AI Based Systemic Disease Using Eye Images and Multi-Agent LLMs

Dr. Dwiti Krishna Bebarta[1]; N. Srujana[2]; R. R. V. Kusuma[3]; A. Lavanya[4]

[1]Professor, Department of CSE (AI&ML), GVP College of Engineering for Women, Visakhapatnam, AP, INDIA
[2,3,4,]Student, Department of CSE (AI&ML), GVP College of Engineering for Women, Visakhapatnam, AP, INDIA

Publication Date: 2026/04/27

Abstract: Diabetic retinopathy, glaucoma and cataracts are the eye diseases that cause serious impairment of the vision in the world. Early signs can avoid the severe deterioration of vision, and in order to do it, it is vital to diagnose a person using only skilled ophthalmologists and specialized equipment, which is not always available. In this paper, the authors introduce an AI-based Systemic Disease screening with the help of eye images and multiagent Large Language Models (LLMs). The system makes use of deep learning model, which is DenseNet121 to extract features of eye pictures, and LightGBM to give precise classification of the disease with probability scores. GradyCAM visualization is also added to show critical areas that affect the model predictions to help in improving interpretability. A multiagent LLM framework is proposed to improve analysis and usability, and the Diagnosis, Validation, Risk Assessment, Explanation, and Report agents make up a multi agent LLM. Such agents offer a complete set of insights, reports that are easy to understand by patients and doctor-level data and enhance decision-making. Multi-language translation, and voice output are also enabled in the system which allows it to be used by a broader group of users. The suggested system shows greater accuracy, interpretability, and accessibility in the screening of eye diseases. It may help medical practitioners to diagnose early and give improved awareness to the patients, particularly in facilities with limited resources.

Keywords: *Eye Disease Detection, Deep Learning, DenseNet121, LightGBM, Medical Image Analysis, Multi-Agent LLMs, Grad-CAM, Artificial Intelligence in Healthcare, Image Processing, Automated Diagnosis.*

How to Cite: Dr. Dwiti Krishna Bebarta; N. Srujana; R. R. V. Kusuma; A. Lavanya (2026) AI Based Systemic Disease Using Eye Images and Multi-Agent LLMs. *International Journal of Innovative Science and Research Technology*, 11(4), 2065-2074. https://doi.org/10.38124/ijisrt/26apr932

I. INTRODUCTION

One of the major causes of vision deficits and blindness in the world is eye diseases. Diabetic retinopathy, glaucoma, and cataracts as well as age-related macular degeneration are usually conditions that progress over time and are not always accompanied by visible symptoms during the initial phase. It is necessary to detect and treat the problem in the early stages of its development to avoid severe vision impairment but the traditional diagnostic interventions are based on the involvement of the experienced ophthalmologists and special medical equipment. This brings about accessibility difficulties particularly in rural and resource stricken regions where access to expert healthcare services is not easily accessible.

Amid the current rapid development of artificial intelligence (AI), the sphere of medical image analysis has been improved significantly. Convolutional Neural Networks (CNNs) and other deep learning methods have proven to be the most effective with regard to retinal and eye image processing to detect diseases. Such models are able to put out complex features automatically and discover patterns that are hard to recognize manually. Although most of them have reached a high accuracy, most of the existing systems only deliver the results of classification, but not with details, validation, and easy to read, which limits their practicality in the real-life healthcare implementation.

In order to overcome these shortcomings, this paper presents a disease screening system based on AI with eye images combined with a multi-agent framework of Large Language Model (LLM). The system makes use of DenseNet121 to extract features during the feature extraction phase, and LightGBM to classify the disease producing predictions and probability scores. Besides this, Grad-CAM visualization is used to indicate which regions are important in the image of the eye, to enhance the understanding of the decisions made by the model.

Moreover, a multi-agent LLM architecture is proposed in the system and includes special agents to diagnose,

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/26apr932

validate, evaluate risks, explain, and generate reports. This will improve the capacity of the system to offer extensive analysis as opposed to mere predictions. The resulting reports are to be used by patients and doctors, and some of its characteristics are multi-language translation and voice-based output to enhance accessibility and usability. On the whole, the suggested system should create an intelligent, understandable, and user-friendly solution to the problem of screening early eye disease. Using the concept of deep learning with multi-agent AI systems, it helps advance the healthcare support, shorten the diagnostic time, and help patients become more aware, particularly in underserved areas.

II. LITERATURE REVIEW

➢ *Artificial Intelligence in Medical Imaging*

Artificial intelligence, when used in medical imaging, has greatly enhanced the process of disease diagnosis and detection. The initial studies were aimed at applying machine learning algorithms to help clinicians examine medical images. Convolutional Neural Networks (CNNs) nowadays are used extensively with the development of deep learning to extract high-level features of medical images including retinal fundus images. Investigations by Ting et al. [1] and Abramoff et al. [9] show that deep learning models are highly accurate in the observation of eye illnesses such as diabetic retinopathy. Such systems save on manual labor, increase diagnostic time and allow mass screening. They however need large data and high computing power to perform well.

➢ *Deep Learning for Retinal Image Analysis*

Deep learning methods have widely been used to analyse retinal images in order to identify various eye conditions. Hierarchical features of fundus images have been extracted in terms of models like ResNet [5] and VGG [6] among others. Also, such datasets as ODIR-5K [4] offer multi-label classification features, which enables the detection of several diseases using a single image. Attention mechanism and visualization tools such as Grad-CAM [8] also contribute to interpretability by indicating important areas in the image. These methods enhance diagnostics and lead to automated screening. Nevertheless, there are still issues of class imbalance, image quality variability and the reliance on labeled data.

