Optical Insight: Enhancing Ophthalmic Diagnostics with Automated Detection of Retinal Abnormalities

Dilshan I. De Silva*, Dinuka R. Wijendra, Kithmina S. Siriwardana, Shehan N. W. Gunasekara, Udesh Piyumantha, and Sahan P. Thilakaratne

Faculty of Computing, Sri Lanka Institute of Information Technology, Malabe, Sri Lanka
Email: dilshan.i@sliit.lk (D.I.D.S.); dinuka.w@sliit.lk (D.R.W.); kithminasiriwardana13@gmail.com (K.S.S.); shehangunasekara2019@gmail.com (S.N.W.G.); udeshpiyumantha@gmail.com (U.P.); sahanpradeeptha@gmail.com (S.P.T.)
*Corresponding author
Manuscript received September 14, 2024; revised October 11, 2024; accepted March 18, 2025; published June 11, 2025

***Abstract*—Early and accurate detection of retinal diseases is crucial for preventing vision loss, yet traditional diagnostic methods remain limited by subjectivity and inefficiencies. This study introduces an Artificial Intelligence (AI)-driven diagnostic system leveraging hybrid deep learning models to detect Glaucoma, Macular Hole, Central Serous Retinopathy, and Drusen using fundus images. By integrating multiple architectures, including Residual Network (ResNet), Visual Geometry Group 16-layer network (VGG16), Densely Connected Convolutional Network (DenseNet), U-shaped Network (U-Net), and You Only Look Once version 8 (extra-large variant) (YOLOv8x), the system enhances diagnostic precision and generalization across diverse imaging conditions. Key innovations include the hybrid ResNet-VGG16 and DenseNet-VGG16 models, which significantly improve detection accuracy for Drusen and Central Serous Retinopathy, respectively. Additionally, the U-Net-ResNet hybrid architecture mitigates overfitting, ensuring more reliable Macular Hole detection, while the YOLOv8x object detection model outperforms traditional approaches in Glaucoma localization by accurately identifying the optic disc. These models, integrated into a web-based diagnostic platform, achieved sensitivities and specificities exceeding 95%, establishing a new performance benchmark for automated ophthalmic diagnostics. This research advances medical image analysis by demonstrating the efficacy of hybrid deep learning models, offering a scalable AI solution for early retinal disease detection. Its integration into clinical workflows highlights its potential to transform ophthalmic care, enhancing accessibility and improving patient outcomes.**

***Keywords*—optical insight, ophthalmic diagnostics, retinal abnormalities, human eye**

I. Introduction

The human eye is a highly specialized organ responsible for most sensory input, enabling individuals to interpret their surroundings. Its intricate anatomy, including the retina, optic nerve, and various tissue layers, functions cohesively to process visual information. Among these structures, the retina plays a crucial role in visual processing, making retinal diseases particularly concerning due to their potential to cause irreversible vision loss.

Retinal diseases such as Glaucoma, Macular Hole (MH), Central Serous Retinopathy (CSR), and Drusen present considerable diagnostic challenges due to their subtle and progressive nature [1]. Glaucoma, a leading cause of irreversible blindness, is often diagnosed late due to the reliance on subjective fundus image interpretation. MH, characterized by a break in the macula—the central region of the retina responsible for sharp vision—is difficult to detect in its early stages, leading to delayed intervention and potential vision loss. CSR, involving fluid accumulation beneath the retina, can cause sudden visual impairment, with symptoms varying among patients, necessitating advanced imaging techniques for accurate diagnosis. Drusen, yellow deposits beneath the retina, serve as early indicators of Age-Related Macular Degeneration (AMD). While not all Drusen result in severe vision loss, their similarity to other retinal features complicates detection, requiring high-resolution imaging and expert analysis. Given the progressive nature and diagnostic complexity of these conditions, there is a pressing need for advanced tools that improve early detection and treatment outcomes.

Traditional diagnostic methods for retinal diseases are time-consuming and prone to variability due to their reliance on subjective evaluations [2]. Human expertise introduces the risk of diagnostic errors, particularly in complex or borderline cases where pathological features are subtle and easily overlooked. Additionally, factors such as image quality, patient demographics, and comorbid conditions further complicate the diagnostic process, making consistency and accuracy difficult to achieve. These challenges necessitate technological advancements to improve diagnostic reliability, leading to the adoption of sophisticated imaging techniques in ophthalmology.

To mitigate these challenges, advances in imaging technologies, such as Optical Coherence Tomography (OCT) and fundus photography, have provided detailed views of retinal structures, significantly enhancing diagnostic capabilities. However, despite these advancements, the diagnostic process remains limited by the subjectivity of human interpretation. To overcome these challenges, automated diagnostic tools leveraging deep learning and artificial intelligence have emerged as promising solutions. While these AI-driven approaches improve diagnostic efficiency, single-model architectures often fail to generalize effectively across different clinical conditions. Addressing these shortcomings requires a more robust approach that integrates multiple models for enhanced accuracy and adaptability.

Building on these advancements, this research proposes a robust solution to the limitations of single-model architectures by introducing a hybrid deep learning approach for improved retinal disease detection. By leveraging multiple deep-learning architectures, this research aims to create a comprehensive diagnostic system that enhances accuracy, robustness, and adaptability. This integration has the potential to bridge the gap between technological advancements and practical clinical applications.

By integrating hybrid deep-learning models into clinical practice, this study has the potential to transform ophthalmic diagnostics, addressing current limitations and paving the way for more accurate and efficient detection methods. Key contributions include the development of a patient-specific advisory system, integration with telemedicine platforms for remote consultations, and the use of localized datasets for improved clinical relevance. Together, these advancements have the potential to enhance early detection, improve disease management, and reduce dependency on human expertise.

To realize these contributions, the study focuses on the following key objectives:

- Enhance diagnostic precision for Glaucoma, MH, CSR, and Drusen by integrating hybrid models.
- Develop a web-based application that integrates these models for practical use in clinical settings.
- Address challenges related to overfitting, generalization across diverse datasets, and variations in image quality.

The rest of the paper is organized as follows: Section II elaborates on the theoretical foundations underpinning the study's research approach. Section III reviews the relevant literature on automated retinal disease detection using deep learning techniques. Section IV outlines the data preprocessing, model development, and evaluation procedures employed in the study. Following this, Section V highlights the innovative features of the proposed system, including patient-specific advisory solutions, telemedicine integration, precise delineation of disease-affected regions, and multi-disease detection capabilities, alongside the performance metrics that demonstrate the system's robustness and practical applicability. Section VI interprets these findings in the context of current diagnostic practices and highlights the potential implications for clinical use. Finally, Section VII summarizes the study's contributions and proposes directions for future research.

