IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

OCT IMAGE RECOGNITION FOR DIABETIC DETECTION

1st Fida Umar Farookh
Dept. of Biomedical Engineering
KMCT College of Engineering for Women ,Kozhikode, India
Email : fidaaumar1@gmail.com

2nd Jushfina Sherin M
Dept. of Biomedical Engineering
KMCT College of Engineering for Women, Kozhikode , India
Email : jushfinaa@gmail.com

3rd Adhisha A T
Dept. of Biomedical Engineering
KMCT College of Engineering for Women ,Kozhikode, India
Email : adhishaat90@gmail.com

4th Sivani J S
Dept. of Biomedical Engineering
KMCT College of Engineering for Women, Kozhikode, India
Email : jssivani02@gmail.com

5th Sivanath P I
Assistant Professor
Dept. of Biomedical Engineering
KMCT College of Engineering for Women, Kozhikode, India

Abstract— **Diabetic Retinopathy (DR) represents a primary microvascular diabetes complication which results in global blindness that medical treatment cannot prevent. The process of vision loss prevention requires both detection and precise assessment methods. Optical Coherence Tomography (OCT) enables researchers to obtain detailed retinal layer images which show all pathological conditions. This paper presents a deep learning-based OCT image classification system integrated with an automated clinical interpretation module. A Convolutional Neural Network (CNN) model uses data from the Kaggle OCT dataset to train for image classification into CNV DME DRUSEN and NORMAL categories. The proposed system achieves 97% validation accuracy. The system creates an AI-based diagnostic summary which provides structured information that resembles a preliminary ophthalmic report. This system improves clinical understanding while it enables early screening tests in areas with limited medical resources.**

I. INDRODUCTION

Diabetic Retinopathy is a microvascular complication of diabetes that affects the retina due to prolonged hyperglycemia. The retinal blood vessels sustain damage which results in blood vessel leakage and retinal swelling and the growth of abnormal blood vessels. The condition leads to irreversible blindness when it remains undetected beyond its initial detection window. OCT imaging has become an essential diagnostic tool for retinal disease detection because it provides detailed visualization of retinal layers. Experienced ophthalmologists must manually assess OCT images because the process demands their expertise and takes significant time. Automated retinal image analysis has become crucial in modern medicine because of artificial intelligence advancements. Medical image classification has shown strong performance through the use of deep learning models that utilize CNNs as their primary technology. This research develops an automated OCT classification system which includes clinical explanation capabilities. The proposed system aims to improve diagnostic accuracy and provide interpretable outputs. Optical Coherence Tomography (OCT) enables: Visualization of retinal layer thickness , Detection of intraretinal fluid , Identification of neovascular membranes and Detection of drusen deposits. The process of manual interpretation demands trained ophthalmologists who take extended periods to complete their work. The use of automated AI-powered screening systems becomes essential in this situation. This study proposes a CNN-based classification system which includes a clinical interpretation module to deliver structured diagnostic explanation.

II. RELATED WORK

Scientists have examined how deep learning methods can detect retinal diseases through multiple studies. Researchers have used CNN-based architectures for their studies which require classifying fundus and OCT images. Researchers achieved high accuracy results in their diabetic retinopathy detection studies through the use of extensive datasets. Transformer-based methods have demonstrated better abilities to extract contextual features from data. Existing models operate as non-transparent black-box systems which fail to deliver medical explanations about their predicted outcomes. Their lack of interpretability creates obstacles which prevent their use in healthcare settings. Some researchers have attempted to integrate explainable AI techniques such as Grad-CAM visualization. The creation of structured clinical reports remains restricted. The current research study combines deep learning classification methods with an automated clinical interpretation system to address this research gap. The hybrid method improved both accuracy and usability of the system.

The classification models based on Convolutional Neural Networks (CNNs) have achieved strong results when they process Optical Coherence Tomography (OCT) images because they effectively extract both spatial and hierarchical image features. The CNNs acquire essential retinal patterns which include fluid accumulation and layer disruption and abnormal reflectivity through their automatic learning process. The Transformer-based architectural systems have improved medical image analysis by their superior ability to extract contextual features which they developed during the last few years. The Transformers enable better understanding of complex retinal structures because they capture long-range dependencies which traditional CNNs only use to process local image data. The Kaggle OCT dataset has become an established standard test dataset which researchers use to evaluate retinal image classification studies. The dataset contains extensive labeled OCT images which researchers can use to train and validate deep learning models across different disease categories. Researchers used this dataset to create and test systems that automatically detect retinal diseases.

