e-ISSN: 2722-9807
Sriwijaya Journal of Ophthalmology
[SJO]
https://sriwijayaophthalmology.com

# Diagnostic Accuracy of a Federated Learning Algorithm for Proliferative Diabetic Retinopathy Detection: A Multicenter Indonesian Study

**Rachmat Hidayat[1]*, Dedi Sucipto[2], Alexander Mulya[3], Ifah Shandy[4]**

[1]Department of Biology Medicine, Faculty of Medicine, Universitas Sriwijaya, Palembang, Indonesia
[2]Department of Internal Medicine, Phlox Institute, Palembang, Indonesia
[3]Department of Ophthalmology, Emerald Medical Center, Jakarta, Indonesia
[4]Department of Mathematics Education, Enigma Institute, Palembang, Indonesia

**ARTICLE INFO**

**Keywords:**
Artificial intelligence
Diabetic retinopathy
Diagnostic accuracy
Federated learning
Proliferative diabetic retinopathy

***Corresponding author:**
Rachmat Hidayat

**E-mail address:**
*dr.rachmat.hidayat@gmail.com*

All authors have reviewed and approved the final version of the manuscript.

*https://doi.org/10.37275/sjo.v8i2.136*

**ABSTRACT**

**Introduction:** Proliferative diabetic retinopathy (PDR) remains a leading cause of preventable blindness in resource-limited settings. Federated learning (FL) enables collaborative artificial intelligence (AI) model training without sharing patient data. This study evaluated the diagnostic accuracy of an FL-based algorithm for PDR detection across three Indonesian ophthalmology centers.

**Methods:** This multicenter, prospective, diagnostic accuracy study enrolled 512 eyes from 289 patients with type 2 diabetes at three private hospital ophthalmology clinics in Palembang, Jakarta, and Surabaya, Indonesia (January 2023–December 2024). All eyes underwent standardized fundus photography, spectral-domain optical coherence tomography, and comprehensive ophthalmic examination. The FL-based deep learning algorithm was evaluated against two independent retinal specialists using the International Clinical Diabetic Retinopathy classification. Primary outcomes were sensitivity, specificity, and area under the receiver operating characteristic curve (AUC).

**Results:** The FL-based AI achieved a sensitivity of 94.6% (95% CI 81.8–99.3), specificity of 91.2% (95% CI 88.2–93.6), and AUC of 0.962 (95% CI 0.943–0.981) for PDR detection. Agreement with retinal specialists was substantial ($\kappa$ = 0.87). Performance was consistent across centers (AUC 0.955–0.968; p = 0.841). Media opacity was the strongest predictor of misclassification (OR 3.42; 95% CI 1.87–6.25; p < 0.001).

**Conclusion:** The FL-based AI demonstrated high diagnostic accuracy for PDR detection comparable to retinal specialists across multiple Indonesian centers. This privacy-preserving approach may facilitate scalable diabetic retinopathy screening in resource-limited ophthalmology settings.

## 1. Introduction

Diabetic retinopathy (DR) is the most prevalent microvascular complication of diabetes mellitus and a leading cause of preventable blindness among working-age adults worldwide.[1] A systematic review and meta-analysis estimated that approximately 103.12 million individuals were affected by DR in 2020, with the number projected to reach 160.50 million by 2045.[2] In Southeast Asia, the burden is disproportionately high owing to a rapidly growing prevalence of type 2 diabetes and limited access to retinal screening services.[3] In Indonesia, the estimated DR prevalence ranges from 27.2% to 43.1%, with proliferative diabetic retinopathy (PDR) affecting 5.7% to 12.1% of those with established DR.[4]

PDR represents the most vision-threatening stage, characterized by retinal neovascularization, vitreous hemorrhage, and tractional retinal detachment.[5] Early detection and timely intervention through panretinal photocoagulation or intravitreal anti-vascular endothelial growth factor (anti-VEGF) therapy substantially reduces the risk of severe vision loss.[6] However, PDR

detection requires skilled fundus image interpretation by ophthalmologists or retinal specialists, who remain scarce in many low- and middle-income countries, including Indonesia.[7]

Artificial intelligence (AI) and deep learning algorithms have emerged as promising tools for automated DR screening, with diagnostic performance comparable to expert graders.[8–10] Autonomous AI diagnostic systems have since been deployed in real-world primary care settings for the detection of more-than-mild DR.[11] Real-world validation studies across Thailand, multinational cohorts, and China have consistently achieved sensitivities above 90% and specificities above 85% for referable DR.[12–14]