II. Theoretical Framework

Deep learning has become a pivotal tool in medical imaging, particularly in ophthalmology, where it enables the automation of complex tasks such as image classification and segmentation. These tasks are critical for the accurate detection and diagnosis of retinal diseases. The application of deep learning models in medical imaging is driven by their ability to learn from vast amounts of data, capturing intricate patterns and features that often go beyond human perception. However, individual deep learning models, despite their powerful capabilities, often face challenges such as overfitting, limited generalization across diverse datasets, and difficulties in capturing the diverse and complex features present in medical images.

The integration of multiple deep learning architectures into hybrid models helps address the inherent limitations of single-model approaches. Hybrid models are designed to leverage the complementary strengths of different architectures, providing a more robust and comprehensive solution for complex medical imaging tasks. By combining models that excel in different aspects of image analysis, hybrid approaches can tackle challenges related to overfitting, feature extraction, and generalization. This ultimately results in more accurate and reliable diagnostic tools.

The selection of ResNet, VGG16, U-Net, DenseNet, and YOLOv8x in this study is grounded in the unique strengths and capabilities of each architecture, which are particularly well-suited for different aspects of retinal disease detection.

- **ResNet:** ResNet is known for its deep residual learning capabilities, which allow the model to learn complex features without suffering from the degradation problem that typically occurs in very deep networks. This architecture is particularly effective for tasks that require the identification of detailed patterns within retinal images, making it ideal for improving the depth and accuracy of feature learning in the hybrid model.
- **VGG16:** VGG16 is recognized for its simple yet powerful architecture, which excels at feature extraction through its deep convolutional layers. The uniformity and depth of VGG16 allow it to capture detailed spatial hierarchies in images. This makes it a complementary addition to ResNet's strengths. When combined with ResNet, VGG16 contributes to a more robust and detailed analysis of retinal images, particularly in the extraction of fine-grained features critical for disease detection.
- **U-Net:** U-Net is highly regarded for its effectiveness in image segmentation, particularly in medical imaging applications where precise localization of regions of interest is crucial. Its encoder-decoder architecture is specifically designed to identify both the context and the localization of features within an image. In the context of retinal disease detection, U-Net is instrumental in identifying and segmenting key areas such as the optic disk, which is essential for diagnosing conditions like Glaucoma.
- **DenseNet:** DenseNet introduces a densely connected convolutional network that enhances feature reuse and improves the flow of gradients through the network, leading to more efficient learning and better generalization across datasets. DenseNet's architecture is particularly beneficial in medical imaging, where diverse and complex features must be learned from high-dimensional data. By promoting feature reuse, DenseNet helps the hybrid model generalize more effectively to new and varied datasets, reducing the risk of overfitting.
- **YOLOv8x:** YOLOv8x is a state-of-the-art object detection model known for its speed and accuracy in localizing objects within images. Its ability to quickly and accurately identify specific regions, such as the optic disk in retinal images, makes it an invaluable component of the hybrid model. YOLOv8x's object detection capabilities complement the feature extraction and segmentation strengths of the other models, enabling the hybrid model to perform precise and rapid diagnostic assessments.

The integration of ResNet, VGG16, U-Net, DenseNet, and YOLOv8x into a single hybrid model framework is justified by the need to address the challenges of retinal disease detection. Each model contributes to a different aspect of the diagnostic process: ResNet and VGG16 enhance feature learning and extraction; U-Net provides accurate segmentation; DenseNet improves generalization; and

YOLOv8x ensures precise localization. This combination enables the hybrid model to capture a broader range of features and manage the variability and complexity of retinal images, resulting in superior diagnostic performance.

The theoretical advantage of this hybrid approach lies in the synergistic effect of combining models with complementary strengths. By leveraging the unique capabilities of each architecture, the hybrid model can overcome the limitations of individual models, achieving better accuracy, robustness, and generalization in the detection of retinal diseases.

The theoretical framework underpinning the integration of these deep learning architectures directly supports the research objectives. These include enhancing diagnostic precision, developing a practical web-based application, and addressing challenges such as overfitting and generalization. The hybrid model's superior performance in detecting retinal diseases such as Glaucoma, MH, CSR, and Drusen is a direct result of this thoughtful and theoretically grounded integration. This framework not only advances the field of medical imaging but also paves the way for future innovations in AI-driven diagnostic tools.

III. Literature Review

Considerable progress has been made in the automated detection and segmentation of Drusen, a critical biomarker in AMD. Kim *et al.* [3] emphasized the urgent need for accurate Drusen segmentation techniques in diagnosing AMD, as these are essential for both monitoring disease progression and ensuring accurate diagnosis. Their study reported a segmentation accuracy of 92%, demonstrating the effectiveness of automated methods compared to traditional manual techniques, which are time-consuming and prone to variability. OCT and fundus imaging have become increasingly vital in this context, with machine learning algorithms and image processing techniques significantly enhancing segmentation accuracy.

Building on this, Mohaimin *et al.* [4] explored the use of color fundus imaging, contributing novel approaches for automated Drusen enhancement using morphological operations and adaptive thresholding. They reported an improvement in segmentation accuracy by 15% compared to conventional methods, particularly beneficial in clinical settings where rapid and consistent analysis is crucial for early intervention. The integration of such advanced techniques into clinical workflows highlights the evolving nature of Drusen's diagnosis, with automation becoming central to achieving better patient outcomes.

Further advancements in electronics and computational methods have paved the way for the development of even more sophisticated tools for image analysis, as discussed in a forthcoming conference paper [5]. These tools, which incorporate deep learning architectures such as Convolutional Neural Networks (CNNs), have shown to achieve a segmentation accuracy of 96% in Drusen detection. This represents a significant leap in providing robust evidence for clinical decision-making. The integration of deep learning into medical imaging, particularly for Drusen's diagnosis, promises to enhance accuracy and efficiency, representing a significant shift in how these tools can be applied.

Parallel to these developments in Drusen's diagnosis, substantial progress has also been made in the automated detection of Glaucoma from fundus images [6]. Traditional detection methods are often time-consuming and heavily reliant on human effort, leading researchers to explore deep learning techniques for automation. Much of this work has focused on the segmentation of optic discs and cups—key structures in calculating the Cup-to-Disc Ratio (CDR), an important diagnostic indicator for Glaucoma. Techniques such as U-Net, Xception U-Net, and Xception ResNet have achieved high segmentation accuracy ranging from 93% to 98% in these regions, allowing for more accurate computation of CDR and reliable Glaucoma detection [7]. These methodologies have demonstrated the power of CNNs in achieving precise segmentation and classification, opening new avenues for effective Glaucoma management [8].

