International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

A Review of Privacy-Preserving Deep Learning for Automated Diabetic Retinopathy Detection Using Hybrid CNN–Transformer Models

[1]Gopi Vivek B
Dept of CSE, Sri Venkatesa Perumal College of Engineering & Technology, Puttur, Andhrapradesh, India, gopiviveksvpp@gmail.com
[2]D Nagaraju
Professor, Dept of CSE, Sri Venkatesa Perumal College of Engineering & Technology, Puttur, Andhrapradesh, India, raj2dasari@gmail.com

Abstract- Diabetic Retinopathy (DR) is a serious retinal disease that occurs due to diabetes and is one of the major causes of vision loss globally. While recent advances in computer vision and deep learning have significantly enhanced automated retinal image analysis for early DR detection, centralized learning methods introduce important considerations of patient privacy, regulatory standards, data leakage and cyber-security risks. In this review, we explore the privacy-preserving deep learning frameworks like Federated Learning (FL), Differential Privacy (DP), Homomorphic Encryption (HE), and Secure Multi-Party Computation (SMPC) for automated diabetic retinopathy (DR) detection using CNN and hybrid CNN-Transformer architectures. This study investigates the present status of decentralized medical imaging systems for collaborative model learning without sharing sensitive patient data and critically analyzes the evolution of the systems. Further, the research delves into complex hybrid models, which integrate Convolutional Neural Networks (CNNs) with Vision Transformers (ViTs), to achieve both local and global context understanding for retinal disease classification. A comparative study of recent state-of-the-art models shows that privacy-preserving federated models achieve over 95% diagnostic accuracy, enjoy formal privacy guarantees, and are resistant to gradient inversion and membership inference attacks and poisoning attacks. Furthermore, several different aspects of communication efficiency, heterogeneity-aware federated optimization, embedding explainable artificial intelligence (XAI) for clinical deployment, and problems faced in multi-institutional healthcare environments are discussed. Finally, the future research directions of large-scale federated learning systems, trustworthy explainable AI, fairness-aware learning, secure medical security using blockchain, and lightweight edge-deployable Transformer models are pointed out for future intelligent ophthalmic diagnostic systems.

Keywords: Diabetic Retinopathy Detection, Federated Learning, Privacy-Preserving Deep Learning, Hybrid CNN–Transformer Networks, Differential Privacy, Homomorphic Encryption, Vision Transformers in Healthcare, Explainable Medical AI

1.Introduction And Overview

Diabetic retinopathy (DR) is one of the most important preventable causes of blindness in the world, especially affecting people of working age, who have diabetes mellitus [1]. Early recognition and proper classification of DR severity is crucial to timely intervention and vision preservation. Automated DR detection using deep learning, especially CNNs, has achieved impressive diagnostic accuracy in fundus image analysis, but there are significant challenges in implementing these methods in clinical practice, including patient privacy, compliance with regulations, and data security. The extensive literature review examines the implementation of privacy-preserving techniques, federated learning solutions, and cutting-edge hybrid CNN-Transformer architectures for creating AI-powered, technically viable, and privacy-aware DR detection systems. As privacy laws and regulations such as the General Data Protection Regulation (GDPR) and Health Insurance Portability and Accountability Act (HIPAA) become prevalent, the need for decentralized and privacy-preserving machine learning in the healthcare sector has never been greater. Traditional, centralised, deep learning techniques involve compiling of patient medical imaging data at central data centres, thereby creating a risk of the data being leaked, re-used and breached by other

parties, and also vulnerable to regulatory issues[2]. Federated learning, which allows for the joint training of a model across different healthcare institutions without sharing patient data, has proved to be a paradigm shift that can pave the way to privacy-preserving AI systems in medical imaging [2].

2. Federated Learning Frameworks For Privacy-Preserving Dr Detection

2.1. Fundamental Principles of Federated Learning in Healthcare

Federated learning (FL), one of the collaborative machine learning paradigm, can train a global model by a group of multiple decentralized clients (healthcare institutions) without sharing any patient information, rather just model parameters or gradients [3]. It can directly address privacy concerns of centralized data aggregation, and it can enable the creation of more robust and more generalizable models with the use of multiple datasets across multiple institutional boundaries. The basic federated learning systems are server-client based when local models are trained locally on institution-specific data and a central server aggregates the model updates which were trained locally using algorithms like Federated Averaging (FedAvg) [4].