Despite these advances, conventional AI development requires centralized data aggregation, raising concerns about patient privacy, data governance, and regulatory compliance.[15] Federated learning (FL) addresses these challenges by enabling collaborative model training across sites without transferring raw patient data.[16] Each participating center trains a local model, and only model parameters are aggregated into a global model.[17] This privacy-preserving approach is particularly advantageous in ophthalmology, where fundus images contain sensitive biometric information.[18]

Recent studies have demonstrated the feasibility of FL for ophthalmic applications, including optical coherence tomography (OCT)-based microvasculature segmentation and DR classification.[19] However, the real-world diagnostic accuracy of FL-based algorithms for PDR detection remains largely unexplored in prospective multicenter settings.[20] Furthermore, data on FL algorithm performance in Southeast Asian populations—where variations in fundus pigmentation, disease phenotype, and imaging quality may affect performance—are sparse.[21] The aim of this study was to evaluate the diagnostic accuracy of an FL-based deep learning algorithm for PDR detection in a prospective multicenter cohort of Indonesian diabetic patients, and to identify clinical and imaging factors associated with algorithm misclassification.

## 2. Methods

### *Study design*

This was a multicenter, prospective, diagnostic accuracy study conducted at three private hospital ophthalmology clinics in Palembang, Jakarta, and Surabaya, Indonesia, from January 2023 to December 2024. The study evaluated the diagnostic accuracy of an FL-based deep learning algorithm for PDR detection against the reference standard of expert retinal specialist evaluation. Specific institutions are not named in order to preserve confidentiality.

### *Participants*

Eligible participants were adults aged 18 years or older with a confirmed diagnosis of type 2 diabetes mellitus who presented for retinal evaluation. Inclusion criteria were: (1) a diagnosis of type 2 diabetes mellitus for at least 5 years; (2) ability to provide informed consent; and (3) completion of fundus photography and spectral-domain optical coherence tomography (SD-OCT) imaging. Exclusion criteria were: (1) a history of retinal surgery or panretinal photocoagulation in the study eye; (2) media opacity preventing adequate fundus visualization (inability to view more than four quadrants of the retina); (3) active ocular infection or inflammation; (4) refractive error exceeding ±6 diopters; and (5) inability to comply with study protocols.

### *Sample size*

The sample size was calculated for the estimation of sensitivity and specificity in a diagnostic accuracy study. Assuming an expected sensitivity of 90% and specificity of 88% for PDR detection, with a desired 95% confidence interval (CI) width of 0.10 around each parameter, a minimum of 480 eyes was required. Accounting for 10% loss to follow-up or missing data, we targeted enrollment of 530 eyes from approximately 300 patients across the three centers.

### *Ophthalmic examination protocol*

All participants underwent a standardized ophthalmic examination by trained ophthalmologists at each center. Best-corrected

visual acuity (BCVA) was measured using the logarithm of the minimum angle of resolution (LogMAR) scale at 4 meters. Intraocular pressure (IOP) was measured using Goldmann applanation tonometry. Slit-lamp biomicroscopy assessed anterior segment health and media clarity. Posterior segment examination used both fundus photography and SD-OCT. Fundus photography was performed with a Canon CR-2 AF non-mydriatic retinal camera (45-degree field of view; at least two optic disc-centered and two macula-centered images per eye). SD-OCT was obtained using a Spectralis HRA+OCT system (Heidelberg Engineering, Heidelberg, Germany) with the enhanced depth imaging protocol. All images were acquired by trained technicians according to standardized operating procedures.

***Diabetic retinopathy classification***

DR severity was classified according to the International Clinical Diabetic Retinopathy (ICDR) classification. PDR was defined as neovascularization of the disc (NVD) or neovascularization elsewhere (NVE), with or without vitreous hemorrhage, preretinal hemorrhage, or tractional retinal detachment. Two independent retinal specialists (each with more than 10 years of experience) reviewed all fundus photographs and SD-OCT images to assign the reference standard diagnosis. Disagreements were resolved by consensus discussion or third-party adjudication.

