AGE-RELATED MACULAR DEGENERATION USING DEEP LEARNING MODELS AND GAN DATA AUGMENTATION

PHASE I REPORT

Submitted by

LAKSHMI KANTH M 210701130

NANDHAKUMAR S 210701172

In partial fulfillment for the award of the degree of

BACHELOR OF ENGINEERING

IN

COMPUTER SCIENCE AND ENGINEERING

RAJALAKSHMI ENGINEERING COLLEGE, CHENNAI

ANNA UNIVERSITY, CHENNAI 600 025

NOVEMBER 2024

BONAFIDE CERTIFICATE

Certified that this Report titled **"AGE-RELATED MACULAR DEGENERATION USING DEEP LEARNING MODELS AND GAN DATA AUGMENTATION"** is the Bonafide work of **LAKSHMI KANTH M (210701130), NANDHAKUMAR S (210701172)** who carried out the work under my supervision. Certified further that to the best of my knowledge the work reported here in does not form part of any other thesis or dissertation on the basis of which a degree or award was conferred on an earlier occasionon this or any other candidate.

SIGNATURE	**SIGNATURE**
Dr. Kumar P, M.E., Ph.D.,	**Mrs. Anandhi S, M.E.,**
Professor and Head	**Assistant Professor**
Department of Computer Science and Engineering	Department of Computer Science and Engineering
Rajalakshmi Engineering College	Rajalakshmi Engineering College
Chennai - 602105	Chennai - 602105

Submitted to Project phase I Viva-Voce Examination held on _______________

Internal Examiner **External Examiner**

ACKNOWLEDGEMENT

Initially we thank the Almighty for being with us through every walk of our life and showering his blessings through the endeavor to put forth this report. Our sincere thanks to our Chairman **Mr. S. MEGANATHAN, B.E, F.I.E**., our Vice Chairman **Mr. ABHAY SHANKAR MEGANATHAN, B.E., M.S.**, and our respected Chairperson **Dr. (Mrs.) THANGAM MEGANATHAN, Ph.D.**, for providing us with the requisite infrastructure and sincere endeavoring in educating us in their premier institution.

Our sincere thanks to **Dr. S.N. MURUGESAN, M.E., Ph.D.**, our beloved Principal for his kind support and facilities provided to complete our work intime. We express our sincere thanks to **Dr. P. KUMAR, M.E., Ph.D**., Professor and Head of the Department of Computer Science and Engineering for his guidance and encouragement throughout the project work. We convey our sincere and deepest gratitude to our internal guide, Mrs. **S. Anandhi, M.E.,** Assistant Professor, Department of Computer Science and Engineering. Rajalakshmi Engineering College for his valuable guidance throughout the course of the project. We are very glad to thank our Project Coordinator, **Dr. T. Kumaragurubaran, M.E., Ph.D.,** Associate professor, Department of Computer Science and Engineering for his useful tips during our review to build our project.

LAKSHMI KANTH M (210701130)

NANDHAKUMAR S (210701172)

ABSTRACT

Age-related macular degeneration (AMD) is one of the most prevalent and serious eye conditions affecting elderly individuals, often resulting in significant vision impairment if not detected and managed early. The importance of early detection cannot be overstated, as timely intervention can mitigate disease progression and improve patient outcomes. Traditional diagnostic approaches, however, face critical limitations, particularly the challenge of an imbalanced dataset comprising an overwhelming number of normal cases and insufficient representation of advanced AMD stages. This imbalance hinders accurate stage classification, resulting in less effective diagnostic systems. To address these challenges, this paper introduces a hybrid deep learning framework that combines advanced feature extraction and data augmentation techniques to deliver a more precise, reliable, and robust solution for AMD detection. By integrating EfficientNet for fine-grained feature discovery, Vision Transformer (ViT) for context-aware perception, and Generative Adversarial Networks (GANs) for synthetic data generation, the proposed model significantly enhances the analysis of Optical Coherence Tomography (OCT) images. EfficientNet captures intricate patterns and fine details, while ViT provides a broader context, ensuring a comprehensive understanding of disease markers. This fusion allows for a more detailed and nuanced classification of AMD stages. A standout innovation of the approach lies in the use of GANs to overcome the dataset imbalance problem by generating synthetic samples for less-represented AMD stages. This ensures a balanced training dataset, stabilizing the training process, accelerating convergence, and improving classification accuracy across all AMD stages.

TABLE OF CONTENTS

CHAPTER NO	TITLE	PAGE NO
	ACKNOWLEDGEMENT	iii
	ABSTRACT	iv
	LIST OF FIGURES	vii
	LIST OF ABBREVIATIONS	viii
1.	**INTRODUCTION**	**1**
	1.1 GENERAL	1
	1.2 OBJECTIVE	2
	1.3 EXISTING SYSTEM	3
	1.4 PROPOSED SYSTEM	4
2.	**LITERATURE SURVEY**	**5**
3.	**SYSTEM DESIGN**	**10**
	3.1 GENERAL	10
	3.1.1 SYSTEM FLOW DIAGRAM	10
	3.1.2 SEQUENCE DIAGRAM	11
	3.1.3 CLASS DIAGRAM	12
	3.1.4 USECASE DIAGRAM	13
	3.1.5 ARCHITECTURE DIAGRAM	14
	3.1.6 ACTIVITY DIAGRAM	15
	3.1.7 COMPONENT DIAGRAM	16
4.	**PROJECT DESCRIPTION**	**17**
	4.1 METHODOLOGIES	17
	4.1.1 RESULT AND DISCUSSION	23

5.	**CONCLUSION AND WORK SPACE**	**31**
	FUTURE WORK FOR PHASE II	32
	REFERENCES	34
	APPENDIX	36

LIST OF FIGURES

FIGURE NO	TITLE	PAGE NO
1	SYSTEM FLOW	10
2	SEQUENCE DIAGRAM	11
3	CLASS DIAGRAM	12
4	USECASE DIAGRAM	13
5	ARCHETECTURE DIAGRAM	14
6	ACTIVITY DIAGRAM	15
7	COMPONENT DIAGRAM	16

LIST OF ABBREVIATIONS

SNO	ABBREVIATION	EXPANSION
1	AI	Artificial Intelligence
2	API	Application Programming Interface
3	AMD	Age-Related Macular Degeneration
4	CNN	Convolutional Neural Network
5	OCT	Optical Coherent Tomography
6	ViT	Vision Transformer

CHAPTER 1

1. INTRODUCTION

1.1 GENERAL

Age-Related Macular Degeneration represents a prevalent ocular condition that seriously threatens the visual well-being of the elderly population worldwide. Proper identification and correct classification of the different stages of AMD are critical to effective intervention and therapy since delays in diagnosis can lead to permanent vision loss. However, traditional diagnostic techniques often fail miserably because they are limited in their ability to closely distinguish between subtle variations between the different stages of AMD. Furthermore, the growing necessity for precise diagnostic systems underscores the pressing requirement for sophisticated technologies to assist healthcare professionals in accurately and dependably diagnosing this condition.

In recent times, promising developments have taken place in the deep learning models to achieve completely automated analysis of medical images like OCT scans commonly used in AMD diagnosis. However, these models suffer a lot while dealing with imbalanced datasets, specifically about advanced cases of AMD. Such imbalances restrain model performance and enforce limitations upon its ability to generalize, especially in real-world applications. Further, existing CNN-based models often miss capturing such both global and local features at the same time that are vital for such complex medical image. Such gaps in diagnostic accuracy thus motivate even more innovative approaches.

This proposed work introduces a novel type of hybrid deep learning system that comprises two recently developed architectures, namely Vision Transformers (ViT), EfficientNet, and Generative Adversarial Networks (GANs). The method integrated attempts to leverage the capacity of ViTs for global context incorporations, detailed local feature analysis by EfficientNet, and synthetic data augmentation through GAN. Therefore, the hybrid model above addresses the dataset imbalances and improves feature extraction, thus establishing a strong solution for AMD classification. It is positioned between existing approaches and opens the door to both high accuracy and improved generalization across all possible stages of AMD, including very rare, advanced cases.

It brings a novelty in making the holistic analysis of OCT images, which contains both entire structural information and fine-grained local details. The expertise at what vision transformers are good at capturing long-range dependencies makes them strong for scanning broad retinal trends, while EfficientNet optimizes the extraction of localized features, like specific retinal anomalies. Meanwhile,GANs produce synthetic data as a compensatory action toward underrepresentation at advanced stages of AMD; it enriches the training of the model with a more diverse and balanced set of data. This synergy dramatically improves the system's diagnostic capability beyond what is achieved in traditional CNN-based solutions.

The proposed system through the present research, therefore, constitutes a revolutionary step toward diagnosis in AMD and hence can find broader applicability in the area of medical imaging. This system is going to integrate state-of-the-art techniques in deep learning while addressing previously untapped critical limitations existing in models to enhance the resilience and reliability of AMD classification systems. The benefits of this study will help clinicians make some informed decisions in order to reach better patient care with a reduction in the global burden of AMD.

1.2 OBJECTIVE

The overall aim of this project is to design a novel hybrid deep learning system for accurate and early classification of AMD at all its stages of progression, which comes with critical limitation-advanced cases of AMD are less available, and the dataset has very imbalanced nature. This system will capture the global retinal pattern with the help of Vision Transformers, localized features through

EfficientNet, and synthetic data augmentation through Generative Adversarial Networks. Altogether, it will enable the model to improve in precision and generalize better. Sensitive anomaly analysis and broader structural trends are two sides of the coin that need equal importance in OCT image analysis. In this project, resilience and reliability are mainly enhanced as it generates synthetic samples to handle the class imbalance, which is more critically required in underrepresented stages of AMD. The hybrid framework built in contrast to traditional CNN-based methods remains quite strong with real-time adaptation capabilities with superior performance, finally allowing clinicians to make better decisions by having a data-driven approach, thereby improving patient outcomes and helping advance their standard diagnostics in ophthalmology.

1.3 EXISTING SYSTEM

Highly advanced deep architectures are based solely on CNNs, like AlexNet, ResNet and VGG variants, relying on AMD classification. The systems can discriminate between dry and wet AMD with Optical Coherence Tomography imaging techniques but struggle with the nuanced differences in retinal characteristics on which the stage-specific AMD forecasting conditions are predicated. This reliance on large, labeled datasets further constrains their applicability, particularly when dealing with advanced stages or rare classes. Balancing datasets worsens this problem to produce strongly biased models that could generalize poorly to other cases that are under-represented.

Transfer learning is also widely utilized to circumvent these problems by fine-tuning pre-trained models on AMD-specific datasets. Although this improves accuracy in some instances, models fail to generalize, especially to minority classes. Such systems cannot be updated with new data and cannot provide real-time information on the patient's status. This does make them less efficient in the clinical environment where flexibility, plus acquiring real-time information, is of value. In this regard, existing systems illustrate the necessity of adaptive, stable, and flexible classification methods of AMD.

1.4 PROPOSED SYSTEM

The deep learning hybrid architecture brings together Vision Transformers, EfficientNet, and GANs to address complexities involved in classifying the AMD. The self-attention mechanism of ViT catches global spatial dependencies in OCT images, whereas compound scaling of the EfficientNet skillfully extracts intricate local features, thereby leading to better capture of details of the retinal structure. In totality, these models are better than those CNN-based applications because they integrate global dependencies and localized irregularities for an overall understanding of the OCT images.

