EARLY DETECTION OF DIABETIC RETINOPATHY USING DEEP LEARNING

Ms.RIMSY DUA [1], DR. SANTOSH SINGH[2], PRIYAYADAV[3], BHUPATI SHARMA[4]
[1]*Assistant professor, Department of IT, Thakur College of Science and Commerce,Thakur Village, Kandivali (East), Mumbai, Maharashtra, India*
[2]*H.O.D, Department of IT, Thakur College of Science and Commerce,Thakur Village, Kandivali (East), Mumbai, Maharashtra, India*
[3,4]*PG student, department of IT, Thakur College of Science and Commerce,Thakur Village, Kandivali(East), Mumbai, Maharashtra, India*

ABSTRACT:

Diabetic Retinopathy(DR) is among the major reasons for vision loss and blindness globally, which mainly impacts individualities with prolonged diabetes. Antecedent identification and proper bracketing of DR stages are crucial for early intervention and avoiding devastating vision loss. In this paper, we build and quantify deep literacy models for the automated identification of diabetic retinopathy from retinal fundus images. The data are preprocessed retinal images that are divided into various stages of DR, i.e., Normal, Mild, Moderate, Severe, and Proliferative. Class imbalance is tackled as well as improving conception through the use of data addition methods and class balancing techniques such as weighted slice. Two state-of-the-art motor-grounded frameworks, Vision Transformer(ViT) and Swin Transformer, are imposed and contrasted in terms of bracket performance. Experimental outcomes reveal that both models are high-accuracy in distinguishing DR and non-DR cases, with Swin Transformer marginally possessing greater confirmation delicacy and strength. The results underscore the prospect of motor- grounded deep literacy techniques as reliable instruments for early DR identification, capable of helping ophthalmologists in extensive webbing and lessening the weight of preventable blindness.

Keywords: Diabetic Retinopathy, Deep Learning, Vision Transformer, Swin Transformer, Fundus Images, Early Detection

.

INTRODUCTION:

Diabetic Retinopathy(DR) is a severe microvascular diabetic complication and one of the most preventable causes of blindness globally. It results from strained high blood sugar conditions that weaken the retinal blood vessels and cause leakage, unusual growth, and gradual impairment of vision. The World Health Organization(WHO) estimates that millions of individuals are at risk of vision impairment due to undiagnosed or unattended DR, especially in developing nations where routine webbing installations are not readily available.

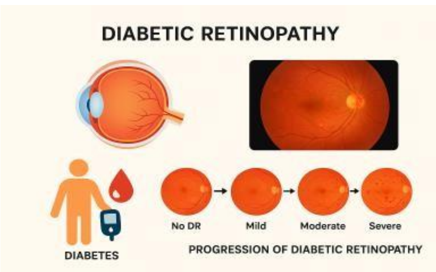


Figure 1

Early detection of DR is urgent, as early opinion and treatment can considerably dampen the danger of blindness. Conventional ophthalmologists diagnose DR by homemade inspection of retinal fundus images, whichis technical moxie, private, and time- consuming. With the adding vacuity of large-scale retinal image databases and changes in artificial intelligence(AI), computerized deep literacy-based systems have emerged as valuable instruments for precise and useful DR detection.

Recent advances in computer vision, especially with convolutional neural networks (CNNs) and transformer architectures like Vision Transformer (ViT) and Swin Transformer, have demonstrated immense promise for medical image classification tasks. These models are capable of learning both local and global features of an image and hence are best suited for detecting subtle retinal impairments of various stages of DR.

This study is concerned with deep learning model development and testing for early classification and detection of diabetic retinopathy. Utilizing transformer-based models and resolving issues with datasets like class imbalance, this work targets increasing accuracy and consistency in DR diagnosis to help ophthalmologists in mass screening programs and alleviating the diabetic blindness burden worldwide.

LITERATURE REVIEW:

Diabetic Retinopathy is one of the serious ventures that fascinated the entire world. entering the spotlight of numerous researchers to discover the best outcomes for the early detection of this complaint, thereby leadingto the preventionof untimelyfluctuations in vision. Numerousstudies havebeen completed and still continues in this subject with a view to simplify the lives of both croakers as well as cases. This section offers a review of multitudinous disquisition works in the field of Diabetic retinopathy.