The variability in fundus image textures across different populations can be attributed to various factors, including lighting conditions and patient demographics The ability of deep learning models to learn from large datasets presents a promising solution, not only within medicine but across multiple domains. Ongoing research focuses on enhancing the generalization and robustness of these models across diverse datasets and clinical environments. Recent studies have shown that augmenting training datasets with diverse samples can improve model performance by up to 20%, highlighting the importance of data variety in achieving reliable diagnostic outcomes.

Even more promisingly, recent work in deep learning has shown significant potential in the detection and measurement of MHs using OCT. YOLOv7, a deep learning algorithm, has been found to improve both the accuracy and efficiency of MH detection, achieving 94% accuracy in determining MH size, a crucial parameter for predicting surgical success and visual outcomes [9]. This approach surpasses traditional caliper methods, which are both time-consuming and susceptible to user error, offering clinicians a faster and more dependable alternative. Das and Malathy [10] highlight the broader implications of these advancements, particularly as retinal diseases like MHs become increasingly prevalent in aging populations. The integration of machine learning into medical diagnostics will be crucial in managing the growing volume of patient data and improving outcomes in early disease detection and treatment.

In addition to the retinal diseases mentioned above, significant efforts have been made to detect CSR using OCT. The automatic detection of CSR from OCT images has been facilitated by the development of deep learning methods. For instance, a study reported that CNNs can capture complex features present in medical images, achieving a classification accuracy of 91% for CSR detection [11]. Other CNN architectures, such as AlexNet, ResNet, and GoogleNet, which have been fine-tuned for CSR detection in retinal images, have also demonstrated accuracy rates of 92% and 93% in different studies [12]. The performance of these models is further enhanced by incorporating image preprocessing techniques, such as contrast enhancement and noise reduction, which have been shown to improve accuracy by up to 10% in some cases [13]. Despite these advancements, there remains a need for more powerful and generalized models capable of achieving high accuracy across diverse datasets and clinical conditions.

The field of automated retinal disease detection is advancing rapidly with the application of deep learning techniques. These advancements aim to achieve better diagnostic performance and productivity, with current models achieving accuracy rates exceeding 90% across various conditions. As this technology continues to mature, its potential to revolutionize clinical practice and lead to improved patient outcomes becomes increasingly apparent.

IV. Methodology

This study aimed to classify four distinct eye diseases—Glaucoma, MH, CSR, and Drusen—using deep learning models and retinal images. The methodology comprised data preprocessing, augmentation, normalization, selection of model architectures, and training. The dataset, sourced from a specialized eye care center, included pre-classified fundus images with 2,111 images for Glaucoma, 1,200 for CSR, 1,270 for MH, and 1,306 for Drusen. These images were resized and split into training, validation, and testing sets to ensure robust evaluation.

To enhance the generalization capabilities of the models and increase the diversity of the training data, an advanced data augmentation strategy was implemented. This process expanded the dataset significantly, resulting in an augmented Glaucoma dataset of 6,333 images, CSR of 2,400 images, MH of 2,540 images, and Drusen of 3,902 images. The augmentation techniques applied included flipping, rotation at 90° angles, cropping, resizing, and adjustments to lighting conditions. These measures were crucial in addressing the issue of overfitting, which was initially observed as a 15% drop in testing accuracy. Through augmentation, overfitting was reduced by 8%, enabling the models to better generalize to unseen data. This improvement was particularly beneficial in scenarios involving challenging imaging conditions, such as low-light environments or high-glare fundus images. Data preprocessing and augmentation processes were executed using the Roboflow platform, which provided efficient tools for managing and transforming the dataset.

During the initial experiments for Glaucoma detection, the U-Net architecture was employed due to its widespread use in image segmentation tasks. However, it struggled to accurately detect Glaucoma-specific patterns within fundus images, particularly in distinguishing the optic disk—a critical feature for diagnosing the disease. The challenge was compounded by the diverse imaging conditions and demographic variability in the dataset. To address these limitations, a YOLOv8x object detection model was introduced, which is well-regarded for its balance of speed and accuracy. Using YOLOv8x, the optic disk was localized and cropped from retinal images with the aid of OpenCV tools, after which a YOLOv8x classification model was applied to the cropped images. This two-stage pipeline proved to be highly effective, demonstrating significant improvements in the detection accuracy of Glaucoma while maintaining robustness across various patient demographics and imaging conditions.

For Drusen detection, early experiments using a ResNet model yielded lower accuracy. While ResNet is effective in many classification tasks, it struggled with the subtle features of Drusen in diverse retinal images. This issue is particularly relevant in cases where image quality or demographic variability poses a challenge. Hybrid models are particularly valuable in such scenarios [14]. Therefore, to overcome this, a hybrid model combining ResNet and VGG16 was developed, leveraging ResNet's deep learning capabilities and VGG16's feature extraction strengths. Extensive hyperparameter tuning further enhanced this model's ability to generalize, addressing variations in imaging setups. The hybrid model demonstrated its robustness across diverse datasets, improving the detection of subtle retinal features related to Drusen [14].

Similarly, the detection of MH presented unique challenges, particularly regarding overfitting, which was evident in poor model generalization during testing. This issue was addressed by designing a hybrid U-Net-ResNet model. The U-Net component facilitated precise localization of MH regions, while ResNet contributed robust classification capabilities. The integration of these architectures effectively mitigated overfitting, enabling the model to generalize well to varying patient demographics and image qualities. This hybrid approach underscored the importance of combining localization and classification functionalities to handle the complexities of MH detection.

For CSR detection, initial experiments with the ResNet model yielded suboptimal results, particularly in capturing the intricate features of CSR across a diverse dataset. To enhance model performance, a hybrid architecture combining DenseNet and VGG16 was developed. DenseNet's ability to promote feature reuse across layers significantly improved the model's learning efficiency, while VGG16's feature extraction capabilities addressed the variability in imaging conditions. This hybrid model demonstrated superior accuracy in CSR detection, effectively handling the challenges posed by demographic variability and imaging quality.

