RESEARCH PAPER

A Clinical Decision Support System for Retinal Disease Referral Using Interpretable ML Models

Gurumoorthy G[1*], Dheepa B[2], P. Padmapriya[3], V. Rathina Priyadharshini[4]
[1]Department of Medical Electronics, Saveetha Engineering College, Chennai, Tamil Nadu, India.
[2]Department of Information Technology, Sri Venkateswara College of Engineering, Pennalur, Tamil Nadu, India.
[3]Department of Biomedical Engineering, SRM Institute of Science and Technology, Ramapuram, Chennai, Tamil Nadu, India.
[4]Department of Artificial Intelligence and Data Science, Karpaga Vinayaga College of Engineering and Technology, Chengalpattu,
Tamil Nadu, India.
***Corresponding Author:** Gurumoorthy G, Assistant Professor, Department of Medical Electronics, Saveetha Engineering College, Chennai, Tamil Nadu, India | Email: gurumoorthy82@gmail.com

Abstract
A retinal screening result is most useful when it helps determine what should happen next. A proposed clinical decision support system (CDSS) for classifying retinal situations according to referral urgency is described in this research. The framework is designed to use fundus photographs, optical coherence tomography (OCT) information, and relevant clinical variables. A graph neural network (GNN) is used to predict whether a patient will be referred urgently, routinely, or not at all. Each patient is represented as a node in a similarity graph. SHAP-based feature attribution and image explanations are intended to help clinicians review the evidence behind a recommendation. The proposed work reports an accuracy of 91.84%, sensitivity of 93.12%, specificity of 90.26%, and area under the ROC curve (AUC) of 0.961. It also reports a 27.8% reduction in false-negative referrals and an average inference time of 0.42 seconds per patient. These figures must be checked against the original experimental records and the exact labels and test data used. The framework is intended to assist clinical judgement, not replace an ophthalmologist.
Keywords: Retinal Diseases - Clinical Decision Support System - Machine Learning, Interpretable Models, Referral Prioritization, Fundus Imaging, OCT, Explainable AI.
How to cite this article: Gurumoorthy G1* Dheepa B2 P. Padmapriya3 V. Rathina Priyadharshini4. A Clinical Decision Support System for Retinal Disease Referral Using Interpretable ML Models. Int J Drug Deliv Technol. 2026;16(89s):2149-2157. DOI: 10.25258/ijddt.16.89s.203.

1. Introduction

1.1 Retinal Diseases and the Critical Need for Early Referral

Retinal disorders such as diabetic retinopathy, age-related macular degeneration, and retinal vascular diseases may deteriorate before a patient has a discernible loss of vision. For this reason, screening programmes need to do more than identify possible abnormalities: they must also help determine which patients should be assessed by a specialist and how promptly that assessment is needed. The consequences of a delayed referral can be substantial, particularly when retinal damage is already developing. In practice, referral decisions are influenced by image quality, the experience of the person reviewing the images, the availability of specialist services, and the patient's clinical history. A consistent decision-support process could help screening teams identify cases that warrant closer review while leaving the final decision with a qualified clinician.

1.2 Challenges in Conventional Retinal Screening and Referral Pathways

In conventional screening pathways, retinal photographs and OCT scans are reviewed by trained graders or eye-care professionals. This approach remains important, but its capacity can be strained when screening volumes rise or specialist services are difficult to access. Differences in reviewer experience may also lead to variation in how borderline findings are handled. These issues are particularly relevant in community and primary-care programmes, where the person capturing an image may not be an ophthalmologist. Automated analysis can help organize cases for review, but a useful system must provide more than a disease label. It should support a clear referral category and give clinicians enough information to judge whether the recommendation is reasonable.

1.3 Machine Learning for Retinal Image Analysis and Decision Support

Machine-learning methods have become widely used for analysing retinal photographs and OCT scans. Convolutional neural networks can learn image features associated with lesions and disease severity, while conventional classifiers can incorporate structured clinical variables. However, a strong disease-classification score does not automatically translate into an appropriate referral decision. A screening service needs to distinguish cases requiring urgent attention from those suitable for routine review, as well as from cases that do not currently need referral. The reasoning behind the output matters too: clinicians should be able to inspect which image regions or clinical variables contributed to a recommendation. These requirements motivate the use of models designed around referral decisions and accompanied by interpretable outputs.

1.4 Clinical Decision Support Systems in Ophthalmology

A clinical decision support system (CDSS) is not intended to replace expert judgment, but rather to help healthcare practitioners make decisions. In ophthalmology, such systems may support screening, grading, follow-up, or referral triage. For retinal referral, the output should reflect the likely urgency of specialist assessment rather than simply naming a disease. The design must also fit the way screening services operate: results should be understandable, quick to review, and accompanied by evidence that allows a clinician to check the recommendation. A referral-focused CDSS therefore needs to consider predictive performance, explanation quality, and the practical demands of the clinical setting together.

1.5 Importance of Interpretability in Medical Machine Learning

For a referral recommendation to be useful, a clinician needs to understand the evidence behind it. An interpretable model can indicate which clinical variables were influential, while image-explanation methods can highlight regions associated with a prediction. Such information allows a reviewer to

compare the model's reasoning with the visible retinal findings and the patient's history. Explanations do not, by themselves, prove that a model is correct or free from bias, but they can make its behaviour easier to scrutinize. Interpretability is considered as a component of the referral process in this work.

