DESCRIPTION OF THE CONTEST PRODUCTS

FOR TABLE D2, D3

PRODUCT INTRODUCTION

1. Idea of the product

Currently, more than **2.2 billion people globally** suffer from diseases that cause vision impairment, of which **45% can be prevented** if detected and treated early. **Fundus photography** is a simple, low-cost, non-invasive and very effective method for early detection of retinal diseases. However, **manual diagnosis from fundus imaging presents many challenges**, especially in areas where there is a shortage of ophthalmologists. In this context, **artificial intelligence (AI),** especially **convolutional neural networks (CNNs),** have been successfully applied to support pathological identification from retinal images. However, most current models **only focus on individual diseases**, reducing their effectiveness when applied to the actual clinical environment. **This study introduces VisionCare** – a comprehensive platform that combines **an eye disease lookup web and an in-depth AI system**, capable of **simultaneously diagnosing multiple retinal diseases** from fundus images. VisionCare uses **advanced CNN models such as EfficientNet, DenseNet,** Explainable AI **integration**, advanced preprocessing techniques, and remote consultation capabilities, helping to **improve the accuracy and accessibility of medical services**, especially in under-resourced areas. **Testing on publicly available retinal datasets** shows that VisionCare achieves high accuracy, fast response speed, and better generalization capabilities than current methods, showing **strong application potential in clinical** and public eye healthcare.

2. Overview

2.1. Questioning

According to the World Health Organization (WHO), there are more than **2.2 billion people in the world have vision problems**, where **45% preventable if detected early**. **Retina** The light-sensitive tissue at the back of the eye – plays an essential role in converting light into visual signals, so examining the retina is important in protecting vision and overall health.

Conditions such as **diabetic retinopathy, macular degeneration (AMD), glaucoma, cataracts, retinal hypertension, and refractive errors** can lead to blindness if not detected in time. However, **Initial symptoms are usually faint**, making it difficult to detect with the naked eye. **Fundus imaging** and **OCT tomography scan** are the two main methods for early detection of eye diseases, in which FUNDUS is more popular due to its low cost and ability to screen extensively. However, the current diagnosis is still **reliance on specialists, time-consuming, and lack of high-quality data**, especially with rare diseases. Therefore, **AI – especially convolutional neural networks (CNNs)** – is being applied to **Automatic accurate and fast diagnosis of eye diseases**, supporting

the reduction of the load for doctors. However, **Most of the current models only focus on individual diseases**, while clinical reality requires a diagnosis **multiple concurrent pathologies on the same retinal image**.

2.2. Research the document

Several previous studies have built automated eye disease diagnosis support systems based on artificial intelligence with different approaches. Koh et al. proposed an automated retinopathy screening system using SURF and PHOG features combined with canonical correlation analysis, achieving an accuracy of up to 96.21%. However, this approach relies heavily on manual characterization, limiting the ability to generalize.

In recent times, the CNN network has been widely used in the classification of eye diseases thanks to its ability to automatically learn deep features from retinal image data. Islam and his colleagues developed the CNN model from scratch on the ODIR dataset, achieving an F-score of 85 percent, but it was still limited when it did not take into account the simultaneous classification of multiple diseases from eye images. Luo et al. used the EfficientNet model with the FC-loss function to classify AMD, glaucoma, and cataracts, but the best effect was only in the form of binary classification.

Most recently, Yang and Yi proposed the DSRA-CNN model that integrates DS, DSR and SE blocks, which yielded positive results with an accuracy of 87.90%, but was not optimal for binocular image analysis. Deng and Ding improved by combining characteristics from EfficientNet-B2 and InceptionResNetV2 along with age and gender information, significantly improving the results, achieving an accuracy of 92-93%.

From the above studies, it can be seen that there are research gaps such as: (1) lack of an effective system for the simultaneous diagnosis of multiple eye diseases from retinal imaging; (2) it is necessary to improve generalization and mitigate the problem of data imbalances; and (3) lack of research directly applied to clinical practice with comprehensive remote support.

2.3. Research objectives

In this study, we introduce a comprehensive eye disease diagnostic system called VisionCare, which uses artificial intelligence technology and a professional website system to solve the above problems. The VisionCare system provides automatic, fast, and accurate diagnosis of various eye diseases simultaneously through retinal imaging, and integrates teleconsultation and consultation features, expanding access to intensive ophthalmic medical services, especially for remote areas as well as in local health clinics and regional hospitals. In addition, VisionCare was created with the aim of reducing the load on the number of people visiting hospitals specializing in ophthalmology and helping people conveniently in examining, preventing and monitoring diseases from the local grassroots level to the higher level.

2.4. Research contributions

2.4.1. Novelty

VisionCare shows many new features when it combines two powerful AI streams simultaneously – Classification and Segmentation – with a model interpretation

mechanism (Explainable AI) in a single web platform. Different from solutions that only focus on one pathology or one method, VisionCare simultaneously applies EfficientNet, Swin Transformer, ResNet-50, Densenet121 and Unet to handle multiple diagnostic scenarios, and integrates Grad-CAM and LIME in parallel to "open the black box" of AI, creating an unprecedented comprehensive approach in the field of retinal diagnostics.

2.4.2. Practicality

The application is developed on a web application platform, allowing immediate deployment with just a few command lines without the need for complex infrastructure. The interface is intuitive, optimized for both doctors and patients, meeting the remote and on-site examination process: from commune health stations, private clinics to provincial general hospitals. The system automatically stores the results into SQLite and exports medical standard PDF reports, saving >70% of diagnosis time, reducing administrative workload, and increasing medical examination capacity by 3–5 times compared to traditional methods.

