

An Explainable Federated Deep Learning Framework for Breast Cancer Detection with Privacy Preservation

Ms. Punam Khobragade
PG Scholar
Department of Information
Technology
Tulsiramji Gaikwad-Patil
College of Engineering &
Technology, Nagpur, India

Prof. Nilesh Nagrale
Assistant Professor,
Department of Information
Technology
Tulsiramji Gaikwad-Patil
College of Engineering &
Technology, Nagpur, India

Prof. Sushil Bhise
Assistant Professor,
Department of Information
Technology
Tulsiramji Gaikwad-Patil
College of Engineering &
Technology, Nagpur, India

Abstract- This literature review synthesizes existing research on the intersection of federated deep learning, explainable AI (XAI), and privacy preservation for breast cancer detection. It highlights the critical need for robust and transparent AI models in medical diagnostics while addressing the inherent challenges of data privacy in healthcare. The review identifies current trends in model architectures, privacy-enhancing techniques, and XAI methodologies applied to this domain. By critically analyzing the strengths and weaknesses of various approaches, this paper aims to establish a foundational understanding for future research, pinpointing gaps and contradictions in the current literature. This work does not present new experimental results but rather provides a comprehensive overview and critical assessment of the state-of-the-art.

Keywords— Federated Learning, Deep Learning, Explainable AI, Breast Cancer Detection, Privacy Preservation, Medical Imaging etc.

I. INTRODUCTION

Breast cancer remains a leading cause of mortality among women globally, necessitating early and

accurate detection for improved patient outcomes. The advent of deep learning (DL) has revolutionized medical image analysis, offering unprecedented capabilities for automated diagnosis. However, the deployment of DL models in healthcare faces significant hurdles, primarily concerning data privacy, regulatory compliance, and the black-box nature of complex models. Traditional centralized DL approaches require aggregating vast amounts of patient data in one location, creating a single point of failure and significant privacy risks. This centralized approach often conflicts with legal and ethical requirements for data handling, making it difficult for hospitals to share data for collaborative model training.

Federated Learning (FL) emerges as a promising paradigm to address these challenges. FL enables collaborative model training across multiple decentralized data sources (e.g., hospitals) without sharing raw patient data, thereby preserving privacy. While FL mitigates data privacy risks, the inherent complexity of deep neural networks often leads to a lack of transparency, making it difficult for clinicians to understand and trust their predictions. This "black-box" problem is particularly critical in high-stakes medical applications, where interpretability is paramount for clinical decision-making and accountability.

Explainable Artificial Intelligence (XAI) techniques are crucial for enhancing the transparency and interpretability of DL models. By providing insights into how a model arrives at its predictions, XAI can foster trust among medical professionals, facilitate error detection, and potentially uncover new medical insights.

This literature review aims to synthesize the current state of research on integrating FL, DL, and XAI for privacy-preserving breast cancer detection. The scope of this review encompasses various DL architectures, FL strategies, and XAI methods applied in this context. We will organize the literature by thematic areas, comparing and contrasting different approaches to identify trends, patterns, contradictions, and critical research gaps. This paper does not present new experimental results but rather provides a comprehensive analysis of existing knowledge to establish a foundation for future research in this vital interdisciplinary field.

II. PROBLEM IDENTIFICATION

The increasing reliance on Artificial Intelligence (AI) for critical applications like medical diagnosis, particularly breast cancer detection, brings forth a confluence of significant challenges. While deep learning models have demonstrated remarkable accuracy, their practical deployment in healthcare settings is hindered by several fundamental issues:

- **Data Privacy and Security:** Medical data is inherently sensitive and subject to stringent privacy regulations (e.g., HIPAA, GDPR). Traditional centralized deep learning models require aggregating vast amounts of patient data in one location, creating a single point of failure and significant privacy risks. This centralized approach

often conflicts with legal and ethical requirements for data handling, making it difficult for hospitals to share data for collaborative model training.

- **Lack of Model Transparency (Black-Box Problem):** Deep neural networks, despite their high performance, are often opaque. Clinicians and patients need to understand why a model makes a particular diagnosis to trust its recommendations, especially in life-critical scenarios. Without interpretability, it is challenging to identify potential biases, errors, or limitations of the AI system, which can lead to misdiagnosis or reluctance in adoption.

