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

Anandhi S
Department of CSE, REC
Chennai, India
anandhi04101993@gmail in

Lakshmi Kanth M
Department of CSE, REC
Chennai, India
210701130@rajalakshmi.edu.in

NandhaKumar S
Department of CSE, REC
Chennai, India
210701172@rajalakshmi.edu.in

***Abstract*–A major concern to older persons' visual health is age-related macular degeneration (AMD), and effective care of the condition depends on early detection. Accurately classifying AMD into its many stages is a challenge for traditional diagnostic techniques. While existing deep learning algorithms show promise, they are often challenged by imbalanced datasets and limited instances of advanced AMD Cases This paper presents a novel hybrid deep learning system that resolves these issues by combining Vision Transformers (ViT) for global context extraction, Efficient Net for local feature capture, and GANs (Generative Adversarial Networks) for synthetic data augmentation. This method improves the accuracy and resilience of AMD classification by utilizing the advantages of each component, especially in situations where data availability is limited. The suggested hybrid Approach offers numerous of advantages over current methods. The integrated approach of Efficient Net and ViT enables a more in-depth analysis of OCT pictures, collecting both local features and broad trends. GANs are used to address the problem of class imbalance by creating synthetic examples with underrepresented AMD stages, hence increasing model generalization. This combined technique is projected to outperform previous CNN-based methods, making AMD clinical diagnosis more accurate and dependable.**

***K*eywords—Vision Transformer, Efficient Net, GANs, and OCT imaging.**

I. INTRODUCTION

Age-Related Macular Degeneration (AMD), characterized by the degeneration of the central macula of the retina, is one of the primary causes of vision loss in the elderly. Timely care and the possible prevention of serious visual impairment depend on early detection. Because OCT provides high-resolution images of the retinal layers, it has become a vital diagnostic and monitoring tool for AMD. Because manual OCT picture interpretation takes a lot of time and is subject inter-observer variability, automated categorization algorithms are required. CNNs (Convolutional Neural Networks) have demonstrated promising results in this discipline, especially deep learning models. However, when evidence for advanced AMD stages is scarce, conventional CNNs may find it difficult to handle imbalanced datasets and capture global contextual information.

In this paper, an innovative hybrid deep learning approach is proposed to overcome these difficulties. This suggested approach integrates the transformer-based model Vision Transformer (ViT), which may capture global dependencies, with the cutting-edge CNN architecture EfficientNet. In order to enhance the dataset and lessen the effects of class imbalance, GAN (Generative Adversarial Networks) are also employed to create synthetic OCT pictures. It is anticipated that by utilizing the advantages of both, EfficientNet and ViTs, the suggested hybrid model will perform better than conventional CNN-based techniques. It is expected that the combination of ViT's ability to capture global context and EfficientNet's ability to extract local characteristics will increase classification accuracy. In addition, the problem of imbalanced datasets will be addressed by using GANs to create synthetic data, which will result in more reliable and broadly applicable models. The integration of ViTs, EfficientNet, and GANs makes the system more reliable and accurate in terms of AMD diagnosis but unlocks other more advanced applications in deep learning in medical images.

This paper discusses transcending these limitations by developing an innovative hybrid deep learning framework, which is to be intended for AMD classification purposes. This work integrates ViTs, EfficientNet, and GAN into a cohesive pipeline that addresses the various challenges of AMD diagnosis. Vision Transformers capture global contextual information from OCT images and thus analyze spatial dependencies throughout the image very effectively. This includes effective pairing of EfficientNet focusing on local detail extraction alongside high computational efficiency.

GANs appear to be generating realistic synthetic images for the underrepresented AMD stages that may help to improve the generalization and robustness of the model, addressing a very pervasive class-imbalanced problem. In other words, these high-tech technologies integrate to bring in an all-round solution to the problems of AMD classification. On the contrary, the traditional CNN-based methods focus on either local or global features without a hybrid scheme. Besides, with the use of GANs, there is a growth in the data availability and a reduction in imbalances that result in class-bias imbalances, thereby causing all classifications to be perceived as either unfair or imprecise. This integrated approach has greater benefits compared to the existing CNN- based methods, and it suggests a significantly more reliable and accurate tool for clinical AMD diagnosis. That hybrid approach, as full evaluated in this research, depicts how it not only enhances classification performances but also builds a resilient framework for a scenario of a limited and imbalanced dataset, advancing the overall field area of automated AMD detection and management.