➢ *Large Language Models and Multi-Agent Systems in Healthcare*

Recent developments in Large Language Models (LLMs) allowed the creation of intelligent systems, which are able to comprehend and produce human-like text. Other models including the one suggested by Brown et al. [7] are proven to be very effective in the task of natural language processing. In addition, multi-agent LLM systems as described by Wang et al. [3] break down complex tasks into smaller units to be tackled by specialized agents like diagnosis, validation and report generation that enhance reasoning, interpretability and report generation in healthcare systems. Such systems are applicable in the production of clinical and patient friendly reports. But, they are in need of careful design, validation and computational assets in order to be deployed in reality.

➢ *Machine Learning and Hybrid Models for Disease Prediction*

Machine learning methods are important in enhancing the quality of prediction in medical systems particularly when it is paired with deep learning. There has been strong performance improvement with hybrid models combining deep feature extraction with conventional machine learning classifiers. Indicatively, tree based models like LightGBM [10] are very effective with structured data, and there is faster training with improved accuracy compared to deep learning models of tabular data. Deep learning models such as CNNs are applied to analyze medical images to obtain features that are subsequently transformed into machine learning classifiers to obtain final prediction. This is a combination that increases the accuracy and efficiency of the system. These hybrid methods are very effective in classification of multiple diseases and on-the-fly usage. Nevertheless, it takes a close selection of features and tuning of these models to get the best performance.

III. SYSTEM ARCHITECTURE

The suggested system is based on a systematic pipeline of automated disease detection with eye images combined with a multi-agent LLM system.

It starts at the user level, in which the patient takes or uploads an eye image by using a camera interface. The acquired image is then forwarded to the image preprocessing phase, where it is subjected to quality enhancement, noise reduction and normalization, aimed at enhancing the image quality and consistency of the image. Following the preprocessing, the picture is passed through the model of dense net-121, an element that entails feature extraction through the identification of key patterns and structures in the retinal picture. These extracted features are then fed to the LightGBM classifier which gives predictions of the possible diseases and their particular confidence.

The Multi- Agent LLM system then processes the prediction results and this system comprises of special agents that include Diagnosis Agent, Validation Agent, Risk Assessment Agent and Explanation Agent. Specific agents are utilized to complete a particular task, such as prediction validation, severity of the disease, and explanations that can be readily understood. specialized agents: Diagnosis Agent, The Validation Agent, Risk Assessment Agent, and Explanation Agent. Agents are also able to do a particular task, such as verification of predictions, disease severity, and produce a clear explanation.

Report Agent - The report generated by all the agents is processed to create a patient friendly report. The system further increases the accessibility by translating the report into different languages with the help of the Translation Module and creating audio output with the help of the Voice Generator (Text-to-Speech).

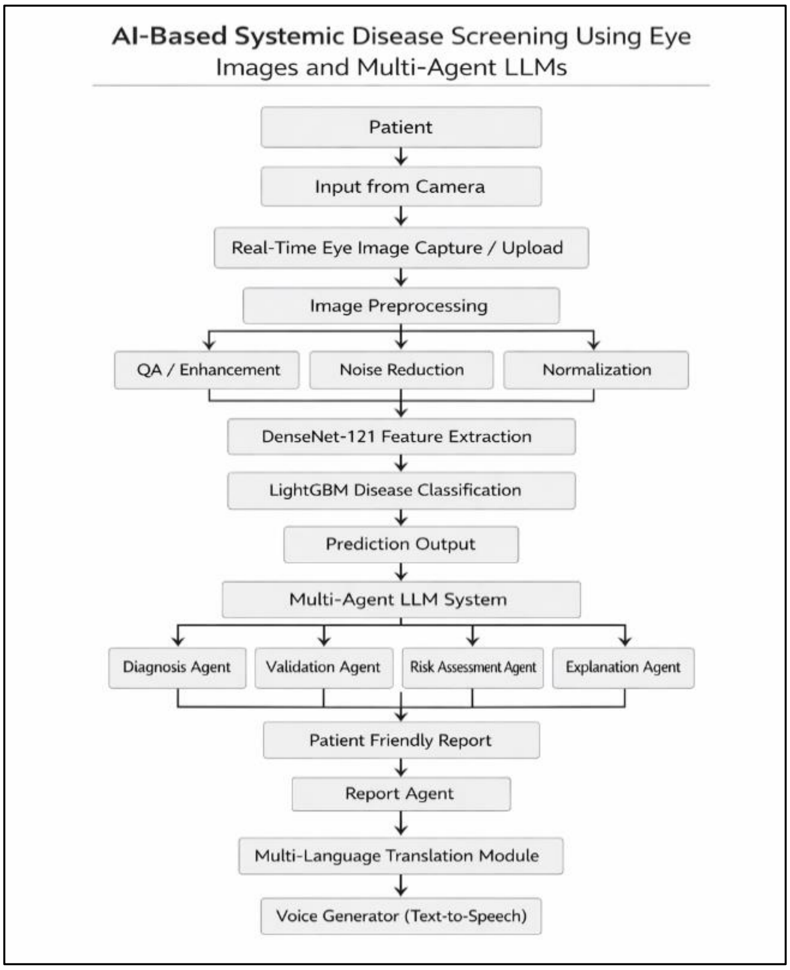

Fig 1 System Architecture

The suggested system is based on a systematic pipeline of automated disease detection with eye images combined with a multi-agent LLM system.