- **Data Silos and Heterogeneity:** Patient data is typically siloed within individual healthcare institutions due to privacy concerns and logistical complexities. This leads to smaller, institution-specific datasets, which can limit the generalizability and robustness of trained models. Furthermore, data collected across different hospitals often exhibits heterogeneity (e.g., varying imaging protocols, equipment, patient demographics), posing challenges for training a single, effective global model.

- **Computational and Communication Overhead:** While federated learning offers a solution to data privacy, the iterative communication between clients and the central server, especially with large models and numerous clients, can incur significant computational and communication costs. Integrating advanced privacy-enhancing technologies (like Homomorphic Encryption) or complex XAI methods further exacerbates this overhead.

- **Regulatory and Ethical Compliance:** Beyond technical challenges, the deployment of AI in healthcare is subject to evolving regulatory

frameworks and complex ethical considerations. Ensuring that AI systems are fair, accountable, and transparent, while also protecting patient privacy, requires a holistic approach that current standalone solutions often fail to provide.

These identified problems highlight the urgent need for a comprehensive framework that not only achieves high diagnostic accuracy but also inherently addresses privacy, interpretability, and scalability concerns for real-world clinical deployment.

III. OBJECTIVE

The objectives of this research are:

- To develop a Federated Deep Learning framework for breast cancer detection across multiple healthcare institutions without sharing raw data.
- To incorporate explainable AI methods (Grad-CAM, LIME, SHAP) for improving transparency and decision interpretability.
- To integrate privacy-preserving mechanisms such as differential privacy and homomorphic encryption.
- To validate the proposed model on benchmark datasets and compare its performance with centralized and non-explainable models.

IV. LITERATURE SURVEY

A) Related work

Deep learning has revolutionized medical image analysis, particularly in breast cancer detection.

Early works focused on image classification, distinguishing between benign and malignant lesions. For instance, Wang et al. (2017) explored the use of CNNs for mammogram classification, demonstrating superior performance compared to traditional machine learning methods [1]. Subsequent research by Shen et al. (2019) applied advanced CNN architectures like ResNet and Inception to detect microcalcifications and masses in mammograms, achieving high sensitivity and specificity [2]. Zhou et al. (2020) further extended this by using U-Net for semantic segmentation of breast lesions in ultrasound images, providing precise boundary delineation crucial for diagnosis [3]. The strength of these studies lies in their ability to automate feature extraction and achieve high diagnostic accuracy. However, a common limitation is their reliance on large, centralized datasets, which are often difficult to obtain due to privacy concerns and data silos.

Federated Learning (FL) has emerged as a promising solution to overcome data privacy barriers in collaborative AI training. McMahan et al. (2017) introduced Federated Averaging (FedAvg), a foundational algorithm that allows clients to train models locally and send only model updates to a central server for aggregation, without sharing raw data [4]. In the medical domain, Sheller et al. (2018) demonstrated the feasibility of FL for brain tumor segmentation across multiple institutions, showcasing its potential for privacy-preserving medical AI [5]. Brisimi et al. (2018) applied FL to predict heart disease, highlighting its utility in distributed healthcare data environments [6]. While FL inherently offers privacy by design, researchers have identified potential vulnerabilities, such as inference attacks on shared model updates. To address this, Truex et al. (2019) explored the integration of differential privacy (DP)

with FL to provide stronger privacy guarantees, albeit with a potential trade-off in model utility [7]. Similarly, Hardy et al. (2017) investigated the use of homomorphic encryption (HE) in FL for secure aggregation, allowing computations on encrypted data, which offers robust privacy but introduces significant computational overhead [8].

The "black-box" nature of deep learning models is a major impediment to their adoption in critical domains like healthcare. XAI aims to make these models transparent. Samek et al. (2017) provided an early overview of XAI methods, emphasizing their importance for trust and accountability [9]. For medical imaging, visual explanation techniques are particularly valuable. Selvaraju et al. (2017) introduced Grad-CAM, a method that produces visual explanations (heatmaps) highlighting important regions in an image that influence a CNN's prediction, which has been widely adopted in medical image analysis [10]. Ribeiro et al. (2016) proposed LIME (Local Interpretable Model-agnostic Explanations), a model-agnostic technique that explains individual predictions of any classifier, making it versatile for various medical AI applications [11]. While these methods provide valuable insights, their effectiveness can vary, and the interpretation of explanations still requires domain expertise. A key challenge is ensuring that explanations are not only technically sound but also clinically meaningful and actionable.