The study has showed that Federated Learning can achieve the performance of diagnosis that is comparable to or better than the centralized baseline approaches and still satisfy the strict data localization and privacy protection. For example, a federated learning approach to DR detection obtained a 94.66% accuracy for five clients across the IDRiD and STARE datasets, demonstrating that the distributed training imparts a high level of privacy on the diagnostic accuracy without sacrificing it [5].

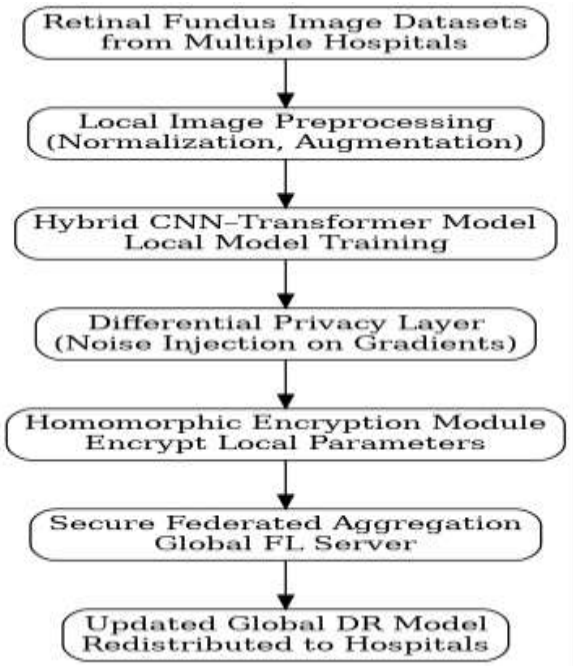


Fig. 1. Privacy-Preserving Federated Learning Framework for DR Detection

As illustrated in Figure 1, the general architecture of the privacy-preserving federated learning system for diabetic retinopathy detection consists of multiple hospitals that train the hybrid CNN–Transformer model on their local data, securely send encrypted model parameters to the centralized aggregation server, and the aggregation server aggregates the received model

parameters. The framework combines the use of differential privacy and homomorphic encryption to ensure patient privacy while enabling collaborative learning.

2.2 Non-IID Data Distribution and Federated Optimization

One of the main problems with federated learning in medical imaging is the presence of Non-Independent and Identically Distributed (non-IID) data in each healthcare institution being shared. This heterogeneity is caused by the differences in the institutions with regard to the patient population, imaging devices, acquisition protocols and incidence of the disease [6]. Traditional federated averaging algorithms may suffer from poor performance and convergence issues under non-IID settings. To overcome these challenges, enhanced federated optimization techniques have been developed. The Federated Proximal (FedProx) algorithm adds proximal regularization to make the convergence more robust and, when paired with momentum driven methods like Adam optimization and Shuffling SGD, performs better in detecting DR. Experimental results show that FedProx with shuffling SGD can reach 92.63% accuracy on the diabetic retinopathy datasets with non-IID data distribution, significantly outperforming the standard FedAvg implementations in [6].

2.3 Multi-Institutional Federated Learning Deployments

Effective federated learning systems have been used in various healthcare institutions for DR detection. In a study of five clients with different fundus image data sets, the prediction accuracy of DR and DME detection with data privacy was 96% on average [7]. The framework used MobileNetV2 as the backbone architecture, which showed that lightweight models can be used to obtain a high diagnostic accuracy in the federated setting without the need for centralized data aggregation..

Algorithm : Federated Learning-Based Privacy-Preserving DR Detection

Input:

Distributed retinal datasets Dk

Global model G

Output:

Optimized global DR model

1: Initialize global model parameters
2: for each communication round do
3: Send global model to hospitals
4: Train local CNN–Transformer model
5: Apply differential privacy noise
6: Encrypt gradients using HE
7: Upload encrypted parameters
8: Aggregate model updates
9: Update global model
10: end for

11: Return optimized DR detection model

This proposed algorithm provides a privacy-preserving federated learning framework for diabetic retinopathy detection in which multiple healthcare institutions can work together to train a world-wide CNN–Transformer model without compromising patient privacy. The algorithm is made to mix differential privacy with homomorphic encryption to allow for secure model aggregation and protection against privacy attacks.