It starts at the user level, in which the patient takes or uploads an eye image by using a camera interface. The acquired image is then forwarded to the image preprocessing phase, where it is subjected to quality enhancement, noise reduction and normalization, aimed at enhancing the image quality and consistency of the image.

Following the preprocessing, the picture is passed through the model of dense net-121, an element that entails feature extraction through the identification of key patterns and structures in the retinal picture. These extracted features are then fed to the LightGBM classifier which gives predictions of the possible diseases and their particular confidence.

The Multi- Agent LLM system then processes the prediction results and this system comprises of special agents that include Diagnosis Agent, Validation Agent, Risk Assessment Agent and Explanation Agent. Agents are also able to do a particular task, such as verification of predictions, disease severity, and produce a clear explanation. Report Agent - The report generated by all the agents is processed to create a patient friendly report. The system further increases the accessibility by translating the report into different languages with the help of the Translation Module and creating audio output with the help of the Voice Generator (Text-to-Speech).

IV. PROPOSED METHODOLOGY

➢ *CNN (Convolutional Neural Network)*

Convolutional Neural Networks (CNNs) are a type of deep learning networks that are particularly developed to process visual information. They are commonly employed for image classification, object detection, and image analysis in medical science. In this research, CNN is employed to extract useful information from eye images for disease detection.

The basic CNN architecture consists of a series of layers, namely convolutional layers, pooling layers, and fully connected layers. In a convolutional layer, filters are applied to the input image to extract local information such as edges and patterns. Mathematically, a convolutional layer can be defined as:

$$(A_{x,y}^{k} = \sigma\left(\sum\sum W_{p,q}^{k} \cdot I(x+p, y+q) + b^{k}\right) \quad (1)$$

Where is the activation at position (x, y) in the k-th feature map, is the weight of the filter, is the input image, is the bias term, and is the activation function. The pooling layer

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/26apr932

is used to reduce the spatial dimensions of the feature maps, keeping in mind that important information is retained. One of the techniques used in this layer is max pooling, given by: Px,y=max(Ri,j)

Where is the pooled value, and is a region in the feature map.

Lastly, fully connected layers are used for classification, where features are combined. This can be given by: $Z_n = \sigma(\sum W_{mn} \cdot F_m + B_n)$

Where Z_n is the output, W_{mn} are weights, F_m represents feature inputs, and B_n is the bias.

In the proposed system, DenseNet121, an advanced CNN architecture, is used to efficiently extract deep features from eye images.

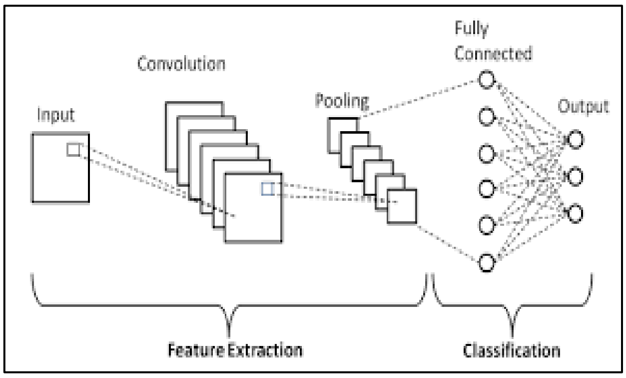

Fig 2 Structure of CNN

➢ *DenseNet121 and LightGBM Classification*

The DenseNet121 is a deep convolutional neural network that connects every layer to every other layer. This facilitates feature reuse and gradient flow. This is helpful in extracting complex features in eye images. These features are critical in identifying various eye diseases.

The feature vector is given to a LightGBM classifier. This is a gradient boosting framework that is efficient and highly performing. LightGBM is based on decision tree learning. This classifier is highly efficient in handling large-scale problems with high accuracy and speed.

$$Y = \sum_{i=1}^{n} f_i\,(X) \tag{2}$$

The classification can be defined as:

Where is the predicted output, is decision trees, and is input features.

➢ *Grad-CAM Visualization*

Grad-CAM (Gradient-weighted Class Activation Mapping) is used to provide visual explanations for model predictions. It highlights important regions in the input image that influence the decision of the model.

The Grad-CAM heatmap is generated using gradients flowing into the final convolutional layer:

$$L^c_{Grad-CAM} = \sum_k \alpha^c_k\, A^k \tag{3}$$

Where α^c_k represents the importance weights and A^k is the activation map.

This helps doctors and users understand the model's reasoning by visually identifying affected regions in the eye image.

➢ *Multi-Agent LLM Framework*

- To enhance interpretability and usability, the system integrates a multi-agent Large Language Model (LLM) framework. Each agent performs a specific task:
- Diagnosis Agent → Identifies the detected disease
- Validation Agent → Verifies prediction confidence
- Risk Agent → Assesses severity level
- Explanation Agent → Provides human-understandable explanation
- Report Agent → Generates structured reports

This modular approach improves reliability and provides comprehensive analysis beyond simple predictions.