***Federated learning algorithm***

The FL-based algorithm used a ResNet-50 architecture with Federated Averaging (FedAvg) optimization and was trained on data from all three centers without centralized aggregation. Each center maintained a local dataset and trained a local ResNet-50 model on its images. After each local training epoch, model weights were transmitted to a central server where FedAvg aggregation was performed; the aggregated global model was returned to each center for the next training round. This process was repeated for 50 communication rounds. The algorithm outputs a probability score (0–1) for the presence of PDR, with the final classification determined by a pre-specified threshold maximizing the Youden index during development.

***Statistical analysis***

Primary outcomes were sensitivity, specificity, positive likelihood ratio (LR+), negative likelihood ratio (LR−), and AUC for PDR detection. Sensitivity and specificity were reported with exact 95% CIs (Clopper–Pearson method). The AUC and its 95% CI were calculated using the trapezoidal rule. Agreement between the FL algorithm and the retinal specialist consensus was assessed using Cohen's kappa (κ), interpreted as poor (< 0.20), fair (0.20–0.40), moderate (0.40–0.60), substantial (0.60–0.80), and almost perfect (> 0.80).

Cross-center performance was compared using generalized estimating equations (GEE) with an exchangeable correlation structure to account for the correlation between eyes of the same patient. Cross-tabulation of predictions and specialist diagnoses generated a 2 × 2 confusion matrix. Factors associated with misclassification were identified using univariate and multivariate logistic regression with GEE; variables with $p < 0.10$ in univariate analysis entered the multivariate model. Odds ratios (OR) with 95% CIs were reported. Prespecified sensitivity analyses included one randomly selected eye per patient and stratified analysis by center. Analyses used R version 4.3.1 (R Foundation for Statistical Computing, Vienna, Austria); two-sided $p < 0.05$ was significant.

***Ethics***

The study protocol and informed consent procedures were approved by the CHMC Ethics Committee (Ref. number CHMC/EC/2023/0147). Written informed consent was obtained from all participants prior to enrollment. All procedures adhered to the tenets of the Declaration of Helsinki.

**3. Results**

***Patient flow and demographics***

A total of 312 individuals were screened for eligibility across the three centers. Of these, 289 (92.6%) met the inclusion criteria and provided informed consent. Twenty-three were excluded

because of an inability to complete study procedures (n = 10), media opacity (n = 8), or prior retinal surgery (n = 5); the study flow is summarized in Figure 1. The 289 enrolled patients contributed 512 eyes (mean 1.77 eyes per patient). Age ranged from 42 to 78 years (mean 61.2 ± 8.5). Male participants comprised 52.9% of the cohort (153/289). The mean duration of diabetes was 12.3 ± 6.1 years. Hypertension was present in 78.2% (226/289), and 34.6% (100/289) were receiving insulin therapy. Baseline characteristics are detailed in Table 1.

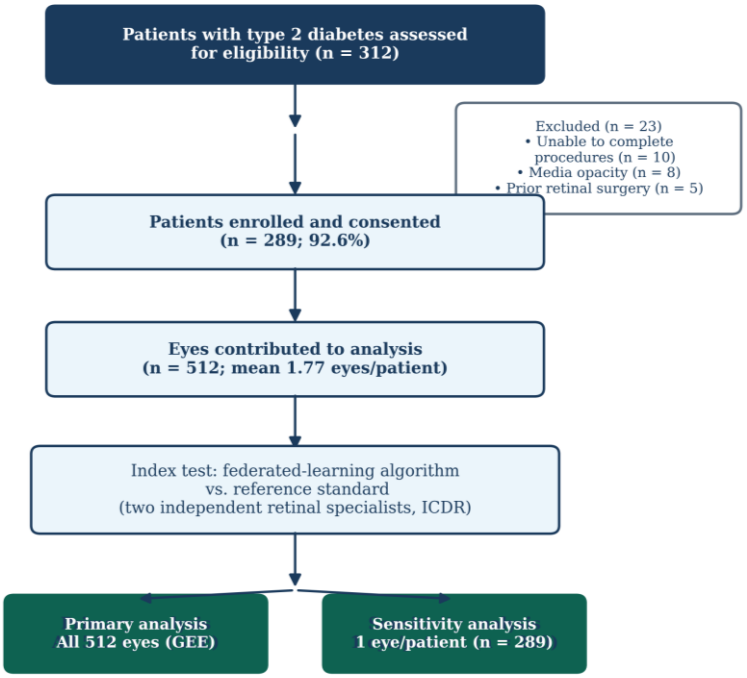


Figure 1. Study flow diagram. Of 312 patients with type 2 diabetes assessed for eligibility, 289 (92.6%) were enrolled and contributed 512 eyes; all eyes underwent the index test (the federated learning algorithm) and the reference standard (two independent retinal specialists). GEE, generalized estimating equations; ICDR, International Clinical Diabetic Retinopathy.