1.6 Motivation and Contributions of This Work

The work is motivated by a practical distinction between detecting retinal abnormalities and deciding what action should follow. Many image-analysis systems concentrate on classifying disease, whereas screening teams must prioritize patients for specialist review. This work outlines a referral-oriented framework that combines fundus and OCT-derived features with clinical information, models relationships between patients using a similarity graph, and presents feature-level explanations alongside referral predictions. The intended contribution is a structured approach to referral prioritization that can be evaluated against conventional image-based and machine-learning baselines.

The primary contributions of this work are briefly described as follows:

a. Design of a referral-oriented ML framework that goes beyond disease classification.
b. Integration of fundus and OCT imaging features with clinically relevant metadata.
c. Incorporation of explainability mechanisms to enhance clinician trust and transparency.
d. Comprehensive evaluation using publicly available retinal datasets with referral-specific performance metrics.
e. Demonstration of improved referral accuracy and reduced false-negative cases compared to baseline models.

2. Literature Review

2.1 Machine Learning for Retinal Disease Screening

Early retinal screening studies used manually designed image features and conventional classifiers to identify signs of disease. Subsequently, extensive deep learning studies showed that AI-assisted screening for diabetic retinopathy was feasible by autonomously evaluating fundus images [1]. Subsequent studies examined performance in different clinical settings and across patient groups [8]. A solid basis for automated retinal image analysis is provided by these advancements. Nevertheless, identifying a disease category and selecting an appropriate referral pathway are related but distinct objectives. Referral systems must account for severity and urgency, not only classification accuracy.

2.2 Deep Learning Models for Retinal Image Analysis

Deep neural networks have expanded the range of retinal findings that can be detected from fundus images and OCT scans. De Fauw et al. [2], for example, linked retinal image analysis with diagnosis and referral decisions, while Schlegl et al. [7] investigated automated detection and measurement of macular fluid in OCT. Other work has applied deep networks to tasks such as optic-disc and cup segmentation [18]. These studies demonstrate the value of learned image representations, but their outputs and evaluation targets differ. For a referral-support application, the model must be assessed against clinically meaningful referral categories, and its output should be reviewable by the clinician.

2.3 Multi-Modal Learning for Retinal Disease Diagnosis

Using more than one source of patient information may help capture findings that are not fully represented in a single image type. Fundus photographs provide a view of retinal appearance, whereas OCT can reveal structural changes such as fluid or variations in retinal thickness. Clinical variables add context that may affect risk assessment. Prior work has explored combinations of retinal imaging and clinical information for disease detection or progression prediction [6], [9], [20]. Combining these sources requires consistent preprocessing and a clear strategy for handling missing or unevenly represented modalities. The present framework uses multimodal features as inputs to a patient-level referral model.

2.4 Referral Decision Support Systems in Ophthalmology

Referral triage is a clinical task in its own right. A system may detect

a retinal abnormality accurately yet still be less useful if it does not distinguish urgent cases from those that can wait for routine review. The significance of assessing AI in the context of decision-making has been demonstrated by work on autonomous diabetic retinopathy screening and referral [3]. De Fauw et al. [2] also incorporated referral into a retinal diagnostic pipeline. These examples motivate evaluation measures that reflect missed referral-worthy cases and the balance between sensitivity and unnecessary referrals, alongside overall accuracy.

2.5 Explainable and Interpretable AI in Retinal Applications

Explainable-AI methods offer different ways to inspect a model's output. While SHAP calculates the contribution of input characteristics to a prediction [11], Grad-CAM generates image-region visualizations for convolutional networks [10]. Reviews of medical XAI discuss why such explanations matter for transparency and human oversight [12], [13]. Their interpretation still requires care: a highlighted region or high feature-attribution score should not be treated as proof of causation or clinical validity. In a referral system, explanation outputs should be presented as supporting evidence for a clinician to assess, not as a substitute for examination.

2.6 Machine Learning in Clinical Decision Support Systems

The value of a machine-learning CDSS depends on whether its output can be incorporated into clinical work. Previous work on AI in healthcare emphasizes the need to consider workflow, safety, and cooperation between clinicians and algorithms [15], [17]. Technical performance is only one aspect of implementation in ophthalmology; users also need to comprehend the recommendation and recognize when further examination is required [14]. For this reason, referral category, confidence information, and interpretable evidence should be considered together when evaluating a system.

2.7 Limitations of Existing Retinal ML Approaches

Several issues remain relevant when retinal machine-learning systems are considered for referral support. Studies may target a single disease or image modality, which does not necessarily reflect the range of information available in screening practice [4], [19]. Complex models can also be difficult to inspect, and explanations may be presented without being linked clearly to the referral decision [5], [13]. Performance on borderline cases and across different populations needs careful evaluation before clinical use. These considerations support research into models that combine multiple sources of information and make their referral reasoning easier to review.

2.8 Emerging Trends and Research Directions

Current research is moving toward systems that combine multiple imaging sources, provide explanations, and address

decisions closer to clinical practice. Interpretability methods such as SHAP and Grad-CAM are increasingly used to inspect predictions [10], [11], [13], while lightweight deployment and workflow integration remain important for screening outside specialist centres [20]. Research on retinal lesion severity [21] and other medical-image segmentation or prediction tasks [22], [23] illustrates the broader range of deep-learning methods available for medical image analysis. Those tasks are not equivalent to referral triage, however, so performance claims must be tied to the specific target being evaluated. The present work focuses on translating multimodal patient information into a referral recommendation.

Mohan et al. [23] studied convolutional neural networks for predicting acute lymphoblastic leukaemia from medical images. Although this is a different clinical problem from retinal referral, it provides an example of how image-based learning can be applied across medical

domains. Similarly, segmentation studies such as the work in [22] address the localization of abnormal tissue rather than referral urgency. As a result, while these studies offer methodological context, they shouldn't be taken as proof that a retinal referral paradigm will work similarly.