2.4.3 Community

VisionCare aims to "vision health for everyone, everywhere." By simplifying technical barriers, any grassroots medical facility can use it only through a web browser; people in remote and remote areas can check and keep records at home. The EyeVision Chatbot feature provides 24/7 preliminary medical advice, promotes public health education, raises early awareness of risk signs, thereby contributing to the prevention of blindness for millions of people.

2.4.4. Scientificity

The VisionCare platform is a testament to the convergence of advanced AI research: using transfer learning on ImageNet and ODIR-5K, applying CosineAnnealing Warm Restarts during fine-tuning, and multitasking loss formula combining FocalLoss and continuous Kappa loss to help optimize classification accuracy and ranking consistency in parallel. Internal test results showed a diagnostic sensitivity of ≥92% and a Kappa coefficient of ≥0.87, which was superior to previous studies. The model is designed to be open to continue researching, correcting and expanding new pathologies, affirming the scientific stature and potential for sustainable development.

PRODUCT DESCRIPTION

1. Main functions of the product

1.1. Clearly describe the features of the product

VisionCare is a dedicated website platform, built to support users to look up and access medical information related to eye diseases. The platform is designed to optimize the user experience with three main features: The first and most important is the innovative novelty of the AI system for diagnosing various eye diseases and Explainable AI and fractionation of vascular images of the macular point of the fundus of the eye, the second is **looking up eye diseases according to the database** and finally automatic

consultation through EyeVision Chatbot is focused in detail for eye diseases as well as ophthalmic information.

1.1.1. System using AI Retinal Diagnostics for specialists and telemedicine patients

We develop a comprehensive AI application platform for diagnosing and supporting the treatment of retinal diseases, including two main components: **Classification** and **Segmentation**, along with auxiliary functions such as **Model Interpretation (Grad-CAM, LIME), Medical Consultation Chatbot**, etc. **Archive – manage records** and **Export PDF reports**. It's all integrated in an interactive web interface, running on **Streamlit** with **the Python** language.

Figure 1: The home page of the AI interface for diagnosing multiple retinal diseases

A. Multi-Disease Classification

The multi-disease classification feature allows the system to automatically identify up to 11 groups of retinal lesions at the same time by uploading images of the fundus taken with specialized equipment. Users don't need in-depth knowledge: after selecting a shape and activating it, the model returns a label and probability with an early warning if an anomaly is detected.

List of 11 retinal diseases

Cataracts, Central Subretinal Vitreous Fluid, Diabetic Retinopathy, Dendritic Edema, Glaucoma, Normal, Macular Scar, Myopia, Flirta, Retinal Detachment, Retinitis Pigmentosa.

Goals & Technology

- **Objective**: Identify 11 retinal lesions simultaneously on a fundus image, providing label and probability results with immediate emergency alerts. Users can select each model individually or run multiple models simultaneously, and then make decisions to increase objectivity.

Model	Accuracy (%)	Data Sources

EfficientNetV2_rw_s	**95.3**	ODIR-5K, aggregated dataset
SwinV2_base_window8_256	91.8	ODIR-5K, aggregated dataset
EfficientNet_B4	92.7	ODIR-5K, aggregated dataset
DenseNet121	**93.9**	ODIR-5K, aggregated dataset

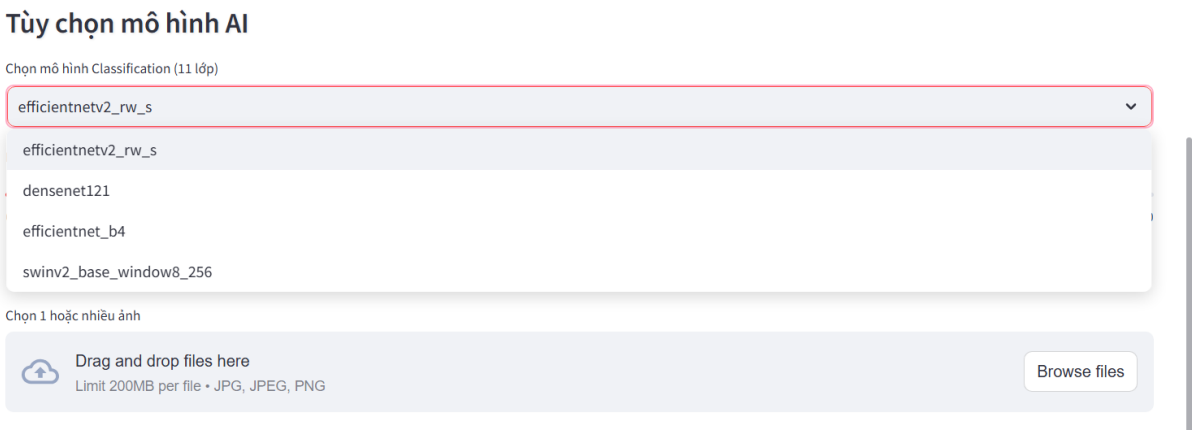
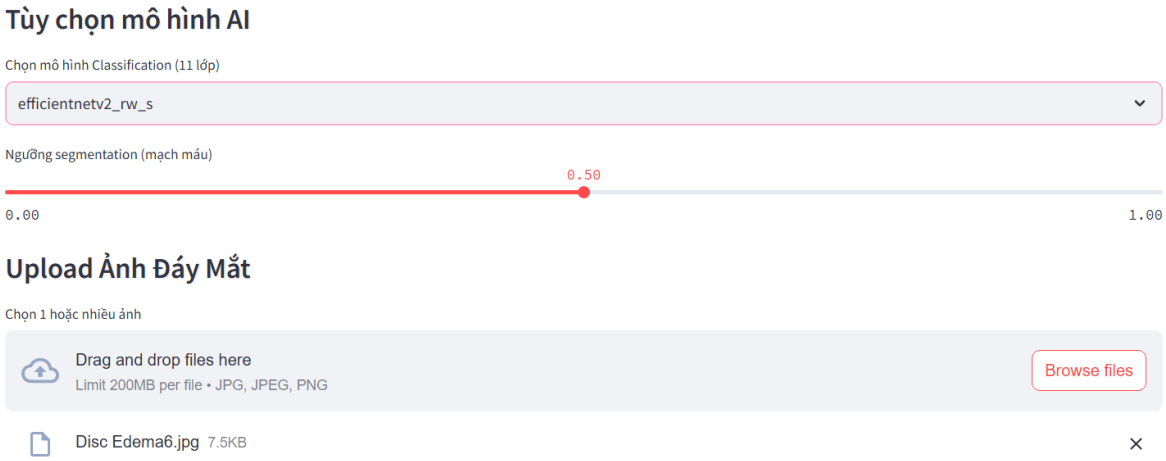