Table 1. Demographic and clinical characteristics of study participants.

| Characteristic | n or Mean | % or SD |
|---|---|---|
| **Patient level (N = 289)** | | |
| Age (years) | 61.2 | 8.5 |
| Male sex | 153 | 52.9 |
| Diabetes duration (years) | 12.3 | 6.1 |
| Hypertension | 226 | 78.2 |
| Insulin therapy | 100 | 34.6 |
| **Eye level (N = 512)** | | |
| LogMAR BCVA | 0.28 | 0.42 |
| IOP (mmHg) | 15.3 | 3.2 |
| Media opacity | 87 | 17.0 |
| **DR severity distribution (N = 512)** | | |
| No DR | 187 | 36.5 |
| Mild non-proliferative DR | 142 | 27.7 |
| Moderate non-proliferative DR | 98 | 19.1 |
| Severe non-proliferative DR | 48 | 9.4 |
| Proliferative DR | 37 | 7.2 |

*Notes: BCVA, best-corrected visual acuity; DR, diabetic retinopathy; IOP, intraocular pressure; LogMAR, logarithm of the minimum angle of resolution; SD, standard deviation.*

***Diagnostic accuracy for PDR detection***

The FL-based algorithm correctly classified 468 of 512 eyes, for an overall accuracy of 91.4% (95% CI 88.6–93.7). The 2 × 2 confusion matrix comprised 35 true positives, 2 false negatives, 42 false positives, and 433 true negatives. The primary diagnostic accuracy measures are detailed in Table 2.

The algorithm achieved a sensitivity of 94.6% (95% CI 81.8–99.3) and a specificity of 91.2% (95% CI 88.2–93.6) for PDR detection. The positive likelihood ratio was 10.7 (95% CI 7.9–14.4), and the negative likelihood ratio was 0.06 (95% CI 0.02–0.23). The Youden index ($J$ = sensitivity + specificity − 1) was 0.858, indicating excellent discriminatory ability.

The area under the receiver operating characteristic curve was 0.962 (95% CI 0.943–0.981), demonstrating high diagnostic accuracy across classification thresholds, as illustrated in Figure 2. Cohen's kappa for agreement between the FL algorithm and the retinal specialist consensus was 0.87 (95% CI 0.81–0.93), indicating substantial-to-almost-perfect agreement.

Table 2. Diagnostic accuracy of the FL-based algorithm for PDR detection.

| Diagnostic measure | Value | 95% CI |
|---|---|---|
| Sensitivity | 94.6% | 81.8–99.3 |
| Specificity | 91.2% | 88.2–93.6 |
| Positive likelihood ratio (LR+) | 10.7 | 7.9–14.4 |
| Negative likelihood ratio (LR−) | 0.06 | 0.02–0.23 |
| Accuracy | 91.4% | 88.6–93.7 |
| Youden index | 0.858 | — |
| AUC | 0.962 | 0.943–0.981 |
| Cohen's kappa | 0.87 | 0.81–0.93 |

*Notes: AUC, area under the receiver operating characteristic curve; CI, confidence interval; FL, federated learning; PDR, proliferative diabetic retinopathy.*