➢ *System Workflow*

The overall system workflow consists of the following steps:

- Image acquisition (upload/capture)
- Image preprocessing (CLAHE, noise removal, normalization)
- Feature extraction using DenseNet121
- Disease classification using LightGBM
- Grad-CAM visualization
- Multi-agent LLM processing
- Report generation and output

This pipeline ensures accurate, interpretable, and accessible disease screening.

V. SYSTEM DESIGN

The proposed system is designed as a comprehensive architecture, which is modular in nature, incorporating various advanced technologies to achieve efficient and accurate disease detection using images of the retinal eyes. The proposed system incorporates various image processing, deep learning, machine learning, and natural language processing techniques to achieve the goal of providing an efficient automated solution. All the components of the proposed system are designed in a specific way to achieve the required functionality.

The proposed system is designed in a layered fashion, where each layer is designed to process one specific task, ranging from inputting the data to the output delivery.

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/26apr932

➢ *Input Design*

The input design phase focuses on acquiring the retinal images from the users in a simple way. The design enables users to either take pictures of the eyes in real-time using the camera or upload any existing pictures they have on their devices. The input design ensures that the images are in the right formats, i.e., JPEG or PNG, and also meet the required standards in terms of resolution. Appropriate validation procedures are put in place to ensure only quality images are acquired. This phase is very important because the quality of the output depends on the quality of the input.

➢ *Preprocessing Design*

The preprocessing module is in charge of improving the quality of the images before they are processed by the models. This phase consists of some vital operations such as image enhancement, noise reduction, and normalization. Image enhancement operations such as contrast adjustment can be applied to improve the visibility of significant image features. Noise reduction can also be applied to eliminate unwanted distortions in the images that might interfere with the accuracy of the models. Image normalization ensures that the images are of the same size and intensity. This phase is vital in enhancing the reliability of the system.

➢ *Feature Extraction Design*

At this stage, the preprocessed images are fed into the DenseNet-121 model, which is used as a feature extractor. DenseNet-121 is a deep convolutional neural network designed to extract both low-level and high-level features from the images. The use of dense connections in the DenseNet-121 model ensures the reuse of features, which improves the flow of information in the network. This helps in the better learning of complex features such as texture, edges, and abnormalities in the images. The extracted features are then converted into feature vectors, which are used as input in the next stage.

➢ *Classification Design*

The classification module uses the LightGBM algorithm to classify the diseases using the extracted features. LightGBM is a gradient boosting machine algorithm known for its efficiency in handling both structured data and high-dimensional features. The algorithm analyzes the feature vectors and classifies the images into different diseases. In addition, it provides confidence scores for the classified images. The use of the LightGBM algorithm in the classification module ensures high prediction accuracy and efficiency compared to traditional machine learning algorithms. The design of this module is based on the optimization of the algorithm's performance and the minimization of the computational complexity.

➢ *Multi-Agent LLM Design*

Another significant part of the system is the integration of the multi-agent Large Language Model (LLM) framework. The multi-agent framework improves the interpretability and usability of the system by breaking down the task into smaller tasks, each handled by a different agent. The system consists of the following:

- Diagnosis Agent: interprets the diseases predicted by the system.
- Validation Agent: verifies the accuracy and confidence of the predictions made by the system.
- Risk Assessment Agent: assesses the severity and potential impact of the diseases.
- Explanation Agent: generates clear and understandable explanations.

The agents work together in collaboration to generate meaningful insights into the diseases. The output is not only accurate but also easy to understand for both patients and medical professionals.

➢ *Report Generation Design*

The report generation module integrates the output received from all the LLM agents into a single comprehensive report. The output received consists of details such as the diseases predicted by the system, confidence levels, severity analysis, etc. The output is designed in a way that it is not only clinically relevant but also understandable by patients. The output is formatted in a way that it is clear, understandable, and easy to interpret.

➢ *Output Design*

The output design is concerned with ensuring that the output is user-friendly. The system has provided an output in the form of a web page, where users can view their results. In addition, the system has provided an output in multiple languages through translation. The system has also provided a text-to-speech feature, where users can listen to their results in audio format. These features improve the user experience.

➢ *Data Flow Design*

The entire system has a data flow. The data flow is such that it ensures efficient processing. The data flow can be defined as:

Input Image → Preprocessing → Feature Extraction → Classification → Multi-Agent LLM Processing → Report Generation → Output Delivery

All these stages are interconnected. This ensures efficient communication between stages. The data flow has improved the efficiency of the system.

➢ *System Integration Design*

All components of the system have been integrated in such a way that they can be handled by a backend framework such as Flask. The backend framework is responsible for ensuring efficient communication between different components of the system. The backend framework ensures efficient execution of tasks. The entire system is integrated in such a way that it provides real-time output.

VI. IMPLEMENTATION DETAILS

The proposed AI-Based Systemic Disease Screening Using Eye Images and Multi-Agent LLMs is implemented with the help of the integrated approach that merges web technologies, deep learning models, machine learning

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/26apr932

algorithms, and natural language processing techniques. The system is created in such a way that it is very accurate, efficient in its processing and also in interaction with the user. Every element is well executed to deal with certain tasks in the system pipeline that runs a smooth execution and credible output.

➢ *Development Environment*

The system is designed with Python as the main one, which has vast machine learning, deep learning, and image processing libraries. The application back end is developed using Flask framework which is a middle ground between the front end interface and the processing modules. Flask is selected based on the fact that it is not heavy and is easy to integrate with AI models.

