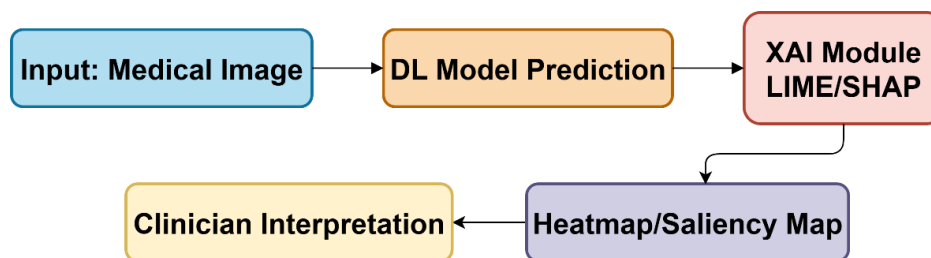


Figure 2. Explainable AI (XAI) Workflow [26], [27].

3.3 | Data Privacy and Security

Medical DL relies on vast sensitive datasets, creating privacy risks.

- **Re-identification Risks:** Fredrikson et al. [31] demonstrated that model outputs could reveal sensitive patient attributes (e.g., genetic markers) via model inversion attacks.
- **Federated Learning:** Sheller et al. [32] showed federated DL (training models locally without sharing raw data) maintained diagnostic accuracy while preserving privacy in multi-institutional studies.
- **Regulatory Compliance:** HIPAA (US) and GDPR (EU) impose strict data protection requirements, complicating data sharing for DL [33].

Privacy-preserving DL techniques (e.g., federated learning, differential privacy) are critical but not foolproof [34].

