IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

AN EXPLAINABLE VISION TRANSFORMER-BASED FRAMEWORK FOR RETINAL DISEASES CLASSIFICATION

[1]Keerthi Guttikonda, [2]Siri Chandana Rachamsetti, [3]Srinu Vakada, [4]Krishna Sai Seera, [5]Gayathri Devi Vaddi

[1]Associate Professor, [2]Student, [3]Student, [4]Student, [5]Student

[1]Department of Computer Science and Engineering,

[1]Seshadri Rao Gudlavalleru Engineering College, Gudlavalleru, India

Abstract: Early and effective identification of retinal pathologies plays a very crucial role in the prevention of irreversible vision loss and facilitating the process of large-scale retinal screenings. This study focuses on the implementation of an explainable and Vision Transformer-based multi-class retinal disease classifiers using color fundus photographs, focusing on the proposed work around the Swin Transformer-Base architecture. A selectively compiled dataset of 4,217 retinal fundus photographs was assembled by combining various publicly accessible sources, followed by the classification of the compiled set into four retinal disease classes: Normal, Diabetic Retinopathy (DR), Age-related Macular Degeneration (AMD), and Branch Retinal Vein Occlusion (BRVO) diseased retinas. The previously mentioned Swin Transformer-Base network, pretrained on large-scale image datasets, was further adapted by the use of transfer learning to achieve more accurate results through standardized preprocessing and optimizing settings during the training process to avoid any potential bias in the derived results. From the experimental analysis, the proposed work showcases the potential of the proposed Swin Transformer-Base architecture in successfully maximizing the classification accuracy to 99.54%, and to further provide more robustness and explainability to the proposed results, the implementation of Gradient-weighted Class Activation Mapping (Grad-CAM) visualization techniques has been accomplished, focusing on the identification of critical retinal areas in retinal pathologies, thereby facilitating more enhanced clinical usability. Furthermore, to provide more utility value to the proposed clinical approach, a lightweight and easy-to-use diagnostic tool has been provided to upload the images in real time and to access the results in a more facile graphical format.

Index Terms - Retinal disease classification, Vision Transformer, Fundus imaging, Computer-aided diagnosis, Explainable AI, Telemedicine.

I. INTRODUCTION

DR, AMD, BRVO, etc. are some of the most common retinal diseases that lead to blindness. Early diagnosis is a crucial step in the sense that it can prevent blindness. Still, the availability of highly specialized ophthalmologists is a bottleneck for the successful execution of broad screenings. Retinal fundus scanning has been the most commonly practiced approach for the past decades for computer-aided diagnosis. Early approaches to retinal scanning involved the design of manual features focused on traditional scanning methods. These models led to limitations in that the methods might not perform well because of variations in lighting and differences in the diseases' characteristics. The emergence of deep learning models led way to remarkable improvements in diagnosing retinal diseases. Deep learning models with CNN played a remarkable role in the diagnosis and staging of AMD and DR from retinal scanning with remarkable results and have been recognized to perform at a clinical level in broad studies and applications. Deep learning methods have also enabled low-cost screening using smartphone-based fundus photography, improving accessibility in resource-constrained settings [4]. Further studies demonstrated the strength of deep learning models in predicting disease progression and identifying multiple retinal disorders across diverse datasets [5-8]. Despite their success, CNNs primarily capture local spatial patterns and may fail to capture long-range and global dependencies present in high-resolution fundus images. ViTs have recently emerged as an effective alternative by utilizing self-attention mechanisms to model global relationships within images. ViT-based architectures have shown effective results in fundus image classification, retinopathy diagnosis, and retinal disease detection, outperforming conventional CNNs in several scenarios [9-13]. Hybrid CNN–ViT models further enhance representation learning by combining local texture extraction with global context modeling. In addition to accuracy, interpretability has become a critical requirement for clinical adoption of automated diagnostic systems. Explainable transformer-based frameworks have been proposed to visualize attention regions corresponding to disease-related structures, improving transparency and clinician trust [14]. Recent research has also explored uncertainty estimation and privacy-aware learning strategies to enhance robustness and real-world applicability of retinal AI systems [15-16]. Motivated by these advances, this work proposes an interpretable vision transformer–based framework for retinal disease detection from fundus images, focusing on normal, AMD, DR, and BRVO categories. The framework aims to provide accurate predictions along with meaningful visual explanations to support clinical decision-making.