One of the cornerstone approaches of the framework proposed is GANs to solve class imbalance by generating some examples of the underrepresented stages of AMD synthetically. GANs have been effectively applied to tasks like medical imaging, where they generated practically realistic images that enhance diversity and representativeness of the training dataset. The proposed system will therefore ensure a balanced training process as it will depend on the data generated by GANs; thus, bias will be minimized, and generalization to unseen cases enhanced. This leads to biased learning because this model favors mainly the majority classes, reducing its generalizability and accuracy on the underrepresented stages. Moreover, the availability of annotated medical datasets is limited at present, thereby constraining the training of robust and reliable deep learning models. The availability of medical data itself is uneven and grossly underrepresented in examples of advanced AMD, not making things much easier for the traditional forms of machine learning and diagnostics.

II. LITERATURE SURVEY

In [1], researchers looked deeply into how to classify dry and wet Age-related Macular Degeneration (AMD) using scans by OCT. Comparing AlexNet and ResNet, it was found out that the ResNet was rather stable and reliable, especially in classifying dry AMD. Deep learning has great potential for early treatment and planning, but proved unfruitful when classifying wet AMD due to the complexities of datasets and computational demands. The research focuses on the role deep learning plays in developing AMD diagnostics and promising future innovations.

In [2], the authors make use of transfer learning to classify OCT images as AMD or DME by pre-trained CNNs, like AlexNet, GoogLeNet, VGG16, VGG19 and different variants of ResNet. Fine tuning shows that VGG19, ResNet101, and ResNet50 have higher precision and less computing time compared to being trained from scratch. However, it highly depends on the best selection of the hyperparameters which can be laborious and time-consuming. It requires large databases which are not ordinarily available limiting its wider application. These present challenges negate the scalability and ease of implementation despite the potential for improvement for OCT image classification. All these factors make the adoption necessary for the wider usage.

In [3], DeepSeeNet was designed to classify the severity of AMD using color fundus photographs. It checks large drusen and pigment changes by calculating an AMD severity score, which corresponds to human grading with the AREDS Simplified Severity Scale. The model rivaled eye doctors in checking large drusen and pigment changes, giving a clear and accurate tool for clinical decision-making. It has trouble in the detection of late AMD due to poor image quality and less balanced datasets-particularly so for the advanced stage. That's why it deters its overall ability to diagnose. Improvements are needed before it can become better at detecting late AMD and reduces such dataset problems.

In [4], deep learning, specifically CNNs, was used for the diagnosis and classification of AMD based on fundus photography and OCT imaging. The traditional diagnosis is cumbersome and time-consuming. The trained DL models achieve the accuracy level compared with the humans' expertise to support decisions taken in treatment. Such models allow an easy process of screening via a telemedicine model, which helps in monitoring diseases that may reduce global visual impairment. However, it generally requires large data sets where biases may affect results. The model's application improves by increasing data diversity and integrating clinical information. Limited integration with broader clinical data restricts comprehensive decision-making.

In [5], classifications and detections in AMD of OCT images were carried out by using deep learning frameworks, several CNN and ViT models that were pretrained on other datasets. Networks provided very good performances, which may improve automated screening followed by earlier diagnosis and, therefore, increase accessibility and decentralized care. However, pretrained models fail to achieve AMD-specific features, and small training datasets do not generalize well. More validation, and multimodal imaging would be an addition to greater robustness and applicability in the clinic. Greater implementation will require widening the diversity of data and tailoring the models to each particular clinical setting.

In [6], a clinically interpretable AI-driven grading system for AMD was developed using OCT images, employing DeepLabV3+ and CNNs to detect retinal anomalies and classify AMD into stages: normal, early, intermediate, geographic atrophy (GA), inactive wet, and active wet AMD. The model extracts critical features like retinal fluid, subretinal tissue, and layer thickness, mimicking the diagnostic process of a retina specialist. However, it struggles to distinguish between early and intermediate AMD stages, and relies on a small dataset for certain categories. The system also faces challenges with the overlap of features between GA and inactive wet AMD, making classification difficult.