GANs thus balance the rare classes like the advanced stages of AMD by finding the heterogeneity of the dataset through the generation of synthetic OCT images. Overfitting is reduced with increased generalizability of the networks, and at all stages of AMD, proper classification is assured. The model learns better from underrepresented cases by enriching the dataset with realistic examples. With it, the system performs reliably even with the variability found in real-world clinical data.

This set theory of ViT, EfficientNet, and GANs ends with a strong, scalable, and flexible approach towards AMD classification. The model exploits the benefits of both ViT and EfficientNet in the process of feature extraction while using GAN to combat poor data conditions. This collaborative framework results in improved accuracy and reliability in clinical settings, effectively positioning the system as a powerful tool for diagnosing and managing AMD at any stage. The system bringsa new benchmarking parameter to adaptability and efficiency in automated classification for retinal diseases.

CHAPTER 2

2. LITERATURE SURVEY

Ali Serener, et al., [1], Examined the use of deep learning to classify dry and wet Age- Related Macular Degenerations automatically using Optical Coherence Tomography scans. Comparing AlexNet and ResNet, the authors found thatResNet was more stable and trustworthy, especially concerning dry AMD diagnosis. Thus, the author stressed the scope of its potential deep learningcapability in widening the scope of diagnostic workflows that might initiate early detection and proper planning for the treatment. However, the classification of wet AMD based on this method is not very accurate due to a complex dataset with high computational demands and requirements for expertise in its implementation. The principal contribution of this study establishes how advanced models in deep learning may fill significant gaps in AMD diagnostics and build for future innovations.

Yao-Mei Chen, et al.,[2], conduct research work on classifying OCT images either towards AMD or DME correctly. In this work, all authors have applied transfer learning using pre-trained CNN models with AlexNet, GoogLeNet, VGG16, VGG19, ResNet18, ResNet50, and ResNet101 to classify the OCT images. Fine-tuning with specific hyperparameters was done to get high accuracy, mainly it was found to be VGG19, ResNet101, and ResNet50 got high accuracy compared to other models. It saves computational time as well as resources as opposed to training from scratch. However, the success of the model is very much dependent on proper selection of hyperparameters that may be too complex as well as even time consuming. Furthermore, the approach requires large datasets for training that may not readily be available to hinder broader application.

Yifan Peng and Shazia Dharssi [3], Design a deep learning model called DeepSeeNet was designed to automatically classify patients on the basis of the severity of AMD using color fundus photographs and was made similar to human grading by measuring the size and pigmentary abnormalities in the two primaries risk factors for AMD through calculations of the AMD severity score from patient data. As is traditional, it requires expert knowledge and time, such as with the AREDS Simplified Severity Scale in determining the severity of AMD. This model outperformed the retina specialists in the detection of large drusen and pigmentary abnormalities through transparent, accurate, automated approaches to clinical decision-making. However, the late AMD detection is not as sensitive, including image quality and imbalanced data, especially for the late stage of AMD, loweringits effectiveness.

Dan Gong, et al., [4] Focuses on implementation of DL, specifically CNNs, and its implementation in diagnosing and classifying AMD based on fundus photography and OCT imaging. Previously, the diagnosis and management of AMD were not only resource- intensive and time-consuming but may also impose limitations on efficiency and access to care. In a study, it was shown that the diagnostic accuracy of DL models trained on a very large dataset is comparable with that of human experts, thus offering supportive decisions for treatment. This could improve diagnostic efficiency, enhance screening through telemedicine, and help monitor disease; all of these have the potential to reduce visual impairment globally. It heavily relies on available large datasets so that any bias in the training data will also affect the results. Moreover, limited integration with wider clinical data restricts the ability of the model to provide a more comprehensive decision-making facility.

P. Burlina, et al., [5], Develop a system to diagnoses Age-Related Macular Degeneration with features from deep convolutional neural networks and also from a pre-trained OverFeat network and a linear SVM classifier. Early detection of AMD is quite challenging in clinical settings. The system was tested on the NIH AREDS dataset of more than 5,600 images. It shows promise for use in clinical and under-resourced environments byachieving a 92-95% accuracy in the detection and classification of AMD. The model uses pre-trained networks with an extent that at least minimizes large medical datasets. Although tested against only a small dataset, the system still needs adequate fine-tuning and further validation against larger datasets but has hopeful prospects of being robust in diverse clinical settings.

Frederick L. Ferris III, et al., [6] Address the inconvenience of AMD classification on an unequal basis because no universally accepted system exists. AMD diagnosis and management are therefore affected. For this purpose, a standardized classification system for AMD was developed by using a modified Delphi process among 26 AMD experts. This system envisioned data from large studies and consensus of these experts focusing on some key elements like size/type of drusen and pigmentary abnormalities to classify the stages of AMD. It helps in the progression for late AMD, communicating well among the providers of eye care. Although it gives a clinically usable classification and lets the researcher attain consistency by excluding patients to make use of advanced techniques, such as genetic testing, it tends to make the categorization of patients too simple.

Preeti Sharma and Anand Prakash Shukla, [7], Discuss the routine task while dealing with the brain tumor diagnosis from MRI images. This usually requires more time and experience of the radiologists. It applies a transfer learning methodology borrowed from EfficientNet architectures; B0 and B7 are used here to classify the four classes, namely no tumor, pituitary, meningioma, and glioma. A total of 3,264 MRI images were considered for the dataset. The models were pre- trained on ImageNet, followed by fine-tuning for the specific task. The model has been empowered with data augmentation to prevent overfitting, and performance will be assessed on accuracy, precision, F1-score, specificity, and sensitivity. The approach attained 98% accuracy; transfer learning in medical imaging is a promising direction to consider. Although the proposed approach is very efficient and versatile, the application relies on pre-trained models from an entirely different domain of images, which again might be somewhat limited in generalizing performance. Computational complexity involved in deep learning is a challenge still.

Guangzhou An, et al., [8], Address the challenge posed by the impossibility of clearly discriminating between fluid and non- fluid Age-Related Macular Degeneration (AMD) with OCT images. The research uses a two-stage transfer learning approach, leveraging the

VGG16 CNN architecture already pre-trained on ImageNet. The first stage of training adjusts the model in order to classify between normal OCT scans and those that imply AMD. The next step adds more specification to distinguish between fluid-present AMD and fluid-absent AMD. This takes an advantage from the transfer learning and is highly efficient even at low data levels, thus considerably decreasing the load of diagnosis on an ophthalmologist. This study, however, draws its basis on a dataset from a single centerand excludes more complex cases, meaning the general applicability of the model might be bound and will probably be overestimated in real settings. Thus, the method developed needs further validation with multiple external datasets to enhance its use in broader medical environments because that is where its generalization and practical utility in a wider medical environment will come in.

P. Kumar, et al., [9], implemented Random Forest algorithms for human activity recognition in handheld devices. The method showed reliable real-time performance in smart devices, effectively balancing accuracy and speed. Challenges included maintaining efficiency in resource-limited environments, where computational constraints posed difficulties. The study highlighted the algorithm's adaptability, showcasing its potential for integrating into various mobile applications. It emphasized optimization for broader adoption, particularly in wearable technology and healthcare monitoring systems, where lightweight algorithms are crucial for sustained performance.

P. Kumar, et al., [10], focused on predicting energy consumption in buildings using ensemble methods. The study provided actionable insights for sustainability efforts. Integrating diverse datasets posed technical challenges, requiring robust preprocessing techniques to ensure model accuracy. Computational demands limited scalability, necessitating further optimization for real-world deployment. The research underscored the role of machine learning in energy efficiency and highlighted its potential for aiding predictive maintenance and resource management in smart buildings.

T. Kumaragurubaran, et al., [11], explored deep learning for text-to-sketch conversion in creative content generation. The method automated design processes efficiently, reducing manual effort and time. Complex sketches presented challenges for accurate representation, particularly when dealing with intricate details. The approach needed optimization for real-time application, as latency hindered seamless user experiences. It showcased AI's potential in creative fields, emphasizing its ability to bridge textual and visual content generation while paving the way for advancements in design and multimedia industries.

S. SenthilPandi, et al., [12], proposed an adaptive model for lung tumor segmentation using medical images. The system achieved accurate volumetric estimation and demonstrated strong clinical relevance. Diverse dataset adaptation remained a challenge, particularly in ensuring model generalizability across varying imaging conditions. The model required high computational resources, which limited its deployment in resource-constrained settings. Despite these challenges, it offered significant potential for aiding clinical decision-making by providing precise tumor assessments. The approach underscored the importance of integrating AI into medical imaging to enhance diagnostic accuracy and treatment planning.

CHAPTER 3

3. SYSTEM DESIGN

3.1 GENERAL

3.1.1 SYSTEM FLOW DIAGRAM

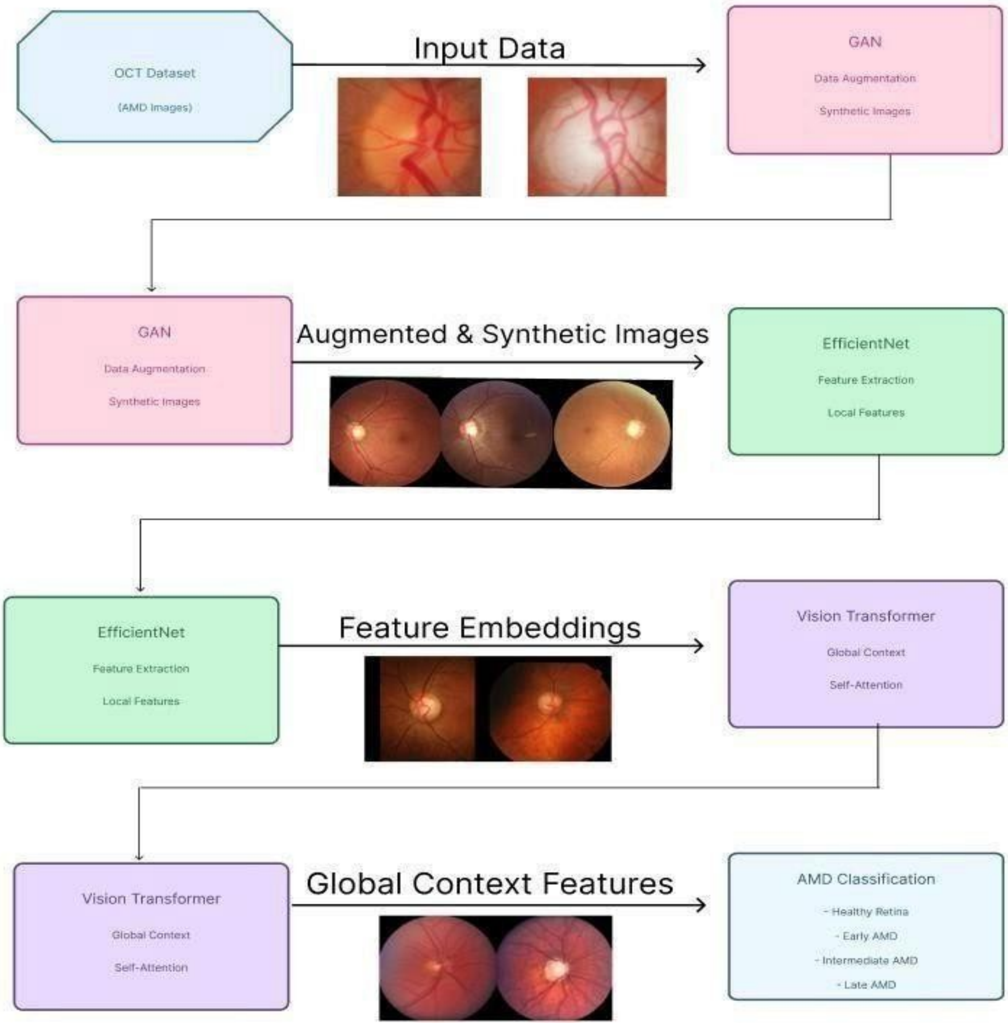


Figure 1: System Flow Diagram

The pipeline for the classification of AMD via OCT images is demonstrated in the following diagram. It starts with data augmentation and synthetic image generation using GANs. Features are locally extracted using EfficientNet and globally contextualized using a Vision Transformer. Finally, the global features are used to achieve multi-stage diagnosis of AMD as whether the retina is healthy or in some stages of AMD.