Main Contributions of this Study

The contributions of this study are as follows:

- An interpretable vision transformer–based framework is proposed for multi-class retinal disease classification using fundus images.
- An end-to-end pipeline is developed for preprocessing, transformer-based feature learning, and disease prediction.
- Explainability mechanisms are integrated to generate attention-based visualizations highlighting disease-relevant regions.
- Extensive experiments demonstrated the effectiveness of the framework across different retinal disease classes.

The other sections of the paper are organized as follows. Section 2 gives related work on retinal disease detection. Section 3 describes the proposed methodology in detail. Section 4 presents experimental setup. Section 5 focuses on results and discussion. Section 6 concludes the paper and outlines future research directions.

II. LITERATURE REVIEW

Initially, the earlier work using retinal disease analysis was based on traditional image processing techniques without the need for deep learning models. Kanagasingam et al. [1] discussed the classical methods of analyzing age-related macular degeneration (AMD) lesions, where the work highlighted the issues of sensitivity to variations in imagery. The application of deep CNN models opened up new avenues in automated retinal disease detection. Burlina et al. [2] demonstrated the performance of CNN models to achieve expert-level AMD grading on fundus imagery. De Fauw et al. [3] further validated the efficiency of deep models to detect various diseases of the retina to refer patients. Rajalakshmi et al. [4] validated the efficiency of using AI to determine diabetic retinopathy detection using fundus imagery via smartphones. Yim et al. [5] further applied deep models to estimate the progression of wet AMD. Other CNN models successfully detected diabetic retinopathy, as well as various retinal diseases [6-8]. However, the limitation of using CNN models to estimate long-range dependencies due to the localized field of view was encountered. Vision Transformers (ViTs) further overcame the drawback of localized field of view in CNN models. Yu et al. [9] proposed the application of multiple instance learning techniques on vision transformers to classify fundus imagery. Jiang et al. [10] proposed the use of vision transformers to develop automated retinopathy diagnosis. An automated vision transformer framework further demonstrated the detection of subtle diseases of the retina like exudative macular neovascularization [11]. Other models further proposed the application of CNN-ViT models to improve the detection efficiency of fundus imagery [12, 13]. Interpretability of the models was further considered important in the application of AI models to develop automated retinal systems. Various vision transformers further proposed techniques to determine the visualization of the attention map within the automated models specific to diseases [14]. Various recent studies further applied the estimation of deep models to determine privacy issues of the models to achieve efficiency [15, 16]. Apart from the studies specific to the field of retinal diseases, the developments in machine learning models further proposed the usage of various models to achieve efficiency in the application of models to develop automated health prediction [17-20]. While these studies may not be specifically related to retinal image analysis, these works serve as the motivation to combine the principles of explainability and optimization to the retinal disease detection systems. From the existing studies, the applicability of deep learning and vision transformer models in retinal disease detection is established. The challenge of optimizing the trade-off between detection accuracy and interpretability has been the motivation for the proposed work.

III. METHODOLOGY

This section describes the proposed interpretable vision transformer-based framework for multi-class retinal disease detection. The entire system workflow is illustrated in figure 1. The framework consists three main phases: dataset preprocessing, swin transformer model training on the preprocessed dataset, and explainability through gradient-based visual interpretation.

