
HbA1c-Integrated Multimodal Deep Learning for Diabetic Retinopathy Severity Classification and Short-Term Progression Risk Stratification

Ananya V S¹, Jasmine Selvakumari Jeya I², Aldrin Joan Pandian W³, Kannaiya Raja N⁴, Indirani M⁵, Saravanan S⁶

^{1,3}*School of Computing Science Engineering and Artificial Intelligence, VIT Bhopal University, Sehore, Madhya Pradesh, India, vsananya02@gmail.com, aldrinjoan6@gmail.com*

^{2,4}*School of Computing Science and Engineering, VIT Bhopal University, Sehore, Madhya Pradesh, India, wjasminejeya@gmail.com, kannaiyaraja@vitbhopal.ac.in*

⁵*Department of Computer Science and, Business Systems, Vel Tech Multi Tech Dr. Rangarajan Dr. Sakunthala Engineering College, Avadi, Chennai, India, indugovind10@gmail.com*

⁶*School of Computing, SRM Institute of Science and Technology, Tiruchirappalli, India, jeyasaraa@gmail.com*

Abstract.

Diabetic Retinopathy (DR) is a primary factor of reducible blindness among adults having diabetes, yet conventional fundus image screening provides only static severity grading and lacks quantitative estimation of progression risk or integration of key biomarkers such as glycated hemoglobin (HbA1c). This study presents a Multimodal DR-HbA1c framework, a lightweight deep learning architecture that performs five class DR severity classification from retinal fundus images while generating patient specific 6 and 12 month progression risk probabilities by combining predicted DR grade with a VTDR inspired rule based risk calculator calibrated from literature distributions. The model uses a pretrained DenseNet121 backbone producing 256 dimensional image features, a compact clinical branch encoding HbA1c into 16 dimensional features, and late fusion to form a 272 dimensional joint representation. The model attains 79.93% validation accuracy on 1,096 unseen APTOS 2019 images, enhancing 6.39% over the image only baseline with macro AUC 0.9185. Sensitivity calculation validates glycemic efficiency, exposing risk escalation with increasing HbA1c. Transparency applying LIME showcases clinically related retinal areas. This model proves the suitability of HbA1c determined personalized risk classification in DR screening for resource finite settings and helps adaptive screening periods for improved early intervention and management outputs.

Keywords. Multimodal Deep Learning, Progression Risk Prediction, Retinal Fundus Image Analysis, HbA1c, Diabetic Retinopathy (DR).

1. INTRODUCTION

Diabetic Retinopathy is a leading microvascular is-sue of diabetes and the foremost effect of preventable vision loss in the employable age population worldwide. As of 2021, an estimated 103 million people globally were living with DR, with prevalence rising sharply in low- and middle-income nations-particularly India-due to the accelerating incidence of type 2 diabetes [1]. Current clinical screening methods depend solely on retinal fundus photo-graphs for static severity grading (No DR, Mild, Moderate, Severe, Proliferative). Although this method is effective for detection, it does not provide a quantitative assessment of the risk of future disease progression and does not con-sider essential clinical biomarkers, including glycated hemoglobin (HbA1c)[2]. To address this critical gap, the present work introduces the Multimodal DR-HbA1c framework -a novel lightweight deep-learning architecture that, performs five-class DR severity classification via a DenseNet121 image backbone and generates patient-specific 6- and 12-month progression risk probabilities through a VTDR-inspired clinical calculator that integrates the predicted DR grade with a single HbA1c input - enabling personalized risk stratification from the same retinal image under varying glycemic control [3]. A key clinical contribution is demonstrating that patients with identical retinal fundus appearances and thus identical DR grades from any image-only sys-tem can receive materially different 12-month progression risk estimates and screening interval recommendations depending on their glycemic control, a distinction that is entirely invisible to convention-al image-based screening [4]. The proposed system employs late-feature fusion of a frozen DenseNet121 image branch with a compact HbA1c clinical branch, a VTDR-inspired dual-horizon risk calculator, and LIME-based visual explainability all integrated into an end-to-end Gradio prototype. As the APTOS 2019 dataset contains no paired HbA1c labels, clinical HbA1c values were generated via literature-derived class-conditional normal distributions and used for both training and sensitivity analysis. The remainder of this paper presents the architecture, experimental results on 1,096 unseen validation images, and a sensitivity analysis demonstrating the clinical value of glycemic context in personalized DR risk stratification.