Figure 2: Optional list of deep learning models for classification and diagnosis



**** Figure 3: Download a photo or take an image of the fundus of the eye and select the model, select the segmentation threshold***

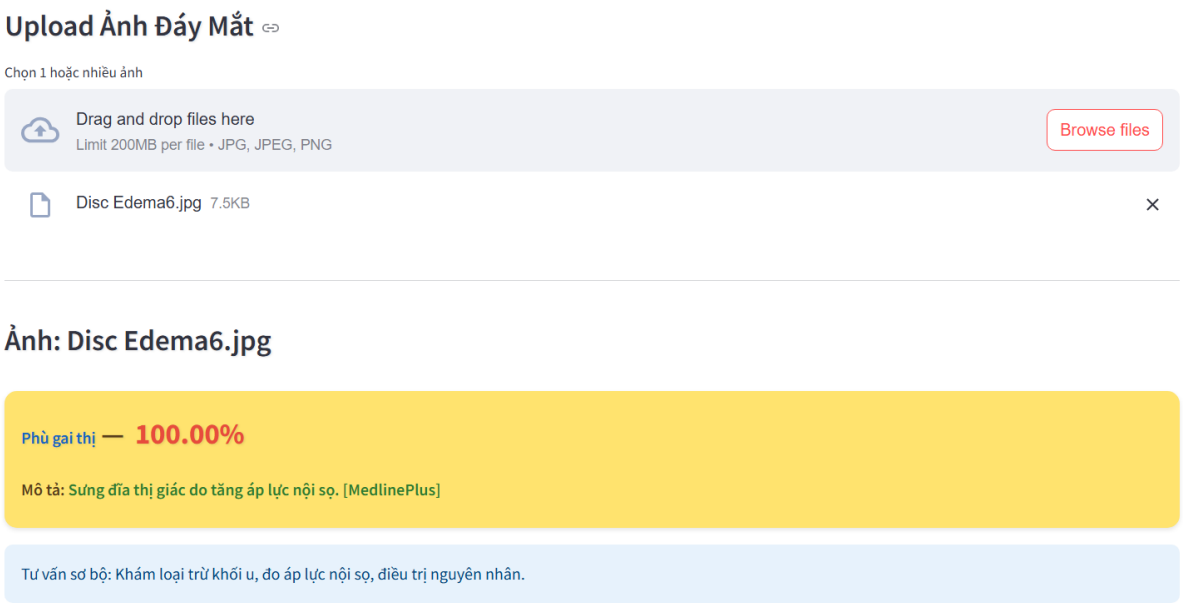
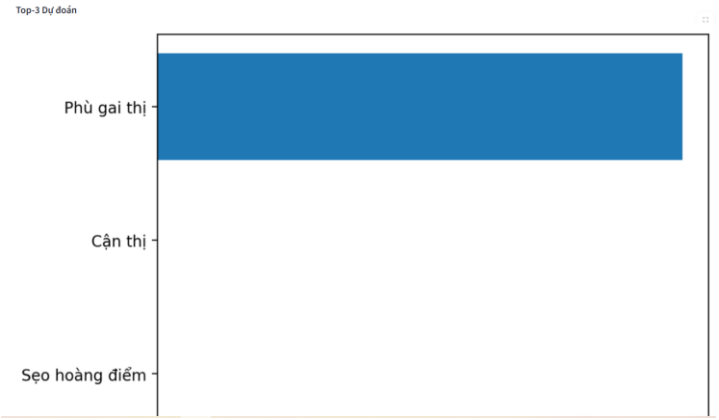


Figure 4: Interface of retinal disease prediction results and description and consultation information.