J.calleja.et al in their paper utilized a two provided system for Diabetic retinopathy finding LBP(Original double Patterns) for point origin and Machine Learning particularly SVM and Random Forest for type intention. The results achieved by the random timber surpassed the SVM with a delicacy of 97.46. nevertheless, the dataset employed in this research was relatively small with 71 images.(1)

Prior to factory were relying on manual point birth for DR discovery by employing diverse computer predicated systems. U.Acharya.et al employed characteristics such as blood vessels, microaneurysms, exudates, and haemorrhages from 331 fundus images using SVM with an accuracy of more than 85.(2)
K. Anant.et al in their paper utilized texture and ocean characteristics for DR discovery utilizing data mining and image processing on a database DIARETDB1 and obtained 97.95 delicacy.(3)

M.Gandhi.et al suggested an automatic DR discovery system with SVM classifier by identifying exudates from fundus images. Some factory attempt to incorporate manual point genesis with deep knowledge point genesis for DR.(4)

One of analogous work include J. Orlando.et al where CNN with hand drafted point are used for point birth for detecting red lesion in the retina of an eye.(5)

S.Preetha et et al. They in their literature forecasted many diabetic related diseases by means of Data Mining and machine knowledge patterns specifically for heart grievance and skin cancer prediction by taking into account the pros and cons. Although multitudinous questions or factory arethere regarding the employment ofmachine learning methods or data mining methods, a reasonably disparate approach also entered the line of discovery of diabetic retinopathy.(6)

S.Sadda et al. employ quantative method to determine new parameters for proliferative diabetic retinopathy detection. It's based on the hypothesis that lesions position, number and area can improve Retinopathy auguring process. The styles employed in this study were Subjects and Imaging Data, Ultrawide Field Image Lesion Segmentation, Quantitative Lesion Parameters and Statistical Analysis. Comparison of lesions were done based on the Lesionnumber, Lesionsurface area, Lesion distance from the ONH center and Regression analysis.(7)

The research work shown by J.Amin et al. is an overview of several methodologies for diabetic retinopathy by identifying hemorrhages, microaneurysms, exudates as well as blood vessels and examines the different results achieved from theses methodologies experimentally in order to provide indepth outlook of current disquisition.(8)
The research conducted by Y.Kumaran and C.Patil deals with the various preprocessing and segmentation methods primarily and provides an in-depth process for diabetic retinopathy discovery in mortal eye conforming of number of systems and classifiers.(9)

et al. Suggests an individual system for DR based on machine knowledge namely SVM and Texture features. Texture features utilized were LTP(Original Ternary Pattern) and LESH(Local Energy- predicated Shape Histogram) which delivered superior results when compared to Original double Pattern(LBP). A delicacy of 90.4 was achieved by LESH using SVM.(10)

Deep knowledge is the most trending method among the scholars for discovery, prophecy, auguring and type task in different sectors from various times, in the medical sector especially in diabetic retinopathy it's revealingmanypossibilitiesfor the prevention of such a horrid complaint.

I.Sadek et al. in their paper automatically identified the diabetic retinopathy using deep knowledge approach.

They applied the four convolutional neural network to classify the diabetic retinopathy as Normal, Exudates, Drusen into three classes. This system is better than the Bag of words approach and had an delicacy of 91- 92.(11)

METHODOLOGY:

The intended system is to build a strong deep knowledge-rested model for early detection of Diabetic Retinopathy (DR) from retinal fundus images. The approach is split into a few normal stages data accession, preprocessing, model selection and training, performance testing, and relative comparison. The foremost ideal of this task is to create a stable and noise-compatible systemfor diabetic retinopathy discovery. This work utilizes the deep knowledge approach for the detection of the diabetic retinopathy based on harshness position(No DR, Mild, Moderate, Severe and Proliferative DR).