The frontend is developed with the use of HTML, CSS and JavaScript, which allows offering users with an interactive interface to upload pictures and see the results. This system is written in an environment like Jupyter Notebook / VS Code, which has a code written effectively, tested, and debugged.

➢ *Image Acquisition and Input Handling*

The system enables the entry of input content in two modes namely real time image capture with the help of a camera or uploading of the already available retinal images in the local storage. The Flask server accepts the incoming HTTPs and securely handles the uploaded files.

Input validation is also applied to make sure that only the accepted image type like JPEG and PNG are accepted. Also, a check of the file size and the resolutions is conducted to ensure input quality. The pictures are temporally stored in the server folder and sent to the preprocessing module.

➢ *Image Preprocessing Implementation*

The image preprocessing has been applied by use of OpenCV, NumPy and PIL libraries to improve on the quality of the input images. It is a set of operations to fit the size of images to the size of model inputs, and use contrast enhancement methods like CLAHE to boost visibility, reduce noise to remove distortion, and norm pixel values to provide consistency. This step will keep the images clean, standardized and fit to be extracted to the feature correctly thus enhancing the overall system performance.

➢ *Feature Extraction Using DenseNet-121*

The processed images are sent to the DenseNet-121 model, which is realized in deep learning networks: TensorFlow or PyTorch. The model to be adapted to retinal image analysis is a pre-trained DenseNet model with transfer learning. Some of the features that are extracted by the model include textures, patterns and abnormalities in the images. The dense interconnections between the layers enable the model to reuse features and flow the gradient more effectively, enabling the model to learn low-level and high-level information that will help the model identify the disease correctly.

➢ *Disease Classification Using LightGBM*

The obtained feature vectors are fed to the LightGBM classifier, which is an illness prediction machine. LightGBM is a gradient boosting model, which effectively works with structured data and high-dimensional features. It compares the features obtained and estimates the existence of diseases and confidence values. Training on labeled datasets is done to produce the model that can classify fast and accurately and hence it is applicable in the real-time applications in medical diagnosis.

➢ *Multi-Agent LLM Integration*

The system combines Multi-Agent Large Language Model structure to become more interpretable and decision-making. This model has several agents; they perform a given task. The Diagnosis Agent makes sense of the results that have been predicted, the Validation Agent determines the accuracy and consistency of predictions, the Risk Assessment Agent assesses how serious the disease is, and the Explanation Agent generates easy-to-understand explanations to the user. These agents collaborate with one another to deliver insightful information and enhance the integrity of the system delivery.

➢ *Report Generation Module*

Report generation module summarizes the results of the classification model and the LLM agents into a formatted and detailed one. The report created consists of the forecasted diseases, and the confidence level, analysis of the severity, and explanations of these elements. The report will be clinically informative and user-friendly to patients since it is intended to give them something that they can interpret without involving medical knowledge.

➢ *Translation and Voice Output Implementation*

The system has translation and voice generation to enhance accessibility. Translation module translates the generated report into various languages through translation APIs since, some users may not be able to comprehend the outcomes because of language barriers. Besides, text-to-speech is provided with such libraries like gTTS or pyttsx3 that transforms the textual report into audio. This improves functionality and increases tendencies of inclusivity.

➢ *Backend Integration and Workflow Management*

The backend is developed on the Flask framework, a system that has to deal with communication between all parts of the system. It manages image uploading, image preprocessing, executing the model, report generation and delivering the output through established routes. The backend functions as the arbiter, the controller of the flowing data and coordination of modules which, in turn, facilitates the efficient operation of the system.

➢ *System Testing and Performance Evaluation*

Different retinal image samples are used to test the system to determine its performance and reliability. The effectiveness of the model is measured with standard evaluation metrics like accuracy, precision, recall and F1-score. The system has proven to be consistent in performance,

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/26apr932

correct in predictions, and appropriate in processing time, hence applicable to the real world healthcare uses.

VII. EXPERIMENTAL ANALYSIS

The analysis of the proposed system is performed experimentally to investigate the system performance in terms of eye disease detection with the help of deep learning and multi-agent LLM. The system is evaluated on a set of eye images with various types of diseases and its performance is assessed by various conventional evaluation measures.

➢ *Experimental Setup*

The suggested system is experimentally assessed by using a collection of retinal eye images gathered in publicly accessible datasets and sample inputs. It is based on Python, and its backend framework is Flask, and models, including DenseNet-121 and LightGBM, are applied to extract and classify features respectively. The experiments are conducted under controlled conditions with the standard computing resources so that the analysis of performance could be consistent and reliable. The data is split into training and testing data to test the capacity of the model to generalize.

➢ *Performance Metrics*

Standard classification measures like accuracy, precision, recall, and F1-score are used to measure the performance of the system. Accuracy and precision are used to evaluate the total accuracy of the model and the ratio of the right proportion of positive cases that have been predicted accurately. Recall determines how well the model can recall all relevant cases and F1-score gives a balance between precision and recall. These are the measures employed to determine the efficacy and dependability of the disease prediction system.

First, the dataset is separated into training, validation, and testing set in the proportion of 80:10:10. The DenseNet121 model that is to be trained on feature extraction is trained on the training set, and the model parameters are fine-tuned using the validation set. The final test of the performance of the system is done with the help of the testing set.