3. Privacy-Preserving Mechanisms And Cryptographic Techniques

3.1. Differential Privacy for Gradient Protection

One approach is to add mathematically provable privacy guarantees to gradients in the federated learning process via differential privacy (DP) [8]. The differentially private stochastic gradient descent (DP-SGD) algorithm can be used in deep neural networks and formally bounds on privacy can be provided as privacy budget (ε and δ parameters). Differential privacy has been successfully incorporated into federated learning systems for detection applications for diabetic retinopathy to guard against gradient inversion attacks, which are attacks that reconstruct raw patient images from intercepted gradients. The advanced differential privacy techniques used for medical imaging are capable of high accuracy (~90% for moderate privacy budgets, $\varepsilon \sim 10$), and have been shown to be particularly effective for imaging modalities with naturally diverse features, such as those used for medical imaging [9]. Homomorphic encryption and differential privacy have been used together to achieve layered privacy protection for the detection of DR, with a classification accuracy of 96.98%, and the defense against multiple attacks such as model inversion and membership inference attacks [10].

3.2 Using Homomorphic Encryption for Secure Aggregation

Homomorphic encryption (HE) is the most powerful encryption method to guarantee data confidentiality in federated aggregation without a need for decryption while performing cryptographic operations [11]. The Cheon-Kim-Kim-Song (CKKS) homomorphic encryption scheme, which can be used to perform approximate arithmetic on encrypted real numbers, has become the choice of scheme for medical imaging applications. To protect the aggregation of model parameters, several federated learning (FL) frameworks for detecting retinal diseases have adopted homomorphic encryption. One multi-key fully homomorphic encryption (FHE) scheme for privacy preserving federated learning (FL) that reached 96.30% accuracy for diabetic retinopathy detection (DRD) protected against collusion attacks and ensured post-quantum security [12]. The multi-key integration enables every health care provider to have a different key, which means there is no single point of failure as in single-key schemes.

3.3 Secure Multi-Party Computation and Zero-Knowledge Proofs.

Secure multi-party computation (SMPC) and zero-knowledge proofs are complementary privacy guarantees to federated learning and differential privacy, respectively. These techniques allow health care institutions to check whether a federated update is valid, or whether a bad participant is injecting updates into the system without revealing the content of the update [13]. A federation system with privacy-preserving methods such as federated learning, secure multi-party computation, SMOTE for class balancing, blockchain for auditability, and IPFS for efficient data storage yielded a diabetic retinopathy detection accuracy of nearly 90% without the exchange of any raw data [14]. This multi-layered defence can protect against gradient inversion, model poisoning and membership inference attacks and does not compromise diagnostic performance.

Table 1: Comparative Analysis Of Privacy-Preserving Techniques In Federated Medical Imaging

Technique	Privacy Level	Cost	Security Strength	Accuracy Impact
Differential Privacy	High	Medium	Strong	Moderate
Homomorphic Encryption	Very High	High	Very Strong	Low
Secure Multi-Party Computation	High	High	Strong	Moderate
Federated Learning	Medium	Medium	Moderate	Very Low

Table 1 provides a comparative analysis of major privacy-preserving techniques used in federated medical imaging systems, highlighting their privacy strength, computational complexity, security level, and impact on classification accuracy.

4. Hybrid CNN-Transformer Architectures For Retinal Image Analysis

4.1 Vision Transformer Fundamentals and Advantages Over Pure CNNsThe Vision Transformers (ViTs) are a paradigm shift from the convolutional architectures, instead relying on the self-attention mechanism to model long-range spatial dependencies globally instead of only with convolutional operations with local receptive fields [15]. The self-attention mechanism handles the computation of relationships among all image patches at the same time, which can help the model to capture contextual information relevant to retinal pathology detection at various scales and spatial distances. The Vision Transformer has been used in detection of DR with 93% accuracy and 0.89 ROC curve, which is comparable with existing CNN based methods and presents an advantage of explainability by using attention visualization for DR detection [16]. Vision Transformers have proven to be especially useful in detecting subtle lesion features that are typical of mild diabetic retinopathy, such as microaneurysms and retinal hemorrhages, which are visually low contrast and thus need an understanding of the global context.