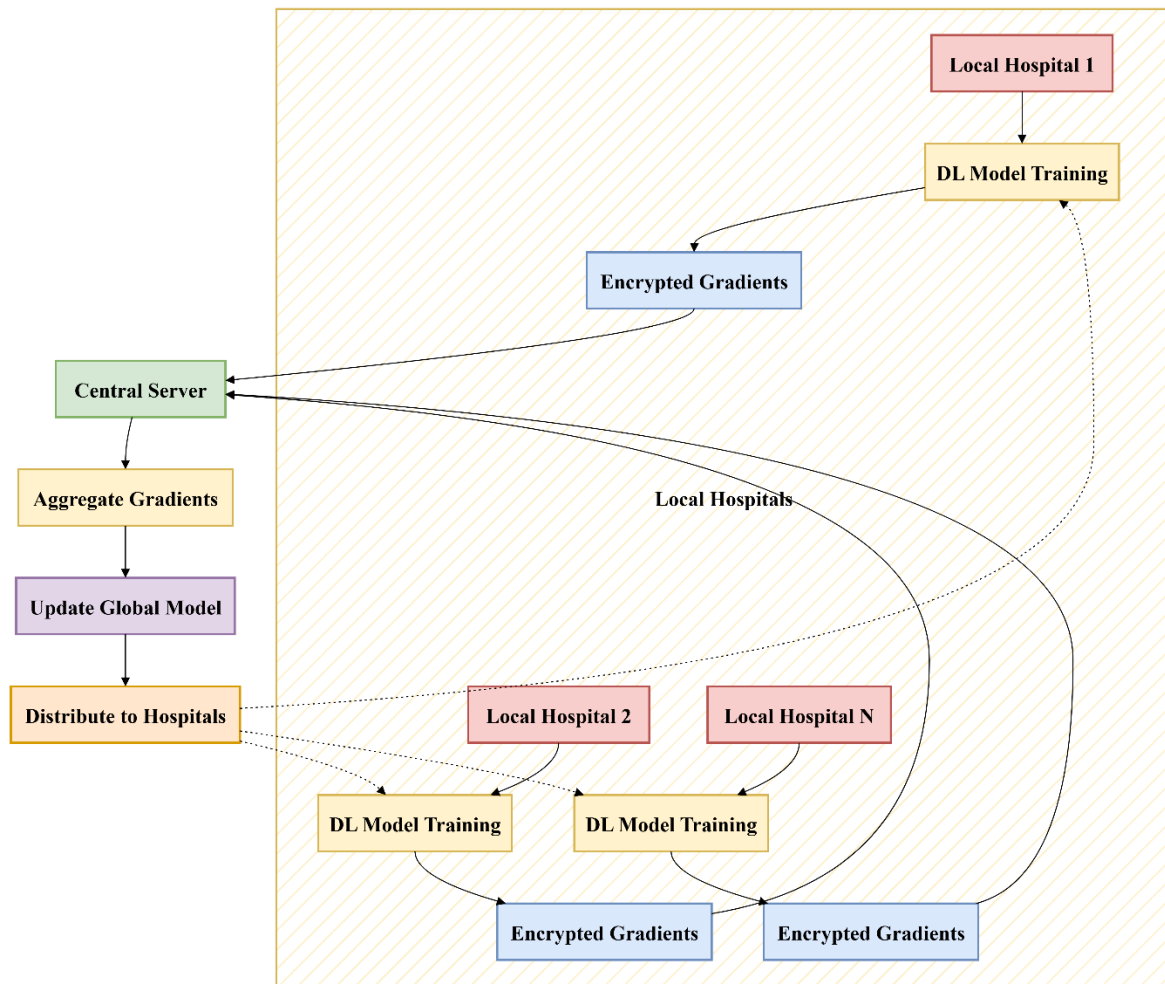


Figure 3: Federated Learning for Medical AI [32].

3.4 | Accountability and Liability

Clear frameworks for assigning responsibility when DL systems fail are lacking.

- **Legal Gaps:** Cohen et al. [35] highlight that malpractice laws are ill-equipped to address AI errors, creating accountability vacuums.
- **Human-AI Collaboration:** Cabitza et al. [13] emphasize that clinicians must remain "in the loop," but unclear decision boundaries can lead to automation bias (e.g., uncritical acceptance of AI outputs).
- **Regulatory Approaches:** The FDA's AI/ML Software as a Medical Device (SaMD) framework requires continuous monitoring but lacks explicit liability guidelines [36].