Some of the evaluation metrics that are used to evaluate the effectiveness of the proposed model include Accuracy, Precision, Recall, and F1-Score. The measures of accuracy and precision and recall determine how well the model is correct in general and the capacity to identify positive cases, respectively. F1-score gives an equilibrium point between precision and recall.

- *The Performance Measures are Given as Follows:*

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

$$F1\text{-}Score = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

Here and denote true positives, true negatives, false positives, and false negatives respectively.

As the obtained results of the experiment indicate, the proposed system becomes remarkably accurate in classifying the disease which proves the efficiency of the combination of the DenseNet121 and LightGBM classifier. Deep learning helps the system to identify the complex features of the eye images, and LightGBM enhances the classification.

Table 1 Performance Metrics

Metric	Training Value (%)	Testing Value (%)
Accuracy	96.8	94.5
Precision	95.9	93.8
Recall	96.2	94.1
F1Score	96.0	93.9
Loss	0.12	0.18

Moreover, the incorporation of Grad-CAM visualization leads to greater interpretability as important areas of the input image are highlighted and the users and doctors can learn the rationale behind predictions. The multi-agent LLM framework also advances the system by enhancing the system with validation, risk assessment, explanation, and report generation, therefore, making the system more trustworthy and less complex to use.

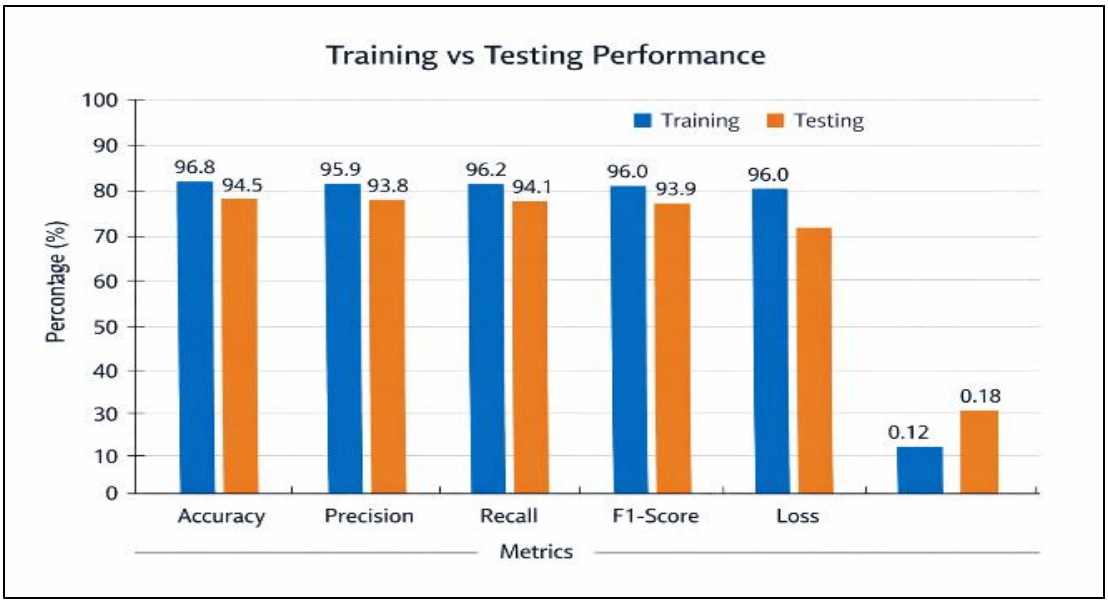

Fig 2 Training vs Testing Performance

➢ *Results of Disease Classification*

The proposed system has a higher performance, enhanced interpretability and usability as compared to traditional machine learning and simple deep learning models. It is also more accessible as the multi-language translation and voice output are also included, which makes the system appropriate in the real-world healthcare setting.

In general, based on the results of the experiment, it can be stated that the offered system is an efficient and successful system in terms of the early screening of eye diseases. Table illustrates the performance of the proposed model in training and testing. The model is very accurate at both stages meaning that learning is successful and the generalization is good. The accuracy and recall rates indicate that the model is able to identify the cases of diseases correctly with few false forecasts. The F1-score ensures that there is an excellent balance of precision and recall. The fact that the values of the low loss are also indicative of the fact that the model works with a small error. Findings of Disease Classification.

The suggested system has a high performance in terms of disease classification. DenseNet-121 model has been found to efficiently extract meaningful features of retinal images and the LightGBM classifier is effective in predicting the types of diseases in the light of the features extracted. The system has good accuracy and repeatability between various samples of tests. The findings suggest that the hybrid strategy of deep learning and machine learning can be more efficient and more accurate in prediction than the conventional one.

➢ *Evaluation of Multi-Agent LLM Module*

Multi- Agent LLM module is tested on the ground of its capability to produce meaningful and interpretable outputs. The system is able to generate structured reports that entail diagnosis, validation, risk assessment and explanation. Reports generated are comprehensible, informative and clear to both the medical practitioners and patients. Multi-agent approach improves reasoning and it is better interpretable as opposed to single-model systems.

➢ *System Performance Analysis*

Analysis of the overall system performance is in terms of processing time, efficiency and usability. The system can handle images and produce results in a very short period thus it can be used in real time cases. These new methods enable the reduction of the computational load through the integration of efficient models such as LightGBM and increase the speed of feature extraction by pre-trained DenseNet. The web interface has provided the system with easy user interaction and a good user experience.