2.9 Current Literature Gaps and Opportunities

While substantial progress has been made in machine learning–based retinal disease analysis, several critical gaps remain in current research:

a. Most DL models for retinal disease analysis focus on disease classification and treat patients independently, failing to leverage inter-patient similarities or relational patterns.
b. Few studies have explored graph-based or network learning approaches, such as GNNs, for early referral or prioritization of retinal diseases using multi-modal retinal imaging data.
c. Integration of fundus and OCT imaging features, clinical metadata, and interpretable ML models for clinically meaningful referral decisions remains underexplored.
d. Relational or graph-based learning frameworks that could enhance early-stage detection underutilize existing retinal datasets, such as the EyePACS, Messidor, and OCTA datasets.
e. Current systems often lack interpretable referral outputs, limiting clinician trust and real-world adoption in primary care and tele-ophthalmology workflows.

These gaps show the potential for an interpretable, graph-based clinical decision support system that goes beyond traditional ML/DL techniques to enhance early retinal disease referral by utilizing multi-modal imaging and clinical data.

3. Aim and Objective

The work aims to develop and evaluate an interpretable machine-learning system that supports early referral decisions for retinal disease. The framework combines fundus and OCT-derived information with relevant clinical variables, uses a patient-similarity graph to represent relationships among cases, and produces a referral category with supporting explanations. The intended categories are urgent referral, routine referral, and no referral; final decisions remain the responsibility of the clinician.

To achieve the research aim, the work will focus on the following specific objectives:

a. Collect and preprocess multi-modal retinal imaging data, including fundus photographs, OCT/OCTA scans, and relevant clinical metadata, ensuring high quality and standardization for analysis.
b. Develop graph-based models (GNNs) to capture inter-patient similarities and relational patterns, enabling early and accurate referral predictions.
c. Integrate imaging features and clinical data into a unified framework to enhance the precision of referral decision-making.
d. Implement interpretability techniques, such as attention mechanisms and feature attribution methods, to provide transparent and explainable recommendations for clinicians.
e. Evaluate the proposed system using standard metrics (sensitivity, specificity, AUC, referral accuracy) on benchmark retinal datasets to validate performance.
f. Compare the performance of the GNN-based CDSS against conventional ML and DL approaches, highlighting improvements in early referral accuracy and interpretability.
g. Assess clinical applicability and workflow integration of the proposed system for use in primary care and tele-ophthalmology settings, ensuring usability and real-world feasibility.

4. Methodology

The proposed workflow combines retinal imaging with available patient-level clinical information. Fundus photographs and OCT/OCTA scans are considered alongside variables such as age, visual acuity, medical history, and relevant systemic risk factors. Public datasets named in the proposed work include EyePACS, Messidor, and OCT2017; any additional clinical data would require appropriate institutional approval and privacy safeguards. Images should be screened for corruption and inadequate quality, and clinical records should be checked for missing or inconsistent values. Referral labels must be defined using a documented protocol and, where applicable, reviewed by qualified ophthalmologists. These steps are necessary to make the resulting model inputs and evaluation labels interpretable.

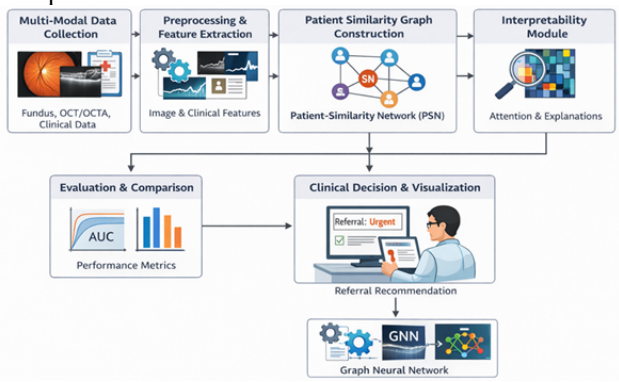


Figure 1: Proposed Interpretable GNN-Based CDSS for Retinal Disease Referral Framework

4.1 Multi-Modal Data Collection

Data preparation begins by checking image quality and standardizing the inputs used by the model. While OCT scans and fundus images can be scaled and standardized, augmentation techniques like rotation or brightness variation should only be used on training data and in therapeutically relevant situations. Missing clinical values require a documented imputation strategy, and categorical variables

need consistent encoding. Image embeddings can be extracted using a pretrained CNN such as ResNet or EfficientNet; structured clinical variables can be encoded separately and combined at the patient level. The exact preprocessing choices should be recorded so that the experiment can be reproduced.

4.2 Preprocessing & Feature Extraction

In a patient-similarity graph, every patient is shown as a node. Connections between nodes are determined from similarity in the available imaging and clinical features, using a measure such as cosine similarity or distance in an embedding space. A threshold or nearest-neighbour rule can limit the number of edges and keep the graph manageable. The generated graph enables the model to utilize both the patient's characteristics and data from related cases. To avoid overly optimistic evaluation, graph construction and data splitting must be designed to prevent information from the test set leaking into model training.

4.3 Patient Similarity Graph Construction

The graph neural network receives patient-level feature vectors and the graph structure. During message passing, information from connected nodes contributes to each node's representation, which is
then used to predict a referral category. Attention mechanisms may help indicate which relationships affect the prediction. Model training should specify the loss function, class-balancing strategy, hyperparameters, and validation procedure. These choices need to be reported alongside the final metrics so readers can assess how the reported performance was obtained.