2. METHODOLOGY

The proposed Multimodal DR-HbA1c framework was developed to perform five-class DR severity classification from retinal fundus images and generate personalized 6- and 12-month progression risk probabilities by integrating the predicted grade with a HbA1c-aware clinical risk calculator. The image branch uses a pretrained DenseNet121 backbone (frozen) followed by global average pooling and a dense layer to extract 256-dimensional features. The clinical branch processes a single HbA1c value through two dense layers ($32 \rightarrow 16$ features). Late fusion concatenates both branches into a 272-dimensional representation, followed by dense layers ($128 \rightarrow 64$) and a 5-class softmax output. Only 306,517 parameters (4.2 % of total) are trainable. The model was trained on 2,566 retinal fundus illustration from the APTOS 2019 vision loss identification dataset, partitioned 70/30 into training and validation sets (2,566 train / 1,096 validation). The training class distribution was: No_DR ($n=1,264$, 49.3%), Moderate ($n=700$, 27.3%), Mild ($n=259$, 10.1%), Proliferative ($n=207$, 8.1%), and Severe ($n=136$, 5.3%). As the APTOS 2019 dataset contains no HbA1c labels, training and validation HbA1c values were synthetically generated using literature-derived class-conditional normal (No_DR: $\mu=6.8\%$, $\sigma=1.2\%$; Mild: $\mu=7.8\%$, $\sigma=1.4\%$; Moderate:

$\mu=8.6\%$, $\sigma=1.5\%$; Severe: $\mu=9.8\%$, $\sigma=1.3\%$; Proliferative: $\mu=10.2\%$, $\sigma=1.4\%$), clipped to the clinical range [4.0%, 14.0%]. These parameters are derived from published HbA1c-DR correlations. Table 1 summarizes performance metrics across all models. The proposed multimodal model achieves the highest accuracy of 79.93 %, representing a +6.39 % improvement over the best single-image baseline (DenseNet121 at 73.54 %) and +4.84 % over the optimized ensemble. Macro F1-score is 0.5962 (driven by imbalance in Severe/Proliferative classes), macroaverage AUC is 0.9185, and micro-average AUC is 0.9589. Inference time is 27.6ms on Tesla T4 GPU.

Table 1. Efficiency analysis of baseline framework and Multimodal DR-HbA1c

Model	Accuracy (%)	Precision (Macro)	Recall (Macro)	F1-Score (Macro)	Inference Time (ms)
ResNet50	66.42	0.421	0.392	0.377	16.0
ResNet18	69.43	0.463	0.429	0.424	6.1
MobileNetV2	69.53	0.509	0.536	0.512	25.9
DenseNet121	73.54	0.552	0.565	0.549	27.5
Simple Ensemble	75.00	0.574	0.556	0.559	~75.5
Optimized Ensemble	75.09	0.573	0.574	0.567	~75.5
Multimodal DR-HbA1c Framework	79.93	0.637	0.583	0.596	27.6

3. RESULT AND DISCUSSIONS

The incorporation of a single clinical biomarker-glycated hemoglobin (HbA1c)-with retinal fundus images meaningfully improves both classification accuracy and clinical utility. The proposed multimodal DR-HbA1c framework achieved 79.93% validation accuracy on the held-out validation set (n=1,096) of the APTOS 2019 dataset, surpassing the image-only DenseNet121 baseline by 6.39 percentage points while training only 4.2% of the total parameters. This gain demonstrates that HbA1c supplies complementary information beyond retinal texture alone. The most important clinical finding is that the same retinal images are linked to significantly different risks of disease progression and screening recommendations, depending on how well blood sugar is controlled. For example, a Mild DR retina shows 12-month development risk of 16.2% (requiring 6-monthly screening) at HbA1c 10.5%, evaluated with 9.9% (annual screening) at HbA1c 6.5%. This impact is sustained by a analytically important spearman correlation between HbA1c and 12 month risk ($\rho = 0.3706$, $p = 0.0438$, $n=30$) and strong within grade pearson associations ($r > 0.97$, $p < 0.01$) for four of five DR classes. Even retinas graded No_DR switch from low risk (9.9%, annual screening) under stable handling to moderate risk (16.2%, 6 monthly screening) under poor control (HbA1c 10.5%), illustrating a clinically practical adaptive by metabolic situation. Collectively, these results demonstrate the framework's capacity to move DR screening from static image-based grading toward personalized, actionable risk stratification [7].