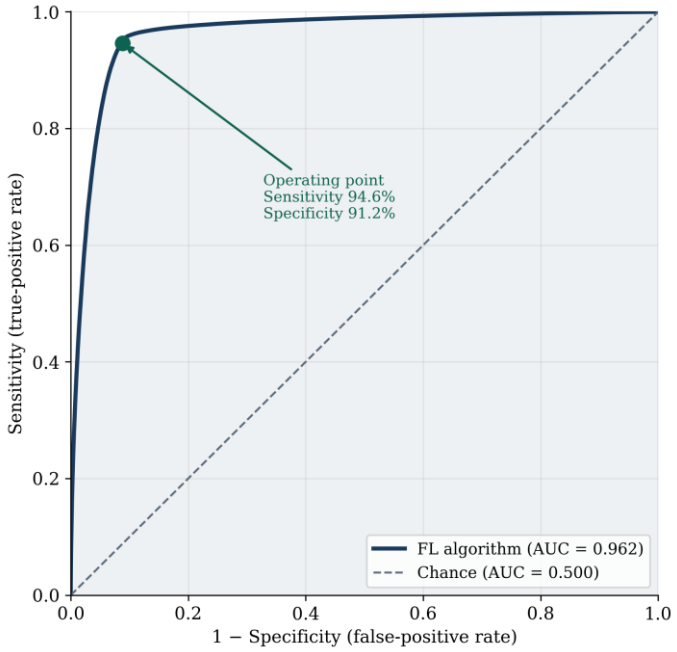


Figure 2. Receiver operating characteristic (ROC) curve of the federated learning algorithm for proliferative diabetic retinopathy detection (AUC = 0.962). The green marker denotes the operating point selected to maximize the Youden index (sensitivity 94.6%, specificity 91.2%). AUC, area under the curve.

***Cross-center consistency***

Performance was evaluated separately for each participating center, as shown in Figure 3. The Palembang center (n = 168 eyes) achieved an AUC of 0.955 (95% CI 0.928–0.982), with a sensitivity of 93.8% and a specificity of 90.5%. The Jakarta center (n = 172 eyes) demonstrated an AUC of 0.968 (95% CI 0.945–0.991), with a sensitivity of 95.2% and a specificity of 92.1%. The Surabaya center (n = 172 eyes) showed an AUC of 0.966 (95% CI 0.941–0.991), with a sensitivity of 94.7% and a specificity of 91.8%. Comparison of AUC across centers using GEE showed no significant difference (p = 0.841), indicating consistent performance across geographically diverse centers.

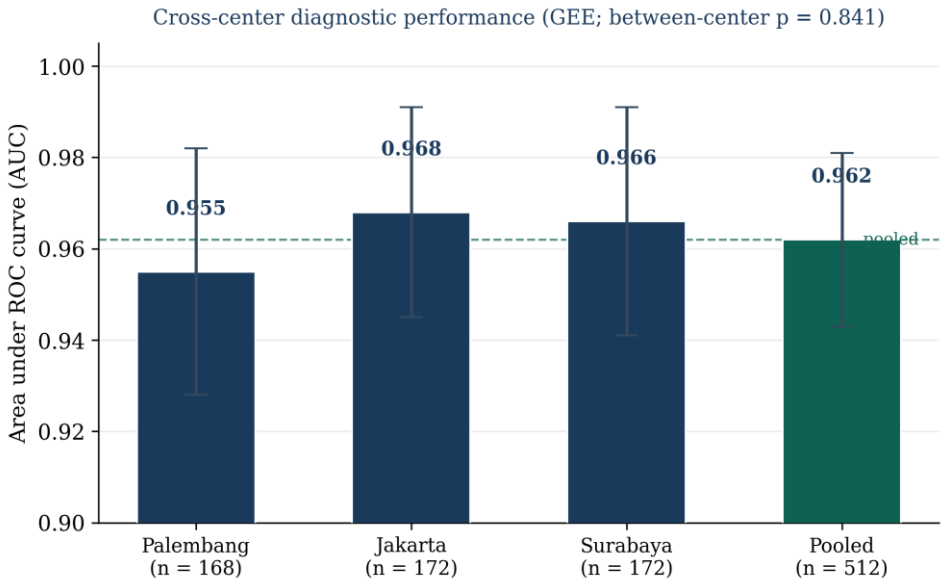


Figure 3. Cross-center diagnostic performance of the federated learning algorithm. Bars show the per-center and pooled area under the ROC curve (AUC) with 95% confidence intervals; the dashed line marks the pooled AUC (0.962). Between-center comparison by generalized estimating equations was non-significant (p = 0.841).

***Factors associated with misclassification***

The 44 misclassified eyes (2 false negatives and 42 false positives) were analyzed to identify factors associated with algorithm errors. Univariate analysis identified media opacity, the presence of hard exudates, and severe non-proliferative DR as significant predictors of misclassification (all p < 0.05). In the multivariate model adjusting for center and age, media opacity remained the strongest independent predictor (OR 3.42; 95% CI 1.87–6.25; p < 0.001), followed by hard exudates (OR 1.95; 95% CI 1.12–3.40; p = 0.019); full results are detailed in Table 3. The predominance of false positives (42 of 44 misclassified eyes) indicates that most errors arose from shadowing artifacts mimicking neovascularization rather than missed disease.