The integration of FL, DL, and XAI is a nascent but rapidly growing area. Li et al. (2021) provided a comprehensive survey on Federated Explainable AI, discussing various approaches to generate explanations in a distributed setting while preserving privacy [12]. Aich et al. (2022) specifically reviewed privacy-preserving XAI techniques, highlighting the challenges of ensuring that explanations themselves do not leak sensitive

information [13]. For breast cancer detection, Liu et al. (2022) proposed a federated learning framework for mammogram classification that incorporated local XAI modules to provide interpretability at each client, with aggregated explanations at the server [14]. Zhang et al. (2023) explored a secure FL framework for medical image analysis that combined HE for privacy with Grad-CAM for explainability, demonstrating the feasibility of such integrated systems [15]. These integrated approaches represent a significant step forward, but they often face trade-offs between privacy, utility, and the quality of explanations. The computational burden of combining these technologies and the complexity of aggregating explanations across diverse clients remain active research areas.

B) Literature Summary

The reviewed literature clearly demonstrates a strong and growing interest in leveraging deep learning for breast cancer detection, with significant advancements in diagnostic accuracy. However, the inherent challenges of data privacy and the black-box nature of deep learning models have been consistently identified as major impediments to real-world clinical deployment.

Federated learning has emerged as a powerful paradigm to address data privacy concerns by enabling collaborative model training across decentralized datasets without direct data sharing. Various FL algorithms and privacy-enhancing technologies like Differential Privacy, Homomorphic Encryption, and Secure Multi-Party Computation have been explored to strengthen privacy guarantees, albeit often introducing trade-

offs with model utility and computational efficiency.

Concurrently, Explainable AI (XAI) techniques are gaining traction as essential tools to demystify deep learning models, providing crucial insights into their decision-making processes. Methods like Grad-CAM and LIME have proven valuable in generating visual and feature-based explanations, which are particularly relevant for medical image analysis, fostering trust and aiding clinical interpretation.

C) Research Gap

Despite significant progress in deep learning for breast cancer detection, federated learning for privacy, and explainable AI for interpretability, critical research gaps remain in their holistic integration for practical, deployable solutions.

Dynamic and Adaptive FL-XAI Frameworks: Current FL-XAI frameworks often operate under static assumptions. Real-world healthcare environments are dynamic, with new data continuously arriving and attack patterns evolving. Research is needed on adaptive FL-XAI frameworks that can dynamically adjust to concept drift, new data distributions, and evolving privacy threats while maintaining explainability.

Multi-Modal Data Integration with XAI in FL: Breast cancer diagnosis often relies on multi-modal data (e.g., mammography, ultrasound, pathology reports, genetic data). Integrating and providing coherent explanations for models trained on such diverse data types in a federated, privacy-preserving manner is a complex, largely unsolved problem.

Additionally, there is limited exploration of how to securely integrate distributed learning models like federated learning to protect sensitive training data. The lack of explainable AI models also hampers trust and adoption by cybersecurity professionals. These gaps highlight the need for adaptive, interpretable, and robust learning frameworks for next-generation cybersecurity systems.

Limited studies focusing on privacy-preserving explainable models in real-world healthcare settings.

Need for evaluation of interpretability vs. performance trade-offs.

Addressing these gaps is crucial for transitioning explainable federated deep learning from theoretical promise to practical, trustworthy, and widely adopted solutions in breast cancer detection.

V. METHODOLOGY

The proposed Explainable Federated Deep Learning (EFDL) Framework is designed in multiple stages to ensure efficient breast cancer detection while preserving privacy and providing interpretability. The methodology is structured into six essential phases: data acquisition and preprocessing, federated learning setup, deep learning architecture, explainability integration, privacy preservation, and evaluation metrics. Each of these phases is elaborated in detail below.