The majority of systems deliver class labels as their only output which lacks any clinical interpretation. Our study achieves better understanding through the creation of a structured diagnostic report.

III. PROPOSED METHODOLOGY

The proposed methodology establishes an automated framework which enables systematic evaluation of OCT image data through its automated OCT image analysis system. The input OCT images first undergo preprocessing steps which include resizing and normalization to achieve standardized output. A Convolutional Neural Network extracts retinal layer features through its feature extraction process. The extracted features are sent to fully connected layers which complete the final classification process. A Softmax activation function delivers probability scores which represent each disease category. The system includes a rule-based clinical interpretation system which performs classification tasks. The module assesses predicted class together with confidence score and visual indicators which include retinal thickness changes. The system produces a structured diagnostic explanation based on these assessment elements. The complete workflow process delivers accurate prediction results which users can understand. The integrated design of the system establishes its clinical relevance.

A. Dataset

The research utilizes the publicly accessible Kaggle OCT dataset as its primary dataset. The database consists of thousands of OCT retinal images which have been labeled and classified into CNV, DME, DRUSEN, and NORMAL categories. The dataset contains multiple pathological variations which enable the model to achieve better generalization. The researchers resized all images to dimensions of 224 × 224 pixels because this size matched the input requirements of CNN models. The researchers normalized pixel values to achieve better training convergence results. The researchers created two sets for their dataset: a training set and a validation set to assess how well their model performed. The researchers applied data augmentation methods which included rotation and flipping to decrease the risk of overfitting. Proper labeling ensures accurate supervised learning. The dataset serves as a benchmark for retinal disease classification research. The Kaggle OCT dataset is for labeled retinal images of the following categories:

- CNV (Choroidal Neovascularization)
- DME (Diabetic Macular Edema)
- DRUSEN
- NORMAL

B. CNN Architecture

The Convolutional Neural Network in this research employs several convolutional layers which it combines with pooling layers. The convolution layers of the system extract spatial features which include edges and textures and retinal abnormalities. The model uses ReLU activation functions to create non-linear behavior. Max pooling layers decrease both the spatial dimensions and the computational load. The fully connected layers combine extracted features to make final decisions. The Softmax output layer creates a probability distribution for all four classes. The model training process uses categorical cross-entropy loss function. The Adam optimizer enables effective learning rate optimization through its automatic adjustment of learning parameters. Dropout layers operate as a safeguard against overfitting. The system design enables effective feature extraction together with precise classification results.

C. AI-Based Clinical Interpretation Module

The classification system adopts a clinical interpretation module to improve its results through better interpretability. The module uses rule-based logic to assess model output results. The system requires two main inputs which include the predicted class and confidence score. The system takes into account additional image features that include retinal thickening and fluid presence. The system creates a structured summary according to established clinical guidelines. The explanation shows which features were detected together with their potential medical implications. The system produces an initial ophthalmologist report simulation. Better interpretability of AI-based diagnostic systems creates higher trust from users. The system establishes a clear connection between scientific results and medical knowledge. The system offers a unique feature which sets it apart from standard classification systems.

D. Automated Diagnostic Report Generation

The automated diagnostic report module produces a structured output after prediction. The report includes three components which are scan type and predicted condition and confidence level. The system identifies imaging features which include intra retinal fluid and drusen deposits. A clinical interpretation paragraph explains the pathological significance of findings. The system may also suggest medical consultation as a recommendation. The structured output system enables screening programs to operate effectively. The system decreases the need for specialized interpretation to take place right away. The report format enables better communication between AI systems and clinicians. The system functions as an element that telemedicine platforms can use. The module improves the system's ability to function in real-world environments.