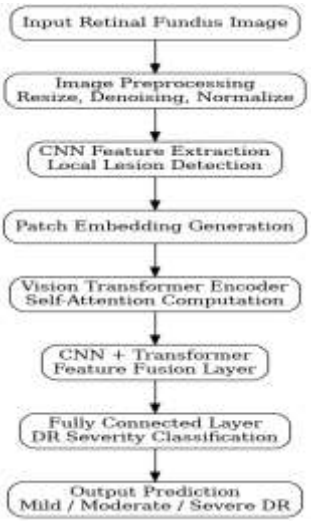


Fig. 2. Hybrid CNN–Transformer Architectures

Figure 2 shows the hybrid CNN-Transformer architecture for retinal image classification, which starts with retinal image preprocessing, followed by feature extraction using CNN and self-attention learning using Vision Transformer. Finally, the fused feature is utilized for predicting the severity of DR by a classification layer.

4.2. Hybrid CNN-Transformer Fusion Strategies

Hybrid CNN-Transformer models leverage the strengths of both architectures: local feature extraction from CNNs and global contextual modeling from Transformers, resulting in synergistic systems that are more effective than either of its individual components [17]. Several fusion methods have been devised such as sequential fusion (CNN features followed by Transformer features), parallel fusion (CNN features and Transformer features in parallel), and hierarchical fusion (multi-scale features integration). Layer-wise feature fusion for encoding, multi-scale feature fusion for skip connections and decoder integration are three fusion mechanisms which were successfully proposed by a parallel symmetric CNN-Transformer hybrid network (PSCT-Net) to achieve better segmentation performance for medical image analysis[18]. These architectures have proven to be especially effective for diabetic retinopathy detection, where they excel at capturing both fine-grained details and boundaries of lesions while preserving the overall context of the retina.

4.3 The State-of-the-Art Hybrid Models for DR Detection

Recently, some hybrid CNN-Transformer models, specifically tailored for the diabetic retinopathy detection problem, have been developed that have demonstrated outstanding performance. With the APTOS 2019 blindness detection dataset, a hybrid model of CNN and Transformers obtained an accuracy of 94.11%, and performed well in separating non-DR and mild DR classes, a key clinical differentiation for screening purposes[19]. An interpretable hybrid CNN-Transformer architecture (CvT-13) achieved a Quadratic Weighted Kappa score of 0.84 and an AUC of 0.93, which is very good performance with improved interpretability by visualizing the attention [20]. The multi-scale cross and self-attention network (MSCAS-Net) with Swin Transformer backbone architecture has demonstrated its effectiveness in handling challenges of imbalanced datasets and inconsistent image quality in diabetic retinopathy detection with accuracy scores of 93.8% on APTOS, 89.80% on DDR, and 86.70% on IDR benchmarks [21].

Table 2. Performance Comparison Of Hybrid Cnn–Transformer Models For Dr Detection

Model	Dataset	Accuracy	Major Advantage
ViT-Based FL Model	APTOS	93.00%	Global Context Learning
MSCAS-Net	IDRID	93.80%	Multi-Scale Attention
CvT-13 Hybrid	DDR	94.11%	Explainability
PSCT-Net	Retinal Dataset	95%+	Feature Fusion

Table 2 compares recent hybrid CNN–Transformer-based diabetic retinopathy detection models in terms of dataset usage, classification accuracy, and key architectural advantages for retinal disease analysis.

Algorithm : Hybrid CNN–Transformer-Based Retinal Image Classification

Input:

Retinal fundus image I

Output:

DR severity class

1: Preprocess retinal image
2: Extract local features using CNN
3: Generate image patches
4: Apply Transformer self-attention
5: Fuse CNN and Transformer features
6: Perform classification
7: Predict DR severity level
8: Return prediction result

The proposed algorithm is described for the retinal image classification process using the hybrid CNN–Transformer for automated prediction of the severity of DR. The classification performance and the contextual understanding of the retinal abnormalities are improved by integrating the CNN-based local feature extraction and the Transformer-based global attention learning.