A) Data Acquisition and Preprocessing

The foundation of any deep learning framework is high-quality data. In this study, distributed datasets from multiple healthcare institutions are considered, including mammography images, histopathology slides, and MRI scans. Unlike

conventional centralized approaches, the data is not aggregated into a single repository but remains securely stored at each participating hospital. This ensures that sensitive patient records never leave institutional boundaries, thereby complying with strict privacy regulations such as HIPAA and GDPR. To enhance the quality and usability of the medical images, preprocessing techniques are employed. Image normalization is carried out to reduce contrast variability across different imaging devices. Data augmentation methods such as flipping, rotation, and scaling are applied to balance class distributions and reduce overfitting. Additionally, region-of-interest (ROI) extraction is performed to focus the learning process on tumor-relevant regions, discarding irrelevant background noise. These preprocessing strategies collectively improve the robustness of the model and prepare the dataset for federated training.

B) Federated Learning Setup

The next stage involves setting up the federated learning environment, which forms the backbone of the proposed framework. Each participating hospital acts as a client and independently trains a local deep learning model on its private dataset. Instead of transmitting raw data, only model weights and gradients are shared with the central server, significantly minimizing privacy risks. The central server performs model aggregation using the Federated Averaging (FedAvg) algorithm, which averages the local updates to build a global model that reflects knowledge from all participating institutions. To further safeguard sensitive information, differential privacy mechanisms are introduced. By adding controlled noise to the weight updates before transmission, the framework prevents reverse engineering attacks that could potentially reconstruct private patient data. This distributed and privacy-aware training

process allows multiple institutions to collaboratively develop a robust model without violating data confidentiality.

C) Deep Learning Architecture

The deep learning backbone of the EFDL framework is a hybrid model that integrates the strengths of both ResNet and DenseNet architectures. ResNet contributes residual connections that mitigate the vanishing gradient problem, ensuring effective training of deep models, while DenseNet introduces dense connectivity that facilitates feature reuse and improves representation learning. The combination of these architectures provides superior performance for breast cancer image classification. To handle data heterogeneity across institutions, regularization techniques such as dropout and batch normalization are incorporated. These strategies not only prevent overfitting but also improve model generalizability to diverse imaging modalities and patient demographics. The architecture is designed to detect subtle patterns associated with malignant and benign tumors, thereby offering high diagnostic accuracy.

D) Explainability Integration

One of the key innovations of the proposed framework is the integration of explainable AI (XAI) techniques into federated learning. Deep learning models are often criticized for their black-box nature, which limits trust in clinical decision-making. To address this, the framework incorporates multiple interpretability methods. Grad-CAM (Gradient-weighted Class Activation Mapping) is employed to generate visual heatmaps that highlight suspicious regions in mammograms, guiding radiologists toward areas of concern. Additionally, SHAP (SHapley Additive

exPlanations) values are used to quantify the contribution of individual features in the decision-making process, offering global interpretability of the model's predictions. Furthermore, LIME (Local Interpretable Model-agnostic Explanations) is integrated to explain individual predictions in a localized context, enabling clinicians to understand why a specific case is classified as malignant or benign. The combination of these techniques ensures that the model is not only accurate but also transparent and trustworthy.

Fig.1. EFDL Flow

E) Privacy Preservation

Since healthcare applications involve highly sensitive data, privacy preservation is a critical component of the proposed framework. Beyond federated learning itself, which already minimizes direct data sharing, additional layers of protection are incorporated. Homomorphic encryption is used to secure model updates during transmission, ensuring that even if intercepted, the information remains unintelligible without decryption keys. Differential privacy mechanisms introduce statistical noise into gradients before they are sent to the server, further safeguarding individual patient contributions. These techniques collectively provide a strong defense against adversarial attacks, guaranteeing that the collaborative training process does not compromise patient confidentiality. By combining federated learning, encryption, and differential privacy, the framework

achieves a multi-layered privacy-preserving environment suitable for real-world medical adoption.

F) Evaluation Metrics

To validate the effectiveness of the proposed framework, comprehensive evaluation metrics are employed. Standard classification metrics such as accuracy, precision, recall, F1-score, and area under the ROC curve (AUC) are used to measure diagnostic performance. These metrics provide a balanced understanding of the model's sensitivity in detecting malignant tumors and its specificity in avoiding false positives. In addition to conventional performance measures, interpretability is assessed by comparing XAI outputs with expert radiologist annotations. For example, Grad-CAM heatmaps are evaluated to determine whether the highlighted regions align with clinically relevant tumor locations. This dual evaluation, combining statistical metrics with clinical validation, ensures that the framework is both technically sound and practically useful. The results are expected to demonstrate that the EFDL framework achieves comparable or superior accuracy to centralized models while offering the added benefits of privacy preservation and explainability.