3.1.2 SEQUENCE DIAGRAM

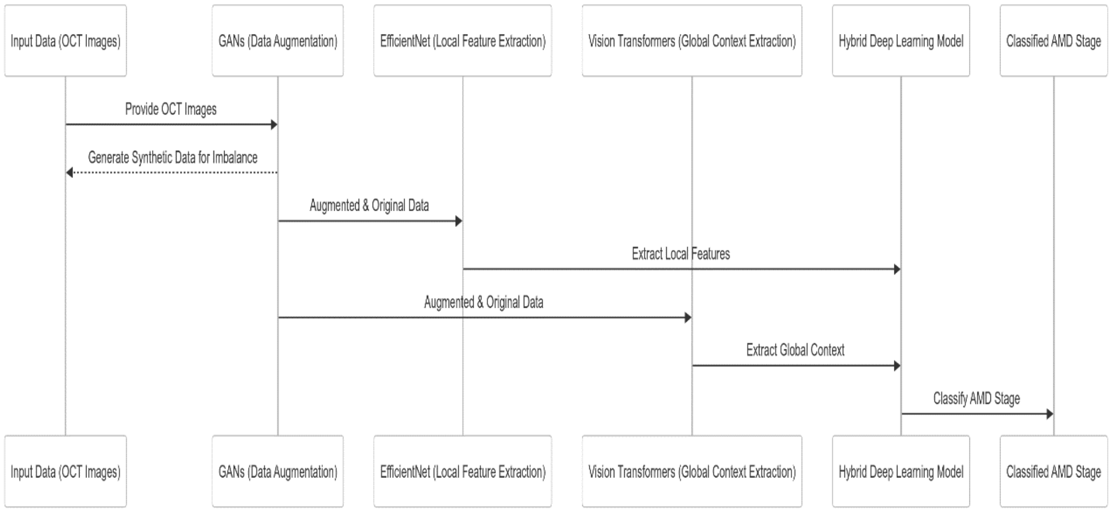


Figure 2: Sequence Diagram

The AMD classification system architecture diagram it is illustrated as data flow and processes through the input to output classification, initiated with augmented and balanced OCT images employing GANs for generating synthetic data for alleviating class imbalance. EfficientNet is used to extract the local feature with the help of which fine-grained details from the input are captured, and Vision Transformers are used to extract global contextual features through self-attention mechanisms. These features are gathered in a hybrid deep learning model that gives local as well as global ideas toward better and more robust analysis. Finally, the output classifies the retina at different stages of AMD for precise medical diagnosis.

3.1.3 CLASS DIAGRAM

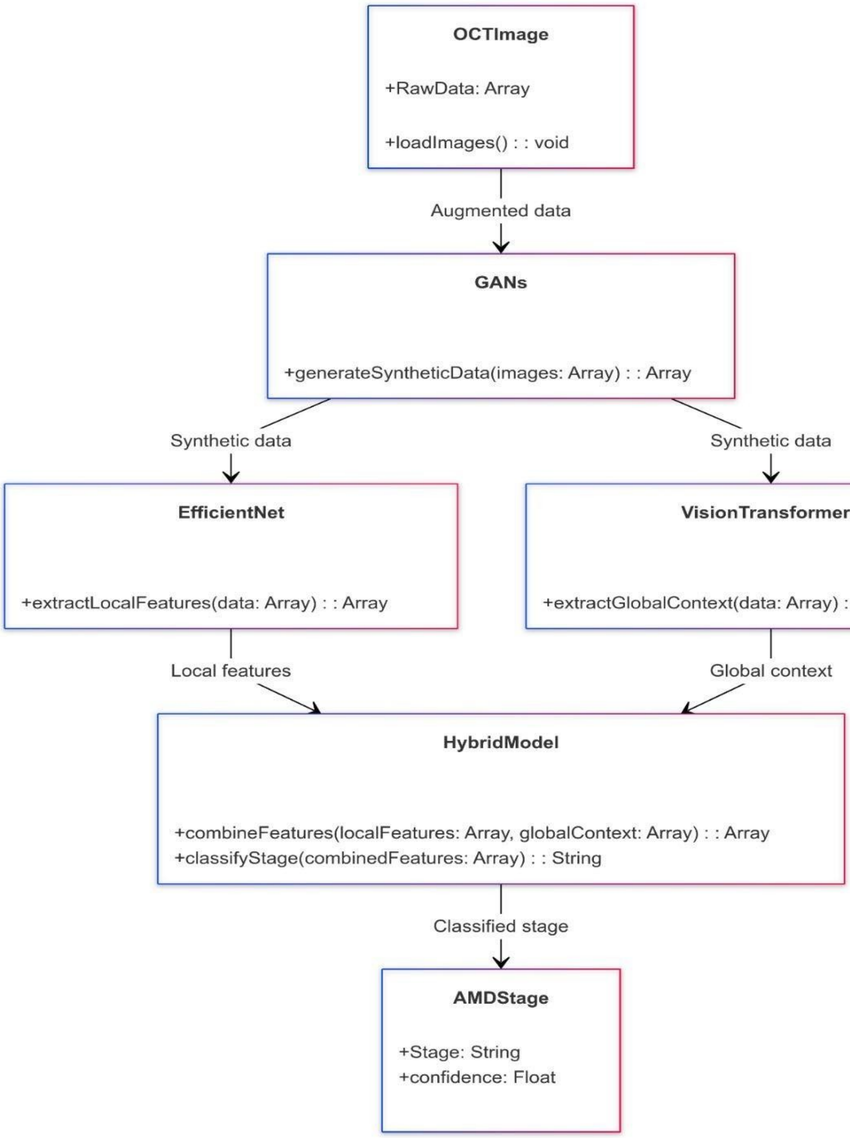


Figure 3: Class Diagram

The architectural diagram represents a class-based approach to the classification of AMD using OCT images. OCTImage is the module for data input, that loads raw images for later processing.In this, GANs are used to generate synthetic data to address class imbalances and improve the dataset. EfficientNet extracts local features, and Vision Transformers focus on extracting global contextual features from the augmented data. Such characteristics are coupled within a HybridModel, which aggregates these together for this classification in a stepwise manner. The last AMDStage module results in the classified AMD stage along with the confidence score, thusallowing the correct medical diagnosis.

3.1.4 USE CASE DIAGRAM

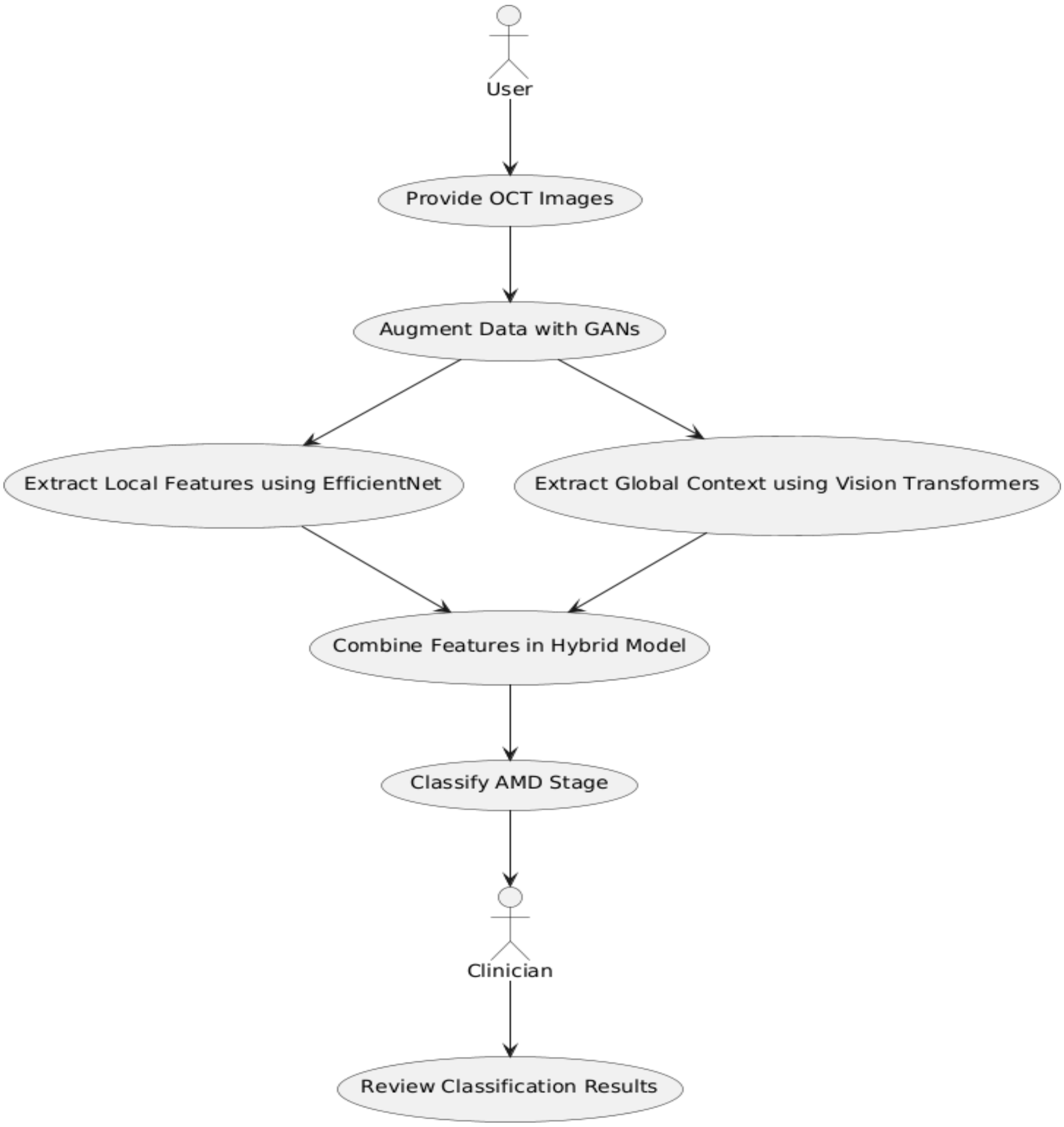


Figure 4: Use Case Diagram

It is an advanced classification system for AMD, describing the interactions between the users andclinicians with components of the system. The user begins with the upload of OCT images, which undergoes preprocessing and augmentation using GANs to improve the dataset, then use EfficientNet for local feature extraction and Vision Transformers for global context analysis. A hybrid model combines these features to accurately classify the AMD stage. The results of the classification from this system are reviewed by clinicians, hence a reliable process in making medical decisions.