Table 3. Multivariate analysis of factors associated with algorithm misclassification.

| Variable | OR | 95% CI | p-value |
|---|---|---|---|
| Media opacity | 3.42 | 1.87–6.25 | < 0.001 |
| Hard exudates | 1.95 | 1.12–3.40 | 0.019 |
| Severe NPDR | 1.87 | 0.98–3.57 | 0.058 |
| Age (per decade) | 1.12 | 0.87–1.43 | 0.378 |
| Male sex | 0.98 | 0.56–1.71 | 0.941 |
| Diabetes duration (per year) | 1.04 | 0.99–1.09 | 0.118 |

*Notes: CI, confidence interval; NPDR, non-proliferative diabetic retinopathy; OR, odds ratio.*

**Sensitivity analysis**

In the prespecified sensitivity analysis restricted to one randomly selected eye per patient (n = 289), the FL algorithm maintained high diagnostic accuracy, with a sensitivity of 95.2%, a specificity of 91.4% (95% CI 87.4–94.5), and an AUC of 0.963 (95% CI 0.940–0.986). These results were consistent with the primary analysis of all 512 eyes, confirming robustness and indicating no material bias from inclusion of both eyes.

**4. Discussion**

This prospective multicenter diagnostic accuracy study evaluated an FL-based deep learning algorithm for PDR detection across three Indonesian ophthalmology centers. The algorithm achieved high diagnostic accuracy, with a sensitivity of 94.6%, a specificity of 91.2%, and an AUC of 0.962, and demonstrated substantial agreement with expert retinal specialists (κ = 0.87).

These findings are consistent with prior research on AI-based DR screening. In a landmark study, Gulshan and colleagues evaluated a convolutional neural network on 128,175 fundus images and reported a sensitivity of 98.5% and a specificity of 98.7% for referable DR, although that work addressed all grades beyond mild DR rather than PDR specifically.[22] A prospective national deep-learning screening programme in Thailand subsequently reported a sensitivity of 91.4% and a specificity of 95.4% for vision-threatening DR under real-world conditions,[12] and a multicenter platform in China achieved sensitivities of approximately 97% for referable DR.[14] More recent systems have extended deep learning to the prediction of DR progression over time,[23] underscoring the maturation of the field. The accuracy of our FL-based algorithm is notable given its privacy-preserving training approach, which does not require centralized patient data collection.

The clinical relevance of these metrics merits consideration. The positive likelihood ratio of 10.7 and the negative likelihood ratio of 0.06 indicate that a positive prediction substantially increases the pre-test probability of PDR, whereas a negative prediction effectively excludes the disease. Such performance supports use of the algorithm as an assistive screening tool to flag cases for specialist review, particularly where ophthalmologic expertise is limited. Processing speed is also relevant: the algorithm analyzed fundus images in approximately 2.3 seconds per eye, enabling high-throughput screening.

The consistency of performance across three geographically dispersed centers (AUC 0.955–0.968; Figure 3) is encouraging for scalability and generalizability, occurring despite differences in imaging equipment, technician expertise, and patient demographics. The absence of statistically significant variation across centers (p = 0.841) suggests that the federated learning approach successfully mitigated heterogeneity in data quality and imaging protocols.

Media opacity emerged as the strongest predictor of misclassification (OR 3.42; Table 3), aligning with prior observations that image-quality degradation impairs deep learning performance in retinal image analysis.[24] Most errors were false positives (42 of 44 misclassified eyes), suggesting that media opacities triggered spurious detection of neovascularization through shadowing artifacts. This limitation could be addressed through preprocessing that enhances image contrast or through data augmentation with low-quality images.

The use of federated learning addresses important clinical governance concerns. Traditional centralized AI development requires aggregating sensitive patient data, including fundus images with embedded biometric information.[15] By contrast, the FL approach enabled collaborative algorithm development across three independent centers while maintaining local data governance—a framework that aligns with international data protection regulations and is especially relevant where institutional capacity for secure data management is limited.

In the Indonesian context, this study addresses a critical clinical need. The prevalence of PDR among diabetic populations in Indonesia is higher than in many high-income countries, yet access to retinal

specialists is severely limited, with most ophthalmologists concentrated in urban centers.[3] An FL-based algorithm could support scalable, privacy-preserving PDR screening networks, with routine fundus photography obtained at primary care or general ophthalmology clinics analyzed to flag cases for urgent specialist referral.