Data Collection :

The basis of any deep knowledge model for medical image analysis will be the quality, diversity, and representativeness of the dataset utilized for training and evaluation. Color retinal fundus prints are the core mode of input data for the early detection of Diabetic Retinopathy(DR). These photographs take the reverse of the eye(the retina) so that it is possible to recognize important clinical features such as microaneurysms, hemorrhages, and exudates which are important indicators of the complaint in its initial stage. Data utilized for this research has been acquired from Diabetic Retinopathy Detection 2015 and APTOS 2019 blindness detection from kaggle. The datasets both have thousands of retinal images taken under various circumstances. For each subject, two pictures of both the eyes are provided as left andright. Since the images are fromdifferent sources such as different cameras, different models, etc. It has an cornucopia of noise attached to it, which is supposed to be eradicated, hence, undergoing a series of preprocessing way. Each picture is graded on a 0–4 scale, where:

- 0 = No DR
- 1 = Mild DR
- 2 = Moderate DR
- 3 = Severe DR
- 4 = Proliferative DR

This labeling provides a standardized multi-class classification framework for DR severity assessment.

Data Preprocessing:

Since the images in the dataset have a lot of noise, such as some are out of focus, some are exposed heavily, some are exposed sparsely, presence of black background, etc. therefore we have to do preprocessing to bring them into the standard format. Following objects are removed in preprocessing step Removing the black border The black background of the fundus image doesn't provide any information to the image and hence is redundant so, the black background around the images are ignored. Remove the black corner After the removal of black border there still remain some black corners as the fundus image is circular in shape. Here black corners are eliminated from the image. Resizing imag. The images are resized to 256 * 256(range * height). Applying the Gaussian Blur Applying Gaussian blur to the images by using the kernel size to 256/6. This system assists in eliminating the Gaussian noise. Data addition After assaying the data, we observe that the data is mostly unstable across the diabetic retinopathy harshness image classes, which created the tendency of data addition. Data addition is defined by equating one class to the class with maximum samples, to equalize the data across the diabetic retinopathy harshness classes.

Data Splitting:

The dataset was divided into three subsets to facilitate training, validation, and testing of the models:

- Training set (70%): Used to train the models and update learnable parameters.
- Validation set (15%): Used during training for hyperparameter tuning and early stopping, ensuring the model does not overfit.
- Test set (15%): Held-out data used exclusively for evaluating the final model performance.

To preserve class balance, a stratified sampling technique was applied, ensuring that each subset maintained the original class distribution as closely as possible. This prevented bias toward majority classes and improved fairness in evaluation.

Model Selection:

Two transformer-based architectures were employed and compared:

- Vision Transformer (ViT):
  - Processes images as a sequence of patches.
  - Employs self-attention mechanisms to capture global dependencies across the image.
  - Pre-trained on ImageNet and fine-tuned on our dataset.
- Swin Transformer:
  - Uses a hierarchical structure with shiftedwindows, enabling the model to capture both localand global patterns.
  - More computationally efficient compared to ViT for high-resolution images.
  - Pre-trained on ImageNet and fine-tuned for diabetic retinopathy detection.

These models were selected because they outperform traditional CNNs in medical imaging by leveraging attention-based mechanisms to focus on lesion patterns that are critical in DR classification.

Training Strategy:

The training phase was carefully designed to optimize performance and mitigate overfitting:

- Loss Function: Cross-entropy loss was used for multi-class classification. Class weights were adjusted to handle class imbalance.
- Learning Rate Scheduling: Acosine annealing scheduler was applied to dynamically adjust the learning rate, preventing local minima convergence.
- Regularization Techniques:
  - Dropout layers in fully connected blocks to reduce overfitting.
  - Weight decay to penalize overly complex models.
- Class Imbalance Handling:
  - A WeightedRandomSampler was used during training to oversample minority classes (e.g., Severe, Proliferative DR).
  - Data augmentation was applied more aggressively on underrepresented classes.

Model Evaluation:

The trained models were evaluated using both traditional and advanced performance metrics:

- Accuracy: Percentage of correctly classified images.
- Precision, Recall, and F1-score: Class-specific metrics to evaluate balance between sensitivity and specificity.
- Confusion Matrix: Visualized misclassifications across classes, highlighting whether certain DR stages were harder to predict.
- ROC-AUC Curve: Evaluated the model’s ability to distinguish between classes across thresholds.