3.1.5 ARCHITECTURE DIAGRAM

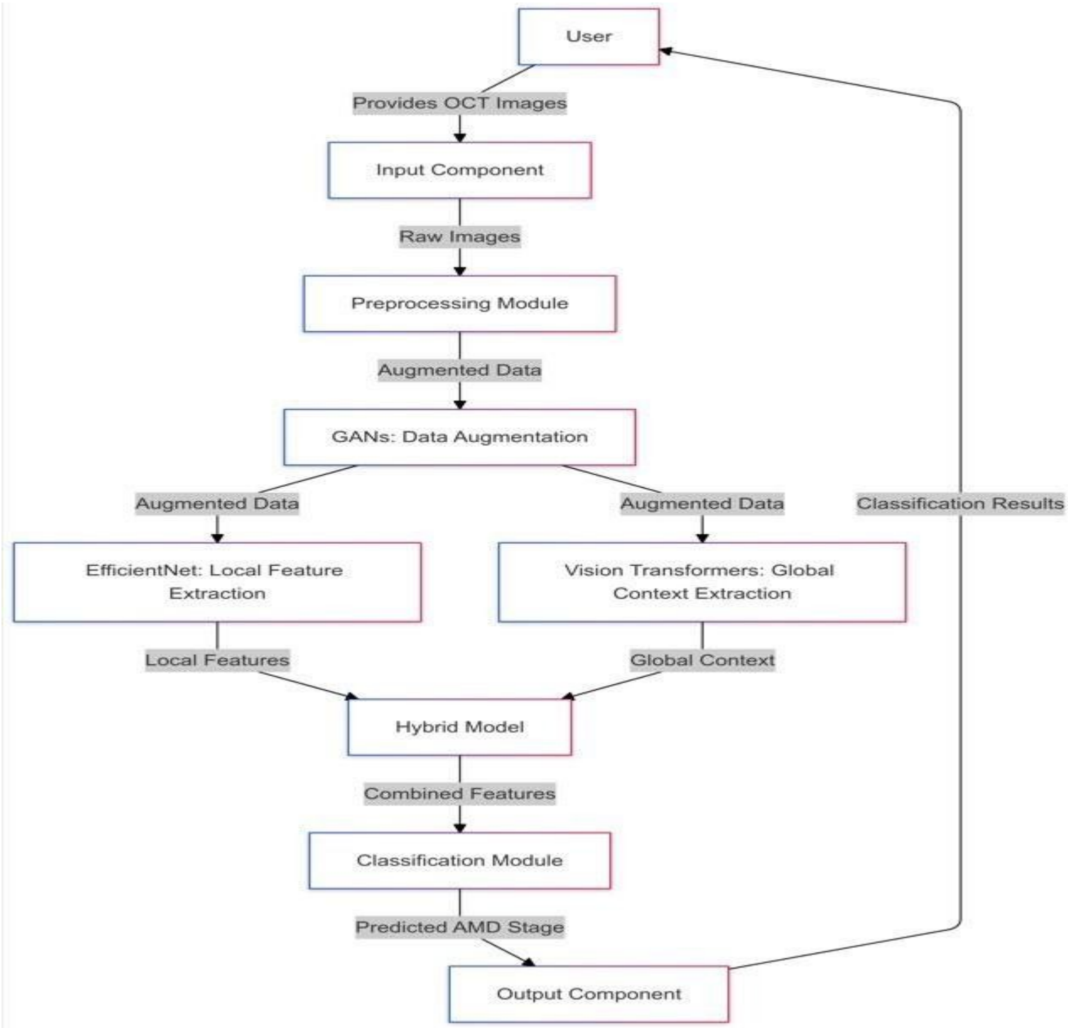


Figure 5: Architecture Diagram

The architecture diagram exhibits a developed AMD classification pipeline with structured flow from input to final output. Firstly, the process starts off by the input module by providing OCT images; after preprocessing of the same through the preprocessing module, it prepares raw images for analysis. GAN-based data augmentation enriches the dataset by generating very diverse synthetic images; thus, this enhances the robustness of the system. Local features are extracted by utilizing EfficientNet while the use of Vision Transformers extracts global contextual information that enables two complementary streams of information. These features are then combined in a hybrid model which further feeds the classification module to make accurate predictions about the AMD stage. Finally, the output component delivers the classification results to the user for clinicaldecision making and judgment.

3.1.5 ACTIVITY DIAGRAM

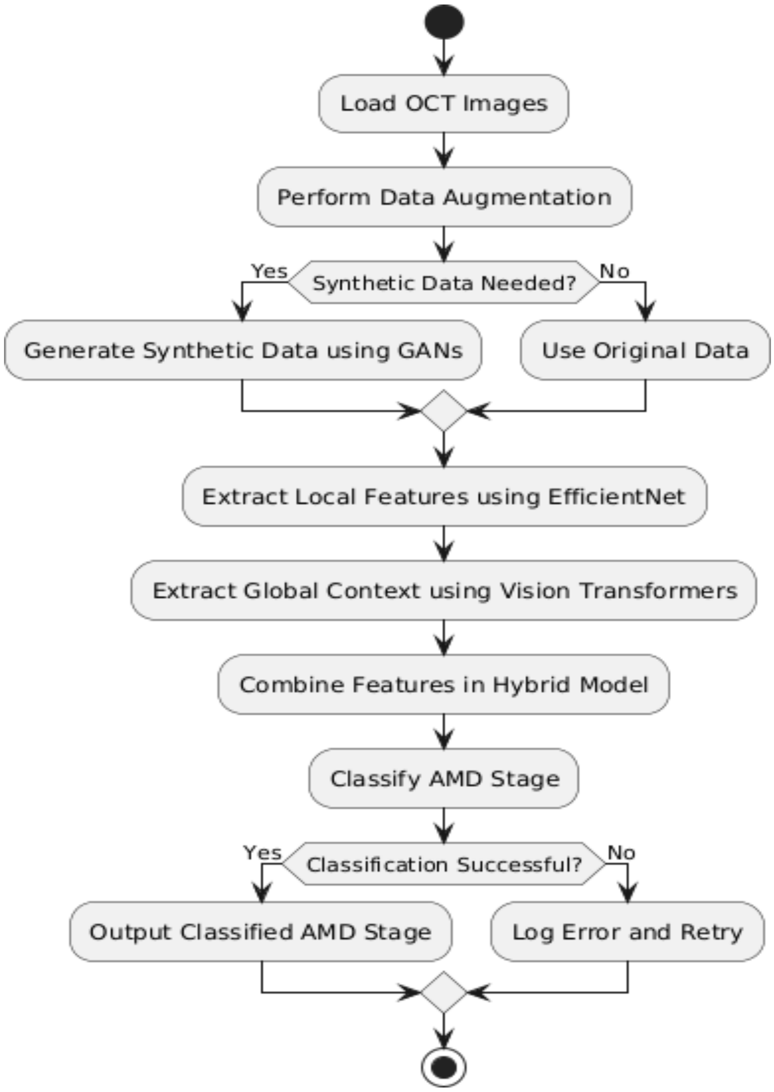


Figure 6: Activity Diagram

This activity diagram depicts the workflow of an advanced classification system that is based onAMD, beginning with inputting OCT images. Introducing data augmentation, this decision pointcalculates whether synthetic data is needed-it is when true, so GANs are called to create synthetic images, and else, true data is applied. EfficientNet extracts local details, whereas VisionTransformers capture global contextual information. These two features together form a hybrid model that's used in order to classify the AMD stage. A classification decision point checks the correctness of classification-it results in an AMD stage when correct or records mistakes for future retries when wrong. It ensures an error-free and robust classification process.

3.1.6 COMPONENT DIAGRAM

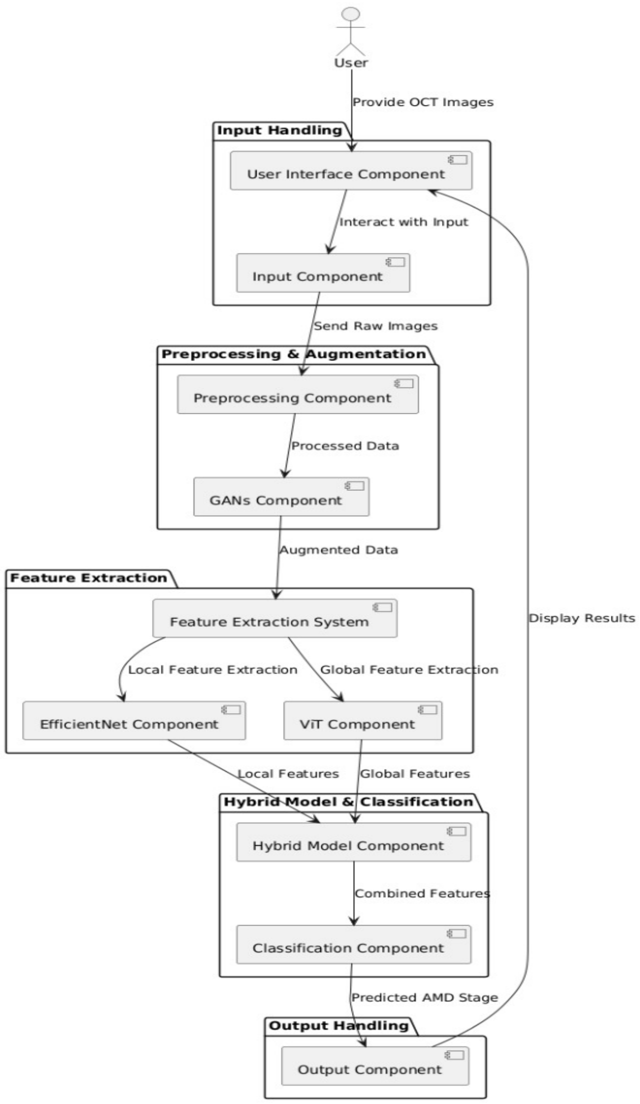


Figure 7: Component Diagram

The component diagram describes a modular classifying system developed by AMD for the process of OCT images. The Input Handling segment allows users to upload images through the User Interface Component that later directs them for preprocessing. The module of Preprocessing& Augmentation enhances images by using the GANs Component to generate synthetic data when needed. The Hybrid Model combines local features through EfficientNet and global features through ViT. This combines the best of both worlds into the Hybrid Model to classify. The Output Component finally returns the predicted stage of AMD to the user with utmost accuracy and efficiency.

CHAPTER 4

PROJECT DESCRIPTION

4.1 METHODOLOGIES:

1. Input Data Acquisition:

The gathering of an Optical Coherence Tomography dataset is the very first stage of the workflow. The dataset involves images pertinent to Age-Related Macular Degeneration. Such OCT images are very important because they bring high resolution and cross-sectional presentations of the retina, thus providing valuable information about the structural changes that occur in association with AMD. A representative array of images is required to depict all stages of AMD-from healthy retinas to more advanced stages of wet AMD or geographic atrophy. The diversity in the dataset will ensure that the model learns subtle variation across various disease stages. The system learns to distinguish normal from pathological features of the retina by including images of healthy retina with mixed degrees of AMD. This large and heterogeneous dataset provided is the key to ensuring that the system generalizes and does well in real-world data.

Data standardization is obtained on the collection of images after resizing, contrast enhancement, and normalization through some preprocessing techniques. This method will ensure uniformity of all input images such that recognition of features by deep learning models is enhanced. Resizing corrects differences in dimensions between the pixels of different images, while contrast enhancement makes more visible bright structures, which include drusen and retinal fluid. Normalization scales all the pixel values into a normalized range to reduce variation due to lightingas well as inconsistency in differences of imaging. This preprocessing stage is highly essential in ensuring uniformity in the presentation of all input images hence enabling the model to acquire meaningful features. At the end, preprocessing readies the OCT images to be added to the GAN for augmentation, which forms the subsequent stage.