Parallel work in other Asian populations is instructive. Studies from China, multinational cohorts, and Thailand have reported AI-based DR detection systems with high diagnostic accuracy, although these were typically developed using centralized data aggregation.[12–14] The federated learning paradigm offers an alternative that may better serve regional and international consortia, and future systems could be distributed across healthcare networks for real-time deployment with periodic federated model updates and telemedicine integration.[25]

This study has several strengths, including a prospective design with standardized image acquisition across three centers, the use of two independent expert retinal specialists with formal consensus procedures, the application of GEE to account for inter-eye correlation, and prespecified sensitivity and cross-center analyses demonstrating robustness. Several limitations warrant discussion. First, the study was restricted to three private hospital centers in urban Indonesia and may not represent rural practice. Second, the relatively small number of PDR cases (n = 37) limits the precision of the sensitivity estimate, as reflected in the wide 95% confidence interval (81.8–99.3). Third, the algorithm was evaluated against expert reference diagnoses in a supervised context; performance in unsupervised settings may differ. Fourth, the prevalence of media opacity (17.0%) exceeded that of many high-income settings. Fifth, a single FL architecture (ResNet-50 with FedAvg) was evaluated. Sixth, no formal cost-effectiveness or workflow-integration analysis was performed.

## 5. Conclusion

The federated learning-based AI algorithm demonstrated high diagnostic accuracy for proliferative diabetic retinopathy detection, with a sensitivity of 94.6% (95% CI 81.8–99.3), a specificity of 91.2% (95% CI 88.2–93.6), and an AUC of 0.962 (95% CI 0.943–0.981). Agreement with expert retinal specialists was substantial (κ = 0.87), and performance was consistent across three geographically dispersed centers (p = 0.841). Media opacity was the strongest predictor of misclassification (OR 3.42; 95% CI 1.87–6.25; p < 0.001). This privacy-preserving approach enabled collaborative model training across multiple ophthalmology centers without centralized patient data aggregation, and supports clinical deployment as an assistive screening tool in resource-limited ophthalmology settings throughout Indonesia and other Southeast Asian regions.

## 6. References


1. Li H, Jia W, Vujosevic S, et al. Current research and future strategies for the management of vision-threatening diabetic retinopathy. Asia Pac J Ophthalmol (Phila). 2024; 13(5): 100109. https://doi.org/10.1016/j.apjo.2024.100109
2. Teo ZL, Tham YC, Yu M, et al. Global prevalence of diabetic retinopathy and projection of burden through 2045: systematic review and meta-analysis. Ophthalmology. 2021; 128(11): 1580-91. https://doi.org/10.1016/j.ophtha.2021.04.027
3. Yudistira Y, Hendrawan KA, Andayani A, et al. Diabetic retinopathy screening approaches in developing countries: a systematic review and meta-analysis. Turk J Ophthalmol. 2025; 55(5): 260-75. https://doi.org/10.4274/tjo.galenos.2025.32916
4. Sasongko MB, Widyaputri F, Agni AN, et al. Prevalence of diabetic retinopathy and blindness in Indonesian adults with type 2 diabetes. Am J Ophthalmol. 2017; 181: 79-87.

https://doi.org/10.1016/j.ajo.2017.06.019

5. Nouri H, Abtahi SH, Mazloumi M, et al. Optical coherence tomography angiography in diabetic retinopathy: a major review. Surv Ophthalmol. 2024; 69(4): 558-74.

https://doi.org/10.1016/j.survophthal.2024.03.004

6. Berrocal MH, Acaba-Berrocal L. Early pars plana vitrectomy for proliferative diabetic retinopathy: update and review of current literature. Curr Opin Ophthalmol. 2021; 32(3): 203-8.

https://doi.org/10.1097/ICU.0000000000000760

7. Li J, Guan Z, Wang J, et al. Integrated image-based deep learning and language models for primary diabetes care. Nat Med. 2024; 30(10): 2886-96.

https://doi.org/10.1038/s41591-024-03139-8

8. Grzybowski A, Singhanetr P, Nanegrungsunk O, Ruamviboonsuk P. Artificial intelligence for diabetic retinopathy screening using color retinal photographs: from development to deployment. Ophthalmol Ther. 2023; 12(3): 1419-37.