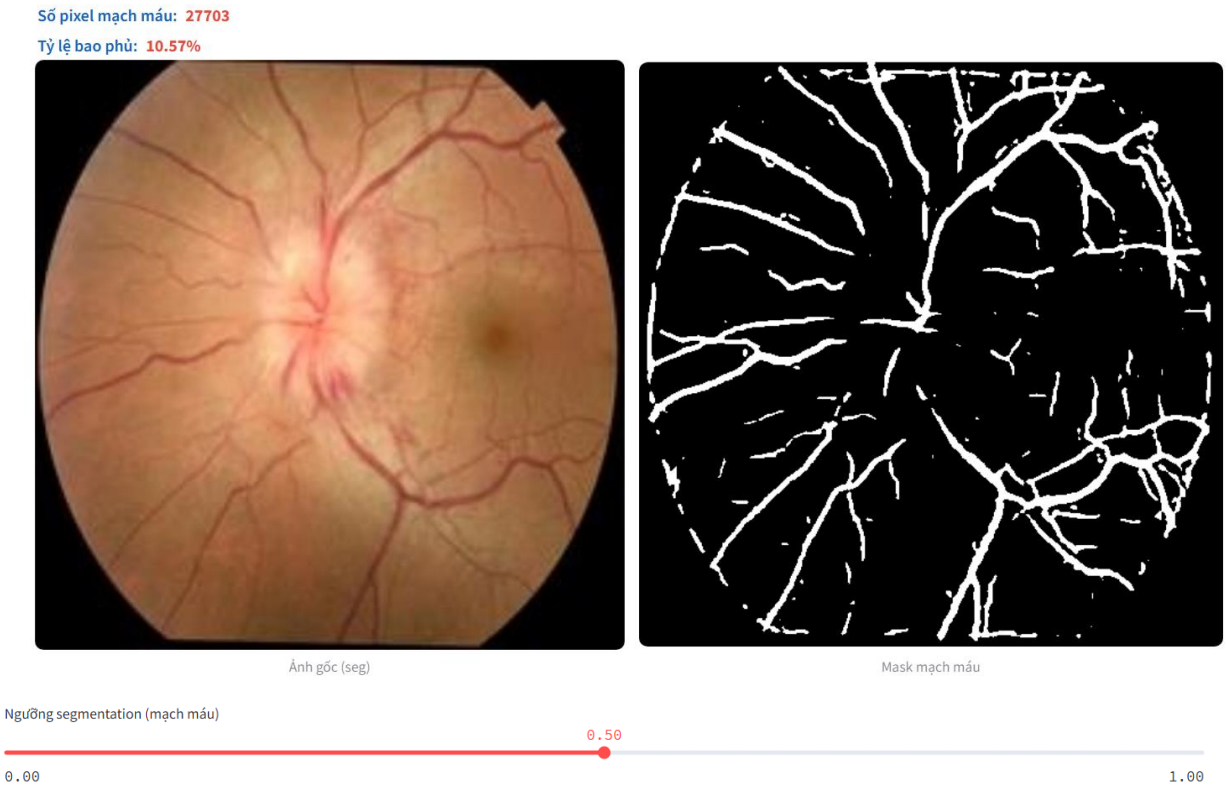
**** Figure 5: Graph of the Top 3 most confident pattern prediction results***

Benefits

- For high accuracy (≥ 92%), while minimizing human errors.
- Early warning of many diseases at the same time, supporting doctors to quickly screen on a large scale.
- Flexibility in selecting models according to context: speed priority (DenseNet121), absolute accuracy (EfficientNetV2), or balance of both (EfficientNetB4, SwinV2).

B. Vessel Segmentation

With the ability to precisely separate the vascular network and the damaged macula as a binary mask, the segmentation allows visualization of the density and extent of the pathology. Users only need to adjust the display threshold, the system will automatically calculate the number of vascular pixels and the coverage ratio, and display a comparison mask next to the original image.



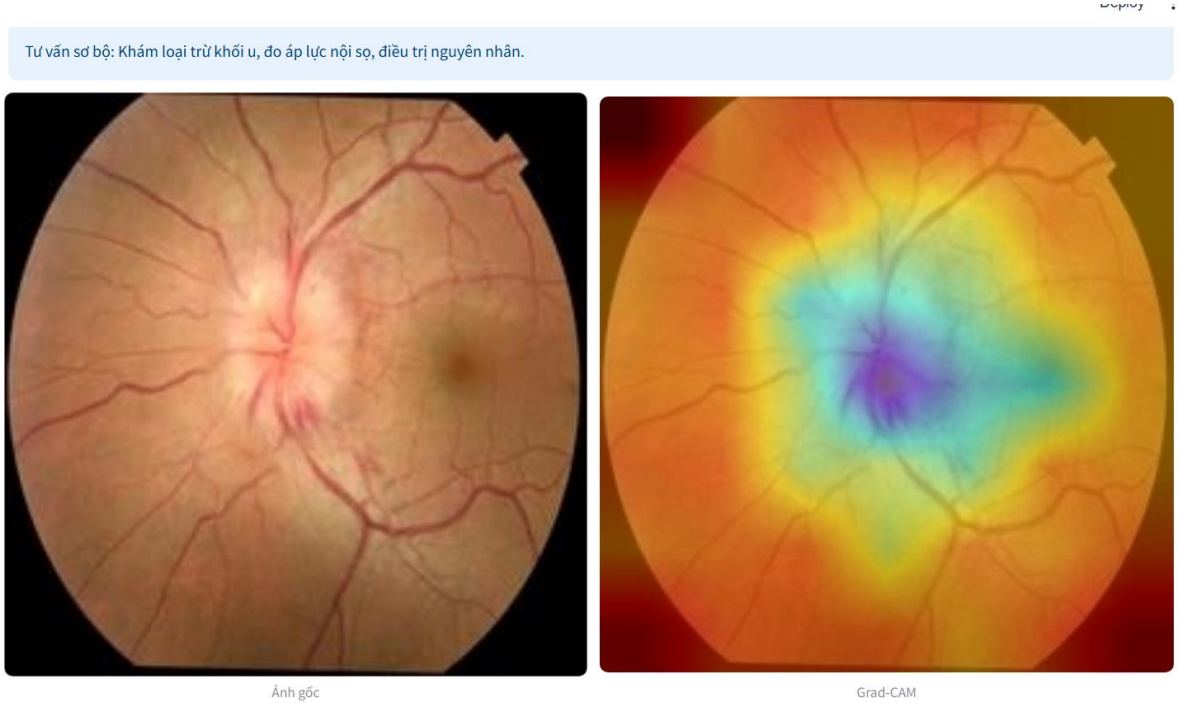
**** Figure 6: Vascular segmentation model of retinal imaging and gives vascular pixel information and coverage ratio based on the initial selected segmentation level***

Goals & Technology

- **Goal**: Precisely separate the vascular network and the damaged macula in the form of a binary mask.
- **Architecture**: U-Net with ResNet50 encoder pretrained on ImageNet.
- **Benefits**
    - Provides quantitative indicators to assess disease progression and treatment effectiveness.
    - When combined with the results of the classification, the doctor can simultaneously review the "what" (type of disease) and "where" (the location of the lesion).
    - Support early intervention planning if the vascular area is extensively damaged.