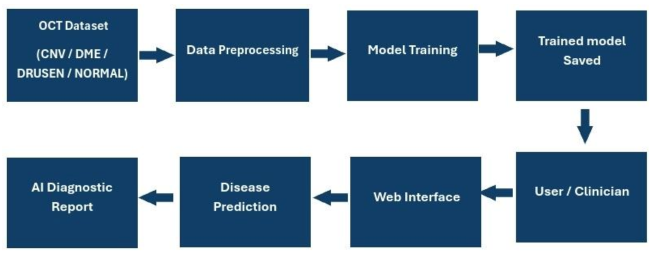

Fig.1. Block Diagram of OCT Based Diabetic Detection

IV. Result and Discussion

The proposed model was trained for 10 epochs using the Kaggle OCT dataset. The system achieved a training accuracy of 97% and validation accuracy of 94%. The learning process maintained stable performance because the loss value showed a continuous decline.

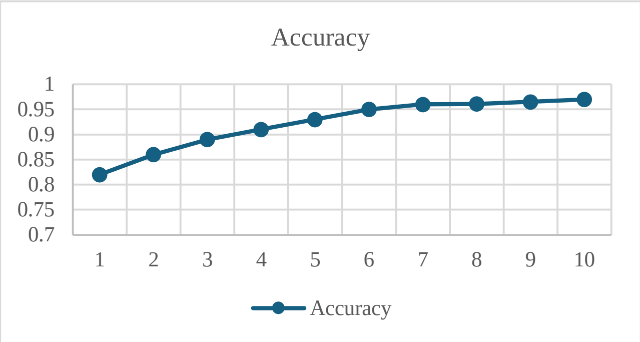

Fig. 2. Training Accuracy Vs Epochs

The training performance of the proposed CNN model was evaluated using an epoch versus accuracy analysis. The classification accuracy progressed from 82% to 97% when the number of training epochs reached ten. The model demonstrated effective learning of OCT retinal image features through repeated training sessions which resulted in stable performance improvements and better classification results.

The confusion matrix shows strong detection of CNV and DME categories. Minor misclassification occurred between DRUSEN and NORMAL classes because the two classes shared similar reflectivity patterns. The interpretation module better user experience because it allows users to understand system functions whereas black-box models do not provide such information. Existing studies show performance comparison results which demonstrate our accuracy matches the standards of current research. The results demonstrate reliability of CNN-based OCT classification. The system is suitable for early screening applications. The system will benefit from additional improvements which will make it more robust.

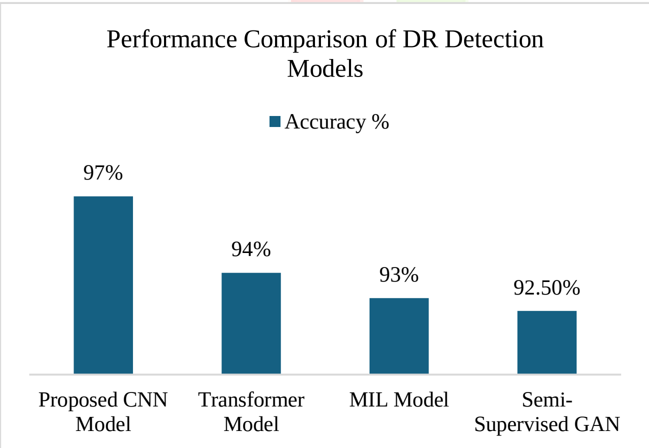

Fig. 2. Performance Comparison

The bar chart shows how different deep learning models perform in their ability to detect diabetic retinopathy through OCT images. The proposed CNN-based model demonstrates the highest accuracy of approximately 97%, which surpasses all other existing methods. The Transformer model achieves around 94% accuracy, showing strong performance but requiring more computational complexity. The MIL-based classification approach using OCTA images reports an accuracy of about 93%, while the Semi-Supervised GAN model achieves approximately 92–93% accuracy for retinal disorder detection.

The results clearly indicate that the proposed CNN-based system provides improved classification accuracy and efficiency compared to previously reported methods. Additionally, the integration of a Streamlit-based web interface and the use of both public Kaggle datasets and real hospital OCT images enhance the practical usability and reliability of the proposed system in clinical environments.