Synthesis: Accountability requires multi-stakeholder governance involving clinicians, developers, regulators, and patients [37].

Table 2. Ethical Challenges and Mitigation Strategies.

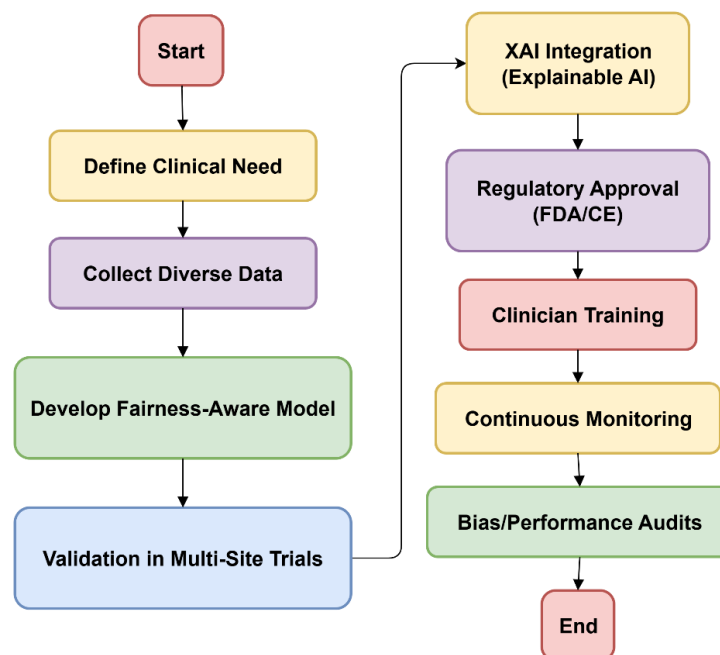
Challenge	Root Cause	Mitigation Strategy	Key Reference
Algorithmic Bias	Non-representative training data	Diverse datasets + fairness-aware algorithms	[22]
Lack of Transparency	"Black-box" model complexity	Explainable AI (LIME, SHAP)	[26], [27]
Privacy Risks	Centralized sensitive data	Federated learning + differential privacy	[32], [34]
Accountability Gaps	Undefined liability frameworks	Clinician-AI collaboration + regulatory oversight	[13], [36]