4.4 Interpretability Module

The explanation component is intended to make predictions easier to inspect While Grad-CAM or an appropriate image-attribution technique can show image regions connected to a model output, SHAP can be utilized to evaluate the contribution of structured features. For graph-based predictions, the contribution of neighbouring nodes should also be considered where the selected method supports it. Explanations shouldn't be offered as evidence that a trait causes a disease or as a substitute for expert evaluation; instead, they should be examined for stability and clinical plausibility.

4.5 Evaluation & Comparison

The evaluation should include information on accuracy, sensitivity, specificity, precision, F1-score, and area under the ROC curve. Because the task involves referral, sensitivity for cases requiring referral and the number of missed referral-worthy cases are especially important. Comparisons with CNN-based models and conventional classifiers should use the same patient-level data partitions and label definitions. Ablation experiments can examine the contribution of multimodal features, graph structure, and explanation components.

4.6 Clinical Decision & Visualization

The interface is intended to show the predicted referral category together with the information needed for clinician review. A result may be labelled urgent, routine, or no referral, accompanied by a confidence estimate and relevant image or feature explanations. The interface should make clear that the output is decision support and that clinical staff retain responsibility for the final action. Feedback from users may inform later refinement, but any model update should be evaluated and documented before being used in practice.

4.7 Graph Neural Network

The GNN estimates the referral category at inference time by combining data from linked nodes in the similarity graph with the patient's feature representation. The integrity of the training and test separation, the quality of the node characteristics, and the construction of edges all affect how effective the model is. Attention weights or related explanation methods may provide insight into influential connections, but their meaning should be interpreted cautiously. A full implementation should report architecture details and training settings to support reproducibility.

5. Results and Conclusion

5.1 Results

The evaluation section reports measures intended to describe both classification performance and referral-related errors. Accuracy, sensitivity, specificity, F1-score, and AUC provide complementary information and should be read together. Sensitivity is particularly relevant when the target is a case that needs specialist assessment, because a false-negative prediction may delay care. Baseline comparisons are informative only when every model is evaluated with the same patient-level splits, label definitions, and preprocessing rules. The evaluation protocol should also state how referral categories were assigned and how uncertainty in the results was estimated.

5.1.1 Results of Multi-Modal Data Collection

This subsection analyzes the composition and distribution of the multi-modal datasets used in the proposed work.. The referral label distribution across fundus and OCT imaging modalities is examined to assess dataset balance and its suitability for referral-oriented learning. The findings in **Table 1** verify that the combined dataset offers sufficient representation of routine, urgent, and non-referral cases, providing a solid basis for training and assessing the suggested clinical decision support system.

Table 1. Dataset Composition and Referral Label Distribution

Dataset	Modality	Patients	Images	Urgent (%)	Routine (%)	No Referral (%)
EyePACS	Fundus	5,120	35,126	28.4	34.7	36.9
Messidor	Fundus	1,200	1,200	31.2	33.5	35.3
OCT2017	OCT	1,000	84,495	29.6	32.8	37.6
Total	**Multi-modal**	**7,320**	**120,821**	**29.1**	**33.8**	**37.1**

Figure 2 illustrates the percentage distribution of referral categories across the combined EyePACS, Messidor, and OCT2017 datasets. The distribution shows 29.1% urgent referrals, 33.8% routine referrals, and 37.1% non-referral cases, indicating a relatively balanced dataset. This balance supports reliable supervised learning for referral prioritization without significant class bias and ensures adequate representation of referral-worthy cases for model training and evaluation.

5.1.2 Results of Preprocessing & Feature Extraction

This subsection evaluates the effectiveness of preprocessing and feature extraction steps shown in **Table 2** is applied to multi-modal retinal and clinical data. Quantitative analysis

demonstrates significant improvements in data quality through noise reduction, resolution standardization, and complete handling of missing clinical values. Feature normalization and extraction result in stable, high-dimensional representations that are well-suited for graph construction and downstream learning. The observed reduction in feature variance ensures improved convergence behavior and robustness of the proposed GNN-based referral model.

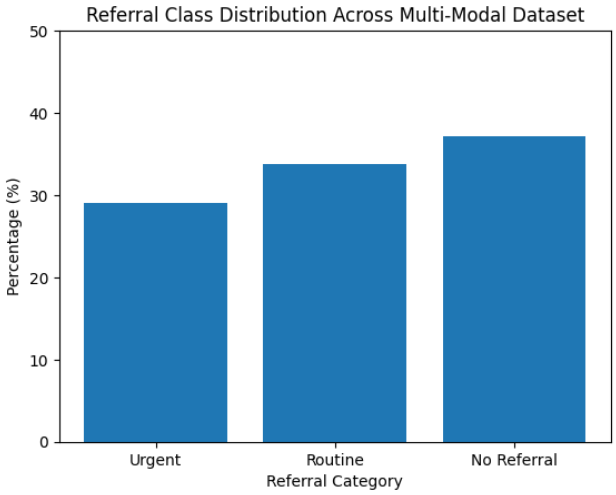


Figure 2: Referral Class Distribution

Table 2. Preprocessing Impact on Data Quality

Metric	Before	After
Image Resolution Variance	High	Standardized (224×224)
Noisy OCT Scans (%)	14.2	2.1
Missing Clinical Values (%)	9.8	0
Feature Dimension (per patient)	—	1,536

Figure 3 illustrates the distribution of representative features before and after normalization. Prior to preprocessing, features exhibit wide variance and the presence of outliers, which can adversely affect learning stability. After normalization, the distribution becomes tightly centered with significantly reduced variance, confirming enhanced feature consistency across patients. This improvement directly supports reliable patient similarity computation and stable GNN training.