4. CONCLUSIONS

Current DR screening primarily depends on retinal fundus images for static severity grading, offering no quantitative estimate of disease progression risk and overlooking key metabolic biomarkers such as HbA1c. This limitation results in uniform screening strategies that fail

to reflect individual patient risk profiles. To address this gap, a lightweight multi-modal deep learning framework was developed by combining a frozen pre-trained DenseNet121 image group with a compact clinical integrating HbA1c data through late-fusion combining. The approach is moreover improved with a VTDR-inspired logistic risk calculator and LIME-based clarity for enhanced interpretability. The suggested framework attained a validation accuracy (79.93%) on a held-out subset ($n=1,096$) of the APTOS 2019 dataset, illustrating a 6.39% enhancement over an image-only baseline, while preserving efficiency with 306,517 trainable features and a 27.6ms inference time. The equal retinal images produce notably distinct 12-month growth risks and screening suggestions over differing HbA1c levels from 6.5% to 10.5%, sustained by a analytically notable Spearman correlation ($\rho = 0.3706$, $p = 0.0438$). However, class imbalance affected recall for Severe and Proliferative DR categories. Since real HbA1c labels were unavailable, synthetic values were used and risk estimates were generated using a transparent rule-based approach rather than longitudinal data-driven modeling.

5. REFERENCES

- [1] M. Kropp, O. Golubnitschaja, A. Mazurakova, L. Koklesova, N. Sargheini, T.K.S. Vo, E. de Clerck, J. Polivka Jr., P. Potuznik, J. Polivka, I. Stetkarova, P. Kubatka, G. Thumann, 'Diabetic retinopathy as the leading cause of blindness and early predictor of cascading complications—risks and mitigation', EPMA Journal, vol. 14, no. 1, pp. 21-42, 2023.
- [2] J.K. Goh, C.Y. Cheung, S.S. Sim, P.C. Tan, G.S. Tan, T.Y. Wong, 'Retinal imaging techniques for diabetic retinopathy screening', Journal of Diabetes Science and Technology, vol. 10, no. 2, pp. 282-294, 2016.
- [3] C.M.P. Jacoba, L.A. Celi, P.S. Silva, 'Biomarkers for progression in diabetic retinopathy: Expanding personalized medicine through integration of AI with electronic health records', Seminars in Ophthalmology, vol. 36, no. 4, pp. 250-257, 2021.
- [4] Z. Li, L. Wang, X. Wu, J. Jiang, W. Qiang, H. Xie, H. Zhou, S. Wu, Y. Shao, W. Chen, 'Artificial intelligence in ophthalmology: The path to the real-world clinic', Cell Reports Medicine, vol. 4, no. 7, 2023.
- [5] M.O. da Silva, A.E.C. do Carmo Chaves, G.C. Gobbato et al., 'Early choroidal and retinal changes detected by swept-source OCT in type 2 diabetes and their association with diabetic kidney disease: A longitudinal prospective study', BMC Ophthalmology, vol. 24, p. 85, 2024.
- [6] S.L. Krishna, I.J.S. Jeya, S.N. Deepa, 'Fuzzy-twin proximal SVM kernel-based deep learning neural network model for hyperspectral image classification', Neural Computing & Applications, vol. 34, pp. 19343-19376.
- [7] I.J. Selvakumari Jeya, S.N. Deepa, 'Lung cancer classification employing proposed real coded genetic algorithm based radial basis function neural network classifier', Computational and Mathematical Methods in Medicine, 2016, pp. 1-15.

Biographies



Ananya V. S. is an undergraduate student at the School of Computing Science, Engineering and Artificial Intelligence, VIT Bhopal University, Madhya Pradesh.



Jasmine Selvakumari Jeya I is currently a Senior Associate Professor with the School of Computing Science and Engineering, VIT Bhopal University, Madhya Pradesh.



Aldrin Joan Pandian W is an undergraduate student at the School of Computing Science, Engineering and Artificial Intelligence, VIT Bhopal University, Madhya Pradesh.



Kannaiya Raja N is currently a Senior Associate Professor with the School of Computing Science and Engineering, VIT Bhopal University, Madhya Pradesh.



Indirani M is currently an Associate Professor, Department of Computer Science and Business Systems, Vel Tech Multi Tech Dr. Rangarajan Dr. Sakunthala Engineering College, Avadi, Chennai, India.



Saravanan S is currently an Associate Professor, School of Computing, SRM Institute of Science and Technology, Tiruchirappalli, India.