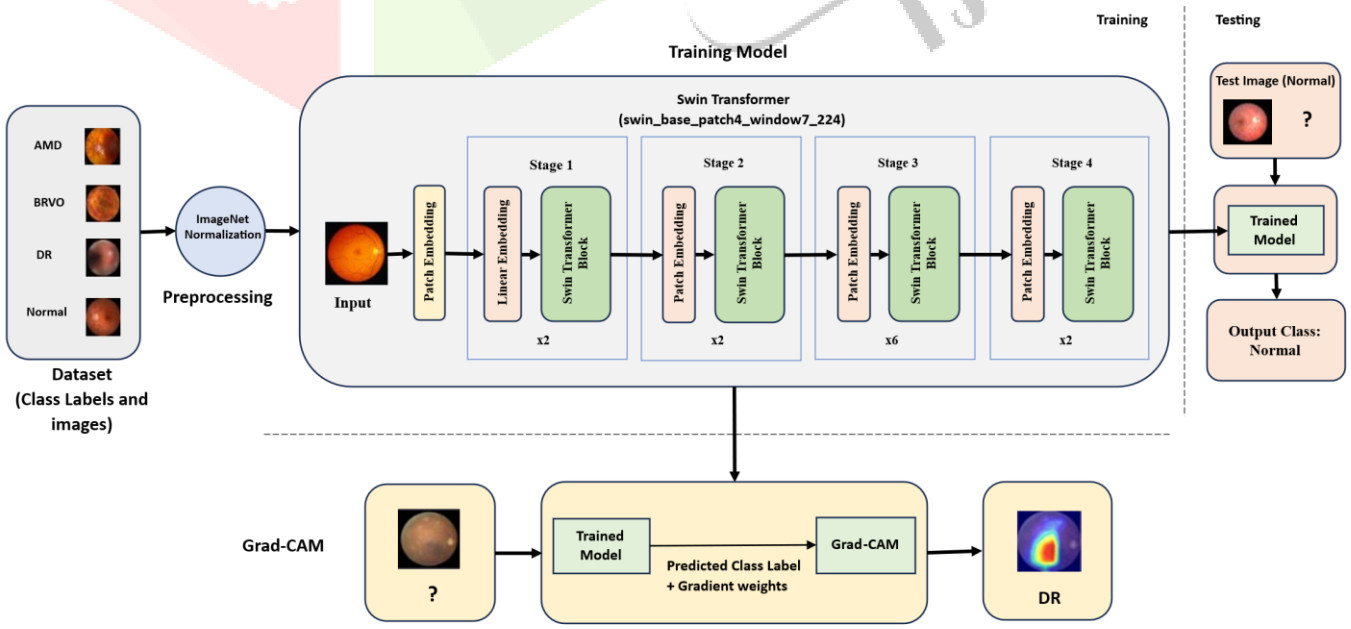

Fig. 1. Overview of framework for Retinal diseases classification.

A. Data Preprocessing

The proposed framework would focus on multi-class retinal disease classifications for colored fundus images. The dataset would include retinal images divided into four clinical retina disease classes, which include Normal, DR, AMD and BRVO. These would include normal as well as pathologic changes in the retina, which could possibly manifest during ophthalmic screenings. All fundus images would have been resized into 224×224 pixels so as to align with the requirements of the transformer-based architecture. Since the fact is noted that fundus images could possibly vary in illumination as well as conditions during which they are acquired, there would also be a standardized image processing routine involved in this framework in order to reduce inter-image variations as much as possible, thereby focusing on clinically important features in the fundus images, which include blood vessels, pathologic regions, or macula regions. This would require image statistics following ImageNet standards so as to match the image distributions to pre-train standards, which would result in extremely effective transfer learning, thereby ensuring faster convergence during training. There would not be harsh manipulations in geometries or intensities during image processing, which could easily damage important patterns in the fundus image.

B. Swin Transformer Architecture

In this work, Swin Transformer-Base(swin_base_patch4_window7_224) architecture is used as a prominent backbone for analyzing the retinal fundus images owing to its inherent superior capacity for representing complex disease manifestations present in the retinal images, along with its computational efficiency. The proposed architecture is highly suitable for analyzing retinal images, as it is capable of capturing both detailed pathological information as well as global knowledge about the image, which is paramount for accurate multi-class classification of various retinal disorders. Unlike standard Vision Transformers, where global self-attention is employed, a window-based self-attention mechanism applied based on a hierarchically stacked architecture is adopted in Swin Transformers, thereby enabling efficient representation of local details, particularly in this work, for retinal images, besides progressively gathering global information about the image. Specifically, for a 224 × 224 input retinal image, it is divided into independent patches of 4 × 4 spatial dimensions, and then these patches are linearly embedded into a embedding space, where all these patches are represented as separate tokens. The entire set of tokens undergo a process involving a series of hierarchically stacked transformer layers, as illustrated in figure 2.