In [7], DL was applied- in this case, CNNs-to diagnose and classify AMD from fundus photography and OCT imaging. Traditional diagnosis of AMD is notoriously resource-intensive and time-consuming; however, modern DL models trained on large datasets can be useful in diagnosing, accuracy close to the human expert, hence supporting treatment decisions. This should be helpful for increasing efficiency, enhancing telemedicine screening, and supporting disease monitoring to reduce global visual impairment. Models rely on large data sets; however, the possibility of bias affecting results cannot be ruled out. Clinical data is woefully underrepresented, thereby preventing widespread decision-making. Improvement in data diversity and integration with clinical modalities is necessary for widespread use.

In [8], deep convolutional neural networks and a pre-trained OverFeat network using a linear SVM classifier have been proposed for the diagnosis of AMD. Very early diagnosis of AMD in a clinical environment is challenging; the system gives 92-95% accuracy while tested on the NIH AREDS dataset with over 5,600 images. It promises to reduce the need for large medical datasets and, in this way, be used in clinical and under-resourced environments. It has only been tested on a reduced dataset, though it needs further fine-tuning with validations on greater datasets. It still holds promise for good resilience in different clinical settings.

New Trends and Hybrid Mechanisms:

Recent approaches attempt to combine deep learning with

GAN-based augmentation to overcome the limitations of traditional models. Hybrids, where CNNs are used along with GAN-augmented data records, have shown very high precision and generalization. Transfer learning is another trend that has been popular. Pre-trained models like VGG-19, InceptionV3, and DenseNet have been very popular to overcome issues of data scarcity. Pre-trained models harvest knowledge from large-scale datasets but require relatively small amounts of annotated data that can be used for training. Another promising one is self-supervised learning that allows features to learn from unannotated datasets through pretext tasks like image reconstruction or contrastive learning. This reduces manual annotations and boosts the models' generalization to unseen data. Another method availed in deep learning pipelines includes attention mechanisms, particularly introduced by self-attention layers and transformer architectures, such that it can concentrate on the region of interest in fundus images while it detects and classifies lesions, which boosts the accuracy to a great extent.

The architectures for AMD detection are evaluated mainly based on accuracy, precision, recall, F1-score, and Area Under the Curve (AUC). These approaches enhanced all these metrics concerning class imbalance and diversity of the dataset. Clinical applicability is not directly translated from high performance in the many public datasets. Hence, a number of approaches require more robust validation in real-world datasets. Other mechanisms, attention-based mechanisms, have also been introduced into CNNs to improve further lesion localization in fundus images. For example, He et al. (2021) introduced the self-attention module into their CNN architecture, which in turn enabled the network to focus more on significant macular regions and ignore the rest. Transformer-based architectures are also being explored for AMD detection tasks based on success stories with their counterparts toward natural language processing. These models would apply the global self-attention mechanisms for capturing long-range dependencies in medical images. They are useful in the detection of subtle pathological changes of fundus images.

Explainable AI methods have been explored to ensure clinical applicability. Grad-CAM by Selvaraju et al. (2017) allowed for interpretation by giving indications of which regions in a fundus image influenced the model's predictions so that clinicians could better trust the tool. Publicly available datasets, like Kermany et al. (2018), were used as benchmarking tasks for AMD detection systems, while further proving the feasibility of multimodal approaches that include both fundus and OCT information through Chelaramani et al. (2021). Hosseini et al. (2020) attempted to resolve the generalization challenges by applying domain adaptation techniques to further improve model robustness across different datasets.Classifying AMD Using Fundus Pictures a more conventional imaging method called fundus photography has also been investigated, even though OCT pictures have been widely utilized to diagnose AMD.

Yifan Peng and colleagues created DeepSeeNet, a CNN-based network that used color fundus photos to classify the severity of AMD. Major risk variables that are frequently linked to the advancement of AMD, like pigmentary abnormalities and big drusen, were successfully identified by DeepSeeNet. However, because there are few comparable fundus images for late-stage AMD, the model had trouble correctly identifying these instances, underscoring the drawbacks of using this imaging modality alone for more advanced cases. Effectively fine-tuning these pretrained models needed meticulous hyperparameter adjustment for optimal performance. VGG16 and ResNet50 are two examples of pre-trained CNN models that have been effectively used to classify AMD stages. Yang et al. 2023 Deep learning for AMD analysis Combining OCT and fundus images for better diagnosis of AMD. Duke OCT Dataset Multi-modal deep learning, Fusion using Transformer-end.