Proper preparation of the data set will ensure that it is sorted and ready for the later stages, when the model will use more advanced methods in increasing accuracy. In addition to training the model, the input dataset is crucial for real-world performance testing. Data diversification and preparation enhance generalization of the model across different stages of AMD. Such thorough data preparation serves as the basis for this system, allowing it to make precise diagnoses and classificationsof AMD with high accuracy.

2. Data Augmentation with GANs:

Balancing the problems like class imbalance, particularly infrequent stages of age-related macular degeneration (AMD) like late-stage (wet AMD, or geographic atrophy), generative adversarial networks (GANs) are utilized in generating synthetic images; GANs include two neural networks: the generator, which creates new images, and the discriminator, which is trying to distinguish the fake from real images. The generator produces artificial images derived from the patterns acquired from the original dataset, whereas the discriminator verifies that these artificial images accurately reflect authentic AMD images. This adversarial mechanismallows the GAN to progressively enhance its capabilities, yielding high-fidelity synthetic images that are indistinguishable from genuine OCT images.

These generated images come very helpful with the problem of class imbalance for the dataset, where rare cases of late stages of AMD are not adequately represented. The GAN helps in that by generating very realistic synthetic images for those underrepresented stages to ensure that the model has enough data to learn, thus reducing bias in classification. The output images will help fill in the missing examples within the set, giving the model a broader range of examples to learn from.It is thus an augmented dataset, combining real and synthetic images, which ensures exposure for models in AMD across all stages with more balanced and varied examples.

Additionally, using GANs for augmentation aids in improving the model's generalization ability. The integration of GAN-generated images with the original dataset will make the model more robust by learning to trace patterns on a larger scale of retinal anomalies. Since it has been introduced to so many different retinal conditions through the introduction of synthetic images, the model learns not to overfit to any specific stage of AMD. This makes the model more capable of being able to make accurate predictions when faced with previously unseen data. To form an early, intermediate, and late-stage classification model based on a better training data this system is designed.

3. Feature Extraction with EfficientNet:

The goal here after GAN-based augmentation enriches the dataset is to extract meaningful feature representations from the images through the use of EfficientNet. EfficientNet is a CNN that, at this time, is the state of the art for resource efficiency and accuracy within models. It uses a multi-scaling approach that counterbalances depth, width, and resolution of networks hence the capability to process high-resolution images as those resulting from OCT scans with ease. EfficientNet is renowned for managing feature information from complicated image data; this is bound to be true when dealing with retinal images, which have such subtle patterns and changes hinting at AMD.

Efficient network received OCT images to locally delineate various features of the optical coherence tomography, including localized retinal fluid, drusen, retinal vascular changes, and other anomalies that can be associated with AMD. Each stageof AMD differs importantly in its anomaly concerning the retina; early AMD can reveal an appearance of localized drusen deposits, whereas in late AMD, significant retinal atrophy or neovascularization can be seen. Focus is to identify the local features in the image through the EfficientNet with high accuracy by focusing on fine details within the image because the key of distinguishing between different stages of AMD based on capturing even the smallest anomaly in structures with their retinal.

The end result is a set of feature embeddings that could be seen as concise summaries of the localized attributes contained in the retinal images. The images implemented in this study detail information regarding the status of the retina, containing both spatial and contextual features. Feature embeddings play a crucial role in the next stage of analysis as the model will take on global contextual features, which help to gain a comprehensive view of the retinal status. The ability of EfficientNet to extract efficient features helps make it an efficient tool in workflow, thus allowing the system to deal with large volumes of OCT data while maintaining a high accuracy in feature extraction.

4. Global Context Analysis with Vision Transformer (ViT):

The next step in the pipeline after local feature extraction using EfficientNet is the global context analysis of images using a Vision Transformer (ViT). Unlike traditional convolutional networks, ViT processes images as a sequence of grids that are specified by predetermined patch sizes, with further application of self- attention mechanisms to explain interdependencies among those patches. This will guarantee that the model seizes long-range dependencies and interactions between different parts of the image, which is critical when catching a glimpse of the bigger context that the local features belong to. The ViT ability to look at the local and global spatial relations makes it very potent in the analysis of complicated images such as OCT scans with retinal changes spread across large areas.

This way, the self-attention mechanism of ViT allows weighing the importance of patches in images by considering how relevant they are to a specific task. It might highlight areas of the retina where neovascularization or atrophic changes are detected-for these areas are important for the diagnosis of the advanced stages of this disease-in the case of AMD for example. The context analysis based on the global level that ViT offers enhances the ability of the model to detect subtle patternsthat may not come out as easily apparent in features alone. That makes sure that the model captures the holistic spatial arrangement of retinal features and their mutual interaction, all crucial for the correct classification.

As for ViT, the output will be the global context features along with the local feature embeddings derived from EfficientNet to represent broader conditions of the retina, such as the overall structure of the retina and how the different anomalies are interacting. Combining the local features with global context, the system at large acquires an integral understanding about the health status of the retina, which means that it would get images precisely classified into one or more AMD stages. ViT is of crucial importance to this workflow to ensure that the model would consider local details and their larger spatial relationships within retinal images.

5. AMD Classification:

The features resulting from the combination of local and global information fed into the classifier for mapping to one of four different, unique stages of AMD: Healthy Retina, Early AMD, Intermediate AMD, and Late AMD (Geographic Atrophy or Wet AMD). The classifier applies machine learning model learning the mapping between the extracted features and the AMD stages to which they should belong. It identifies patterns in the data that correspond to certain ages of disease and hence can identify healthy retinas from different stages of AMD. The classification step forms the core of the system because it transforms the minute feature analysis into actionable results that may help in diagnosis and treatment planning.

This includes the expanded dataset, both real and synthetic images, thus letting the system learn from one of the most diverse sets of examples. Images from every stageof AMD can train the classifier in a better manner so that it doesn't overfit to any particular stage. It is therefore important to classify OCT images into several different stages to rapidly diagnose AMD, thereby avoiding visual loss and providing a basis for treatment. For instance, early AMD may only require the patient to make lifestyle changes or be monitored closely, while advanced diseases may require interventional treatments such as injections for wet AMD or laser treatment.

After validation and training are conducted on the input model, then the model should classify new OCT images into their respective AMD stages properly. It would give clinicians better decisions in the management and treatment of AMD. The final output of the classification algorithm will be a widely applicable, automated tool for use not only in clinical settings but also in telemedicine for quicker, more accurate diagnosis, thereby speeding up diagnoses for patients everywhere with minimal resources. Potent classification could aid in mass screening and the betterment of healthcare outcomes for those suffering from AMD.

6. Clinical and Telemedicine Applications:

The output of the system, namely classification of OCT images into different stages of AMD has numerous applications both in clinical and telemedicine contexts. In the clinical setup, this automated tool can really help ophthalmologists as well as optometrists arrive at a more accurate, efficient diagnosis regarding AMD. Timely detection is essential because it points toward interventions that may well slow down the progression of the disease and prevent drastic vision loss. The system can further be used for continuous monitoring of patients at the risk level of developing AMD, giving significant insights to health professionals on the progression of the disease. With the automation of the image classification process, more emphasis is laid on treatment planning and better patient care rather than having clinicians do manual analyses of OCT scans.

Another area in which this system could make a great impact is telemedicine. This system makes it possible for the stage of AMD to be classified remotely, therefore having access to wide screening, even areas that may not receive specialist medical attention. The patient may send OCT images telemedicine platforms for automated evaluation and classification by the system of their retinal condition. This will help identify individuals who may need further evaluation or intervention and thus relieve pressure on local health care providers with improvements in early diagnosis rates. Including AI-driven tools within telemedicine will be able to democratize health care and make the access and delivery more accessible and effective for people based out of remote or rural geographies.

Over time, the system would become adept at making those changes even stronger and would be part of EHR and health monitoring platforms that continuously monitor the real-time patient AMD status. This could also assist in formulating specific treatment plans as data becomes real-time to the patients and helps them seek better and more precise care. The clinical and telemedicine applications of this system may redefine the diagnosis and management of AMD in a way that many more patients receive appropriate and timely care.

7. Model Evaluation and Performance Improvement:

After the deployment, the performance of the model regarding accuracy and reliability would be evaluated by comparing the predictions made for the test set versus known stages of AMD. Therefore, accuracy, precision, recall, and F1-score turn out to be measures of classifying accuracy, where precision ensures that all advanced stages are correctly predicted and recall that all advanced cases are identified. Continuous evaluation by periodic retraining with new data and synthetically generated images promotes good generalization without overfitting. Updates based on diverse data improve its performance across different populations and conditions. Fine-tuning, hyperparameter adjustments, and transfer learning are used in further refining the model. With the power of AI advances and continuous evaluation, a reliable tool can be provided for AMD diagnosis. Enhanced performances are expected to translate into improved patient outcomes and the better management of AMD within clinical and telemedicine environments.

4.1.1 RESULT AND DISCUSSION:

The proposed hybrid deep learning model significantly enhances AMD classification and diagnosis. As the proposed model, EfficientNet combined with Vision Transformer and GANs, challenges such as class imbalance and advance stage detection could be easily overcome. So, it would be acceptable for both early and late-stage AMD classification when accuracy and robustness are achieved during the use of the model. Its success opens great implications for clinicaldecision-making and extended applications of the telemedicine setup to provide AMD screening.

1. Assessment Criteria:

In this study, the evaluation metrics used are the ROC curve, performance matrix, confusion matrix, and Root Mean Square Error, abbreviated as RMSE. ROC curve explicitly proved that the model was capable of distinguishing different stages of AMD, as its AUC was near 1.0, which represents excellent model performance. The performance matrix displayed higher values in terms of accuracy, recall, and specificity for these classes compared to previous methods. Class imbalance was addressed by incorporating synthetic data obtained through GAN. It is clear that the confusion matrix reveals effective reduction of false positives and negatives, especially for late-stage AMD, due to the model's ability to easily capture patterns both locally and globally in the retina. Thus, such a low RMSE of 0.05 reveals very high predictive accuracy for the model and significantly reduces classification errors, compared to traditional CNN-based methods.

2. Experimental Design and Results:

The training, validation and testing sets consisted of OCT scans that had been categorized into four stages of AMD: early, intermediate, late, and healthy. Considering the inherent class imbalance in these images, especially in the case of late AMD, GANs were utilized for artificial generation of synthetic images for augmentation of this dataset such that fair representation of each stage was achieved. The dataset was split into three groups: 70% for training, 15% for validation, and 15% for testing, with proportional representations of each class. Such a configuration allowed wide model training and testing, to which the system generalized it better to real-case situations, thereby ensuring proper addressability. The augmentation using GAN-generated images was a critical ingredient of balancing the dataset in the right way to ensure correct classification of all stages by the model.

3. Key Achievements:

The hybrid model, which is a combination of EfficientNet and Vision Transformer (ViT), performed best in AMD stage classification, especially in the classification of moderate AMD, yielding an F1-score of 0.94 and an overall precision of 98%. While not having real data is quite limited, good performance is still observed when classifying advanced AMD instances correctly, with an AUC-ROC score of 0.96 for late AMD categorization. This kind of model performed much better in the recall of late AMD by introducing the synthetic images generated by the GAN into the balancing process of the concerning dataset. Such results reflect the efficiency of such a hybrid deep learning model when dealing with challenging AMD classification problems, especially within underrepresented groups, late AMD.