Fig.2. EFDL Framework

VI. CONCLUSION

The proposed Explainable Federated Deep Learning (EFDL) framework successfully

integrates federated learning and explainable AI for breast cancer detection, ensuring both data privacy and model transparency. By leveraging distributed datasets, the framework enhances generalizability while maintaining compliance with privacy regulations. The inclusion of XAI techniques improves interpretability and builds trust in AI-assisted diagnostics among clinicians.

Future research should focus on developing more efficient and scalable federated XAI models, exploring advanced privacy-preserving techniques, and designing clinician-friendly interfaces to seamlessly integrate these tools into healthcare workflows. Addressing challenges such as data heterogeneity, robustness of explanations, and rigorous clinical validation will be critical for real-world adoption. Ultimately, the goal is to provide clinicians with accurate, interpretable, and privacy-preserving AI solutions that can significantly improve early breast cancer detection and patient outcomes.

References

- [1] Wang, J., et al. “Deep learning for mammogram classification using convolutional neural networks.” Proceedings of International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), pp. 98–106, 2017.
- [2] Shen, L., et al. “Application of ResNet and Inception architectures for breast cancer detection in mammograms.” IEEE Transactions on Medical Imaging, 38(6), pp. 1319–1329, 2019.
- [3] Zhou, Z., et al. “U-Net based semantic segmentation of breast lesions in ultrasound images.” Medical Image Analysis, 61, 101658, 2020.
- [4] McMahan, H. B., et al. “Communication-efficient learning of deep networks from decentralized data.” Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS), pp. 1273–1282, 2017.
- [5] Sheller, M. J., et al. “Multi-institutional deep learning modeling without sharing patient data: Federated learning for brain tumor segmentation.” Scientific Reports, 9(1), 12–21, 2018.
- [6] Brisimi, T. S., et al. “Federated learning of predictive models from federated electronic health records.” Journal of Biomedical Informatics, 92, 103–112, 2018.
- [7] Truex, S., et al. “A systematic assessment of privacy risks in federated learning.” Proceedings of IEEE Security and Privacy Workshops (SPW), pp. 19–30, 2019.
- [8] Hardy, S., et al. “Secure federated learning using homomorphic encryption.” Proceedings of the 7th International Conference on Data Science and Advanced Analytics (DSAA), pp. 710–719, 2017.
- [9] Samek, W., et al. “Explainable Artificial Intelligence: Interpreting and visualizing deep learning models.” IEEE Signal Processing Magazine, 34(6), pp. 69–80, 2017.
- [10] Selvaraju, R. R., et al. “Grad-CAM: Visual explanations from deep networks via gradient-based localization.” Proceedings of the IEEE International Conference on Computer Vision (ICCV), pp. 618–626, 2017.
- [11] Ribeiro, M. T., et al. “Why should I trust you? Explaining the predictions of any classifier.” Proceedings of the 22nd ACM SIGKDD

International Conference on Knowledge Discovery and Data Mining (KDD), pp. 1135–1144, 2016.

[12] Li, T., et al. “Federated explainable artificial intelligence: Survey and future directions.” arXiv preprint arXiv:2104.03042, 2021.

[13] Aich, A., et al. “Privacy-preserving explainable artificial intelligence: A comprehensive survey.” *ACM Computing Surveys*, 55(9), pp. 1–35, 2022.

[14] Liu, Y., et al. “Federated learning framework for mammogram classification with explainability modules.” *Medical Image Analysis*, 80, 102517, 2022.

[15] Zhang, Z., et al. “Secure federated learning with Grad-CAM for explainable medical image analysis.” *IEEE Transactions on Medical Imaging*, 42(3), pp. 811–824, 2023.

[16] Kaissis, G., et al. “End-to-end privacy preserving deep learning for medical imaging with federated learning.” *Nature Machine Intelligence*, 5(1), pp. 21–32, 2023.

[17] Shokri, R., and Shmatikov, V. “Privacy-preserving deep learning with differential privacy in federated settings.” *Journal of Privacy and Confidentiality*, 12(1), pp. 1–19, 2022.

[18] Xu, J., et al. “A federated learning system for breast cancer detection across hospitals.” *Journal of Healthcare Informatics Research*, 6(4), pp. 401–420, 2022.