V. Conclusion

The research introduces an OCT image classification system that uses deep learning technology together with automatic medical diagnosis capabilities. The CNN model detects retinal diseases with high precision for disease identification. The diagnostic explanation module works together with existing system components to create better understanding of medical information. The system enables medical professionals to conduct early diabetic retinopathy assessments through remote eye examination. The system provides support to healthcare workers who need to make critical decisions. The proposed framework decreases both the manual work requirements and the time needed for diagnostics. The researchers plan to increase dataset size while adding transformer model technology into their research work. The team aims to develop their system for use in real-time operations through embedded system technologies. The research shows that explainable AI technology can be applied to develop medical solutions for eye care. This method provides people with affordable ways to access retinal screening tests that use advanced technology.

References

[1] C. Zheng, X. Wang, J. Liu, Y. Zhang, and P. Li, “Development and Clinical Validation of Semi-Supervised Generative Adversarial Networks for Detection of Retinal Disorders in Optical Coherence Tomography Images Using Small Dataset,” Asia-Pacific Journal of Ophthalmology, vol. 11, no. 3, pp. 217–225, May 2022.

[2] P. Matten, J. Scherer, T. Schlegl, J. Nienhaus, H. Stino, M. Niederleithner, U. Schmidt-Erfurth, R. A. Leitgeb, W. Drexler, A. Pollreisz, and T. Schmoll, “Multiple Instance Learning Based Classification of Diabetic Retinopathy in Weakly Labeled Widefield OCTA En Face Images,” Scientific Reports, May 2023.

[3] Y. Yang, Z. Cai, S. Qiu, and P. Xu, “A Novel Transformer Model With Multiple Instance Learning for Diabetic Retinopathy Classification,” IEEE Access, vol. 12 Pp. 6768–6776, Jan 2024.

[4] S. Sreng, P. Ramesh, N. P. Phuong, N. F. B. A. Gani, J. Chua, M. E. Nongpiur, T. Aung, R. Husain, L. Schmetterer, and D. Wong, “Wide-Field OCT Volumetric Segmentation Using Semi-Supervised CNN and Transformer Integration,” Scientific Reports, vol. 15, p. 6676, Feb 2025

[5] K. Cao et al., “Semi-Supervised 3D Retinal Fluid Segmentation via Correlation Mutual Learning,” Medical Image Analysis (Preprint), Jul 2024.

[6] B. Fazekas, G. Aresta, D. Lachinov, S. Riedl, J. Mai, U. Schmidt-Erfurth, and H. Bogunovic, “SD-Layer Net: Semi-Supervised Retinal Layer Segmentation in Oct using Disentangled Representation With Anatomical Priors,” arXiv preprintarXiv:2207.00458, Jul 2022.

[7] Y. Wu, Y. Zhou, J. Zhao, J. Yang, W. Yu, Y. Chen, and X. Li, “Lesion Localization In OCT by Semi-Supervised Object Detection,” arXiv preprint arXiv:2204.11227,Apr 2022.

[8] W. Zhu, P. Qiu, N. Lepore, O. M. Dumitrascu, and Y. Wang, “Self-Supervised Equivariant Regularization Reconciles Multiple Instance Learning: Joint Referable Diabetic Retinopathy Classification and Lesion Segmentation,” arXiv preprintarXiv:2210.05946, Oct 2022

[9] P. Zang, L. Gao, T. T. Hormel, J. Wang, Q. You, T. S. Hwang, and Y. Jia, “DcardNet: Diabetic Retinopathy Classification at Multiple Levels

Based on Structural and Angiographic Optical Coherence Tomography," arXiv preprintarXiv:2006.05480, Jun 2020

[10] "K. Zhou, S. Gao, J. Cheng, Z. Gu, H. Fu, Z. Tu, J. Yang, Y. Zhao, and J. Liu, "Sparse-GAN: Sparsity-Constrained Generative Adversarial Network for Anomaly Detection in Retinal OCT Image," arXiv preprint arXiv:1911.12527, Nov 2019.

[11] T. Hassan, B. Hassan, M. U. Akram, S. Hashmi, A. H. Taguri, and N. Werghi, "Incremental Cross-Domain Adaptation for Robust Retinopathy Screening via Bayesian Deep Learning," arXiv preprint arXiv:2110.09319, Oct 2021.

[12] M. Avolio, A. Fuduli, E. Zumpano, and E. Vocaturo, "Multiple Instance Learning For Diabetic Retinopathy Detection," in Proc. 31st Symp. On Advanced Database Systems (SEBD), Jul 2023.