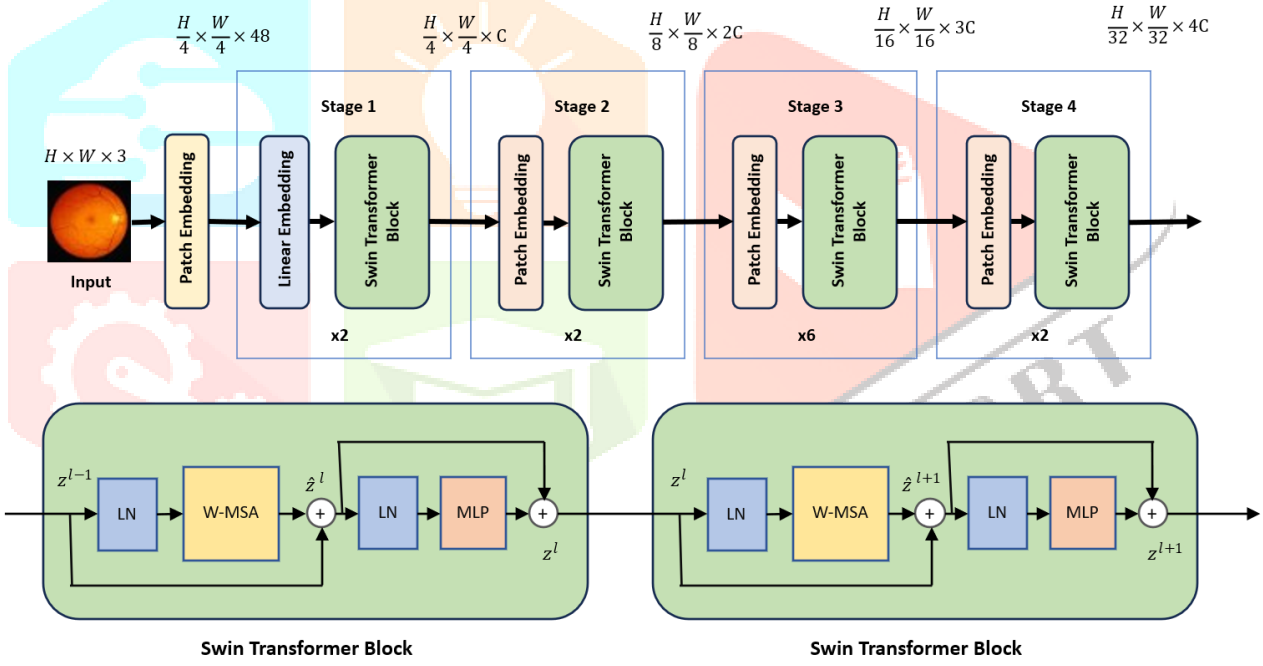

Fig. 2. Architecture of the Swin Transformer-Base model and the internal architecture design of the Swin Transformer blocks.

From Fig. 2, the Swin Transformer-Base model has four hierarchical levels, each of which is made up of several Swin Transformer blocks and patch merging layers for basic units.

- Stage 1: Patch embedding followed by Swin Transformer blocks (×2): This stage begins with patch embeddings, followed by two blocks of
- Stage 2: Merging of Patches, Swin Transformer blocks (Pyramid Vision Transformer, PVT, -).
- Stage 3: More powerful Swin Transformer blocks to capture challenging disease-related information.
- Stage 4: Final Patch Merging and Swin Transformer Layers (×2) to obtain the semantic representation.

The patch merging layers successively decrease the spatial resolution but increase the dimensionality of the features, resulting in a hierarchical feature pyramid to effectively learn representations at multiple scales. Inside swin transformer blocks, windowed multi-head self-attention (W-MSA) operations are alternated along with shifted window self-attention (SW-MSA) which is shown in Figure 2. In the shifted window approach, neighboring windows can communicate with each other to mitigate the locality issue in windowed attention mechanisms while maintaining efficiency in computations. Finally, the last feature representation learned by the Swin Transformer-Base architecture will be put to use for the classification of the retinal disease through the classification head.

C. Equations

The learned feature representations from the final stage of Swin Transformer-Base are passed to a multi-class classification head. Inside, the final-stage hierarchical feature representations produced via shifted window self-attention are aggregated and fed into a linear classification head with four output neurons representing Normal, DR, AMD, and BRVO. Following that, a softmax activation layer is added to get a normalized probability distribution over the four classes, and a class predicted by its maximum probability value is chosen. Since this network predicts a mutually exclusive single class, ambiguity of multi-class to multi-label does not exist. The neural network is trained with an optimal loss function, which is the categorical cross-entropy loss function, to ensure incorrect Class prediction is discouraged in the training process. Transfer learning is also employed in this research, where the initial weights for the Swin Transformer-Base model are set as pre-trained weights. This allows the model to harness very generalized features in images, alongside adapting to specialized features in the images depending on the specific domain. The training process is carried out in a supervised learning framework, where pre-processed fundus images as inputs and their respective class labels as outputs are employed in iteratively adjusting the necessary parameters in the backpropagation process. Optimization and regularization methods are applied in the training process to ensure convergence and avoid overfitting. During testing, the same pre-processing techniques as in training are employed in ensuring consistency in the testing process.