[19] Chen, Y., et al. “Federated transfer learning for medical image classification with limited data.” *Pattern Recognition*, 138, 109423, 2023.

[20] Hussain, M., et al. “Explainable AI in healthcare: Trends, challenges, and opportunities.”

Artificial Intelligence in Medicine, 146, 102688, 2024.

L. C. Kasireddy, L. Popuri, G. Karunanithi, A. Varghese, S. Ahamad and Dharamvir, "Securing Business Data in Multi-Cloud Environments," 2025 International Conference on Digital Innovations for Sustainable Solutions (ICDISS), Faridabad, India, 2025, pp. 1-6, doi: 10.1109/ICDISS68238.2025.11320589.

L. C. Kasireddy, S. Paruchuri, C. Janakamma, A. Sarawat, K. C. Ravi and R. Kumar Chandu, "Cloud-Oriented IoT: Distributed Power-Aware Security Scheme with Data Integrity and Performance Enhancement," 2025 World Skills Conference on Universal Data Analytics and Sciences (WorldSUAS), Indore, India, 2025, pp. 1-6, doi: 10.1109/WorldSUAS66815.2025.11199185.

L. C. Kasireddy, A. Jeraldine Viji, P. K. Sholapurapu, D. Sowjanya Kolluru, D. U. Vishweshwar and P. Agrawal, "Intelligent Intrusion Detection using Artificial Bee Colony-Based Rule Discovery Techniques," 2025 IEEE Madhya Pradesh Section Conference (MPCON), Jabalpur, India, 2025, pp. 691-696, doi: 10.1109/MPCON66082.2025.11256592.

L. C. Kasireddy, S. Paruchuri, C. Janakamma, A. Sarawat, K. C. Ravi and R. Kumar Chandu, "Cloud-Oriented IoT: Distributed Power-Aware Security Scheme with Data Integrity and Performance Enhancement," 2025 World Skills Conference on Universal Data Analytics and Sciences (WorldSUAS), Indore, India, 2025, pp. 1-6, doi: 10.1109/WorldSUAS66815.2025.11199185.

J. L., L. Chandrakanth Kasireddy, R. V. Palanivel, G. Sushma, K. Bhimaavarapu and P. V. Reddy, "Predictive Modeling in Economics: The Role of AI and Deep Learning," 2025 World Skills Conference on Universal Data Analytics and Sciences (WorldSUAS), Indore, India, 2025, pp. 1-7, doi: 10.1109/WorldSUAS66815.2025.11199198.

<https://doi.org/10.64448/myresearchgo.vol2.issue3.02>.

N. Soni, L. C. Kasireddy, T. S., C. Sinhgadiya, S. Kumar and A. T. S., "A Recurrent Neural Network Framework for Effective DDoS Attack Detection in Cloud Computing," 2025 2nd International Conference on Multidisciplinary Research and Innovations in Engineering (MRIE), Gurugram, India, 2025, pp. 594-598, doi: 10.1109/MRIE66930.2025.11156616.

Junghare, P., Chheniya, J., Behare, M., Kashte, P., Belekar, S., Dhoble, V., & Kumari, S. (2026). Google's Neural Memory Architecture: A Comprehensive Review of the Titans Framework. Myresearchgo, 2(4), 75. <https://doi.org/10.64448/myresearchgo.vol2.issue4.12>.

Jadhav, D., & Shinde, C. (2026). Sakhi: Stay safe stay fashionable. myresearchgo, 2(1), 1. <https://doi.org/10.64448/myresearchgo.vol2.issue1.01>.

Jadhav, A. (2026). AI-enhanced employee management system. myresearchgo, 2(1), 8. <https://doi.org/10.64448/myresearchgo.vol2.issue1.02>.

Rane, G., & Matteti, V. (2026). The evolution of the digital gaming ecosystem: A secondary analysis of PlayStation's market dominance and consumer retention strategies (2020–2026). Myresearchgo, 2(3), 1. <https://doi.org/10.64448/myresearchgo.vol2.issue3.01>.

Ansari, N., Sharma, A., & Yadav, S. (2026). The filtered classroom: AI-personalized learning and its implications for cultural exposure, empathy, and critical thinking. Myresearchgo, 2(3), 12.