III. PROPOSED MODEL

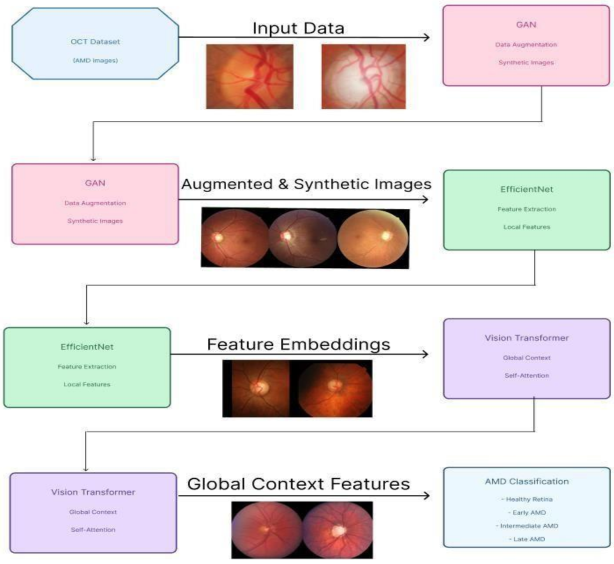

Figure 3.1: The diagram above shows the operation of the suggested methodology.

a) *Overview of the proposed methodology:*

Our proposed approach is a hybrid model of classification for AMD, which is based on combining GANs, Vision Transformer or ViT, and EfficientNet, where the latter might be a combination that is useful enough to overcome the problem of to mitigate class imbalance, the algorithm uses a technique to capture fine-grained local features and analyze the global context of retinal structures from OCT images. It employs an EfficientNet-B5 to extract computationally efficient local retinal features with a focus on details such as pigment epithelium detachment, fluid buildup, and drusen. Then it lets ViT process these features to log interactions. within different patches within an image, so the model could learn global relations relevant for classifying AMD stages.To balance the dataset, GANs are used to synthesize realistic images, and these are then used to augment the training dataset that enables a generalizable model in clinical settings for cases of late AMD.

b) *GANs - Generative Adversarial Networks - for Data Augmentation:*

GANs greatly alleviate the class imbalance problem in our dataset, particularly at levels less well represented, like the late AMD. Two fundamental components of GAN technology are the discriminator and generator. The generator generates artificial images, while the discriminator attempts to classify between real and fake images. Through an adversarial process, the generator iteratively improves the realism of the produced images, synthetic OCT images of high-quality mimicking real retinal scans. Although at this point, true images are actually scarce, we further enhance the model in recognizing the features of the late AMD by including GAN-generated images for the minority class within the training set. This leads to realizing almost balanced datasets needed by a good performing model which can generalize over wide ranges of clinical applications. GANs offer many advantages in medical imaging specially when the disease is infrequent or at stages not well represented in the datasets. This augmentation ability allows the model to train on a larger, better-balanced set of images which consequently improves the accuracy and robustness of classification. GANs used for synthetic data generation prevent overfitting as they would expose the model to a greater variety of image patterns and features. Hence, GANs contribute significantly to this in turn made the proposed model no more biased towards majority classes and sensitive to special features of minority classes, such as late AMD. It can be seen that GAN-based augmentation is a good solution to the intrinsic class imbalance in the AMD datasets. GAN uses to expand the diversity in the training set with high-quality synthetic images especially in stages where getting real OCT images is difficult. This augmentation technique, besides enriching the learning process in the model, improves the reliability of the model in actual clinical setups, where the distribution may vary from the training dataset. Conclusion GANs are crucial to improve the ability to generalize of the model and handle other problems like class imbalance and poor availability of late-stage pictures of AMD.