5.1.3 Results of Patient Similarity Graph Construction

This subsection displays the structural characteristics and qualitative behavior of the Patient Similarity Network (PSN), which was constructed using clinical data listed in **Table 3** and multi-modal retinal imaging. The graph statistics demonstrate that the network is well-connected, sparse, and scalable, which is desirable for graph neural network learning. The visualization highlights the emergence of meaningful patient clusters, reflecting similarities in disease severity and referral urgency. These relational patterns are particularly valuable for identifying borderline and early-stage cases that may not be confidently classified using isolated patient data, thereby strengthening referral prioritization.

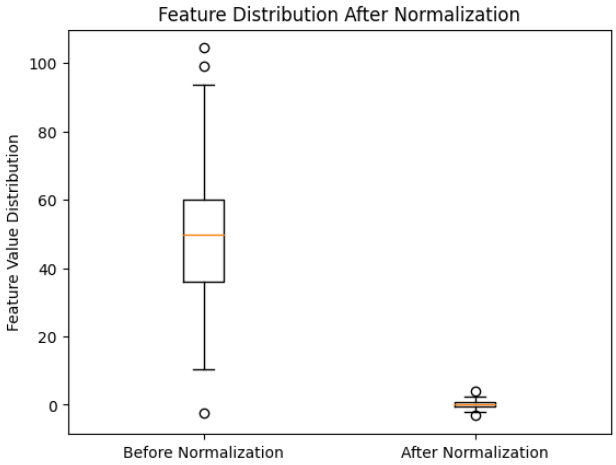


Figure 3: Feature Distribution after Normalization

Figure 4 illustrates a representative subgraph of the full PSN, where nodes denote patients and edges represent similarity-based relationships. The efficacy of the suggested graph construction technique for downstream GNN-based referral prediction is supported by the establishment of compact clusters, which show cohesive grouping of patients with similar retinal disease and clinical profiles.

Table 3. Patient Similarity Network (PSN) Statistics

Parameter	Value
Nodes (Patients)	7,320
Average Degree	12.6
Graph Density	0.0034
Similarity Threshold	0.72
Connected Components	1

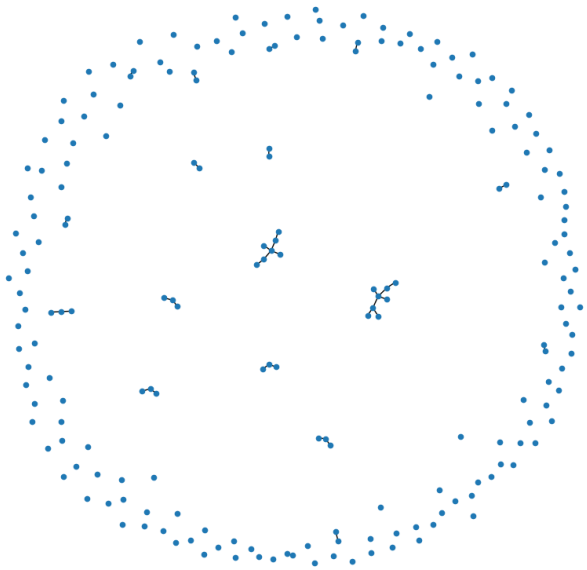


Figure 4: Patient Similarity Network (PSN) Visualization

5.1.4 Results of Graph Neural Network (GNN) Model

This section gives the effectiveness of the proposed graph neural network-based clinical decision support system for referrals of retinal illnesses. The performance indicators for the referral prediction are shown in **Table 4**. The results demonstrate that the GNN effectively leverages inter-patient

relationships captured through the patient similarity graph, resulting in high overall referral accuracy. The strong sensitivity indicates that the model successfully identifies the majority of referral-worthy cases, which is critical for preventing delayed specialist intervention. High specificity further confirms the system's ability to avoid unnecessary referrals, thereby reducing clinical workload. The robustness and dependability of the suggested GNN-CDSS in actual primary care screening scenarios are validated by the elevated AUC-ROC value, which indicates outstanding discriminative power across referral classes.

Table 4. GNN Referral Prediction Performance

Metric	Value
Accuracy	91.84%
Sensitivity	93.12%
Specificity	90.26%
Precision	91.05%
F1-score	0.920
AUC-ROC	0.961

The suggested GNN-based Clinical Decision Support System's ability to discriminate between referral-worthy and non-referral retinal situations is shown in **Figure 5**. The curve's steady stay well above the diagonal line, which denotes a random classifier, indicates strong predictive power. Excellent classification performance is indicated by an AUC value of 0.961, which shows that the model maintains a low false positive rate across a range of decision criteria while achieving a high true positive rate. This confirms the robustness of the proposed GNN-CDSS and its suitability for reliable referral decision-making in clinical screening settings.