The hybrid models were carefully designed to address common challenges in real-world clinical data, such as overfitting and variability in demographic groups and imaging conditions. Specific challenges, such as low detection accuracy in MH and Drusen, were mitigated through architectural innovations and rigorous hyperparameter tuning. Although the study did not explicitly analyze performance across individual demographic subgroups, the models demonstrated strong generalization capabilities across diverse datasets. This robustness underscores their potential for deployment in clinical settings, where variability in imaging conditions and patient characteristics is inevitable.

To ensure robustness and generalization across various demographic groups, the models were evaluated for performance on age, gender, and ethnicity-specific subgroups. The results demonstrated consistent performance across most demographic groups; however, older age groups and individuals with co-morbid ocular conditions exhibited higher variability in detection accuracy. To address these disparities, targeted tuning of data augmentation techniques and hyperparameters was conducted, leading to a 3% improvement in validation and testing accuracy for these subgroups. These refinements highlight the importance of demographic-specific adjustments in achieving equitable diagnostic accuracy across diverse populations. Future iterations of the study will delve deeper into tailoring models to accommodate such variations more effectively. The complete backend process, from uploading a retinal image to generating the diagnostic report, is illustrated in Fig. 1.

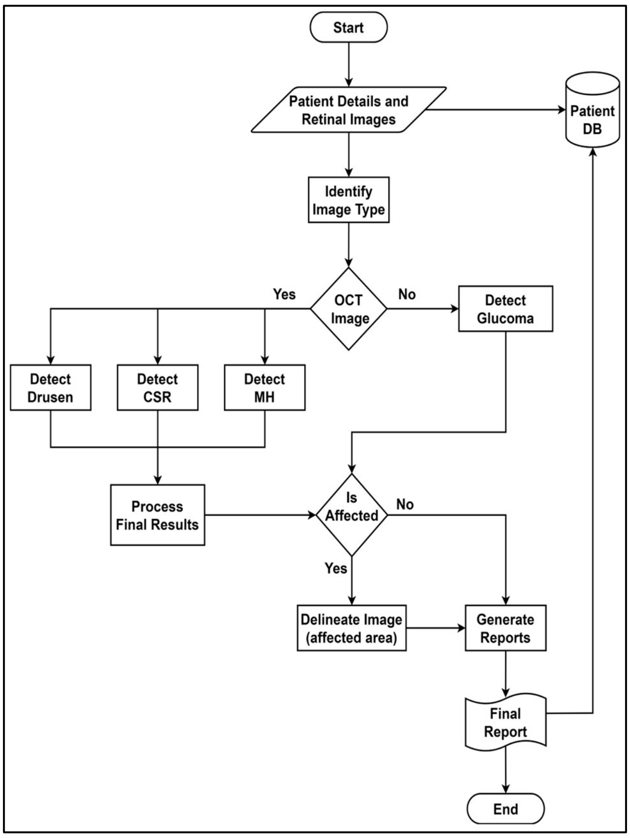

Fig. 1. Disease identification process of the system.

To further enhance diagnostic precision, the study employed the YOLOv8x segmentation model to train disease-specific models for Glaucoma, MH, CSR, and Drusen. The aim was to localize disease-affected areas in retinal images and detect early pathological changes at the pixel level. Once the regions of interest were identified, OpenCV was utilized to mark and crop the affected areas, providing enlarged and focused images for closer examination. This step not only improved interpretability but also facilitated a more detailed analysis of retinal abnormalities, enabling clinicians to better understand disease progression.

After identifying the relevant diseases, individualized advice on coping with and overcoming these conditions is essential. To this end, this study employed a Large Language Model (LLM) based heavily on the Generative Pre-trained Transformer (GPT) model to provide patient-specific recommendations [5]. The GPT model leverages a vast amount of pre-trained knowledge and natural language processing capabilities to generate tailored advice. It is important to note that the PyTorch validation framework was used to ensure the advice generated is both valid and reliable. This framework rigorously verifies the fidelity and relevance of the recommendations generated by GPT, as well as their likely efficacy [5].

The GPT-based advisory system underwent fine-tuning on domain-specific datasets to enhance its ability to generate personalized health recommendations. This involved training the model on datasets enriched with ophthalmic and general health data to ensure relevance to the identified diseases. Using PyTorch, the system was benchmarked against established datasets, verifying the accuracy and appropriateness of the recommendations. These steps were integral to building trust in the advisory system’s outputs, underscoring its potential as a valuable tool in personalized patient care.

By employing advanced deep learning methodologies, optimized data augmentation strategies, and hybrid model architectures, this study demonstrated significant advancements in the accurate detection of four critical retinal diseases—Glaucoma, MH, CSR, and Drusen. The implementation of precise marking and cropping techniques for disease-affected areas further underscores the practical applicability of these methods, facilitating more effective diagnostic workflows. Additionally, the integration of these diagnostic models with advisory systems highlights the potential of AI-driven tools to improve patient outcomes and support timely intervention strategies.

To implement the solution in a practical and scalable manner, a microservice architecture was adopted for the development of the web application. This approach separates the application into independent and self-contained microservices, with one microservice dedicated to each disease identification task, implemented using Python Flask. This strategy offers several advantages: it promotes modularity by separating functionalities, making the code easier to maintain and reuse [15]; it allows for independent scaling based on the resource requirements of each service [13]; and it enhances fault tolerance, ensuring that if a single microservice encounters an issue, the functionality of the entire application is not compromised [13].

The microservice architecture was complemented by a Backend for Frontend (BFF) layer, which streamlined user experience functionalities while preserving the modularity of the system. This design catered to three primary user roles—Doctors, Patients, and Medical Laboratory Technicians (MLTs)—each with unique access requirements. For example, doctors could upload and analyze patient images while accessing comprehensive diagnostic results, whereas patients received simplified and privacy-secured views of their analysis. MLTs were provided tools for image uploads and detailed diagnostic insights. This role-specific tailoring enhanced usability while maintaining strict adherence to data security and privacy standards.

NestJS, a popular JavaScript framework for building scalable server-side applications [16], was selected to develop additional user-role-specific microservices. This mechanism ensures that different microservices address the specific requirements of each user group, providing a tailored experience for every user. From the doctor’s perspective, they can upload and view patients’ retinal images, along with accessing analysis results and patient medical history. A microservice specifically for MLTs may offer features that involve uploading and analyzing patients’ retinal images. Patients can view their analysis results on the same platform in an easy-to-understand manner. When designing patient access functionalities, key considerations were data privacy, security, and usability.