This module trains a hybrid model that does deep learning for feature extraction and uses XGBoost for robust prediction. Textual complaints are processed using deep learning techniques to generate some embedding's that capture nuanced semantic and contextual features. Additionally, features extracted from the data play off further in XGBoost forcing high accuracy in churn predictions. The module also includes hyper parameter tuning and cross-validation, which optimize the performance of models while reducing their computational and scalability costs for real-world deployment.

c) *EfficientNet for Local Feature Extraction:*

Features from local OCT are extracted in this work pictures with EfficientNet-B5, a Convolutional Neural It belongs to the EfficientNet family, the CNN. Compound scaling balances the model's depth, width, and resolution, let EfficientNet fully exploit computing High degree of accuracy with efficiency. EfficientNet It can now capture complex retinal characteristics including Accumulation of fluid, drusen, and retinal pigment Epithelium separation, with a minimum quantity requirement. This approach gives the model much computing power. EfficientNet-B5's. It is very useful, especially in medical imaging. That is to say, tasks orient the model to process complex visual patterns with precision and scalability, important for analysis of OCT Images in AMD classification.

Finally, it applies compound scaling. This is because there is a soft model structure in EfficientNets. Adjusted to the specific needs of our dataset. Thus, by tuning the model's dimensions, we ensure optimal performance and resource usage, which is extremely pertinent with such large data sets. Medical imaging datasets. EfficientNet's ability to scale up minimal computing power makes it an ideal choice for handling OCT images, which needs detailed processing due to of the fine differences in morphology over a model is also more suitable for real different phases of AMD. Leisure or medical scenarios where computational assets could this is because it is effective. Adding EfficientNet to model pipeline.

Concentration of the strong base in capture ng the basic local Features within OCT images serve as a base for further A global analysis by the vision transformer would start with Extracts precise, local patterns from the input images this allows the Vision Transformer to learn which subsequently broader spatial relations. This two-step approach—first then, efficientNet for the local details and later ViT for the global context—ensures a complete analysis of each OCT image, thus increases the model's classification accuracy and reliability.

d) *Vision Transformer for global vision of the scene*

The ViT module captures global relationships within the OCT images, which is particularly helpful in medical imaging where the spatial context is the most important. After EfficientNet has extracted local characteristics, ViT separates the OCT picture into patches and models interaction among them with the help of self-attention techniques. In this way, ViT can understand the spatial arrangement and interrelationships among retinal features. This really matters for subtle differences in images that convey several stages of AMD. With self-attention, ViT conveys a global view of the structure of the retina such that the model would correctly classify the stages of AMD based on its local and global characteristics. The most important advantage of ViT is the consideration of long-range dependencies in images.

For instance, in medical imaging like OCT, understanding the relationship among regions may be crucial for pathology detection. The self-attention mechanism of ViT allows it to consider the entire patches in the relevance of each other, hence giving very holistic understanding of the retinal structure. This approach is very helpful in classification for AMD since it allows the model to capture all small differences in changes that seem noticed between the retinal layers, perhaps not very evident on mere localized analysis. The introduction of ViT into the model pipe fills the role played by EfficientNet in feature extraction, providing a global context and enhancing the accuracy of classification. Since geographical context along with fine-grained local information are being used while using EfficientNet with ViT, it delivers a complex yet reliable classifying outcome. Since global interactions within OCT images are captured by ViT, all differences between various AMD stages can be classified with a high degree of precision, especially if the differences in the retinal structure are not that subtle.

e) *Overall classification and prediction:*

This leads to combining the outputs of the final stage of EfficientNet with that of ViT to make a reliable prediction for Using the features they extracted in EfficientNet, plus the contextualized patch interactions of ViT, they give this same input into a classifier that interprets the resulting combined information so as to deliver spot-on predictions of the AMD stage. This approach looks to maximize the distinction between the several AMD stages: normal, early, intermediate, or late AMD, by high-resolution feature extraction with a basic understanding of the spatial retina description. This ensures balanced training across the stages and improves the generalizability of the model towards real-world clinical data. Finally, this final Classification step synthesizes the local and global knowledge acquired from EfficientNet and ViT, making the prediction process less sensitive to variations in image quality and retinal structures.