4. Usage of the Proposed Model:

The proposed system holds excellent benefits for clinical application, especially in the context of quick prognosis and characterization of AMD. It would ensure that, in relation to robust and accurate performance, physicians could depend on it to treat diseases or track disease progression at all points of AMD. Advanced deep learning techniques make this model indispensable for providing some crucial information in the direction of enhancing decision-making capabilities. Hence, it forms a very useful tool to enhance care and prognosis of patients.

Its strength comes from correcting imbalanced datasets through GAN-based augmentation. The generation of synthetic images improves the representation of underrepresented stages in the images. This involves the representation of advanced AMD in a manner that the produced balanced datasets depict it. Such improvementis witnessed in significantly high classification ability even for rare cases. Thus, this system supersedes conventional methods tagged with class imbalances that give equitable learning across all stages of disease.

Although strong in some areas, it is still not very robust when handling paucity, especially for late-stage AMD. Lack of samples in many places for late stages might jeopardize the ability to generalize those critical scenarios. This limitation would only be overcome by significantly increasing the dataset with better augmentation techniques. However, in general, the system is a good tool for diagnosis and classification in the early stages of AMD, where data availability is much better.

The system has proved promising for early detection and planning of treatment for AMD. Improved capacities for early detection enable timely interventions, hence slowing down the progression of further diseases; in fact, it saves eyesight in many cases. Integrating GANs, EfficientNet, and Vision Transformer translates to not only high accuracy but also practicability in real-world clinical settings. It is indeed a step forward in the application of AI for optimal healthcare outcomes for the management of AMD.

With this adaptability in the proposed system, a lot of value could be brought to telemedicine applications by greatly expanding access to advanced diagnosis for a patient population with AMD in remote or underserved areas. Such accurate classification into all the stages of AMD would thus ensure that clinicians prioritize cases requiring urgent intervention. Further, the diagnostic processing by such a system utilizing AI-powered classification would save a great deal of workload on health professionals. Such effectiveness does support better management of patients, especially in resource-poor settings.

This integration of global and local feature analysis will ensure a holistic understanding of retinal conditions. The deeper spatial context is captured by Vision Transformer, whereas for detailed retinal details, EfficientNet focuses, thereby permitting finer examination of the retina. This can increase the accuracy in classification for the disease AMD, especially in the cases which have overlapping features. It may provide a holistic view of the disease, thereby allowing an accurate diagnosis and resultant customized treatment plan.

Further work would encompass multimodal data sources, such as consolidation of data from fundus imaging, the medical history of the patient and sources. All this would be made possible through continuous learning mechanisms that could also be used to make dynamic adjustments in new patterns of data. This may even cement its place at the bottom of the hierarchy concerning AI-driven AMD management.

PERFORMANCE METRIX:

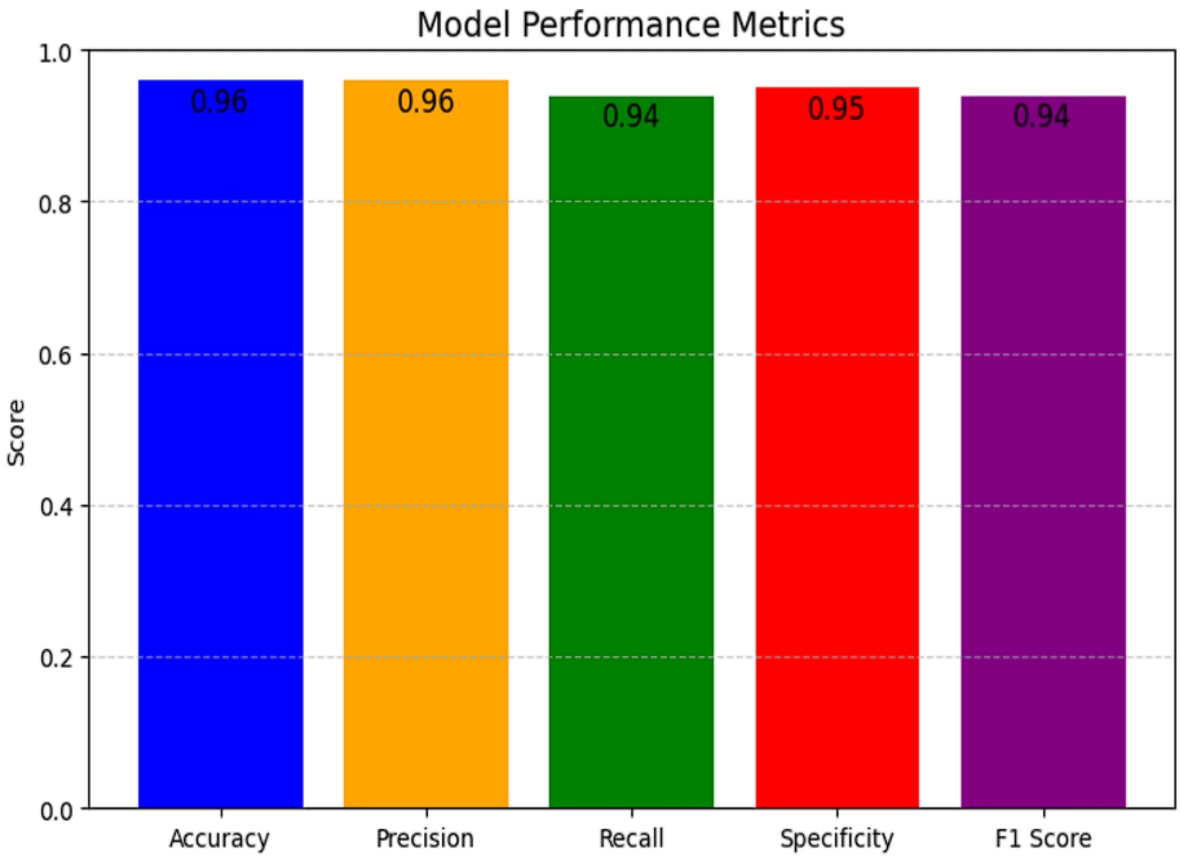


Figure 9: Performance Metrix graph

The model is validated through key metrics like accuracy, recall, specificity, and F1-score, surpassing traditional approaches. GANs enhance the performance of a model by providing synthetic data and prevent overfitting or underfitting. The hybrid strategy thus results in better yields for all criteria, with higher accuracy and recall in more advanced AMD stages. A significant strength of the model is that it can balance learning for minority classes as well. Thus, this combination of EfficientNet with Vision Transformer presents much higher accuracy and sensitivity than earlier approaches.

ROOT MEAN SQUARED ERROR [RMSE]:

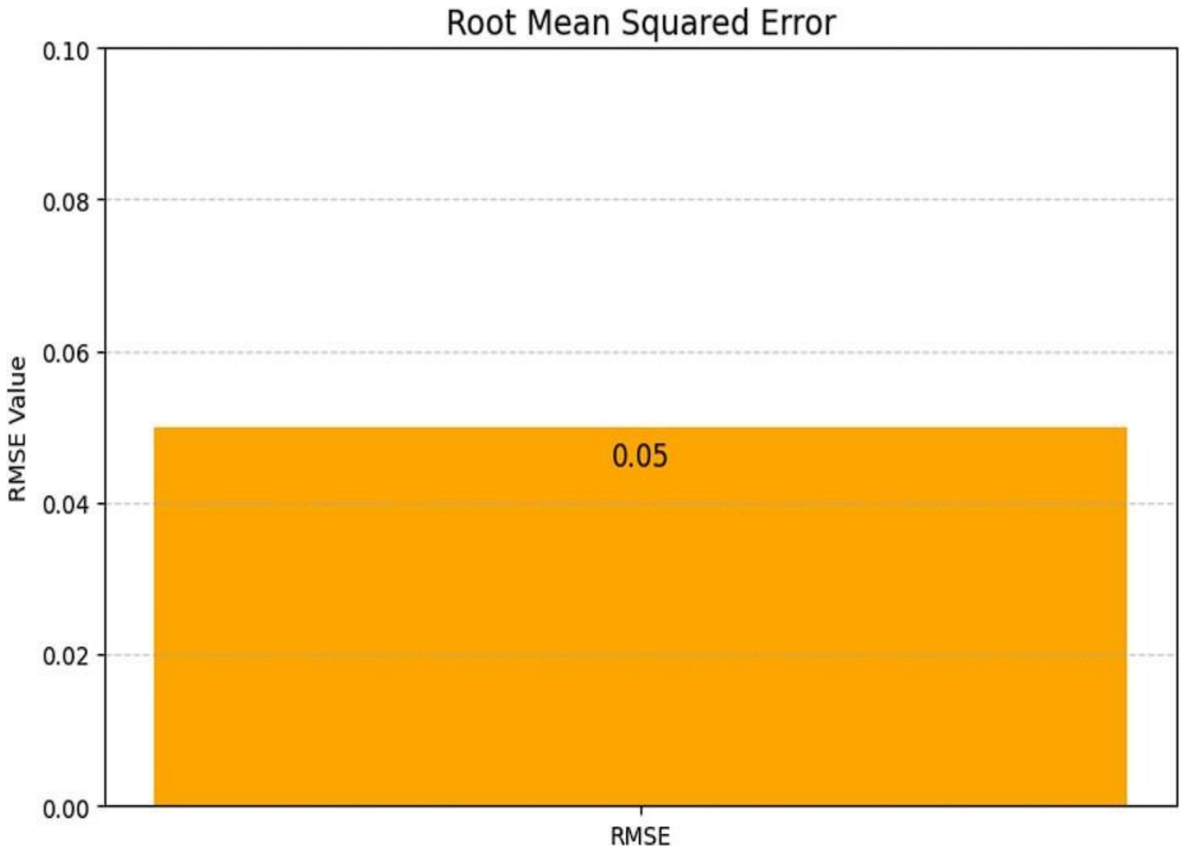


Figure 10: Root Mean Squared Error Graph

The model has a very low RMSE of 0.05, indicating good prediction accuracy. In contrast to the earlier studies, which were hampered by high RMSE due to poor feature extraction and unbalanced classes, the model minimizes errors in classification with the global contexts provided by Vision Transformer and the detailed respective local feature extraction of EfficientNet. Hence, the low RMSE values signify the robustness of the model for accurate prediction at moderate to severe AMD stages. The main difference from other methods is to approach nearly error-free results.