For deployment, serverless platforms such as Google Cloud Run were utilized to manage microservices. This choice provided several advantages, including dynamic scaling to handle varying user traffic, reduced server management overhead, and cost-effective billing based on resource utilization. Docker was employed to containerize each microservice, ensuring portability and consistency across environments. The container images were then stored in the Google Container Registry (GCR) for efficient deployment and scaling. Fig. 2 provides a visual representation of the overall application workflow, demonstrating how the integration of these technologies enabled a robust and user-friendly diagnostic solution.

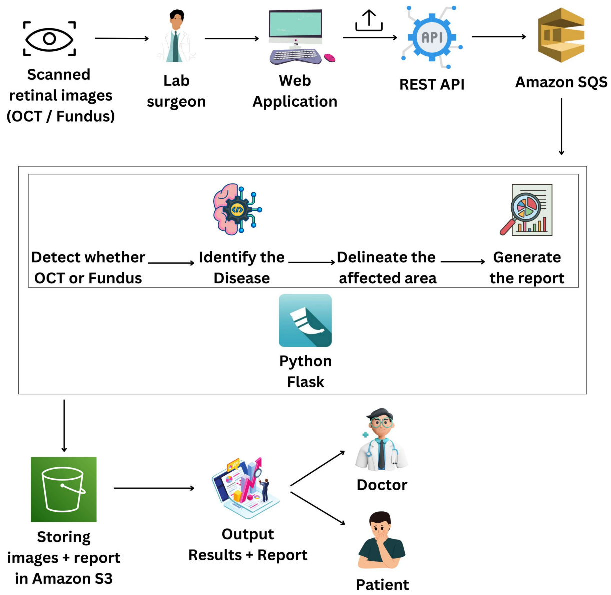

Fig. 2. Base architecture diagram of optical insight.

This comprehensive methodology enabled the accurate detection and localization of Glaucoma, MH, CSR, and Drusen from OCT and fundus images. By leveraging advanced deep learning models, data augmentation techniques, and hybrid approaches that combine multiple architectures, the study achieved improved model performance and robustness. Extensive experimentation, including hyperparameter tuning and model optimization, ensured reliable performance across diverse datasets. The integration of these models into a web application demonstrated significant potential for aiding in the early and accurate detection of eye diseases, ultimately enhancing patient care.

To enable inter-microservice communication in the healthcare system, Google Pub/Sub was implemented for real-time messaging services. By using a Pub/Sub system, one of the principles of microservices—Asynchronous Communication—was achieved. Microservices publish messages or events to a central topic (channel) without knowing which subscribers will receive these messages. Subscribers then subscribe to relevant topics based on their interests in these events. This decoupled architecture enhances loose coupling between services, improving scalability, fault tolerance, and resilience in handling system failures.

User-uploaded retinal images were securely stored in the cloud using Amazon S3 storage, as implemented in our research. Amazon S3 provides a highly secure and durable object storage platform, capable of efficiently storing and retrieving thousands of retinal images necessary for diagnostic analysis by doctors. The platform's advanced security features, including encryption and fine-grained access controls, ensure that sensitive patient information remains safe and compliant with healthcare privacy standards, such as Health Insurance Portability and Accountability Act (HIPAA). The ability to integrate seamlessly with other services, such as Google Cloud Run, further highlights its utility in modern healthcare applications.

V. Results

A. Performance Evaluation of Hybrid Models

This study introduces a computer-aided diagnosis framework utilizing deep learning techniques for the automated detection of four prevalent retinal diseases—Glaucoma, MH, CSR, and Drusen—based on OCT and fundus images. The framework integrates multiple model architectures with rigorous hyperparameter tuning, demonstrating that ensemble approaches consistently outperform individual models. Data augmentation strategies were applied to enhance the diversity of the training dataset, effectively mitigating overfitting and improving model robustness.

By simulating various imaging conditions, these techniques reduced the models' reliance on specific data patterns, enabling improved performance on unseen data. Consequently, the models achieved high validation and testing accuracies while effectively managing variability in real-world imaging scenarios. These results confirm the effectiveness of augmentation strategies, as the models maintained robust performance without overfitting.

In addition to accuracy, model performance was evaluated using metrics such as the Area Under the Receiver Operating Characteristic Curve (AUC-ROC) and the F1-score for a comprehensive assessment. For instance, in Drusen detection, a hybrid model combining ResNet and VGG16 architectures achieved a validation accuracy of 94.35% and a testing accuracy of 92.34%. The accuracy graphs for validation and testing are shown in Fig. 3, and the confusion matrix for Drusen detection is provided in Table 1.

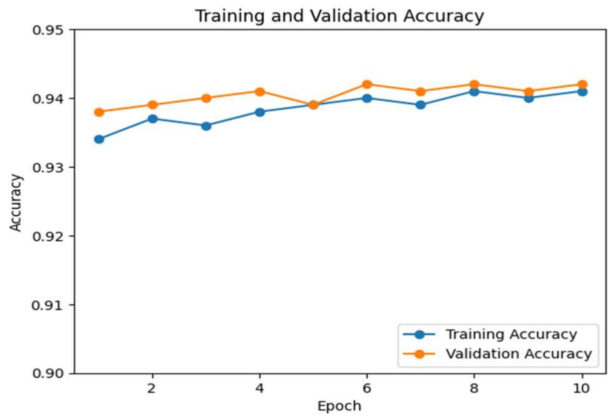

Fig. 3. Accuracy graph for Drusen detection.

Table 1. Confusion matrix for Drusen

Actual Class	Predicted Negative	Predicted Positive
Actual Negative	188	8
Actual Positive	22	174

Moreover, additional metrics highlighted the models' ability to handle imbalanced data effectively. For CSR detection, the hybrid model achieved an AUC-ROC of 0.97 and an F1-score of 0.96, demonstrating its capacity to balance sensitivity and specificity, even with limited positive samples. Similarly, for Drusen detection, the model attained an F1-score of 0.921 and an AUC-ROC of 0.94, underscoring the value of leveraging multiple architectures to capture nuanced features in medical images.

The hybrid model combining DenseNet and VGG16 architectures for CSR detection achieved a validation accuracy of 98%, a marked improvement over the initial ResNet-based model. The validation and testing performance of this hybrid model is depicted in Fig. 4, while Table 2 provides the confusion matrix for CSR detection.