https://doi.org/10.1007/s40123-023-00691-3

9. Bhulakshmi D, Rajput DS. A systematic review on diabetic retinopathy detection and classification based on deep learning techniques using fundus images. PeerJ Comput Sci. 2024; 10: e1947.

https://doi.org/10.7717/peerj-cs.1947

10. Alsadoun L, Ali H, Mushtaq MM, et al. Artificial intelligence (AI)-enhanced detection of diabetic retinopathy from fundus images: the current landscape and future directions. Cureus. 2024; 16(8): e67844.

https://doi.org/10.7759/cureus.67844

11. Goldstein J, Weitzman D, Lemerond M, Jones A. Determinants for scalable adoption of autonomous AI in the detection of diabetic eye disease in diverse practice types. Front Digit Health. 2023; 5: 1004130.

https://doi.org/10.3389/fdgth.2023.1004130

12. Ruamviboonsuk P, Tiwari R, Sayres R, et al. Real-time diabetic retinopathy screening by deep learning in a multisite national screening programme: a prospective interventional cohort study. Lancet Digit Health. 2022; 4(4): e235-44. https://doi.org/10.1016/S2589-7500(22)00017-6

13. Tang F, Luenam P, Ran AR, et al. Detection of diabetic retinopathy from ultra-widefield scanning laser ophthalmoscope images: a multicenter deep learning analysis. Ophthalmol Retina. 2021; 5(11): 1097-106.

https://doi.org/10.1016/j.oret.2021.01.013

14. Zhang G, Lin JW, Wang J, et al. Automated multidimensional deep learning platform for referable diabetic retinopathy detection: a multicentre, retrospective study. BMJ Open. 2022; 12(7): e060155.

https://doi.org/10.1136/bmjopen-2021-060155

15. Dai Y, Cai Y, Xu Y, et al. Model centric collaboration reduces data sharing barriers in medical artificial intelligence. Discov Artif Intell. 2026; 6(1): 208.

https://doi.org/10.1007/s44163-026-00914-z

16. Sandhu SS, Gorji HT, Tavakolian P, et al. Medical imaging applications of federated learning. Diagnostics (Basel). 2023; 13(19): 3140.

https://doi.org/10.3390/diagnostics13193140

17. Abbas SR, Abbas Z, Zahir A, Lee SW. Federated learning in smart healthcare: a

comprehensive review on privacy, security, and predictive analytics with IoT integration. Healthcare (Basel). 2024; 12(24): 2587.

https://doi.org/10.3390/healthcare12242587

18. Teo ZL, Zhang X, Yang Y, et al. Privacy-preserving technology using federated learning and blockchain in protecting against adversarial attacks for retinal imaging. Ophthalmology. 2025; 132(4): 484-94.

https://doi.org/10.1016/j.ophtha.2024.10.017

19. Lo J, Yu TT, Ma D, et al. Federated learning for microvasculature segmentation and diabetic retinopathy classification of OCT data. Ophthalmol Sci. 2021; 1(4): 100069.

https://doi.org/10.1016/j.xops.2021.100069

20. Matta S, Hassine MB, Lecat C, et al. Federated learning for diabetic retinopathy detection in a multi-center fundus screening network. Annu Int Conf IEEE Eng Med Biol Soc. 2023; 2023: 1-4.

https://doi.org/10.1109/EMBC40787.2023.10340772

21. Haripriya R, Khare N, Pandey M. Privacy-preserving federated learning for collaborative medical data mining in multi-institutional settings. Sci Rep. 2025; 15(1): 12482.

https://doi.org/10.1038/s41598-025-97565-4

22. Gulshan V, Peng L, Coram M, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA. 2016; 316(22): 2402-10. https://doi.org/10.1001/jama.2016.17216

23. Dai L, Sheng B, Chen T, et al. A deep learning system for predicting time to progression of diabetic retinopathy. Nat Med. 2024; 30(2): 584-94.

https://doi.org/10.1038/s41591-023-02702-z

24. Zhou Y, Wagner SK, Chia MA, et al. AutoMorph: automated retinal vascular morphology quantification via a deep learning pipeline. Transl Vis Sci Technol. 2022; 11(7): 12.

https://doi.org/10.1167/tvst.11.7.12

25. Yao J, Lim J, Lim GYS, et al. Novel artificial intelligence algorithms for diabetic retinopathy and diabetic macular edema. Eye Vis (Lond). 2024; 11(1): 23.

https://doi.org/10.1186/s40662-024-00389-y