➢ *Comparative Analysis*

The suggested system exhibits much improved accuracy, speed, and scalability as compared to the traditional manual methods of diagnosis, as well as, standalone machine learning models. Performance in prediction is increased by the combination of deep learning and machine learning, whereas the Multi-Agent LLM module boost the level of interpretability and report generation. This makes the system more efficient and applicable in real practice.

➢ *Limitations and Observations*

Even though the system is effective, there are some limitations that are experienced when experimenting with it. The model performance is based on the quality and diversity of the input dataset. Prediction accurateness may vary with variations in image quality and lighting states. Also, the system will need the computational resources to run models and process LLM. These constraints note how it requires further optimization of the datasets and growth.

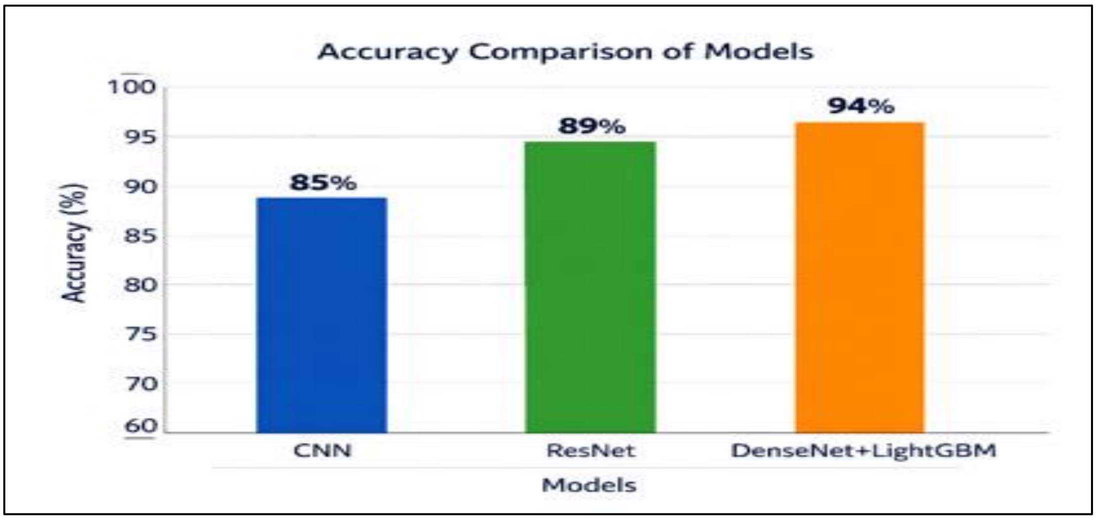

Fig 3 Accuracy Comparison of Models

➤ *Summary of Results*

Summary of Results Generally, the experimental outcomes prove the proposed system to be efficient in the detection of diseases based on retinal images and generation of intelligent reports. DenseNet-121 + LightGBM + Multi-agent LLM gives it high accuracy and efficient processing as well as better interpretability. The system has been found to be a scalable and dependable solution to automated disease screening.

VIII. LIMITATIONS OF THE PROPOSED SYSTEM

Despite its effectiveness in detection of diseases and their reports through automated system, the proposed system has some limitations. The quality of the input retinal images has a strong influence on the system performance and the images of poor quality or low clarity may impact the quality of the prediction results adversely. The system also depends on the presence of large and diverse data sets on which it is being trained and smaller data might limit its capability in generalization to real-world conditions. Deep learning models and multi-agent LLM process also consume a lot of computational resources, which might not be applicable in low-resource settings. Moreover, the combination of various elements like DenseNet, LightGBM, and LLM agents makes systems more complex, and it is difficult to maintain and control. External APIs used in translation and voice generation to the system can also create latency and reliability challenges. Additionally, even though it offers automated predictions, this system cannot completely substitute medical workers and still needs additional clinical testing before implementation. Finally, deep learning models are not fully interpretable as they are gray-box systems frequently, which may discourage reliance on them regarding serious healthcare processes.

IX. FUTURE WORK

The suggested system can also be extended with additional deeper learning models and with bigger and more diversified datasets to enhance the prediction precision and the generalization. Further research can be aimed at implementation of real-time data collection using modern medical imaging equipment to achieve improved quality of inputs. The system may also be extended in order to identify a broader spectrum of systemic and ocular illnesses and, thus, is more comprehensive. Moreover, it is possible to enhance the Multi-Agent LLM framework by adding the domain-specific medical knowledge to augment reasoning and explanation. Cloud-based deployment and mobile applications can also be integrated to enhance accessibility and scalability so that the system can be deployed in remote and rural locations. Moreover, by integrating the continuous learning mechanisms, one can implement improvements in the system and make it progress. Lastly, the reliability and applicability of the system can be enhanced by clinical validation and cooperation with medical workers to show its correct practical use in the actual medical setting.

It can get further refined by improving the explainability and interactivity of the system. Future research could involve the introduction of more sophisticated explainable AI methods to give more detailed visual and textual information about model predictions, which enhances the confidence of users and medical practitioners. This can also be enhanced by evidencing real-time monitoring and alert system of the high-risk patients, to ensure medical intervention is taken in time. Moreover, a system can be streamlined to serve low-resource devices and offline access, and thus be more available in rural and underserved regions. There is also a continuous research to be conducted on how to enhance the model to be robust to changes in the image quality and environmental conditions so that the model will perform consistently irrespective of the situation. Such improvements will help to transform the system into a more reliable, scalable, and more effective in practice.