4 | Discussion: Toward Responsible Deep Learning in Healthcare

The literature reveals a tension: DL offers transformative potential but amplifies ethical risks. Key insights include:

1. **Technical-Ethical Interdependence:** Bias mitigation cannot be an afterthought; it must be embedded in data collection, model design, and deployment [38].
2. **Regulatory Evolution:** Static regulations are inadequate for adaptive DL systems. "Living guidelines" and post-market surveillance are essential [3].
3. **Interdisciplinarity:** Collaboration between AI researchers, clinicians, ethicists, and patients is non-negotiable [39].
4. **Global Equity:** DL must not widen the global health divide. Capacity building in low-resource settings is imperative [40].

Few studies address ethical challenges in real-world clinical workflows [41]. Most research remains in siloed technical or ethical domains.

**Figure 4:** Ethical Deployment Framework for Medical AI[3], [36].

5 | Conclusion and Future Directions

Deep learning is reshaping medical systems, but its ethical integration requires urgent and coordinated action. On the technical side, we recommend prioritizing artificial intelligence (XAI), fairness-sensitive learning, and

privacy-preserving structures. On the clinical side, we recommend developing human-centered AI workflows with clear interfaces between physicians and AI. On the regulatory side, we recommend establishing adaptive and risk-appropriate governance frameworks. On the educational side, we recommend integrating AI ethics into medical and computer science curricula.

As deep learning advances, the medical community must balance innovation with the fundamental ethical imperative: "First, do no harm."

Funding

This research has no funding source.

Data Availability

The datasets generated during and/or analyzed during the current study are not publicly available due to the privacy-preserving nature of the data but are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare that there is no conflict of interest in the research.

Ethical Approval

This article does not contain any studies with human participants or animals performed by any of the authors.

References

- [1] A. Esteva et al., "A guide to deep learning in healthcare," *Nat. Med.*, vol. 25, no. 1, pp. 24–29, Jan. 2019, doi: 10.1038/s41591-018-0316-z.
- [2] E. Topol and G. Winton, *Deep Medicine: How Artificial intelligence can make healthcare Human again*. 2019. Available: https://openlibrary.org/books/OL27754269M/Deep_Medicine
- [3] J. He et al., "The practical guide to medical AI," *Nat. Med.*, vol. 25, no. 9, pp. 1347–1355, Sep. 2019, doi: 10.1038/s41591-018-0542-9.
- [4] V. Gulshan et al., "Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs," *JAMA*, vol. 316, no. 22, pp. 2402–2410, 2016, doi: 10.1001/jama.2016.17216
- [5] S. M. McKinney et al., "International evaluation of an AI system for breast cancer screening," *Nature*, vol. 577, no. 7788, pp. 89–94, Jan. 2020, doi: 10.1038/s41586-019-1799-6.
- [6] G. Campanella et al., "Clinical-grade computational pathology using weakly supervised deep learning on whole slide images," *Nat. Med.*, vol. 25, no. 8, pp. 1301–1309, Aug. 2019, doi: 10.1038/s41591-019-0508-1.
- [7] P. Wang et al., "Automatic brain tumor segmentation using multimodal 3D U-Net," *Med. Image Anal.*, vol. 67, p. 101823, Feb. 2021, doi: 10.1016/j.media.2020.101823.
- [8] G. Litjens et al., "A survey on deep learning in medical image analysis," *Med. Image Anal.*, vol. 42, pp. 60–88, Dec. 2017, doi: 10.1016/j.media.2017.07.005.
- [9] A. Rajkomar et al., "Scalable and accurate deep learning with electronic health records," *NPJ Digit. Med.*, vol. 1, no. 1, p. 18, May 2018, doi: 10.1038/s41746-018-0029-1.
- [10] R. Poplin et al., "Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning," *Nat. Biomed. Eng.*, vol. 2, no. 3, pp. 158–164, Mar. 2018, doi: 10.1038/s41551-018-0195-0.
- [11] S. Nemati et al., "An interpretable machine learning model for accurate prediction of sepsis in the ICU," *Crit. Care Med.*, vol. 46, no. 12, pp. e1265–e1275, Dec. 2018, doi: 10.1097/CCM.0000000000003368.
- [12] K. W. Johnson et al., "Artificial intelligence in cardiology," *J. Am. Coll. Cardiol.*, vol. 71, no. 23, pp. 2668–2679, Jun. 2018, doi: 10.1016/j.jacc.2018.03.521.
- [13] F. Cabitza et al., "Unintended consequences of machine learning in medicine," *JAMA*, vol. 318, no. 6, pp. 517–518, Aug. 2017, doi: 10.1001/jama.2017.07804.
- [14] R. Poplin et al., "A universal SNP and small-indel variant caller using deep neural networks," *Nat. Biotechnol.*, vol. 36, no. 10, pp. 983–987, Oct. 2018, doi: 10.1038/nbt.4235.
- [15] A. Aliper et al., "Deep learning applications for predicting pharmacological properties of drugs and drug repurposing using transcriptomic data," *Mol. Pharm.*, vol. 13, no. 7, pp. 2524–2530, Jul. 2016, doi: 10.1021/acs.molpharmaceut.6b00248.
- [16] J. Jumper et al., "Highly accurate protein structure prediction with AlphaFold," *Nature*, vol. 596, no. 7873, pp. 583–589, Jul. 2021, doi: 10.1038/s41586-021-03819-2.