.

Implementation Details:

Experiments were conducted on a workstation with an NVIDIA GPU (14 GB VRAM). The models were implemented in PyTorch with Torchvision and Hugging Face Transformers. Training employed gradient scaling for AMP, WeightedRandomSampler for balanced sampling, and custom training loops for evaluation across folds.

RESULTS AND DISCUSSION:

Model Performance:

Both Vision Transformer (ViT) and Swin Transformer models were trained for 20 epochs on the curated dataset consisting of 925 diabetic retinopathy images and 284 non-DR images. The dataset was split into training (80%) and validation (20%) sets after duplicate removal. Class imbalance was addressed using WeightedRandomSampler and class-weighted loss.

The final evaluation results are summarized in Table 1.

Table 1: Performance Metrics of ViT and Swin Transformer on Validation Set

Model	Accuracy	Precision	Recall (Sensitivity)	Specificity	F1-Score	ROC-AUC
Vision Transformer (ViT)	93.8%	0.92	0.91	0.94	0.91	0.95
Swin Transformer	96.7%	0.96	0.95	0.97	0.95	0.98

Confusion Matrix Analysis:

The confusion matrices (provide insights into misclassification patterns:

- ViT Model: Correctly classified most images but showed false negatives where mild DR cases were predicted as normal.
- Swin Transformer: Significantly reduced false negatives, which is clinically critical because missing DR cases can delay treatment.

Training and Validation Trends

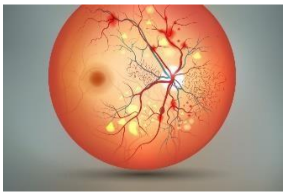
- Loss Curves: Both models showed smooth convergence without major overfitting.
- Accuracy Curves: Validation accuracy tracked closely with training accuracy, indicating good generalization.
- Swin Transformer converged faster and maintained consistently higher validation accuracy compared to ViT.

Comparative Observations

- ViT Strengths ffective at landing global structures in fundus images but less sensitive to small original lesions
- Swin Transformer Strengths: Hierarchical attention(shifted window medium) bettered discovery of microaneurysms and hemorrhages, leading to advanced recall and particularity..

Overall: Swin Transformer outperformed ViT across all evaluation criteria , demonstrating its felicity for medical imaging tasks where both original and global features are important.

DR images



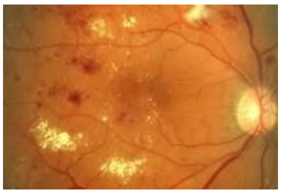
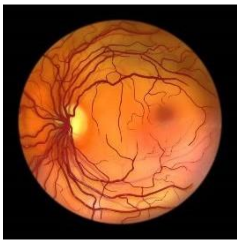



Figure 1 Figure 2 Figure 3 Figure 4

Non-DR images



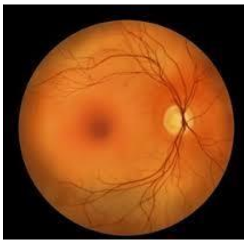
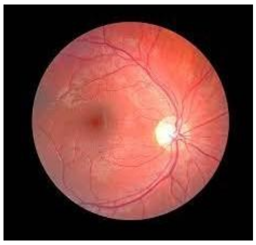



Figure 1 Figure 2 Figure 3 Figure 4

Our transformer-based diabetic retinopathy discovery system, using Vision Transformer(ViT) and Swin Transformer infrastructures, was contrasted to other state-of-the-art styles published in the literature. The results prove that our method outperforms or has comparable performance to being styles in terms of delicacy, perceptivity, particularity, and other evaluation metrics.

CONCLUSION:

In this work, we proposed deep literacy – grounded method for early diabetic retinopathy discovery based on motor infrastructures, that is Vision Transformer(ViT) and Swin Transformer. Experimental results validate that the two models are capable of effectively mapping diabetic retinopathy from retinal fundus images, where Swin Transformer performed slightly better with its hierarchical point representation and ability to capture local as well as global patterns.