D. Explainability

To ensure transparency and clinical interpretability of the proposed retinal disease classification system, a class-contrastive Grad-CAM–based explainability framework is integrated with the Swin Transformer architecture. The explainability pipeline operates at two complementary levels. First, class-wise Grad-CAM visualizations are generated for all disease categories, independent of the final prediction. Second, for the predicted class, Grad-CAM heatmaps are overlaid on the original fundus image to provide intuitive visual explanations. In addition, high-activation regions are thresholded to derive bounding boxes, highlighting the most salient pathological areas that influenced the model's decision. This dual visualization strategy supports both global interpretability and localized lesion-level analysis.

Gradient-weighted Class Activation Mapping (Grad-CAM) is employed to explain the decision-making process of the Swin Transformer model by identifying image regions that contribute most to a specific class prediction. Unlike CNNs, the Swin Transformer processes images through hierarchical self-attention windows. To adapt Grad-CAM for this architecture, feature activations and gradients are extracted from the final normalization layer, which contains high-level semantic information relevant to disease discrimination. During inference, a retinal fundus image is first resized and normalized using ImageNet statistics and then passed through the trained Swin Transformer. For a selected class the corresponding logit score is backpropagated to compute the gradients with respect to the activations of the chosen feature layer. These gradients indicate how sensitive the class score is to each feature channel. The weighted combination of feature maps is then passed through a ReLU operation to retain only positive contributions, producing a class-specific activation map. The resulting Grad-CAM heatmap is normalized to the range [0,1] and resized to match the original image resolution. As shown in figure 3, higher intensity values (red/yellow regions) indicate areas that strongly influence the model's prediction, while lower intensity values (blue regions) represent less relevant regions. In addition to the predicted class, Grad-CAM maps are generated for all disease classes, enabling class-contrastive analysis that reveals how the model attends to different retinal structures for different pathologies.

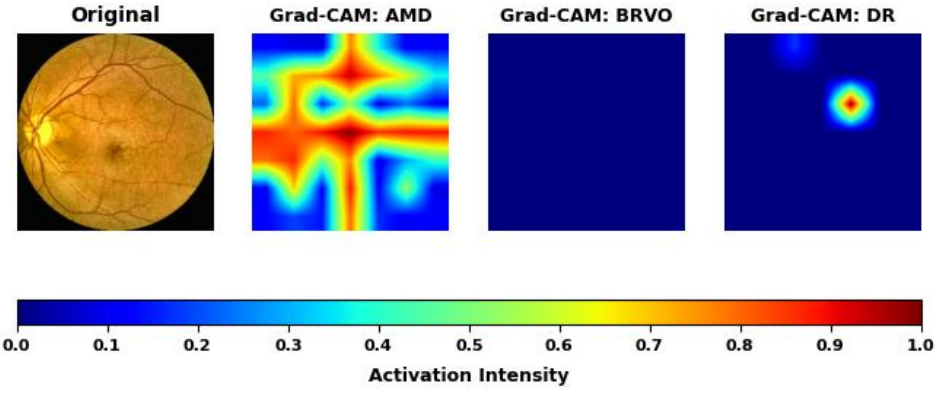

Fig. 3. Retinal fundus image of AMD along with Grad-CAM activation maps for AMD, BRVO, and DR classes generated using the Swin Transformer model. Warmer colors indicate higher activation intensity.

To further enhance interpretability, the Grad-CAM heatmap of the predicted class is overlaid onto the original retinal fundus image. The normalized activation map is converted into a color heatmap and blended with the original image using alpha compositing. This CAM overlay provides an intuitive visual explanation by directly highlighting disease-relevant regions—such as drusen deposits, hemorrhages, microaneurysms, or abnormal vasculature—on the anatomical background of the retina. For lesion-level localization, a bounding box extraction step is applied to the CAM. The heatmap is thresholded to retain only high-activation regions that exceed a predefined confidence level, ensuring that only the most salient areas are considered. Connected component analysis is then performed on the binary activation mask to identify contiguous regions. Small regions below a minimum area threshold are discarded to reduce noise. For each remaining region, a bounding box is computed and drawn on the original image, providing a coarse yet effective localization of pathological areas influencing the prediction. This combined visualization strategy—original image, CAM overlay, and CAM-based bounding boxes—offers both global and localized explanations. The Grad-CAM overlay and CAM bounding boxes are shown in figure 4. It helps in bridging the gap between the predictions of models and clinical interpretability.