5. Advanced Privacy-Preserving Techniques And Defensive Mechanisms

5.1 Gradient Inversion Attack Mitigation

One major threat in federated learning systems is the possibility of gradient inversion attacks, where an adversary can recover the original patient data from leaked model gradients [22]. A recent study showed that sensitive patient information can potentially be extracted from gradient information provided to the aggregator in federated aggregation with sophisticated reconstruction algorithms. The success of such attacks depends on the gradient resolution, the complexity of the model architecture, and the type of input data. To counteract the gradient inversion attack in federated medical imaging, various defense strategies have been devised. In medical imaging, the sensitivity-aware differential privacy outperforms the uniform differential privacy methods by 13.5% in performance using medical imaging datasets [23] due to measured privacy leakage risks from gradient inversion attacks. In particular, vision transformer CLS tokens can be encrypted instead of full gradients and still provide 30-fold communication reduction, preventing near-perfect image reconstruction attacks, while encrypted domain classification can be obtained with an accuracy of 90.02% for the histopathology data [24].

5.2 Membership Inference Attack Prevention

Membership inference attack is the ability to infer whether a particular patient data was used in model training, without directly reconstructing the images, which also results in violation of patient privacy [22]. The attacks rely on patterns in the statistics of model predictions and internal representations to be able to infer the membership with high confidence. The privacy-enhancing federated learning systems show significant improvement in resisting membership inference attacks. An ethical federated learning system using federated learning, differential privacy and safe aggregation achieved only 6.9% attack success rate with 88.6% classification accuracy, whereas base systems had attack success rate of more than 42.5% [25]. This is an important step towards clinical deployable privacy preserving DR detection systems.

5.3 Defense Against Byzantine and Poisoning Attacks

Poisoning attacks in federated learning are types of Byzantine and model poisoning attacks, which involves malicious clients in the federated learning system sending corrupted updates to affect the integrity of the global model [13]. In healthcare environments, where the accuracy of a diagnosis can have life and death impacts, such attacks could have devastating results.

Federated aggregation mechanisms, such as geometric medians, Krum filtering, and multi-sign gradient aggregation, have been incorporated in medical imaging federated frameworks. A zero-knowledge proof framework can be used to prove legitimate training processes to the client and defend against Byzantine attacks, while preserving privacy [13].

6. Clinical Implementation And Real-World Deployment Challenges

6.1 Communication Efficiency and Bandwidth Optimization

The regular exchange of information between the decentralized clients and the central servers is a fundamental requirement of federated learning, which can pose challenges in healthcare settings where bandwidth is limited and there are many participating institutions [2]. The amount of communication overhead during gradient aggregation can grow significantly as the size of the data set increases and the model architectures become more complex.

Communication costs have been significantly reduced by compression methods such as gradient quantization, sparsification and parameter reduction. A multi-deep learning framework for DR detection effectively classified DR with 99.06% accuracy with significantly lower learnable parameters than the baseline methods to enable efficient deployment in a resource-constrained clinical setting [26].Further enhancements in convergence efficiency for multi-institutional federated settings have been achieved using advanced communication scheduling and adaptive sampling methods.

6.2 Model Heterogeneity and Device Constraints

In federated learning, healthcare institutions can often have a mix of powerful medical computing systems equipped with GPUs and other resource-constrained point-of-care diagnostic devices [27]. To tackle these challenges, a personalized federated learning framework with device capability-aware client clustering and scaled neural architecture assignment (PFedCDP) is proposed to provide robust personalization under the non-IID and heterogeneous setting. In the context of diabetic retinopathy detection, this heterogeneity allows smaller rural clinic to be able to take part in federated learning without the need for expensive hardware infrastructure. The framework preserves diagnostic accuracy and allows for a variety of hardware deployment scenarios.

6.3 Regulatory Compliance and Explainability Integration

The systems must fully adhere to regulatory frameworks and incorporate explainability features for clinical deployment. Visualization techniques, such as grad-CAM and attention rollout, facilitate visual explanation of diagnostic decisions, improve clinician trust and facilitate regulatory approval processes [20]. Federated DR detection systems have been modified with the incorporation of Class Activation Mapping algorithms such as Grad-CAM, Grad-CAM++ and Score-CAM to make the model decision making transparent [3].