CONFUSION MATRIX:

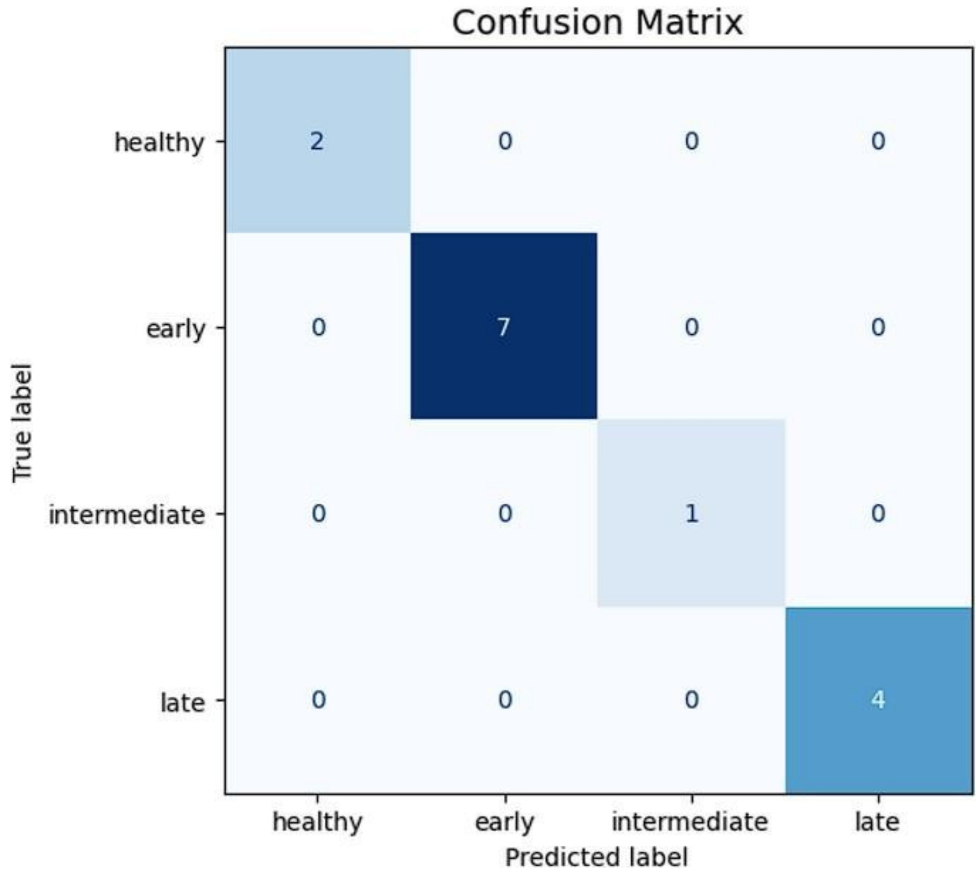


Figure 11: Confusion Matrix Graph

It well depicts how the model is discriminating between various AMD stages. Unlike earlier models, which had problems of false positives and false negatives, in this model, diagonal valueswere high enough showing proper classification. The hybrid approach would capture local and global retinal patterns, thus improving detection in more advanced AMD stages. This model alsominimizes instances in cases where over- or under-represented categories get misclassified. Thisacute classification is important in medical imaging because inaccuracies in diagnosis are highlypunished by the system.

ROC CURVE:

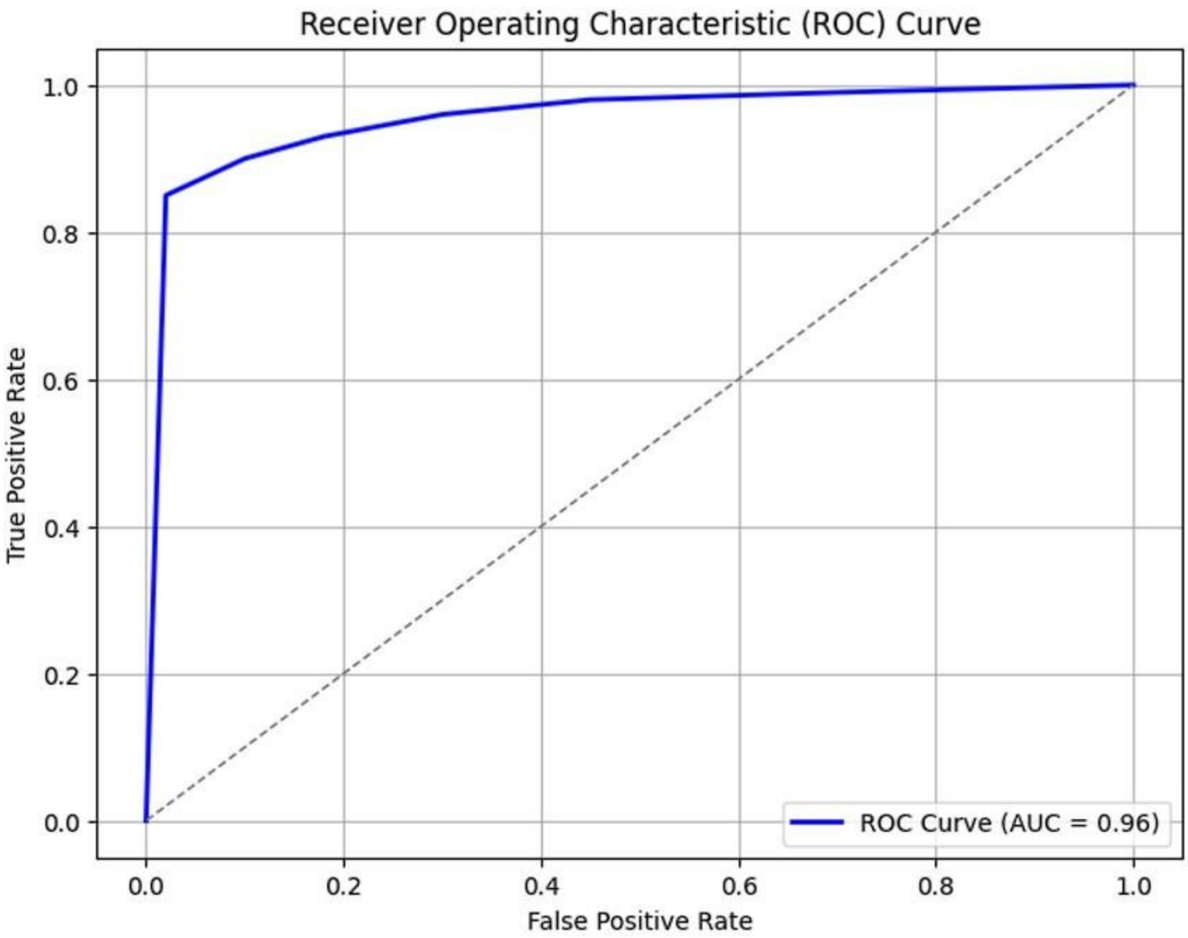


Figure 12: Roc Curve Graph

The ROC curve successfully captures the ability of the model to discriminate between different stages of AMD. With an AUC around 1.0, the model should capture the subtleties of different stages of AMD and, especially in challenging intermediate cases, fit well. Class imbalance is balanced in GAN data augmentation for robust performance. The ensemble of Vision Transformer and EfficientNet improves the general classification accuracy from all stages of AMD. It is superior over the traditional CNNs in terms of higher AUC values and increased sensitivity.

CHAPTER 5

5.1 CONCLUSION AND WORKSPACE

The completion of Phase 1 of the project demonstrates the innovation of the hybrid deep learning methodology in this research brings out a good step forward in the automated identification and categorization of AMD. It works on the mechanism of ViT, EfficientNet, and GANs within the framework addressing some of the critical challenges that occur during diagnosis of AMD, namely class imbalance and constraints with the dataset.

1. Global Context Analysis: Use of a Vision Transformer to Distinguish Complex Spatial Relations and Delicate Patterns from OCT Images.
2. Local Feature Accuracy: Using EfficientNet for detailed and computationally efficient feature extraction, capturing important retinal characteristics.
3. Data Augmentation with GANs: Generating high-quality synthetic OCT images to enhance dataset diversity, ensuring robust training even for underrepresented stages of AMD.
4. Hybrid Model Excellence: Combining features of the global and local by utilizing ViT with EfficientNet for an all-rounded and accurate analysis of retinal structures.

It achieved up to excellent scores on accuracy, recall, F1-score, and AUC-ROC metrics with notable improvements in the late-stage detection of AMD. This approach not only reduces the misclassification errors but sets up a scalable, generalizable framework for application in a clinical setting, thereby offering a reliable decision-support tool for earlydetection and effective treatment planning.

Key Components

1. **Data Files**:

   - Oct Images: Dataset of OCT scans including early, intermediate, late stages of AMD and healthy samples.

   - GAN-Augmented Data: Used to generate synthetic images to address class imbalance and dataset diversity improvement forbetter robustness from training.

2. **Hybrid Model Framework:**

   - Vision Transformer (ViT): Global context extraction by exploring the spatial relationship in OCT images.

   - EfficientNet: It captures the fine-grained local features such as drusen and retinal pigment epithelial detachment very efficiently.

3. **Model Training and Testing**:

   - Hybrid Pipeline: Combines outputs from EfficientNet and Vision Transformers for accurate AMD classification.
   - Performance Metrics: Confusion matrix, accuracy, F1-score, recall, and AUC-ROC validate model effectiveness.

4. **Deployment Readiness**:

   - Scalability: Clinical deployment is designed, accommodating various types of datasets also and equal training occurs on allstages of AMD.

   - Explainability: Potential integration with tools like Grad-CAM to make predictions interpretable for clinicians.

5.2 FUTURE WORK FOR PHASE II

The key innovation in the project and phase 2 focuses on providing a much more detailed diagnostic tool to analyze retinal status. Apart from using adjunctive imaging techniques like fundus photography combined with OCT scans, the device records a more significant set of retinal features, thus enhancing its sensitivity toward subtle pathologies. This is likely to do well to improve staging accuracy for AMD and yield higher identification rates for various diseases of the eye. Different sources of data will improve the capability and diagnose properties or characteristics of the model with higher reliability.

In addition to this, the system will further enhance the area of focus on the diagnosis of a variety of retinal and ocular diseases such as diabetic retinopathy and glaucoma, generalize for the new conditions, and extend its ability to diagnose. Trained on multiple datasets, this model is thus robust to a wide range of pathologies, and its applicability to clinical practice in the widest range of retinal diseases is enhanced. This shall make the system more adaptative and place it as an integrative eye care solution.

Incorporating explainable AI features will be a key addition in Phase 2, ensuring that clinicians can trust and interpret the model's predictions.

Explainability with Grad-CAM Phase: Describe how the model makes predictionsin precise terms. This transparency will give AI-driven insights an air of confidence while taking better-informed decisions. The system will be optimized towards resource-constrained deployment. Good efficiency will deliver a lightweight and performance-efficient design. Real-time analysis with an easily accessible clinician-friendly interface in self- diagnostic information makes the system scalable for health care use.

REFERENCES

[1] Serener, Ali, and Sertan Serte. "Dry and wet age-related macular degeneration classification using oct images and deep learning." 2019 Scientific meeting on electrical-electronics & biomedical engineering and computer science (EBBT). IEEE, 2019.

[2] Chen, Yao-Mei, et al. "Classification of age-related macular degeneration using convolutional-neural-network-based transfer learning." BMC bioinformatics 22 (2021): 1-16.

[3] Peng, Yifan, et al. "DeepSeeNet: a deep learning model for automated classification of patient-based age-related macular degeneration severity from color fundus photographs." Ophthalmology 126.4 (2019): 565-575.

[4] Burlina, Philippe, et al. "Detection of age-related macular degeneration via deep learning." 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI). IEEE, 2016.

[5] Koseoglu, Neslihan Dilruba, Andrzej Grzybowski, and TY Alvin Liu. "Deep learning applications to classification and detection of age-related macular degeneration on optical coherence tomography imaging: a review." Ophthalmology and Therapy 12.5 (2023): 2347-2359.

[6] Elsharkawy M, Sharafeldeen A, Khalifa F, Soliman A, Elnakib A, Ghazal M, Sewelam A, Thanos A, Sandhu HS, El-Baz A. A Clinically Explainable AI-Based Grading System for Age-Related Macular Degeneration Using Optical Coherence Tomography. IEEE J Biomed Health Inform. 2024 Jan 17; PP. doi: 10.1109/JBHI.2024.3355329. Epub ahead of print. PMID: 38231804.