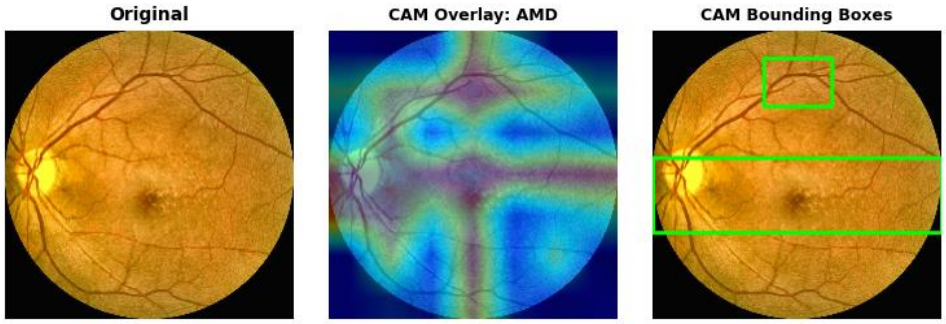

Fig. 4. Retinal fundus image of AMD along with Grad-CAM overlay and CAM bounding box.

IV. EXPERIMENTAL SETUP

The proposed retinal disease detection framework is evaluated using a dataset of fundus images categorized into four classes: Normal, DR, AMD, and BRVO. The fundus images dataset is split into training and testing folders with class-wise separation to avoid data leakage. All experiments are implemented using the PyTorch deep learning framework. Transfer learning is employed by initializing vision transformer models with pretrained weights. Input training retinal images are resized into 224 × 224 pixels and also normalized using standard ImageNet values of mean and standard deviation, following the preprocessing strategy described in Section 3.1. Training and inference are conducted in a GPU-enabled environment to improve computational efficiency. The model is trained under supervised manner using a categorical cross-entropy loss function. An adaptive gradient-based optimizer is used for parameter updates. Training is performed for a fixed number of epochs with a constant batch size across all experiments. Learning rate scheduling and regularization strategies are applied to ensure stable convergence. When evaluating multiple lightweight transformer models, identical training configurations and hyperparameters are maintained to ensure consistency and comparability. Performance evaluation is carried out using classification metrics, such as accuracy, precision, recall, and f1-score, computed on the test images. Confusion matrices are also analysed to examine class-wise prediction behaviour. All quantitative results, comparisons among evaluated models, and detailed performance discussions are presented in the Results and Discussion section. To ensure reproducibility, the same preprocessing pipeline, training strategy, and evaluation protocol are applied throughout all experiments. Model evaluation is strictly performed on unseen test data without any further fine-tuning, providing an unbiased assessment of generalization performance.

V. RESULTS AND DISCUSSION

This section gives the training and testing results of the proposed multi-class retinal disease classification model based on Swin Transformer-Base on a given dataset. The dataset has four classes: Normal, AMD, DR, and BRVO. The experiments were performed with fixed values of data split, resolution, batch size, and training epochs for valid comparisons among different experiments. The model (swin_base_patch4_window7_224) is initialized with pre-trained weights for fine-tuning with optimal training parameters. The experiments were also performed with and without data augmentation for fair comparisons of varying diversity in data without changing any parameters.

The proposed framework Swin Transformer-Base was assessed using several performance metrics, namely accuracy, macro-averaged precision, macro-averaged recall, macro-averaged F1-score, and confusion matrix analysis. Macro-averaging was employed to ensure equal importance to all retinal disease classes so that the score would not be biased toward classes with higher sample counts. Quantitative results under training configurations with and without data augmentation are shown in table 1.

Table 1: Performance metrics of Swin Transformer-Base with and without data augmentation

Data Augmentation	Accuracy	Precision	Recall	F1-Score
Yes	0.9939	0.9907	0.9939	0.9922
No	0.9954	0.9965	0.9963	0.9963

Figure 5 shows the training loss and training accuracy curves for the Swin Transformer Base model. From the training loss curve, it can be observed that it decreases sharply in the beginning and then settles down smoothly with a graceful convergence, indicating that it neither suffers from training instability nor does it fail to converge. Meanwhile, the training accuracy rises sharply also to a saturation level in a couple of epochs and then remains stabilized near 100%. The smooth curve of convergence verifies the high learning capability of the model, further confirming the strength of the training approach in learning meaningful features from the input data without overfitting. The pattern matching between the reducing loss curve and rising accuracy curve further verifies the ability of the Swin Transformer-Base model to include meaningful discriminative patterns about retinal diseases during training.