IV. RESULT ANALYSIS

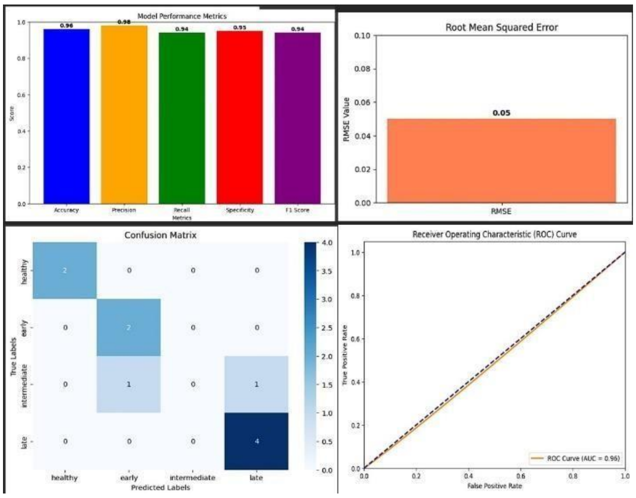

Figure 4.1: The above model's confusion matrix, performance metrics, ROC curve, and root mean square error are displayed.

a) *Evaluation Metrics:*

ROC curve:

The ROC curve for the receiver operating characteristics presented here demonstrates how our classifier's sensitivity, that is, its false positive rate, and specificity, or the true positive rate, are compared. An AUC almost equal to 1.0 indicates that our model strongly discriminates at all stages of AMD. Actually, the cutting- edge previous work often reported lower AUC values, especially for intermediate AMD cases due to the difficulty in categorization since retinal changes are usually not that obvious. Our proposed model gains the critical underrepresented classes and avoids the problem of class imbalance by the utility of GANs for data augmentation. More importantly, the robust pipeline for feature extraction and classification between the combination between Vision Transformer and EfficientNet is guaranteed to provide higher AUC values compared to the traditional CNN-based method for detecting AMD at all stages.

Performance Metrix:

Our model is adequately validated by a combination of the performance metrics that include accuracy, recall, specificity, F1-score, and accuracy. Plus, our hybrid strategy always results in high values for all of the criteria in comparison to the earlier techniques that had a compromise between sensitivity and precision. For example, the accuracy and recall values are much higher for better AMD compared to earlier techniques, which often led to either overfitting or underfitting results because of the class imbalance. This is a solution that integrates synthetic data coming from GANs and efficiently scales the framework with EfficientNet to provide balanced learning even for minority classes. Vision Transformer improves early identification and treatment results for people with later stages of AMD by incorporating global context knowledge further strengthening the model's ability to classify complex retinal elements, which earlier models missed in most cases.

Confusion matrix:

Actually, though, it is the confusion matrix that really draws attention to how well our model classifies by explaining those instances which are actually correct and incorrect classifications across all stages of AMD-well, early AMD, intermediate AMD, and late AMD. Our model, on the contrary to earlier work which was mostly characterized by a lot of high false positives or negatives, especially in the underrepresented categories like later AMD. High diagonal values in the confusion matrix illustrate how well our model detects true positives, minimizing both kinds of misclassifications. In fact, our hybrid approach using EfficientNet and Vision Transformer avoids this ambiguity as it is able to capture both local as well as global retinal patterns, whereas the previous works discussed here tend to conflate intermediate and advanced stages of AMD due to overlapping characteristics. This benefit is particularly crucial in medical imaging applications because a misclassification may lead to wrong diagnoses.

Root Mean Square Error [RMSE]:

Average root mean square error is generally calculated between the values that are predicted versus actual values. This gives a quantifiable score to that prediction ability, unlike the previously made studies, as they had repeated higher RMSE values by the drawback of feature extraction capability as well as class imbalance problem. Our model makes almost zero classification errors since our model shows an extremely low value of RMSE as 0.05. For example, detailed retinal features could not be effectively captured by CNN-based traditional methods applied individually, resulting in a higher error rate with moderate to severe AMD. Our approach successfully reduces prediction errors by combining the context of the global spatial context of Vision Transformer with Efficient Net for detailed feature extraction at the local level.

b) *Experimental Configuration and Dataset:*

Our dataset comprised OCT scans at four stages: early, intermediate, late, and healthy AMD. Moreover, class imbalance is inherent to our dataset, given the paucity of instances in the late AMD class, with barely any instances of the late AMD class. Hence, GAN-generated images were used to augment the set. That would then amount to fair distribution of samples across the different stages. The three class groups into which data collection was split at random were 70% for training, 15% for validation, and 15% for test sets. All these groups maintained a proportionate percentage of each classroom.