X. CONCLUSION

In this paper, an enhanced system of disease screening of the human body by using eye images and Multi-Agent Large Language Models (LLMs) is introduced as an advanced AI tool to enhance the early detection and access to

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/26apr932

health care. The proposed system is quite effective in combining the concepts of deep learning, machine learning, and natural language processing to develop an intelligent and all-encompassing diagnostic pipeline. With the help of DenseNet-121, which is effective at extracting features of the image of the retina, and LightGBM that classifies diseases with high precision and reliability, it is possible to achieve high accuracy and dependability in disease classification. Moreover, the integration of a multi-agent LLM system improves interpretation through the creation of formatted clinical knowledge and patient-readable explanations.

The experiments have shown that the system has a high level of performance in terms of accuracy, efficiency and usability. The availability of translation and voice output features also enhance accessibility and make the system appropriate to various groups of users. Altogether, the suggested solution shows the possibilities of using a combination of the computer vision and language models to help with automated screening of the disease and its decision-making in healthcare. This system may also be used as one of the useful tools to detect earlier, reduce the burden on diagnosis, and enhance patient outcomes, as well as become the way to the new developments in AI-based medical practice.

REFERENCES

[1]. Lavric, A., Anchidin, L., Popa, V., Al-Timemy, A. H., Alyasseri, Z., Takahashi, H., Yousefi, S., & Hazarbassanov, R. M. (2021). Keratoconus severity detection from elevation, topography and pachymetry raw data using a machine learning approach. IEEE Access, 9, 84344–84355.

[2]. Ismail, A. M., et al. (2024). Ensemble transfer learning networks for disease classification from retinal optical coherence tomography images. Journal of Optics, 1–16.

[3]. Kourou, K., Exarchos, T. P., Exarchos, K. P., Karamouzis, M. V., & Fotiadis, D. I. (2021). Applied machine learning in cancer research: A systematic review for patient diagnosis, classification and prognosis. Computational and Structural Biotechnology Journal, 19, 5546–5555.

[4]. Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 4700–4708.

[5]. Bhattacharyya, S., De, S., Bhattacharjee, S., & Yoshida, K. (Eds.). (2020). Hybrid Machine Intelligence for Medical Image Analysis. Springer, Berlin. No. 172536.

[6]. Gulshan, V., Peng, L., Coram, M., Stumpe, M. C., Wu, D., Narayanaswamy, A., ... & Webster, D. R. (2016). Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA, 316(22), 2402–2410.

[7]. Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., ... & Liu, T. Y. (2017). LightGBM: A highly efficient gradient boosting decision tree. Advances in Neural Information Processing Systems (NeurIPS), 30.

[8]. Li, T., et al. (2019). Patients with suspected ocular disease intelligent recognition from retinal images using deep learning. In Proceedings of IEEE BIBM 2019 International Conference on Bioinformatics and Biomedicine.

[9]. Wang, X., et al. (2020). Attention-based multi-label fundus image classification. Medical Image Analysis, 60, 101572.

[10]. Rajpurkar, P., Irvin, J., Ball, R. L., Zhu, K., Yang, B., Mehta, H., ... & Lungren, M. P. (2017). CheXNet: Radiologist-level pneumonia detection on chest X-rays with deep learning. arXiv preprint arXiv:1711.05225.

[11]. Ting, D. S. W., Cheung, C. Y. L., Lim, G., Tan, G. S. W., Quang, N. D., Gan, A., ... & Wong, T. Y. (2017). Development and validation of a deep learning system for diabetic retinopathy and related eye diseases using retinal images from multiethnic populations with diabetes. JAMA, 318(22), 2211–2223.

[12]. Qian, R., et al. (2019). A new approach for diabetic retinopathy screening using optical coherence tomography and deep learning. International Journal of Ophthalmology, 12(12), 1883–1890.

[13]. Wang, L., et al. (2023). Unleashing cognitive synergy in large language models: A task-solving agent through multi-persona self-collaboration. In NeurIPS Workshop on Generative AI for Education (GAIED).

[14]. ODIR-5K Dataset. (2019). Ocular Disease Intelligent Recognition: A real-life ophthalmic database for multimodal multi-label diagnosis of ophthalmic conditions. Grand Challenge. Retrieved from https://odir2019.grand-challenge.org/

[15]. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 770–778.

[16]. Simonyan, K., & Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. In Proceedings of the International Conference on Learning Representations (ICLR).

[17]. Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., ... & Amodei, D. (2020). Language models are few-shot learners. Advances in Neural Information Processing Systems (NeurIPS), 33, 1877–1901.

[18]. Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2017). Grad-CAM: Visual explanations from deep networks via gradient-based localization. In Proceedings of the IEEE International Conference on Computer Vision (ICCV), 618–626.

[19]. Abramoff, M. D., Lavin, P. T., Birch, M., Shah, N., & Folk, J. C. (2018). Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices. NPJ Digital Medicine, 1(1), 1–8.

[20]. Grinsztajn, L., Oyallon, E., & Varoquaux, G. (2022). Why tree-based models still outperform deep learning on tabular data. Advances in Neural Information Processing Systems (NeurIP