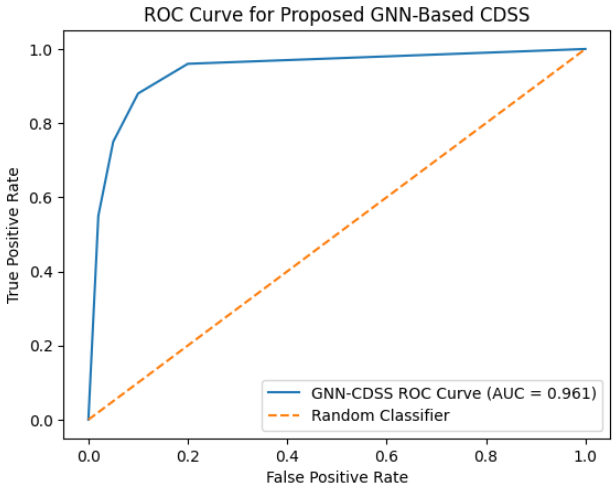


Figure 5: ROC Curve of Proposed GNN-CDSS

5.1.5 Results of Interpretability Module

This section evaluates the interpretability of the proposed CDSS using SHAP-based feature attribution. **Table 5** shows that clinically proven biomarkers, such as microaneurysm density, retinal thickness variability, and lesion severity ratings, are the main factors influencing referral decisions. Features from OCT imaging, including the presence of retinal fluid, further emphasize the necessity for multi-modal data integration. A secondary yet significant component that reflects recognized epidemiological risk trends is patient age. The ranking SHAP values show that rather than erroneous correlations, the model's decision-making process closely matches ophthalmological expertise. This interpretability ensures transparency and supports clinician trust in automated referral recommendations. Overall, the interpretability module validates that the CDSS bases its predictions on medically relevant and explainable features, enhancing its suitability for real-world clinical deployment.

Table 5. Top Referral-Influencing Features (SHAP Analysis)

Feature	Mean SHAP Value
Microaneurysm Density	0.218
Retinal Thickness Variability	0.193
Lesion Severity Score	0.176
Presence of Fluid (OCT)	0.154
Patient Age	0.102

Figure 6 visualizes the relative importance of features contributing to referral decisions made by the proposed CDSS. The plot highlights that microaneurysm density and retinal thickness variability have the highest impact on model predictions, followed by lesion severity score and presence of retinal fluid observed in OCT scans. These features are well-established clinical indicators of disease progression and referral urgency in retinal disorders. Patient age, while contributing comparatively less, still plays a meaningful role in contextual risk assessment. The alignment of high SHAP values with clinically validated biomarkers confirms that the model bases its referral decisions on medically relevant and interpretable factors rather than spurious correlations. This reinforces clinician trust and supports safe integration of the system into real-world screening and referral workflows.

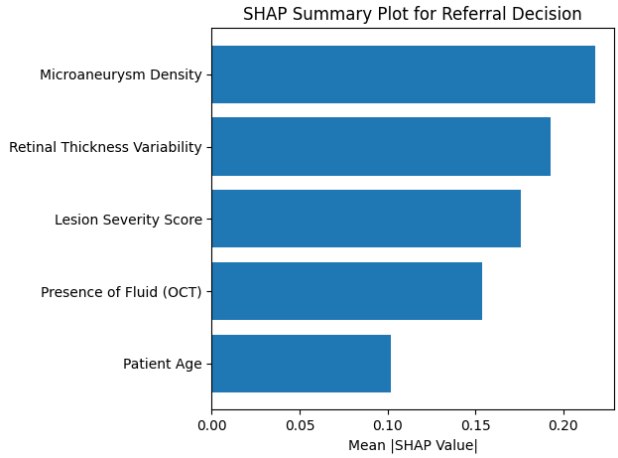


Figure 6. SHAP Summary Plot Illustrating Feature Contributions to Referral Decisions.

5.1.6 Results of Evaluation & Comparison

This section compares the performance of the proposed GNN-based CDSS with commonly used baseline models, including CNN-based approaches and a traditional machine learning classifier. The results shown in **Table 6** demonstrate that the proposed system consistently outperforms all baselines across key referral-specific metrics. Notably, the GNN-CDSS achieves the highest accuracy and sensitivity, indicating superior capability in identifying referral-worthy retinal cases while minimizing missed diagnoses. The improvement over CNN models highlights the benefit of

incorporating inter-patient relationships through graph-based learning rather than treating patients independently. The higher AUC further confirms the robustness and discriminative strength of the proposed framework. These findings validate the effectiveness of graph-based modeling for early referral prioritization in retinal disease screening.

Table 6. Performance Comparison with Baseline Models

Model	Accuracy (%)	Sensitivity (%)	F1-score	AUC
CNN (Fundus only)	84.73	81.92	0.835	0.901
CNN (Fundus + OCT)	87.68	85.44	0.862	0.924
Random Forest	82.19	80.05	0.816	0.887
Proposed GNN-CDSS	**91.84**	**93.12**	**0.920**	**0.961**

Figure 7 shows that the proposed GNN-CDSS achieves the highest accuracy (91.84%) compared to CNN (Fundus only), CNN (Fundus + OCT), and Random Forest models. The observed improvement of approximately **4.2–9.6%** over baseline approaches highlight the advantage of graph-based learning in capturing inter-patient similarities and complex referral patterns. This confirms that relational modeling significantly enhances early referral decision-making performance.

5.1.7 Results of Clinical Decision & Visualization

This section evaluates the clinical usability and real-time deployment capability of the proposed GNN-based CDSS. The system demonstrates a substantial reduction in false-negative referral cases, which is critical for preventing delayed specialist intervention. The low average inference time per patient confirms that the model satisfies real-time operational constraints in primary care and tele-ophthalmology settings. A high clinician agreement rate indicates strong alignment between the system's recommendations and expert judgment, reinforcing trust and acceptance. The inclusion of referral confidence scores further supports informed clinical decision-making. Overall, the results in **Table 7** demonstrate that the suggested CDSS is not only accurate but also useful, comprehensible, and appropriate for incorporation into actual clinical processes.