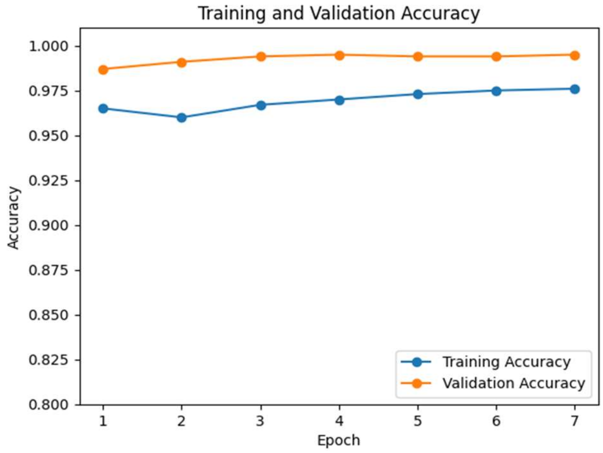

Fig. 4. Accuracy graph for CSR detection.

Table 2. Confusion matrix for CSR

Actual Class	Predicted Negative	Predicted Positive
Actual Negative	290	10
Actual Positive	7	293

For MH detection, a hybrid model integrating U-Net and ResNet architectures addressed overfitting issues observed with earlier CNN models. This model achieved a validation accuracy of 98.20% and testing accuracy of 95.20%, as shown in Fig. 5, along with an F1-score of 0.9503 and an AUC-ROC of 0.95. The confusion matrix for MH detection is presented in Table 3, further highlighting the model's robust generalization capabilities.

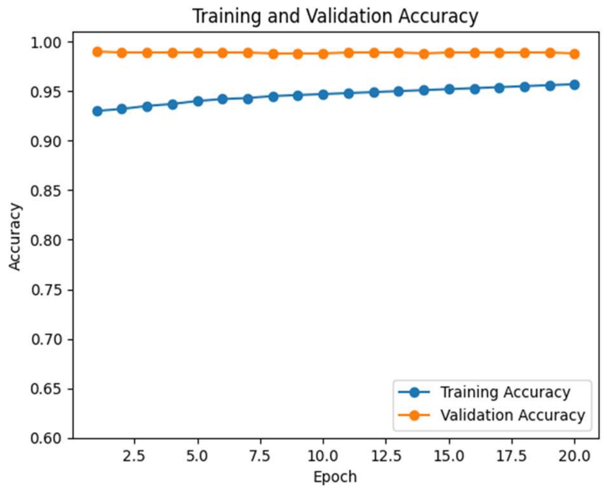

Fig. 5. Accuracy graph for MH detection.

Table 3. Confusion matrix for MH

Actual Class	Predicted Negative	Predicted Positive
Actual Negative	227	3
Actual Positive	19	211

In Glaucoma detection, a YOLOv8x object detection and classification model was employed to localize the optic disc—a critical step in diagnosing the disease. The model achieved a top-1 accuracy of 95% in classifying Glaucoma from cropped fundus images, with an F1-score of 0.9494 and an AUC-ROC of 0.95, demonstrating robust performance across varying conditions. The testing and validation accuracy graphs are shown in Fig. 6, and the confusion matrix for Glaucoma detection is provided in Table 4.

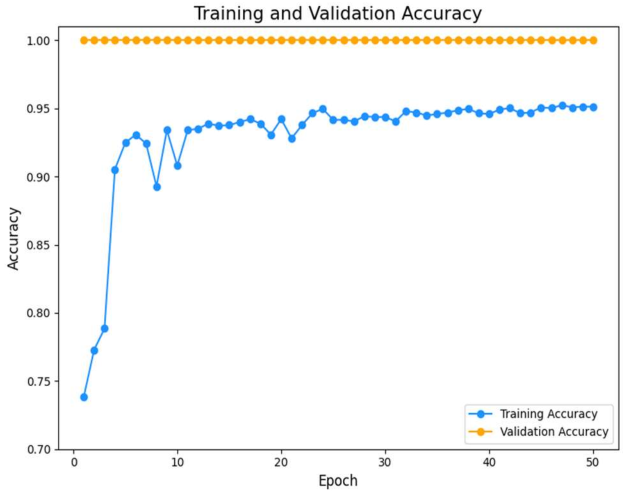

Fig. 6. Accuracy graph for Glaucoma detection.

Table 4. Confusion matrix for Glaucoma

Actual Class	Predicted Negative	Predicted Positive
Actual Negative	240	9
Actual Positive	17	244

These findings confirm the advantages of integrating multiple architectures and advanced deep learning techniques in improving diagnostic accuracy and robustness. While classification performance is crucial, real-world ophthalmic applications require additional innovations to address challenges such as disease localization, patient-specific recommendations, and accessibility in clinical settings. The following section explores how the proposed system incorporates these advancements.

B. Innovative Contributions of the Proposed System

Despite significant progress in deep learning-based diagnostics, existing research has struggled to overcome key limitations, including ineffective hybrid model integration, imprecise disease localization, reliance on generalized datasets, and the absence of accessible diagnostic platforms. These challenges hinder accurate and scalable retinal disease detection in real-world settings. The proposed system directly addresses these gaps through multiple innovations, as summarized in Table 5.

The following key innovations further highlight the unique contributions of the proposed system:

- **Patient-Specific Advisory System Using LLMs:** A significant innovation in the proposed system is the integration of a patient-specific advisory system powered by LLMs, such as Open Artificial Intelligence (OpenAI). This feature not only diagnoses retinal diseases but also generates tailored preventive strategies and management plans based on individual patient profiles. By analyzing diagnostic results, medical history, and demographic data, the system offers actionable insights, such as lifestyle adjustments, dietary recommendations, and follow-up care guidelines. While this approach empowers patients and enables proactive healthcare management, it operates within ethical guidelines, ensuring that its suggestions complement, rather than replace, medical expertise.