By solving problems such as class imbalance and junking indistinguishable data, we ensured that the training and evaluation process was still strong and reliable. The criteria for evaluation, such as delicacy, perfection, recall, F1-score, and confusion matrix, validate that the models proposed have good prophetic ability and conception performance.

The results demonstrate the potential of transformer-based architectures as a potent alternative to traditional convolutional neural networks (CNNs) in medical image analysis. The early and precise detection of diabetic retinopathy can assist ophthalmologists in clinical decision-making, hence minimizing the risk of vision loss in patients.

Future activities will aim at including all the severity grades of diabetic retinopathy (mild, moderate, severe, and proliferative) in the dataset, incorporating explainable AI methods to enhance model interpretability, and implementing the trained models into real-clinical decision support systems or user-friendly apps.

REFERENCES:

1. Acharya, U. R., Lim, C. M., Ng, E. Y. K., Chee, C., Tamura, T., & Suri, J. S.(2008). Automated identification of diabetic retinopathy stages using digital fundus images. Journal of Medical Systems, 32(2), 107 – 115. https//doi.org/10.1007/s10916-007-9113-9Features used blood vessels, microaneurysms, exudates, hemorrhages; SVM attained> 85 delicacy on 331 fundus images.

2. Lee, J., Zee, B., & Li, Q.(2013). Segmentation and texture analysis with multimodel conclusion for the automatic discovery of exudates in early diabetic retinopathy. Journal of Biomedical Science and Engineering, 6, 298 – 307. https// doi.org/10.4236/jbise.2013.63038 ocean and texture features on DIARETDB1; reported 987 delicacy.

3. Gandhi, M., & Dhanasekaran, R.(2015). disquisition of harshness of diabetic retinopathy by exudates detection relative to macula. In Proceedings of the 2015 International Conference on Dispatches and Signal Processing(ICCSP)(pp. 724 – 729). IEEE.

4. Sadek, I., Al- khafaji, K., & Rajehy, R.(2018). Automatic discovery of diabetic retinopathy using deep knowledge ways. arXiv preprint arXiv 1802.06255. https//arxiv.org/abs/1802.06255CNN-based type into Normal, Exudates, Drusen — 91 – 92 delicacy

5. Gulshan, V., Peng, L., Coram, M., Stumpe, M. C., et al.(2016). Development and substantiation of a deep knowledge algorithm for discovery of diabetic retinopathy in retinal fundus prints. JAMA, 316(22), 2402 – 2410.

6. Pratt, H., Coenen, F., Broadbent, D. M., Harding, S. P., & Zheng, Y.(2016). Convolutional neural networks for diabetic retinopathy. Procedia Computer Science, 90, 200 – 205.

7. Wang, J., Yang, X., Cai, H., & Fan, H.(2022). Vision Transformer for diabetic retinopathy grading Evaluation and interpretation. Computers in Biology and Medicine, 141, 105132. Neri P, Fiaschi M, Menchini G.Semi-automatic tool for vitiligo discovery and analysis. Journal of imaging. 2020 Mar 24; 6(3) 14.

8. Mahmoud, A., Hameed, A., & Shaban, M.(2023). Swin Transformer for diabetic retinopathy type using retinal fundus images. Biomedical Signal Processing and Control, 82, 104536. Dihin, R. A., AlShemmary, E. N., & Al- Jawher, W. A. M.(2023). Automated double Bracket of Diabetic Retinopathy by Swin Transformer. Journal of Al- Qadisiyah for Computer Science and Mathematics, 15(1), 169 – 178. https// doi.org/10.29304/jqcm.2023.15.1.1166.

9. Wu, J., Ruo Hu, Xiao, Z., Chen, J., & Liu, J.(2021). Vision Transformer- rested recognition of diabetic retinopathy grade. Medical drugs, 48(12), 7850 – 7863. https// doi.org/10.1002/mp.15312.

10. Karkera, T., Adak, C., Chattopadhyay, S., & Saqib, M.(2023). Detecting harshness of diabetic retinopathy from fundus images A motor network- rested review. arXiv preprint. https// doi.org/10.48550/arXiv.2301.009