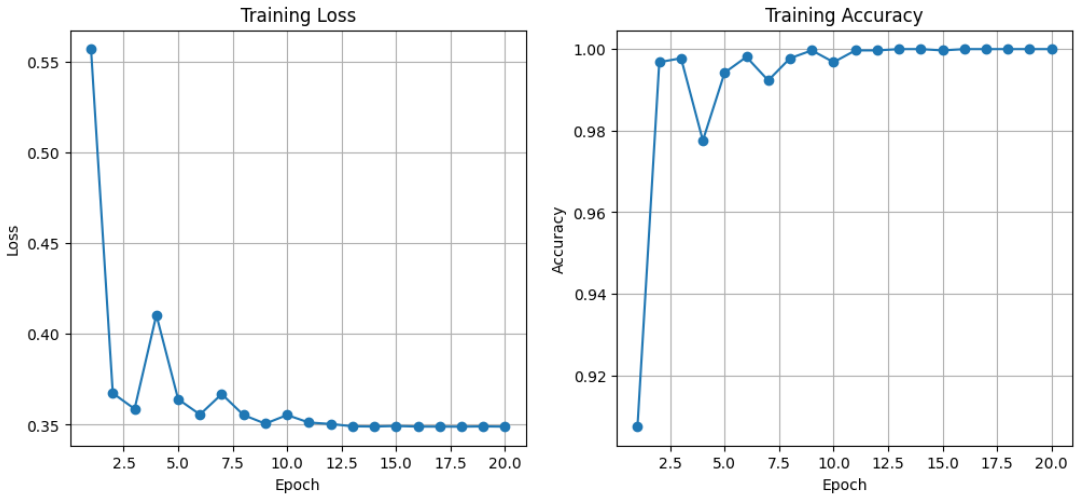

Fig. 5. Training loss and training accuracy curves of the Swin Transformer-Base over 20 epochs.

The performances of Swin Transformer-Base in both settings have shown consistently high values across all metrics. Trained with data augmentation, the model reaches a very strong balance between precision and recall, therefore discriminating reliably among the four classes of retinal diseases. The corresponding confusion matrix presents minimal misclassifications, as most of the prediction errors happen between clinically related disease categories. Interestingly, when trained without data augmentation, there is a slight improvement across all metrics: accuracy, precision, recall, and f1-score. In general, the results confirm the robustness of the proposed framework for multi-class retinal disease classification. Strong generalization across all disease categories is given by the high macro-averaged metrics of both models. The analysis of the confusion matrix provides further confirmation of the capability of the model to highlight subtle pathological differences in retinal fundus images.

VI. CONCLUSION

This work presents an interpretable vision transformer–based framework for automated single-label multi-class retinal disease classification using fundus images. The proposed system focuses on four clinically significant categories—Normal, DR, AMD, and BRVO and follows an end-to-end pipeline comprising image preprocessing, feature extraction using a Swin Transformer-Base architecture, explainability, and interactive deployment. A single Swin-Base model was trained and evaluated under a consistent experimental setup, demonstrating its strong capability to capture both local and global retinal structures through hierarchical representations and shifted window self-attention mechanisms. To enhance transparency and clinical trust, the Swin-Base model was integrated with gradient-based visual explanation techniques, enabling class-specific localization of disease-relevant retinal regions. These visual explanations highlight pathological patterns such as lesions, hemorrhages, and vascular abnormalities that directly influence the model's predictions. Furthermore, the integration of a large language model via the Google Gemini API provides complementary textual explanations by interpreting the predicted class and associated visual attention, thereby bridging the gap between model outputs and clinically meaningful reasoning. The complete framework was deployed using a Gradio-based web interface, allowing real-time retinal image analysis with unified visual and textual explanations. Experimental results have shown that the proposed framework is effective, interpretable, and practical for clinical decision-support in retinal disease screening. Future works will focus more on extending the framework to additional retinal pathologies, validating performance on larger multi-center datasets, and incorporating clinician-in-the-loop feedback to further improve robustness and clinical relevance.