- [17] J. Zou et al., "A primer on deep learning in genomics," *Nat. Genet.*, vol. 51, no. 1, pp. 12–18, Jan. 2019, doi: 10.1038/s41588-018-0295-5.
- [18] J. M. Stokes et al., "A deep learning approach to antibiotic discovery," *Cell*, vol. 180, no. 4, pp. 688–702, Feb. 2020, doi: 10.1016/j.cell.2020.01.021.
- [19] A. Kadur et al., "Natural language processing for clinical trial eligibility screening: A systematic review," *J. Am. Med. Inform. Assoc.*, vol. 28, no. 7, pp. 1411–1421, Jul. 2021, doi: 10.1093/jamia/ocab019.
- [20] Y. Zhou et al., "Network-based drug repurposing for novel coronavirus 2019-nCoV/SARS-CoV-2," *Cell Discov.*, vol. 6, no. 1, p. 14, Mar. 2020, doi: 10.1038/s41421-020-0153-3.
- [21] J. Vamathevan et al., "Applications of machine learning in drug discovery and development," *Nat. Rev. Drug Discov.*, vol. 18, no. 6, pp. 463–477, Jun. 2019, doi: 10.1038/s41573-019-0024-5.
- [22] Z. Obermeyer et al., "Dissecting racial bias in an algorithm used to manage the health of populations," *Science*, vol. 366, no. 6464, pp. 447–453, Oct. 2019, doi: 10.1126/science.aax2342.
- [23] D. B. Larson et al., "Performance of a deep-learning neural network model in assessing skeletal maturity on pediatric hand radiographs," *Radiology*, vol. 287, no. 1, pp. 313–322, Apr. 2018, doi: 10.1148/radiol.2017170296.
- [24] I. Y. Chen et al., "Algorithmic bias in health care," *Science*, vol. 373, no. 6558, pp. 1019–1021, Sep. 2021, doi: 10.1126/science.abj4481.
- [25] P. Rajpurkar et al., "What is machine learning?," *JAMA*, vol. 328, no. 9, pp. 838–839, Sep. 2021, doi: 10.1001/jama.2021.13066.
- [26] M. T. Ribeiro, S. Singh, and C. Guestrin, "'Why should I trust you?': Explaining the predictions of any classifier," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discov. Data Min.*, 2016, pp. 1135–1144, doi: 10.1145/2939672.2939778.
- [27] S. M. Lundberg and S. I. Lee, "A unified approach to interpreting model predictions," in *Adv. Neural Inf. Process. Syst. 30 (NIPS 2017)*, 2017, pp. 4765–4774, doi: 10.48550/arXiv.1705.07874.
- [28] A. Adadi and M. Berrada, "Peeking inside the black-box: A survey on explainable artificial intelligence (XAI)," *IEEE Access*, vol. 6, pp. 52138–52160, 2018, doi: 10.1109/ACCESS.2018.2876022.
- [29] European Commission, "Proposal for a regulation laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)," COM(2021) 206 final, 2021. [Online]. Available: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0206>
- [30] A. B. Arrieta et al., "Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI," *Inf. Fusion*, vol. 58, pp. 82–115, Jun. 2020, doi: 10.1016/j.inffus.2019.12.012.
- [31] M. Fredrikson et al., "Model inversion attacks that exploit confidence information and basic countermeasures," in *Proc. 22nd ACM SIGSAC Conf. Comput. Commun. Secur.*, 2015, pp. 1322–1333, doi: 10.1145/2810103.2813677.
- [32] M. J. Sheller et al., "Multi-institutional deep learning modeling without sharing patient data: A feasibility study on brain tumor segmentation," in *Med. Image Comput. Comput. Assist. Interv. (MICCAI 2020)*, 2020, pp. 665–674, doi: 10.1007/978-3-030-59728-3_65.
- [33] W. N. Price and I. G. Cohen, "Privacy in the age of medical big data," *Nat. Med.*, vol. 25, no. 1, pp. 37–43, Jan. 2019, doi: 10.1038/s41591-018-0272-6.
- [34] M. Abadi et al., "Deep learning with differential privacy," in *Proc. ACM SIGSAC Conf. Comput. Commun. Secur.*, 2016, pp. 308–318, doi: 10.1145/2976749.2978318.
- [35] I. G. Cohen, P. Amarasingham, and A. Shah, "The regulatory and quality landscape for diagnostic algorithms in medicine," *Sci. Transl. Med.*, vol. 6, no. 233, p. 233cm3, Aug. 2014, doi: 10.1126/scitranslmed.3008722.
- [36] FDA, "Artificial intelligence/machine learning (AI/ML) software as a medical device action plan," U.S. Food and Drug Administration, Silver Spring, MD, USA, Jan. 2021. [Online]. Available: <https://www.fda.gov/media/145022/download>
- [37] B. D. Mittelstadt, "Principles alone cannot guarantee ethical AI," *Nat. Mach. Intell.*, vol. 1, no. 11, pp. 501–507, Nov. 2019, doi: 10.1038/s42256-019-0114-4.
- [38] K. Holstein et al., "Improving fairness in machine learning systems: What do industry practitioners need?," in *Proc. CHI Conf. Hum. Factors Comput. Syst.*, 2019, pp. 1–16, doi: 10.1145/3290605.3300607.
- [39] K. H. Yu, A. L. Beam, and I. S. Kohane, "Artificial intelligence in healthcare," *Nat. Biomed. Eng.*, vol. 2, no. 10, pp. 719–731, Oct. 2018, doi: 10.1038/s41551-018-0302-4.
- [40] E. Beede et al., "A human-centered evaluation of a deep learning system deployed in clinics for the detection of diabetic retinopathy," in *Proc. CHI Conf. Hum. Factors Comput. Syst.*, 2020, pp. 1–12, doi: 10.1145/3313831.3376790.
- [41] Y. Liu et al., "Ethical challenges in deploying AI in clinical practice," *Lancet Digit. Health*, vol. 4, no. 3, pp. e186–e188, Mar. 2022, doi: 10.1016/S2589-7500(21)00251-0.

Disclaimer/Publisher's Note: The perspectives, opinions, and data shared in all publications are the sole responsibility of the individual authors and contributors, and do not necessarily reflect the views of Sciences Force or the editorial team. Sciences Force and the editorial team disclaim any liability for potential harm to individuals or property resulting from the ideas, methods, instructions, or products referenced in the content.