C. Explainable AI

To enhance reliability, VisionCare integrates Grad-CAM, Grad-CAM++, and LIME, which help visualize the area of the retinal image that the model "pays attention to". Grad-CAM heatmaps pathological structures and the more advanced Grad-CAM++, LIME displays the most important superpixels that influence decision-making.



**** Figure 7: The model explains by giving Grad-CAM and LIME images to show where the model is focused and make predictions***

Goals & Technology

- **Goal**: To be explanatory and transparent, helping doctors and patients understand "where the AI has been focused" when making predictions.
- **Tools**:
  - **Grad-CAM, Grad-CAM++**: Hook to the last convolution layer, collect gradients and activation maps, create weighted heatmaps.
  - **LIME:** Superpixel separation, random imaging experiment, estimation of the importance of each area.
- **Operation Process**

  1. After dividing the lot ạ i, h ệ th ố NG CH ạ y Grad-CAM, Grad-CAM++ ⟶ overlay heatmap lên she Nh G ố c.

  2. If the user enables LIME, the system initiates an explainer, generating thousands of samples to evaluate the superpixel effect → display the highlight area.
- **Benefits**
  - Increase the reliability of AI results, making it easier for doctors to verify critical lesions.
  - Facilitate doctor-patient consensus when explaining diagnoses.

D. Severity Scoring

Figure 8: of the diabetic retinopathy severity classification interface and where the GRAD-CAM model is focused to understand what the model is predicting

Not only stopping at detecting yes or not, the system also provides a scale of levels from mild – moderate – severe – proliferative, helping to prioritize examination and intervention. This result is displayed immediately after the classification, with a color alert to round up the doctor's workflow.

Goals & Technology

- **Objective**: Rank the severity of retinopathy according to 5 levels: No – Mild – Moderate – Severe – Proliferative.
- **Multi-task architecture**: On top of backbone classification, add head regressor to output continuous values, combined training:
    - **Focal Loss** cho classification
    - **Continuous Kappa Loss** cho severity regressing
- **Operation Process**
    - The classification and severity model run in parallel.
    - The severity results are displayed as colored bars (blue→yellow, → red) and the confidence percentage.
    - Color warning prioritizes handling severe cases first.
- **Benefits**
    - Support patient flow at the hospital: priority is given to examination and early intervention.
    - It allows to monitor the progress of the disease over time, compare the severity score through each follow-up visit.

E. Record & Reporting & Data Collection

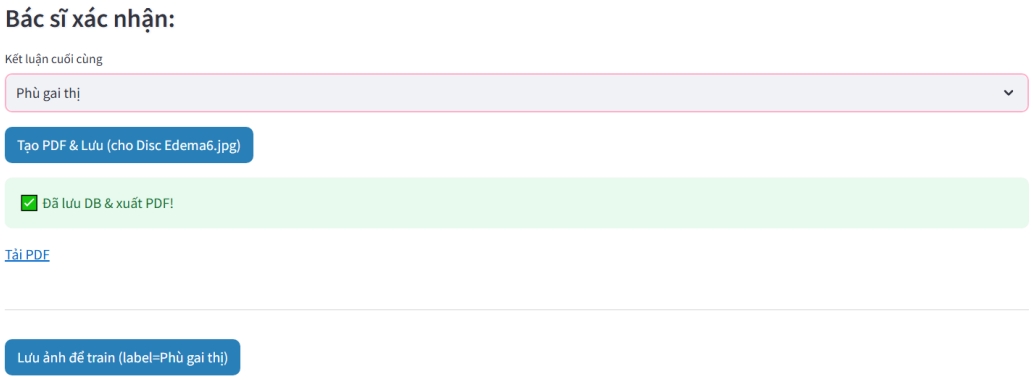


Figure 9: The interface of the feature of creating doctors' medical records and saving images according to medical security to train models and improve data