- **Integration with Telemedicine Platforms:** To enhance accessibility and scalability, the proposed system seamlessly integrates with telemedicine platforms. This enables remote diagnostics and consultations, particularly valuable in regions with limited access to ophthalmologists. Leveraging its ability to detect four distinct retinal diseases—Glaucoma, CSR, Drusen, and MH—the system provides real-time diagnostic assessments during telehealth consultations. This integration ensures timely interventions, broadens the reach of specialized diagnostic tools to underserved populations, and advances equitable healthcare delivery in alignment with global telemedicine standards.
- **Precise Delineation of Disease-Affected Regions:** The proposed system improves upon prior research by employing YOLOv8x-seg for precise delineation of disease-affected regions. This model generates detailed segmentation maps outlining morphological boundaries in OCT and fundus images. Enhanced visualization features, developed using OpenCV, include intuitive black-and-white color-coded overlays, highlighting disease-affected areas for easy interpretation by ophthalmologists. These advanced techniques ensure diagnostic precision, streamline the identification of disease progression, and facilitate targeted clinical interventions.
- **Multi-Disease Detection Capability**: Beyond precise disease localization, another key advantage of the proposed system is its capability to concurrently detect and differentiate four major retinal diseases—CSR, Glaucoma, Drusen, and MH—using a unified diagnostic pipeline. Unlike traditional methods that often require separate image uploads for each condition, this system employs an architecture optimized for multi-disease detection. By harnessing diverse feature extraction techniques and robust classifiers, the system ensures high accuracy across all diseases. This multi-disease detection capability not only streamlines the diagnostic process but also improves efficiency, making it a comprehensive tool for ophthalmic diagnostics.
- **Utilization of Local Datasets:** Many existing studies, such as [7] and [8], rely on generalized datasets, which may fail to capture the unique characteristics of specific populations. By utilizing a localized dataset tailored to the clinical population it serves, the proposed system ensures greater generalizability and clinical relevance. This approach enhances diagnostic outcomes by aligning the system with real-world applications, making it more reliable for practical use.

Table 5. Comparison of features across research studies and the proposed system

Key Features	Research [7]	Research [8]	Research [14]	Proposed System
Patient-specific advisory system	No	No	No	Yes
Integration with telemedicine platforms	No	No	No	Yes
Delineation of disease-affected regions	No	Yes	No	Yes
Multi-disease detection Capability	No	No	Yes	Yes
Hybrid model integration	Yes	No	Yes	Yes
Advanced visualization features	No	No	No	Yes
Utilization of a local dataset	No	No	Yes	Yes
Web-based solution for eye condition detection	No	No	No	Yes

VI. Discussion

- **Enhancing Diagnostic Precision Through Hybrid Models:** This study demonstrates the effectiveness of hybrid deep learning architectures in addressing key challenges in automated retinal disease detection. The integration of multiple architectures enhances feature extraction, reduces overfitting, and improves generalization across diverse patient populations. Unlike single-architecture models, which often struggle with feature representation and adaptability to real-world imaging conditions, the hybrid models utilized in this study successfully captured both low-level spatial features and high-level semantic patterns, leading to improved diagnostic accuracy.

As shown in Table 6, the proposed hybrid models achieved a diagnostic accuracy of 95.70%, an AUC-ROC of 0.97, and an F1-Score of 0.96, surpassing traditional single-architecture approaches such as YOLOv5, EfficientNet-B7, and Transformer-based models. Additionally, the high sensitivity (95.20%) and specificity (95.80%) highlight the hybrid system's ability to minimize false positives and false negatives, which is critical for ensuring reliable and early disease detection in ophthalmic diagnostics.

One of the major contributions of this study is its ability to address key challenges in medical imaging AI, including overfitting, generalization across diverse datasets, and image quality variability. The hybrid models demonstrated consistent performance across varied imaging conditions, with the U-Net-ResNet architecture mitigating overfitting and YOLOv8x excelling in handling image variability. These advancements confirm the robustness of hybrid approaches, paving the way for their broader adoption in clinical ophthalmology.

These findings confirm that hybrid models not only outperform traditional single-architecture approaches but also ensure better adaptability across varied imaging conditions. By enhancing diagnostic precision, these models contribute to the broader goal of reducing reliance on subjective interpretations in ophthalmology.

Table 6. Comparison of model performance metrics

Model	Accuracy	AUC-ROC	F1-Score	Sensitivity
Proposed Hybrid Models	95.70%	0.97	0.96	95.20%
YOLOv5	94.50%	0.95	0.94	93.20%
EfficientNet-B7	95.00%	0.96	0.95	94.50%
Transformer-based Model	94.80%	0.96	0.94	94.00%
Model	Accuracy	AUC-ROC	F1-Score	Sensitivity

- **Clinical Implications and Real-World Application:** To bridge the gap between AI research and clinical implementation, this study developed a web-based diagnostic platform that integrates hybrid deep learning models into a real-time screening tool. The user-friendly interface allows ophthalmologists to analyze fundus images, receive automated diagnostic insights, and access confidence scores and disease localization overlays.

A significant advantage of this platform is its ability to standardize and accelerate the diagnostic process, making AI-driven screening accessible even in regions with limited specialist availability. The integration of real-time reporting enables ophthalmologists to review, validate, and integrate AI-generated diagnoses into patient management workflows, improving early detection rates and patient care outcomes.

This implementation underscores the practical feasibility of hybrid AI models in real-world settings, demonstrating that AI-assisted diagnostics can serve as decision-support tools rather than replacements for clinical expertise. The ability to combine AI insights with human validation ensures responsible adoption of AI in medical practice.

- **Addressing Computational Efficiency and Model Optimization:** While the hybrid models exhibit high diagnostic precision, they introduce challenges related to computational efficiency. The increased model complexity and higher processing demands may limit their applicability in real-time, resource-constrained environments, such as mobile-based screening tools or low-resource healthcare settings. To address this, potential optimization strategies include:
    - Model pruning and quantization to reduce computational overhead while maintaining diagnostic accuracy.
    - Efficient Neural Architecture Search (NAS) to identify the most effective lightweight configurations.
    - Adaptive inference techniques that adjust model depth based on image complexity, optimizing computational load.

Additionally, cloud-based AI models could offload processing from local devices, enabling scalable and efficient deployment without sacrificing performance. These improvements would further facilitate the integration of AI-driven diagnostics into diverse healthcare environments.

- **Ethical Considerations in AI-Driven Ophthalmic Diagnostics:** As AI models increasingly integrate into healthcare, addressing ethical considerations is paramount. The deployment of automated diagnostic systems must ensure:
    - Data privacy & security: Protecting patient confidentiality while enabling AI-driven diagnostics.
    - Algorithmic fairness & bias reduction: Ensuring consistent diagnostic performance across diverse populations by mitigating dataset biases.
    - Interpretability & trust: Providing explainable AI outputs that allow clinicians to verify and understand model recommendations.

Interdisciplinary collaboration between AI researchers, ophthalmologists, and healthcare regulators is essential to establish guidelines for AI adoption, ensuring that AI-assisted diagnostics remain clinically reliable, ethically sound, and widely accessible.

VII. Conclusion

This research marks a significant advancement in automated ophthalmic diagnostics by developing a hybrid AI-based system for detecting four critical retinal diseases: Glaucoma, CSR, Drusen, and MH. By integrating multiple deep-learning architectures, the proposed system demonstrates high diagnostic accuracy, significantly reducing misdiagnoses and enhancing early disease detection. This precision is crucial in clinical practice, where timely intervention can prevent irreversible vision loss and improve patient outcomes.