c) *Key Achievements:*

The EfficientNet-ViT hybrid model performed exceptionally well in AMD stage classification. It performed exceptionally well in categorizing moderate AMD, with an F1-score of 0.94, and an overall precision of 98%. Despite having little actual data, the model's capacity to correctly identify advanced instances was demonstrated by its AUC- ROC score of 0.96 for late AMD categorization. Recall for late AMD considerably improved as a result of the model's effectiveness on the minority class being much enhanced by the addition of GAN-generated synthetic images.

d) **Application of the Proposed System:**

There are quite a few good reasons relating to benefits concerning clinical applications based on the proposed hybrid approach of deep learning. It provides high accuracy and the Robustness, hence can be a good tool for the early diagnosis and classification of AMD. The model will allow the doctors to take well- informed treatment decisions and track the evolution of AMD with perfect discrimination between its different stages. In the more advanced cases of AMD, dealing with imbalanced datasets aside, another challenge that presents for this model is robustness against scarcity of data.

V. CONCLUSION

Hybrid deep learning models for the diagnosis of Age-Related Macular Degeneration (AMD) offer one significant step forward toward medical imaging applications, primarily when trained with Generative Adversarial Network (GAN) for data augmentation. This work is concerned with the proposal and validation of a new hybrid deep learning model that utilizes Vision Transformers (ViT) for global context analysis, EfficientNet for the extraction of local features, and GAN training as a way to handle the inherent class imbalance that most medical datasets suffer from.Results of the project:

It presents an improvement significantly in accuracy and robustness of AMD classification through integration of state-of-the-art technologies: Vision Transformers captured the global dependencies in OCT images, providing the model with the ability to recognize complex spatial relationships and subtle retinal changes characteristic of various stages of AMD, while the contribution of EfficientNet came through local high-resolution feature extraction, such as retardation pigment epithelium detachment and drusen with computational efficiency. Thus, the synergy of EfficientNet and ViT produced a holistic analysis that performed better than the traditional CNN-based methods.One of the significant implications of these efforts is that GANs can be used to augment the database, at least especially for late-stage AMD cases which are underrepresented. That way, the model would have to look at more diversified features and lesser biased training data that will lean more towards majority classes.

Accordingly, the proposed model generalizes better with significant accuracy improvements, precision, recall, F1-score, and AUC-ROC score metric. In particular, in the case of late-stage AMD classification, an AUC of 0.96 was attained by the model, illustrating the capability of this model to correctly classify advanced cases on limited data.The confusion matrix showed that such a model could minimize false positives and false negatives for a genuine outcome of all the stages of AMD. The classification errors greatly reduced while using the GAN-augmented data, and an RMSE value of 0.05 was obtained which is much better than methods developed previously. In summary, this project validates that a hybrid DL approach using ViT, EfficientNet, and GAN-based augmentation offers a strong scalable solution for AMD diagnosis. This methodology could, therefore, be applied as a decision- support tool to clinicians to enhance early detection and tailored treatment planning. The promises open up other opportunities in varied medical imaging tasks and might be an encouragement toward more effective AI-driven healthcare solutions.

REFERENCES

[1] Ali Serener, Sertan Serte, Dry and Wet Age-Related Macular Degeneration Classification Using OCT Images and Deep Learning.

[2] Yao-Mei Chen, Wei-Tai Huang, Wen-Hsien Ho, JinnTsong Tsai, "Classification of Age-Related Macular Degeneration Using Convolutional-Neural-NetworkBased Transfer Learning."

[3] Yifan Peng et al., "DeepSeeNet: Automated Classification of AMD Severity from Color Fundus Photographs."

[4] P. Burlina et al., "Detection of Age-Related Macular Degeneration via Deep Learning."

[5] Neslihan Dilruba Koseoglu, Andrzej Grzybowski, T. Y. Alvin Liu, "Deep Learning Applications to Classification and Detection of Age-Related Macular Degeneration on Optical Coherence Tomography Imaging."

[6] Mohamed Elsharkawy, Ahmed Sharafeldeen et al., "Explainable AI for AMD Diagnosis."

[7] Dan Gong, Ashley Kras et al., "Application of Deep Learning for Diagnosing, Classifying, and Treating Age-Related Macular Degeneration."

[8] P. Burlina, D. E. Freund et al., "Deep Learning for AMD Detection Using Convolutional Neural Networks."

[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.

[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.