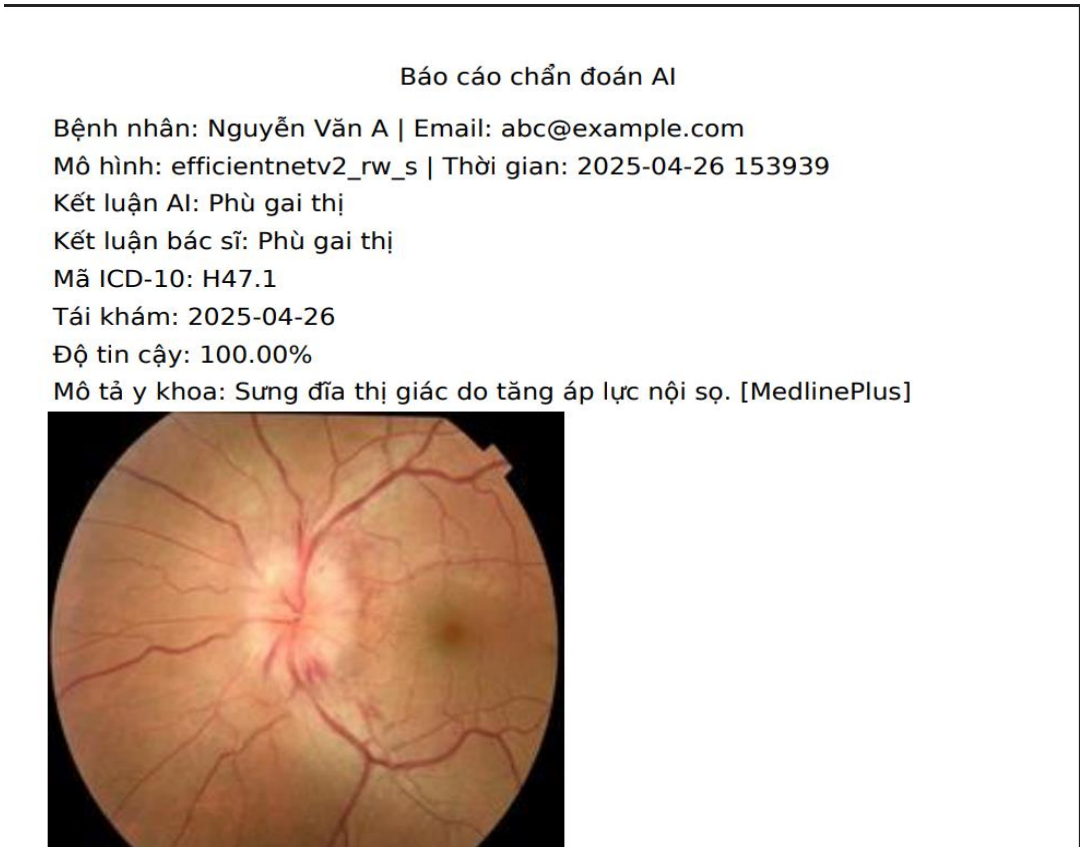


Figure 10: A PDF file of the patient's medical record generated by AI.

Goals & Technology

- **Objective**: Store all metadata (patient, AI results, doctor's conclusions, ICD-10, original photos, Grad-CAM, mask segmentation, severity score) and export PDF reports, create a data collection pipeline for the next training.
- **Operation Process**
  - The doctor corrects the final conclusion, presses "Create PDF & Save" → record to SQLite + export the PDF file.

- The PDF download link is now available on the web, and patients/doctors can save, print or email.
- "Save photos for training" option to collect standard labels in a folder, serving for the next round of model training.

- **Benefits**
  - Ensure end-to-end workflow: upload → diagnostics → explain → manage records → reports.
  - Collect standard label data with moderation to improve the model, creating a "Human-in-the-loop" feedback loop.
  - Fast storage and retrieval, ensuring compliance with Security and Medical Data rules.

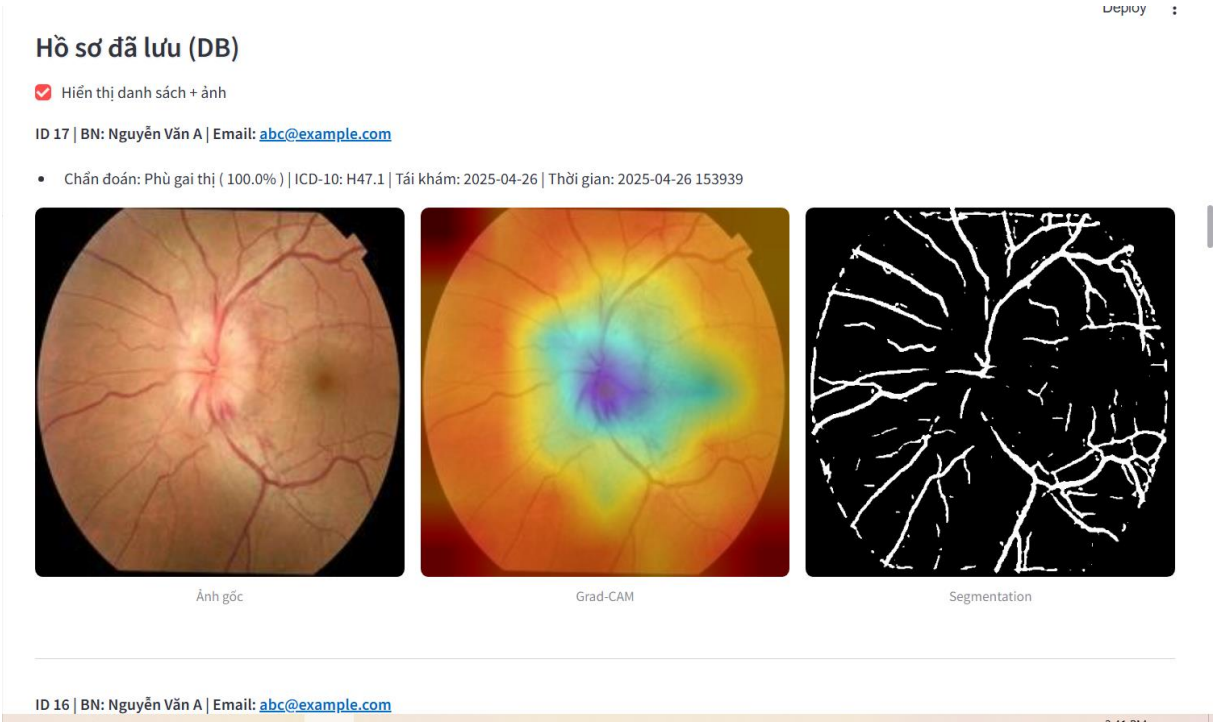


Figure 11: Patient profile examined and showing 3 important photos: original photo, grad-cam photo (AI explained) and vascular segment photo of Phu Gai Thi