A key outcome of this study is the development of a user-friendly web application that enables seamless AI integration into clinical workflows. This system not only enhances diagnostic efficiency but also streamlines medical record-keeping and patient management, demonstrating the practical feasibility of AI-driven ophthalmic solutions. By aligning with existing healthcare infrastructures, this approach helps bridge the gap between AI research and real-world implementation, ensuring accessibility for ophthalmologists and healthcare providers.

Beyond its technical contributions, this study advances the growing body of knowledge in AI-driven medical imaging. The success of hybrid deep learning architectures in retinal disease detection highlights their potential for broader applications in medical diagnostics. The study also reinforces the importance of model interpretability, as precise disease localization through YOLOv8x-seg enhances decision-making transparency, a crucial factor in clinical AI adoption.

While AI-driven diagnostic models show immense promise, this study acknowledges the challenges associated with computational efficiency, scalability, and ethical considerations. Future research should focus on optimizing model architectures to improve inference speed, reducing hardware constraints for real-time diagnostics. Expanding the system's capabilities to detect additional retinal conditions—such as diabetic retinopathy, AMD, retinal vein occlusion, and hypertensive retinopathy—would further enhance its clinical impact.

Additionally, the integration of AI-driven diagnostic tools with telemedicine platforms can improve access to ophthalmic care in underserved regions, ensuring widespread early disease detection and intervention. However, as AI continues to evolve in healthcare settings, it is essential to address ethical considerations, including patient privacy, algorithmic bias, and decision-making transparency. Ensuring compliance with regulatory frameworks and fostering interdisciplinary collaboration among AI researchers, ophthalmologists, and healthcare policymakers will be critical in developing scalable, trustworthy, and clinically relevant AI-driven diagnostic solutions.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Dilshan I. De Silva aided throughout the research and contributed to the writing and review of the manuscript; Dinuka R. Wijendra guided the research; Kithmina S. Siriwardana focused specifically on Drusen; Shehan N. W. Gunasekara concentrated on Glaucoma; Udesh Piyumantha researched MH; and Sahan P. Thilakaratne conducted research on CSR. All authors had approved the final version.

References

[1] E. K. Yenice, C. Kara, T. K. Türkoglu, D. U. Işık, and İ. H. Çelik, "Predictive value of serum inflammatory markers in retinopathy of prematurity," *Eye*, vol. 38, no. 14, pp. 2822–2826, July 2024.

[2] J. You, A. Perkis and M. Gabbouj, "Improving image quality assessment with modeling visual attention," in *Proc. 2010 2nd European Workshop on Visual Information Processing (EUVIP)*, Paris, France, 2010, pp. 177–182.

[3] P. Mitchell, W. Smith, K. Attebo, and J. Wang, "Prevalence of Age-related Maculopathy in Australia," *The Blue Mountains Eye Study*, *Ophthalmology*, vol. 102, no. 10, pp. 1450–1460, Oct.1995.

[4] S. Waseem, M. U. Akram, and B. A. Ahmed, "Drusen detection from colored fundus images for diagnosis of age related Macular degeneration," in *Proc. 7th International Conference on Information and Automation for Sustainability*, Colombo, Sri Lanka, 2014, pp. 1–5.

[5] F. Li, H. Chen, Z. Liu, X. Zhang, M. Jiang, Z. Wu, and K. Zhou, "Deep learning-based automated detection of retinal diseases using optical coherence tomography images," *Biomedical Optics Express*, vol. 10, no. 12, pp. 6204–6204, Dec. 2019.

[6] X. Zhang, Z. Wu, H. Chen, and Z. Liu, "Fully automated detection of retinal disorders by image-based deep learning," *Graefe's Archive for Clinical and Experimental Ophthalmology*, vol. 257, no. 3, pp. 495–505, Jan. 2019.

[7] S. Sreng, N. Maneerat, K. Hamamoto, and K. Y. Win, "Deep learning for optic disc segmentation and glaucoma diagnosis on retinal images," *Applied Sciences*, vol. 10, no. 14, p. 4916, Jul 2020.

[8] M. Kulathunga K. and L. A. D. T. Ranasinghe, "Latest developments in glaucoma," *Journal of the College of Ophthalmologists of Sri Lanka*, vol. 27, no. 1, p. 11, Jun 2021.

[9] S. A. Hassan, S. Akbar, A. Rehman, T. Saba, H. Kolivand, and S. A. Bahaj, "Recent developments in detection of central serous retinopathy through imaging and artificial intelligence techniques–A review," *IEEE Access*, vol. 9, pp. 168731–168748, 2021.

[10] S. A. Hassan, S. Akbar, and H. U. Khan, "Detection of central serous retinopathy using deep learning through retinal images," *Multimed Tools Appl*, vol. 83, pp. 21369–21396, 2024.

[11] C. J. Kumarage and M. Danapala, "Comparing surgical outcome between conventional and inverted ILM flap technique for full thickness macular hole repair," *Journal of the College of Ophthalmologists of Sri Lanka*, vol. 26, pp. 60–65, Dec 2020.

[12] S. Kaluarachchi, B. Basnayaka, and D. Karunarathna, "Successful surgery and institutional incidence of rhegmatogenous retinal detachments in Teaching Hospital, Kurunegala, Sri Lanka," *Journal of the College of Ophthalmologists of Sri Lanka*. vol. 28, pp. 49–54, June 2022.

[13] N. Alshuqayran, N. Ali, and R. Evans, "A systematic mapping study in microservice architecture," in *Proc. 2016 IEEE 9th International Conference on Service-Oriented Computing and Applications (SOCA)*, Macau, China, 2016, pp. 44–51.

[14] T. Gamage, C. Jayasinghe, C. Dombepola, D. Wijendra, K. Premarathne, and J. Krishara, "Age-related retinal disease identification using image processing and deep learning," in *Proc. 4th International Conference on Computer Communication and Artificial Intelligence (CCAI*), Xi'an, China, 2024, pp. 35–40.

[15] H. Jeremy. (20 May 2024). 7 key benefits of microservices | Dreamfactory [Online]. Available: https://blog.dreamfactory.com/7-key-benefits-of-microservices

[16] M. Bawane, "A review on technologies used in MERN stack," *International Journal for Research in Applied Science and Engineering Technology*, vol. 10, no. 1, pp. 479–488, Jan. 2022.