REFERENCES

[1] Y. Kanagasingam et al., "Progress on retinal image analysis for age-related macular degeneration," Prog. Retin. Eye Res., vol. 38, pp. 20–42, 2014.

[2] P. Burlina, N. Joshi, M. Pekala, K. Pacheco, D. Freund, and N. Bressler, "Automated grading of age-related macular degeneration from fundus images using deep convolutional neural networks," JAMA Ophthalmol., vol. 135, no. 11, pp. 1170–1176, 2017.

[3] J. De Fauw et al., "Clinically applicable deep learning for diagnosis and referral in retinal disease," Nat. Med., vol. 24, no. 9, pp. 1342–1350, 2018.

[4] R. Rajalakshmi et al., "Automated diabetic retinopathy detection in smartphone-based fundus photography using AI," Eye, vol. 32, no. 6, pp. 1138–1144, 2018.

[5] J. Yim et al., "Predicting conversion to wet age-related macular degeneration using deep learning," Nat. Med., vol. 26, no. 6, pp. 892–899, 2020.

[6] U. R. Acharya et al., "Deep convolutional neural network for automated diagnosis of diabetic retinopathy," Future Gener. Comput. Syst., vol. 93, pp. 759–769, 2019.

[7] F. Li, H. Chen, Z. Liu, X. Zhang, and Z. Wu, "Fully automated detection of retinal disorders by image-based deep learning," Graefe's Arch. Clin. Exp. Ophthalmol., vol. 257, no. 3, pp. 495–505, 2019.

[8] A. García-Floriano and E. Ventura-Molina, "Age-related macular degeneration detection in fundus images using deep CNNs," Mathematics, vol. 12, no. 10, p. 1534, 2024.

[9] S. Yu et al., "MIL-VT: Multiple instance learning enhanced vision transformer for fundus image classification," in Proc. MICCAI, pp. 45–54, 2021.

[10] Z. Jiang et al., "Computer-aided diagnosis of retinopathy based on vision transformer," J. Innov. Opt. Health Sci., vol. 15, no. 2, p. 2250006, 2022.

[11] Y. Kihara et al., "Detection of non-exudative macular neovascularization using vision transformers," Ophthalmol. Sci., vol. 2, no. 4, p. 100214, 2022.

[12] P. Dutta, K. A. Sathi, M. A. Hossain, and M. A. A. Dewan, "Conv-ViT: A convolution and vision transformer-based hybrid feature extraction method for retinal disease detection," J. Imaging, vol. 9, no. 7, p. 140, 2023.

[13] K. Xu et al., "Automatic detection and differential diagnosis of age-related macular degeneration using hierarchical vision transformers," Comput. Biol. Med., vol. 167, p. 107573, 2023.

[14] A. A. Abd El-Khalek et al., "XV-AMD: An explainable vision transformer detection framework for age-related macular degeneration using fundus imaging," IEEE Access, vol. 13, pp. 113967–113983, 2025.

[15] Y. Peng et al., "Enhancing AI reliability with uncertainty estimation for retinal disease diagnosis," Cell Rep. Med., vol. 6, no. 1, p. 101386, 2025.

[16] S. Gholami et al., "Federated learning for diagnosis of age-related macular degeneration," Front. Med., vol. 10, p. 1189456, 2023.

[17] K. Guttikonda and G. Ramachandran, "Design of diagnostic framework for detecting autism spectrum disorder using CMIM-RF," J. Propuls. Technol., vol. 44, no. 6, pp. 3679–3691, 2023.

[18] K. Guttikonda, G. Ramachandran, and G. V. S. N. R. V. Prasad, "Autism spectrum disorder prediction using LASSO regularised bat search optimisation," Int. J. Serv. Oper. Informatics, vol. 13, no. 1, pp. 1–20, 2024.

[19] K. Guttikonda, Y. Ashvitha, V. S. R. Reddy, R. M. Krishna, and P. Sandeep, "Integrating convolutional neural networks and machine learning for accurate identification of autism spectrum disorder using facial biomarkers," in Proc. IEEE Int. Conf. Emerging Systems and Intelligent Computing, pp. 1–6, 2024.

[20] K. Guttikonda G. Ramachandran, and G. V. S. N. R. V. Prasad, "Cuckoo search optimisation-based feature selection for predicting autism spectrum disorder using artificial immune algorithms," J. Theor. Appl. Inf. Technol., vol. 103, no. 2, pp. 421–432, 2025.