Together with federated learning, a comprehensive framework in conjunction with counterfactual explanation generation was found to be 23% more valid and 18% closer to the proximity scores for explanation quality, allowing clinicians to visualize and comprehend the model decision and noteworthy feature contribution for each patient's diagnosis [28].

7. Comparative Analysis Of Privacy-Utility Trade-Offs

7.1 Quantifying Privacy-Performance Relationships

The classical compromise between the level of privacy protection and the accuracy of the diagnosis by the model has been well studied in various publications. With moderate privacy budget ($\varepsilon \approx 10$), differential privacy achieves > 90% accuracy in medical imaging tasks, and with strict privacy ($\varepsilon \approx 1$), there is usually a 15-25% drop in accuracy [9]. The methods using homomorphic encryption provide higher guarantees of privacy but come with higher computational costs, and need careful optimization for clinical deployment.

In the context of detection methods for DR, the best performance of privacy-enhanced federated learning systems has demonstrated 90.8% accuracy using differential privacy and 98.69% using federated averaging for IID data setting [29]. The choice of privacy mechanism includes institution-specific factors, such as the sensitive nature of the patient population, regulatory environment and computational resources.

7.2 Performance Across Heterogeneous Data Distributions

The performance of federated learning can vary greatly depending on the characteristics of data heterogeneity, and Dirichlet non-IID distributions are a much better representation of real-world medical datasets than uniform distributions[30]. FedProx shows the best performance in the most extreme clinical heterogeneity, with 95.21% accuracy in detecting pneumonia and 85.95% in blood cell classification, and best convergence speed and probability calibration [30].

To address the problem of diabetic retinopathy detection across multiple institutions, performance preservation under non-IID conditions has been validated and models deployed in federated setting have been shown to have comparable or better accuracy than centralized setting, while preserving privacy. This shows that privacy preserving federated methods can improve both privacy and diagnosis performance.

7.3 Scalability and Multi-Institutional Considerations

Practical deployment is important as well because it needs to be scalable to the number of institutions participating. Typical successful federated learning applications for medical imaging have involved 3-10 client institutions, with work underway to develop systems that support 50-100+ client institutions. With adaptive communication scheduling and hierarchical aggregation structures, scalability has been enhanced while guaranteeing privacy and diagnostic accuracy.

Although there have been remarkable progress in privacy-preserving diabetic retinopathy detection, there are still some unsolved issues such as high communication overhead in federated learning, high computational complexity of Transformer architectures, low explainability of encrypted environments and the lack of multi-institutional benchmarking datasets. Additionally, current systems lack in scalability, fair-aware learning, and real-time deployment in resource-constrained healthcare systems.

8. Conclusion

In an era marked by rapidly advancing technology, clinical efficacy, and regulatory compliance, privacy-preserving deep learning for automated detection of DR stands as a pivotal point. Thanks to federated learning paradigms, cryptographic privacy mechanisms such as differential privacy and homomorphic encryption, and advanced hybrid CNN-Transformer architectures, clinically deployable DR detection systems can achieve diagnostic accuracy >95% while also providing formal privacy guarantees and preventing information leakage from multiple attack vectors [2].

The existing environment shows that privacy-preserving approaches do not have to sacrifice the quality of the diagnostics: often, collaborative federated learning is able to yield better models than other models, yet still respect the data localization demands and regulatory limitations of institutions. The hybrid CNN-Transformer architectures offer the flexibility in design to balance computational efficiency and diagnostic precision, while advanced cryptographic techniques and robust aggregation mechanisms address evolving threat landscapes [31].

Future research directions are to create more advanced privacy accounting systems that are compatible with the timeframe of clinical decision making, incorporate fairness auditing to ensure performance parity among different demographic groups, and create a set of common metrics for privacy preserving medical imaging systems, and perform prospective deployments of privacy-preserving systems to investigate their efficacy in clinical environments. Once the whole process of privacy-preserving deep learning for detecting diabetic retinopathy is complete, it can serve as a template for other medical imaging use cases and healthcare AI systems worldwide.