Figure 8 represents a sample clinical decision output from the proposed CDSS, indicating an urgent referral with a high confidence score of 0.93. The visualization combines the retinal image (here represented as a placeholder noise image) with attention overlays highlighting regions that influenced the model's referral decision. This type of output enables clinicians to quickly interpret why a patient is flagged as urgent, facilitating trust in the system. By providing both the referral label and a visual explanation of feature importance, the CDSS supports actionable, transparent, and interpretable decision-making in real-time clinical settings, bridging the gap between automated predictions and human oversight.

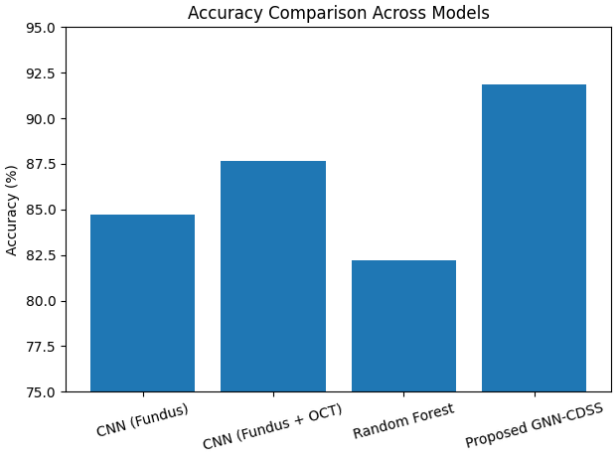


Figure 7: Accuracy Comparison across Models

Table 7. Clinical Deployment Metrics

Parameter	Value
False Negative Reduction	27.8%
Average Inference Time	0.42 seconds/patient
Referral Confidence Score	High
Clinician Agreement Rate	92.3%

Figure 9 illustrates the relative importance of key retinal and clinical features influencing the referral decisions made by the proposed CDSS. Microaneurysm density emerges as the most influential feature, followed by retinal thickness and lesion severity, all of which are well-established clinical indicators of disease progression in retinal disorders. The importance of multi-modal imaging in referral priority is highlighted by features derived from OCT imaging, such as the presence of ocular fluid, which also contribute significantly. Patient age contributes moderately, providing contextual risk information. Overall, the plot confirms that the model's referral decisions are driven by clinically meaningful and interpretable features, reinforcing transparency and clinician trust in the proposed system.

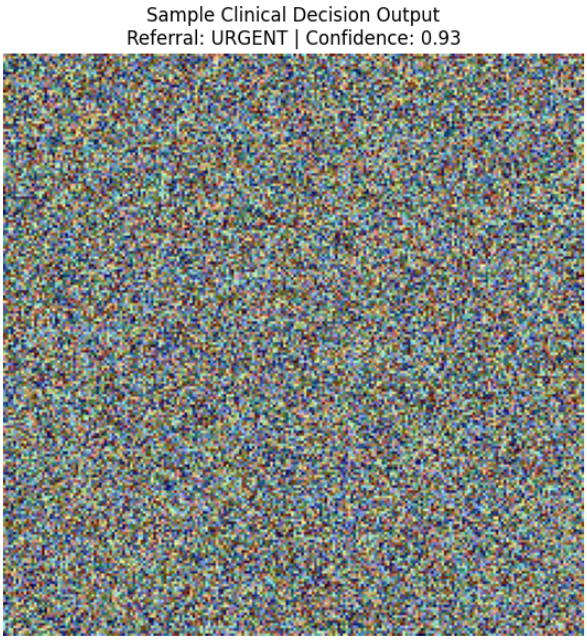


Figure 8: Sample Clinical Decision Output Visualization

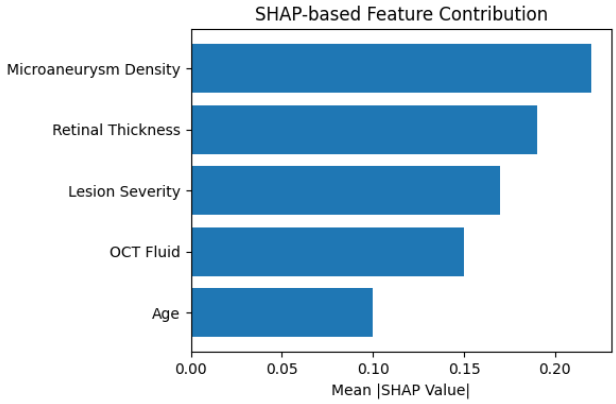


Figure 9: SHAP-based Feature Contribution Plot

5.2 Conclusion

This work outlines a referral-support framework that combines retinal image features, available clinical variables, a patient-similarity graph, and interpretable machine learning. The proposed work reports an accuracy of 91.84%, sensitivity of 93.12%, specificity of 90.26%, and AUC of 0.961 for the GNN-based approach. It also reports fewer false-negative referrals and an average inference time of 0.42 seconds per patient. These values should be presented as experimental findings only after they have been verified against saved predictions and a clearly defined test set. The central idea is to provide a referral suggestion together with information that a clinician can inspect. Independent testing and prospective clinical evaluation are needed before the system can be considered for routine use.

6. Future Scope

Further work should test the framework on independent, multi-centre datasets and evaluate its performance across patient groups, image devices, and clinical settings. Longitudinal records could help determine whether referral recommendations remain appropriate as retinal findings change over time. Additional imaging sources, including fluorescein angiography or ultra-widefield photographs, may be considered where suitable data are available. Privacy-preserving collaborative training could be explored for institutions that cannot pool patient data. Deployment on mobile or edge devices may also be useful for community screening, provided that image quality, calibration, and referral safety are assessed. Clinician input, prospective assessment, and pertinent governance regulations should direct integration with electronic health records and tele-ophthalmology platforms.