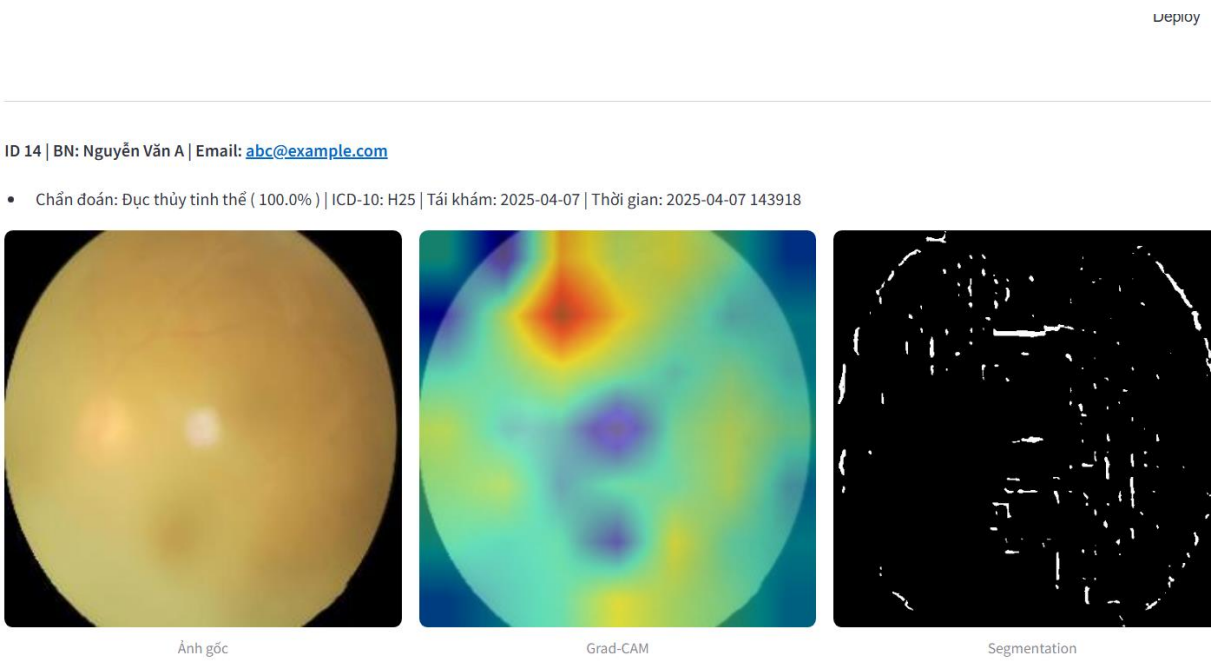


**Figure 12: Patient profile with Cataracts*

F. System Interaction and Usage Instructions

To make the most of the features of our retinal diagnostic AI platform, we designed a clear flow of interaction for two main groups: **Doctors** and Patients. Below is a detailed step-by-step guide with a description of each feature.

+ Login and role selection

1. **Open the browser** and go to the app address.
2. **The Sidebar** allows you to select roles:
   - **Doctor**: Enter a password (e.g. 123) to enable the extension interface.
   - **Patients**: View only AI results without saving DBs or exporting PDFs.

+ Doctor Interface

1. **Patient Information**
   - Fill in **the patient's name**, **email**, select **Follow-up Visitation Date**, and write **Doctor's Notes** (write in PDF).
2. **Choose an AI model**
   - **Classification**: Chọn giữa efficientnetv2_rw_s, densenet121, efficientnet_b4, swinv2_base_window8_256.
   - **Segmentation**: Only Unet+ResNet50 is the default; adjust **the threshold for** the vascular mask with a slider. **Upload a photo of the fundus of the eye**
   - Press **"Select 1 or more photos"**; .jpg/.png format is supported.
   - Automated Application:
     1. Show the original photo.

2. Ch ạ Y Worried ạ i → print out **Result card** with AI label, % rate (below class .result-box).
3. Calculate **Grad-CAM, Grad-CAM++,** and **LIME** (if checkbox is enabled) and display an overlay.
4. Run the segment → show the circuit mask and pixel & coverage statistics.

3. Censorship and storage

- Select **"Final Conclusion"** from the dropdown (calibration doctor).
- Press **"Create PDF & Save"**:
  1. The system writes to DB (SQLite), exports PDF files with information and images.
  2. Show PDF download link instantly.

4. Save photos for training

- You can click **"Save image for training"** to export the labeled image to the train_data/<label>/ folder.

5. View saved records

- Enable the "Show list + photos" checkbox to browse the entire record: original photo, Grad-CAM, mask segmentation, and metadata (name, email, ICD-10, trust rate, etc.).