REFERENCES

1. Dasari, B., & Rajput, D. (2024a). A systematic review on diabetic retinopathy detection and classification based on deep learning techniques using fundus images. PeerJ Computer Science. https://doi.org/10.7717/peerj-cs.1947

2. Matta, S. S., & Bolli, M. (2025). Federated learning for privacy-preserving healthcare data sharing: Enabling global AI collaboration. American Journal of Scholarly Research and Innovation. https://doi.org/10.63125/jga18304

3. Ciaramella, G., Cesarelli, M., Martinelli, F., Agnello, P., Petyx, M., Santone, A., & Mercaldo, F. (2026). A privacy-preserving and explainable method for diabetic retinopathy classification and localization. IEEE Access. https://doi.org/10.1109/ACCESS.2026.3678201

4. Mane, A., & Shekapure, S. (2025). Detection and classification of diabetic retinopathy using a federated learning system ensuring data privacy. Proceedings of the 5th Asian Conference on Innovation in Technology. https://doi.org/10.1109/ASIANCON66527.2025.11281095

5. Dasari, B., & Rajput, D. (2024b). FedDL: Personalized federated deep learning for enhanced detection and classification of diabetic retinopathy. PeerJ Computer Science. https://doi.org/10.7717/peerj-cs.2508

6. Morajkar, S., & Sk, M. (2026). Hybrid FedProx optimization for privacy preserving diabetic retinopathy classification. Journal of Intelligent & Fuzzy Systems. https://doi.org/10.1177/18758967261443688

7. Gulati, S., Guleria, K., & Goyal, N. (2025). Privacy-preserving and collaborative federated learning model for the detection of ocular diseases. International Journal of Mathematical, Engineering and Management Sciences, 10(1).

8. Ziller, A., Usynin, D., Braren, R., Makowski, M., Rueckert, D., & Kaissis, G. (2021). Medical imaging deep learning with differential privacy. Scientific Reports, 11(1). https://doi.org/10.1038/s41598-021-93030-0

9. Mohammadi, M., Vejdanihemmat, M., Lotfinia, M., Rusu, M., Truhn, D., Maier, A. K., & Arasteh, S. T. (2025). Differential privacy for medical deep learning: Methods, tradeoffs, and deployment implications. npj Digital Medicine. https://doi.org/10.1038/s41746-025-02280-z

10. Bhulakshmi, D., & Rajput, D. (2024). Privacy-preserving detection and classification of diabetic retinopathy using federated learning with FedDEO optimization. Systems Science & Control Engineering. https://doi.org/10.1080/21642583.2024.2436664

11. Wang, Q. (2025). The integration of fully homomorphic encryption and machine learning: Technical paths, model adaptation, and practical applications. Applied and Computational Engineering. https://doi.org/10.54254/2755-2721/2026.tj30594

12. Antwi-Boasiako, E., Zhou, S., Liao, Y., Obiri, I., Kuada, E., Danso, E. K., & Mensah, A. E. (2024). Enhanced multi-key privacy-preserving distributed deep learning protocol with application to diabetic retinopathy diagnosis. Concurrency and Computation: Practice and Experience. https://doi.org/10.1002/cpe.8263

13. B. M. B., A. S. S., M. M., S. S., K. Dilip, & M. B. A. (2025). Zero-knowledge privacy-preserving federated learning for cross-institutional medical imaging diagnostics. Proceedings of the International Conference on Emerging Computation and Information Technologies. https://doi.org/10.1109/ICECIT67774.2025.11450988

14. Dib, O. (2025). A decentralized privacy-preserving framework for diabetic retinopathy detection using federated learning and blockchain. Results in Engineering. https://doi.org/10.1016/j.rineng.2025.105456

15. Nie, X., & Du, M. (2026). Vision transformer architectures for high-resolution medical image segmentation and disease localization: Architectural advances, localization strategies, and benchmark evaluation. Clinical Medicine and Health Research Journal, 6(2). https://doi.org/10.18535/cmhrj.v6i02.579

16. Swapna, M., Ravirala, T., & Reddy, N. (2025). Diabetic retinopathy detection using federated learning and vision transformers. International Journal of Interpreting Enigma Engineers, 2(1). https://doi.org/10.62674/ijiee.2025.v2i01.002

17. Hathurusinghe, I. (2025). Hybrid CNN–transformer architectures for computer vision-based medical image segmentation. International Journal of Electrical, Electronics and Computer Systems, 14(2).