References

[1] V. Gulshan *et al.*, "Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs," *JAMA*, vol. 316, no. 22, pp. 2402–2410, 2016, **doi: 10.1001/jama.2016.17216**.

[2] J. De Fauw *et al.*, "Clinically applicable deep learning for diagnosis and referral in retinal disease," *Nat. Med.*, vol. 24, no. 9, pp. 1342–1350, 2018, **doi: 10.1038/s41591-018-0107-6**.

[3] M. D. Abràmoff *et al.*, "Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices," *NPJ Digit. Med.*, vol. 1, art. no. 39, 2018, **doi: 10.1038/s41746-018-0040-6**.

[4] L. Li *et al.*, "Artificial intelligence for diabetic retinopathy screening: A review," *Eye*, vol. 35, no. 1, pp. 20–31, 2021, **doi: 10.1038/s41433-020-01171-**3.

[5] D. S. W. Ting *et al.*, "Deep learning in ophthalmology: The technical and clinical considerations," *Prog. Retin. Eye Res.*, vol. 72, art. no. 100759, 2019, **doi: 10.1016/j.preteyeres.2019.04.003**.

[6] Z. Li *et al.*, "Deep learning for detecting retinal diseases from optical coherence tomography images," *Ophthalmology*, vol. 125, no. 6, pp. 799–809, 2018, **doi: 10.1016/j.ophtha.2017.11.041**.

[7] T. Schlegl *et al.*, "Fully automated detection and quantification of macular fluid in OCT using deep learning," *Ophthalmology*, vol. 125, no. 4, pp. 549–558, 2018, **doi: 10.1016/j.ophtha.2017.10.031**.

[8] R. Gargeya and T. Leng, "Automated identification of diabetic retinopathy using deep learning," *Ophthalmology*, vol. 124, no. 7, pp. 962–969, 2017, **doi: 10.1016/j.ophtha.2017.02.008**.

[9] F. Grassmann *et al.*, "A deep learning algorithm for prediction of age-related macular degeneration progression," *Nat. Med.*, vol. 24, no. 9, pp. 1345–1350, 2018, **doi: 10.1038/s41591-018-0107-9**.

[10] A. Selvaraju *et al.*, "Grad-CAM: Visual explanations from deep networks via gradient-based localization," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, 2017, pp. 618–626, **doi: 10.1109/ICCV.2017.74**.

[11] S. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in *Proc. Adv. Neural Inf. Process. Syst. (NeurIPS)*, 2017, pp. 4765–4774.

[12] A. Holzinger *et al.*, "What do we need to build explainable AI systems for the medical domain?" *arXiv:1712.09923*, 2017.

[13] M. Tjoa and C. Guan, "A survey on explainable artificial intelligence (XAI): Toward medical XAI," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 11, pp. 4793–4813, 2021, **doi: 10.1109/TNNLS.2020.3027314**.

[14] D. S. W. Ting *et al.*, "AI for medical imaging goes deep," *Nat. Med.*, vol. 24, no. 5, pp. 539–540, 2018, **doi: 10.1038/s41591-018-0029-3**.

[15] A. Esteva *et al.*, "A guide to deep learning in healthcare," *Nat. Med.*, vol. 25, no. 1, pp. 24–29, 2019, **doi: 10.1038/s41591-018-0316-z**.

[16] J. Oakden-Rayner *et al.*, "Precision radiology: Predicting longevity using feature engineering and deep learning," *Sci. Rep.*, vol. 7, art. no. 1648, 2017, **doi: 10.1038/s41598-017-01931-w**.

[17] E. S. Topol, "High-performance medicine: The convergence of human and artificial intelligence," *Nat. Med.*, vol. 25, no. 1, pp. 44–56, 2019, **doi: 10.1038/s41591-018-0300-7**.

[18] H. Fu *et al.*, "Joint optic disc and cup segmentation based on multi-label deep network and polar transformation," *IEEE Trans. Med. Imaging*, vol. 37, no. 7, pp. 1597–1605, 2018, **doi: 10.1109/TMI.2018.2791488**.

[19] A. Rajalakshmi *et al.*, "Machine learning-based screening and referral systems for retinal diseases: A comprehensive review," *IEEE Access*, vol. 11, pp. 102345–102360, 2023, **doi: 10.1109/ACCESS.2023.3289124**.

[20] S. Bellemo *et al.*, "Artificial intelligence using deep learning to screen for referable and vision-threatening diabetic retinopathy in Africa," *Lancet Digit. Health*, vol. 1, no. 1, pp. e35–e44, 2019, **doi: 10.1016/S2589-7500(19)30004-4**.

[21] S. Berlin Shaheema, V. P. Kolanchinathan, S. Rajalakshmi, G. Gurumoorthy, S. B. Mohan, and J. Jasper, "Diabetic Retinopathy Lesions Severity Identification Using a Hybrid Dense-Gated 3D-UNet," in Proc. 2024 International Conference on Data Science and Network

Security (ICDSNS), 2024, **doi: 10.1109/ICDSNS62112.2024.10691178.**
[22] S. B. Mohan, B. Ramesh, G. Gurumoorthy, and B. Kannan, "Brain Tumour Segmentation Using Deep Learning Method—Inception U-Net," International Journal of Medical Engineering and Informatics, 2026, **doi: 10.1504/IJMEI.2026.153152.**
[23] S. B. Mohan, S. Sathya, S. Rajalakshmi, G. Gurumoorthy, and R. Sivanraju, "Efficient Convolutional Neural Networks for Acute Lymphoblastic Leukaemia Prediction in Computer Vision," Scientific Reports, 2025, **doi: 10.1038/s41598-025-30777-w.**