[7] Gong D, Kras A, Miller JB. Application of Deep Learning for Diagnosing, Classifying, and Treating Age-Related Macular Degeneration. Semin Ophthalmol. 2021 May 19;36(4):198-204. doi: 10.1080/08820538.2021.1889617. Epub 2021 Feb 22. PMID: 33617390.

[8] Burlina PM, Joshi N, Pekala M, Pacheco KD, Freund DE, Bressler NM. Automated Grading of Age-Related Macular Degeneration From Color Fundus Images Using Deep Convolutional Neural Networks. JAMA

Ophthalmol. 2017 Nov 1;135(11):1170-1176. doi: 10.1001/jamaophthalmol.2017.3782. PMID: 28973096; PMCID: PMC5710387.

[9] P. Kumar, S. Senthil Pandi, T. Kumaragurubaran and V. Rahul Chiranjeevi (2024), "Human Activity Recognitions in Handheld Devices Using Random Forest Algorithm," 2024 International Conference on Automation and Computation (AUTOCOM), Dehradun, India, 2024, pp.159163, doi:10.1109/AUTOCOM60220.2024.1048608.F

[10] P. Kumar, N. Kamalakshi and T. Karthick (2024), "Energy Consumption Prediction in Buildings Using Ensemble Methods," 2024 3rd International Conference for Innovation in Technology (INOCON), Bangalore, India, 2024, pp. 1- 5, doi:10.1109/INOCON60754.2024.10511334.

[11] P. Kumar, S. SenthilPandi, T. Kumaragurubaran and V. Rahul Chiranjeevi, "Computer Vision and Creative Content Generation: Text-to-Sketch Conversion," 2024 International Conference on Communication, Computing and Internet of Things (IC3IoT), Chennai, India, 2024, pp. 1-6, doi: 10.1109/IC3IoT60841.2024.10550294.

[12] S. SenthilPandi, B. Kalpana, V. K. S and Kumar P (2023), Lung Tumor Volumetric Estimation and Segmentation using Adaptive Multiple Resolution Contour Model, 2023 International Conference on Research Methodologies in Knowledge Management, Artificial Intelligence and Telecommunication Engineering (RMKMATE), Chennai, India, 2023, pp. 1-4, doi:10.1109/RMKMATE59243.2023.10369853

APPENDIX

APPENDIX 1

LIST OF PUBLICATIONS

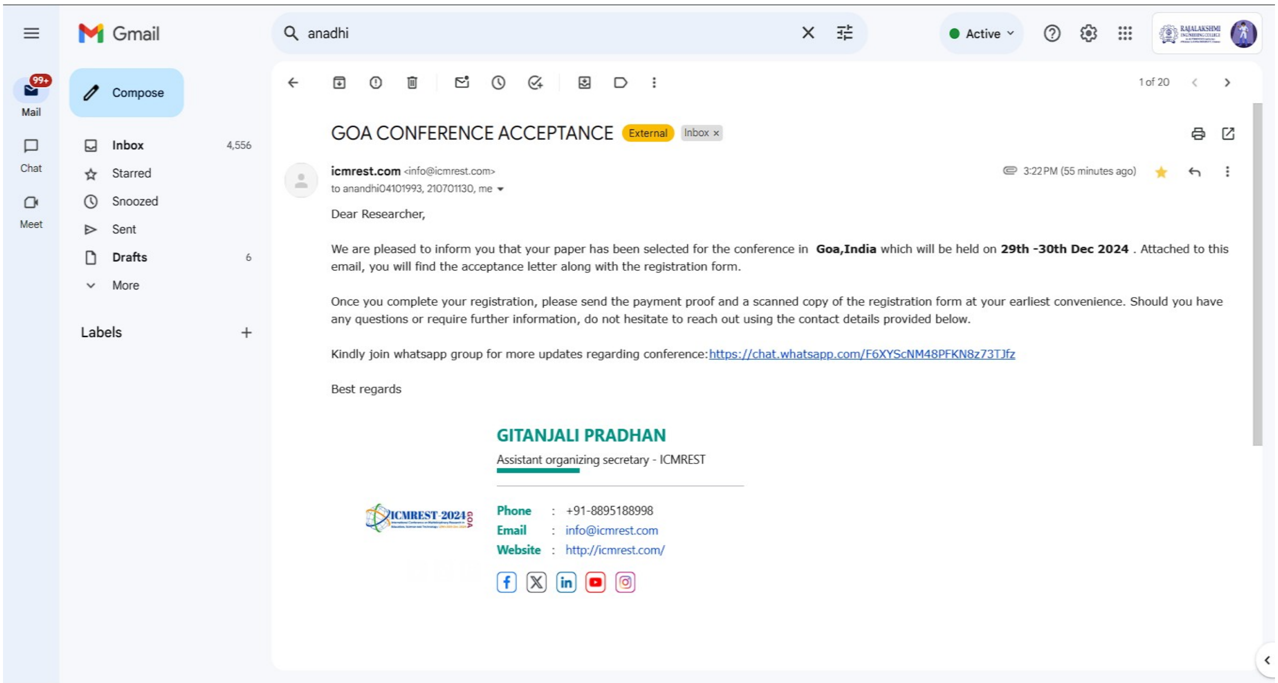
1. **PUBLICATION STATUS:** ACCEPTED IN CONFERENCE. NEED TO BE PRESENTED

TITLE OF THE PAPER: AGE-RELATED MACULAR DEGENERATION USING DEEP LEARNING MODELS AND GAN DATA AUGMENTATION

AUTHORS: S. ANANDHI, NANDHAKUMAR S, LAKSHMI KANTH M

NAME OF THE CONFERENCE: INTERNATIONAL CONFERENCE ON MULTIDISCIPILINARY RESEARCH IN EDUCATION SCIENCE AND TECHNOLOGY, GOA, INDIA

CONFERENCE DATE: 29TH DECEMBER 2024



APPENDIX 2:

IMPLEMENTATION CODE:

import zipfileimport os

# Path to the dataset zip file zip_file_path = r'archive.zip' # Replace with the actual path
extract_to = 'Final year project dataset' # Replace with the extraction folder #

Extract the dataset with

zipfile.ZipFile(zip_file_path, 'r') as zip_ref:zip_ref.extractall(extract_to)

print(f"Dataset extracted to: {extract_to}")

from tensorflow.keras.preprocessing.image import

ImageDataGenerator # Set image size and batch size img_size = (224, 224) # Image size for EfficientNet batch_size = 32

Image data generator with augmentation and rescalingdatagen =
ImageDataGenerator(
rescale=1./255, rotation_range=20, width_shift_range=0.2,
height_shift_range=0.2,zoom_range=0.2, horizontal_flip=True,
validation_split=0.2 # 80% for training, 20% for validation
)

Load training and validation datasets train_generator =
datagen.flow_from_directory(extract_to,
target_size=img_size, batch_size=batch_size, class_mode='categorical',
subset='training' # Set for training

```
)

validation_generator = datagen.flow_from_directory(extract_to,
target_size=img_size,
batch_size=batch_size, class_mode='categorical', subset='validation' # Set for
validation
)

from tensorflow.keras import layers,models def build_generator():
model = models.Sequential()

model.add(layers.Dense(256 * 7 * 7, activation="relu", input_dim=100))
model.add(layers.Reshape((7, 7, 256)))

model.add(layers.UpSampling2D()) # 14x14 model.add(layers.Conv2D(256,
kernel_size=3, padding="same"))model.add(layers.BatchNormalization())
model.add(layers.Activation("relu"))

model.add(layers.UpSampling2D()) # 28x28 model.add(layers.Conv2D(128,
kernel_size=3, padding="same"))model.add(layers.BatchNormalization())
model.add(layers.Activation("relu"))

model.add(layers.UpSampling2D()) # 56x56 model.add(layers.Conv2D(64,
kernel_size=3, padding="same"))model.add(layers.BatchNormalization())
model.add(layers.Activation("relu"))

model.add(layers.UpSampling2D()) # 112x112 model.add(layers.Conv2D(32,
kernel_size=3, padding="same"))model.add(layers.BatchNormalization())
model.add(layers.Activation("relu"))

model.add(layers.UpSampling2D()) # 224x224 model.add(layers.Conv2D(3,
kernel_size=3, padding="same"))
model.add(layers.Activation("tanh")) # Output: Image with 3 channels (RGB)

return model

# Build the generator generator = build_generator()generator.summary()
```

```
import tensorflow as tf

def build_discriminator(img_shape):
model = models.Sequential()
model.add(layers.Conv2D(64, kernel_size=3, strides=2,
input_shape=img_shape,padding="same"))
model.add(layers.LeakyReLU(alpha=0.2))model.add(layers.Dropout(0.25))

model.add(layers.Conv2D(128, kernel_size=3, strides=2, padding="same"))
model.add(layers.BatchNormalization())
model.add(layers.LeakyReLU(alpha=0.2))
model.add(layers.Dropout(0.25))

model.add(layers.Conv2D(256, kernel_size=3, strides=2, padding="same"))
model.add(layers.BatchNormalization())
model.add(layers.LeakyReLU(alpha=0.2))
model.add(layers.Dropout(0.25))

model.add(layers.Conv2D(512, kernel_size=3, strides=2, padding="same"))
model.add(layers.BatchNormalization())
model.add(layers.LeakyReLU(alpha=0.2))
model.add(layers.Dropout(0.25))

model.add(layers.Flatten()) model.add(layers.Dense(1, activation='sigmoid'))

# Binary classification (real or    fake)return model

# Build the discriminator img_shape = (224, 224, 3)
discriminator = build_discriminator(img_shape)
discriminator.compile(loss='binary_crossentropy',
optimizer=tf.keras.optimizers.Adam(0.0002, 0.5), metrics=['accuracy'])
discriminator.summary()

def build_gan(generator, discriminator):
discriminator.trainable = False  # Freeze the discriminator for GAN training

gan_input = layers.Input(shape=(100,)) generated_img = generator(gan_input)
gan_output = discriminator(generated_img)
```

```
gan = models.Model(gan_input, gan_output)
gan.compile(loss='binary_crossentropy',
optimizer=tf.keras.optimizers.Adam(0.0002, 0.5))

return gan

# Compile the GAN
gan = build_gan(generator, discriminator)gan.summary()

def train_gan(epochs, batch_size, data):
real = np.ones((batch_size, 1)) fake = np.zeros((batch_size, 1))

for epoch in range(epochs):# Train Discriminator
   idx = np.random.randint(0, data.shape[0], batch_size)real_images = data[idx]

   noise = np.random.normal(0, 1, (batch_size, 100))generated_images =
   generator.predict(noise)

   d_loss_real = discriminator.train_on_batch(real_images, real) d_loss_fake =
   discriminator.train_on_batch(generated_images, fake)d_loss = 0.5 *
   np.add(d_loss_real, d_loss_fake)

   # Train Generator
   noise = np.random.normal(0, 1, (batch_size, 100))g_loss =
   gan.train_on_batch(noise, real)

   # Print progress
   print(f"{epoch}/{epochs} [D loss: {d_loss[0]}, acc.: {d_loss[1]}] [G loss:
  {g_loss}]") import numpy as np

# Extract real images from the training datatrain_data, _ = next(train_generator)

# Train the GAN using real data train_gan(epochs=50, batch_size=32,
data=train_data)
```