18. Wang, B., Shi, H., Zhao, Z., & Zhang, S. (2025). PSCT-net: A parallel symmetric CNN-transformer hybrid network for medical image segmentation. Medical Engineering & Physics. https://doi.org/10.1016/j.medengphy.2025.104442

19. Lawate, A., Parab, P., Kokate, O., & Surti, N. (2025). Early detection of diabetic retinopathy using deep learning. Proceedings of the International Conference on Computing Convergence Technology. https://doi.org/10.1109/ICCCT63501.2025.11020520

20. Zhang, W., Belcheva, V., & Ermakova, T. (2025). Interpretable deep learning for diabetic retinopathy: A comparative study of CNN, ViT, and hybrid architectures. Computers, 14(5), Article 187. https://doi.org/10.3390/computers14050187

21. Badar, D., Abbas, J., Alsini, R., Abbas, T., Wang, C., & Daud, A. (2025). Transformer attention fusion for fine-grained medical image classification. Scientific Reports. https://doi.org/10.1038/s41598-025-07561-x

22. Koutsoubis, N., Yilmaz, Y., Ramachandran, R., Schabath, M., & Rasool, G. (2024). Privacy-preserving federated learning in medical imaging with uncertainty estimation. arXiv. https://arxiv.org/abs/2406.12815

23. Zheng, L., Cao, Y., Yoshikawa, M., Shen, Y., Rashed, E. A., Taura, K., Hanaoka, S., & Zhang, T. (2025). Sensitivity-aware differential privacy for federated medical imaging. Sensors. https://doi.org/10.3390/s25092847

24. Amin, A., Hasan, K., Hong, L., & Ullah, S. (2025). Privacy-preserving federated vision transformer learning leveraging lightweight homomorphic encryption in medical AI. Proceedings of the International Conference on Computing, Networking and Communications. https://doi.org/10.1109/ICNC68183.2026.11416875

25. Umamaheswari, B., & Sivaraman, K. (2025). Federated and differentially private deep learning with explainability for ethical medical imaging applications. Proceedings of the IEEE 5th International Conference on ICT in Business Industry & Government. https://doi.org/10.1109/ICTBIG68706.2025.11323843

26. Zafar, A., Kim, K., Ali, M. U., Byun, J., & Kim, S. (2025). A lightweight multi-deep learning framework for accurate diabetic retinopathy detection and multi-level severity identification. Frontiers in Medicine. https://doi.org/10.3389/fmed.2025.1551315

27. Shadin, N. S., Zhang, X., Wang, J., & Pan, M. (2025). Heterogeneity-aware private personalized federated learning for medical imaging via contrastive distillation. Proceedings of the IEEE BigData Congress. https://doi.org/10.1109/BigData66926.2025.11401860

28. Acheampong, E. M., Zhou, S., Liao, Y., Antwi-Boasiako, E., Agbley, B. L., & Ayekai, B. J. (2025). Blockchain-enabled multi-key fully homomorphic encryption for privacy-preserving federated learning with counterfactual explanation in medical image analysis. Proceedings of the 22nd International Computer Conference on Wavelet Active Media Technology and Information Processing. https://doi.org/10.1109/ICCWAMTIP68645.2025.11352603

29. Gulati, S., Guleria, K., Goyal, N., AlZubi, A. A., & Castilla, Á. K. (2024). A privacy-preserving collaborative federated learning framework for detecting retinal diseases. IEEE Access. https://doi.org/10.1109/ACCESS.2024.3493946

30. Riyadi, M. A. Q., Dewi, A. M., Mukhlishin, Z. A. N., & Arep, Z. R. A. (2026). Comparative evaluation of federated learning algorithms in Dirichlet non-IID medical imaging. Jurnal Sains, Nalar, dan Aplikasi Teknologi Informasi, 5(1). https://doi.org/10.20885/snati.v5.i1.44597

31. Jahns, E., Stojkov, M., & Kinsy, M. A. (2025). Privacy-preserving deep learning: A survey on theoretical foundations, software frameworks, and hardware accelerators. IEEE Access. https://doi.org/10.1109/ACCESS.2025.3561721