+ Patient Interface

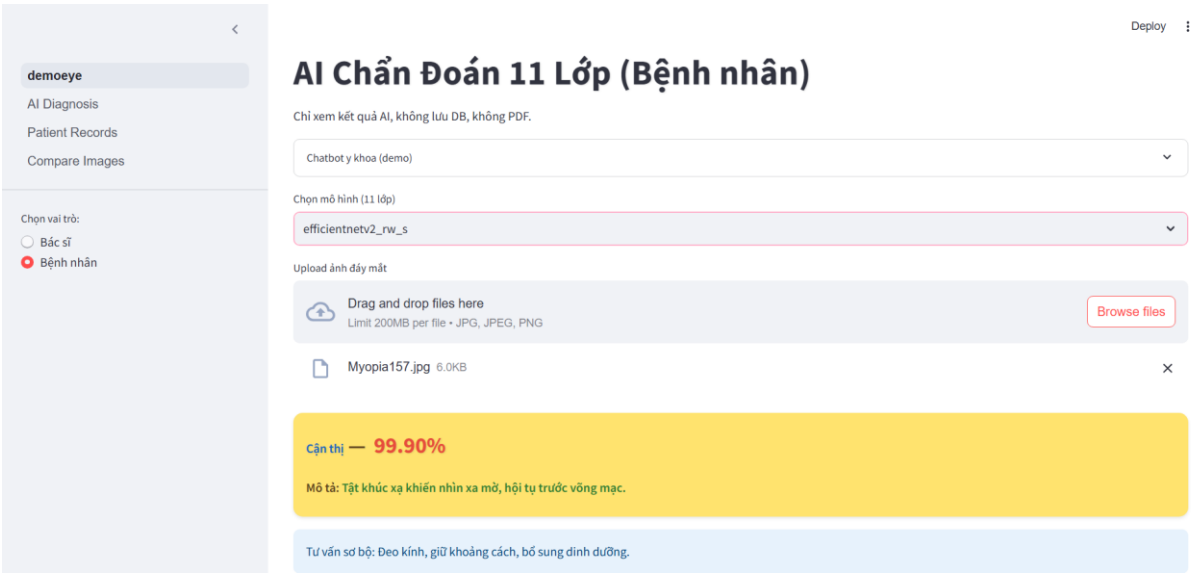


Figure 13: Patient interface used to self-diagnose retinopathy

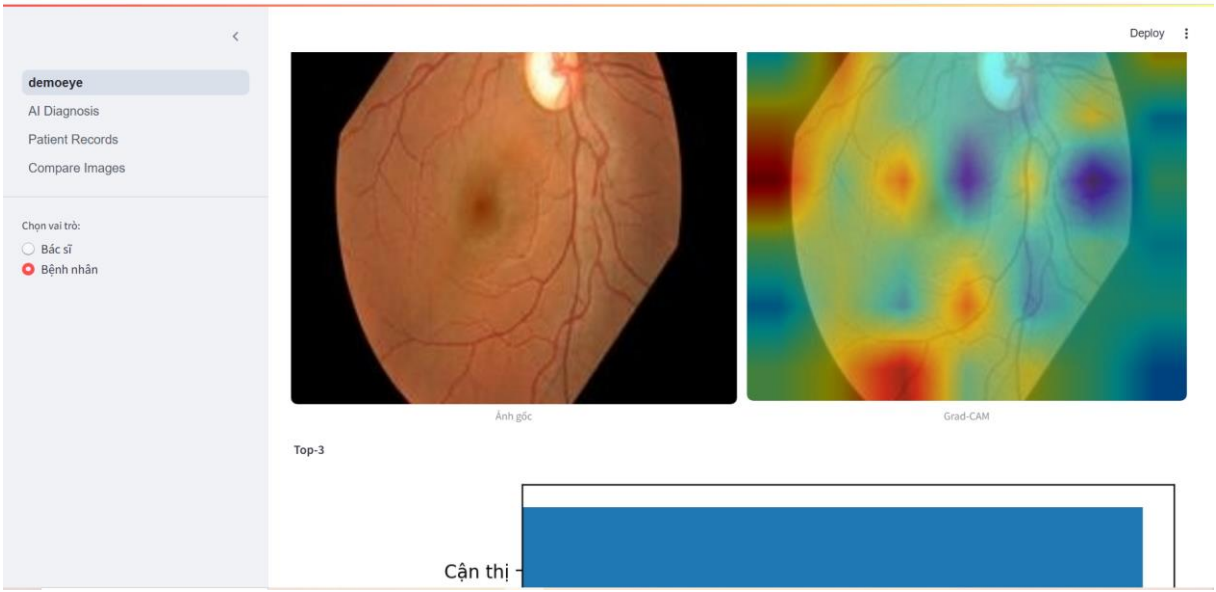
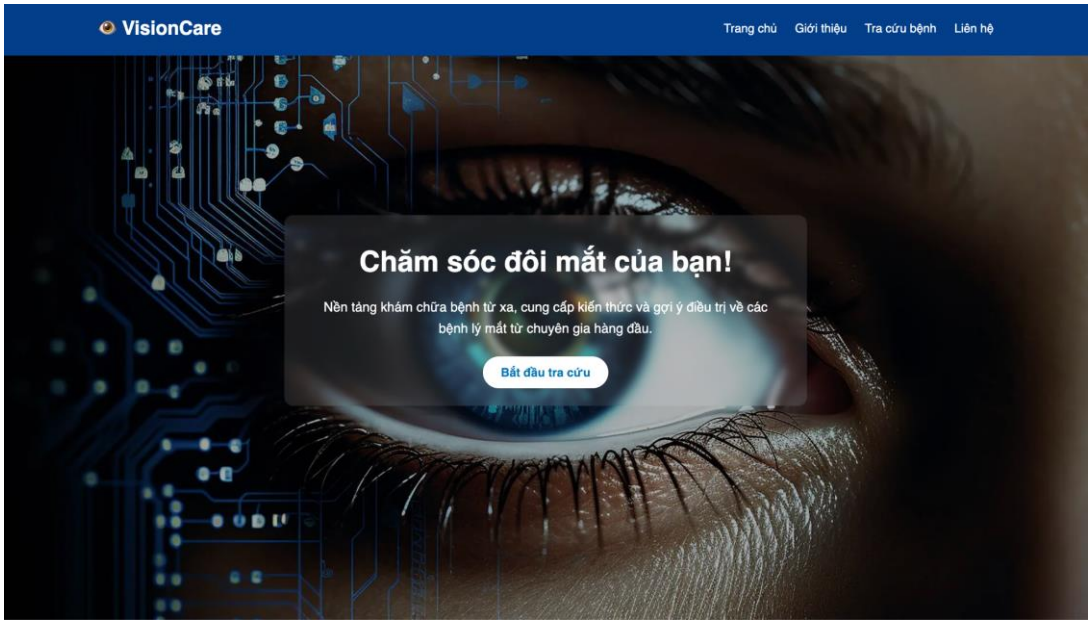


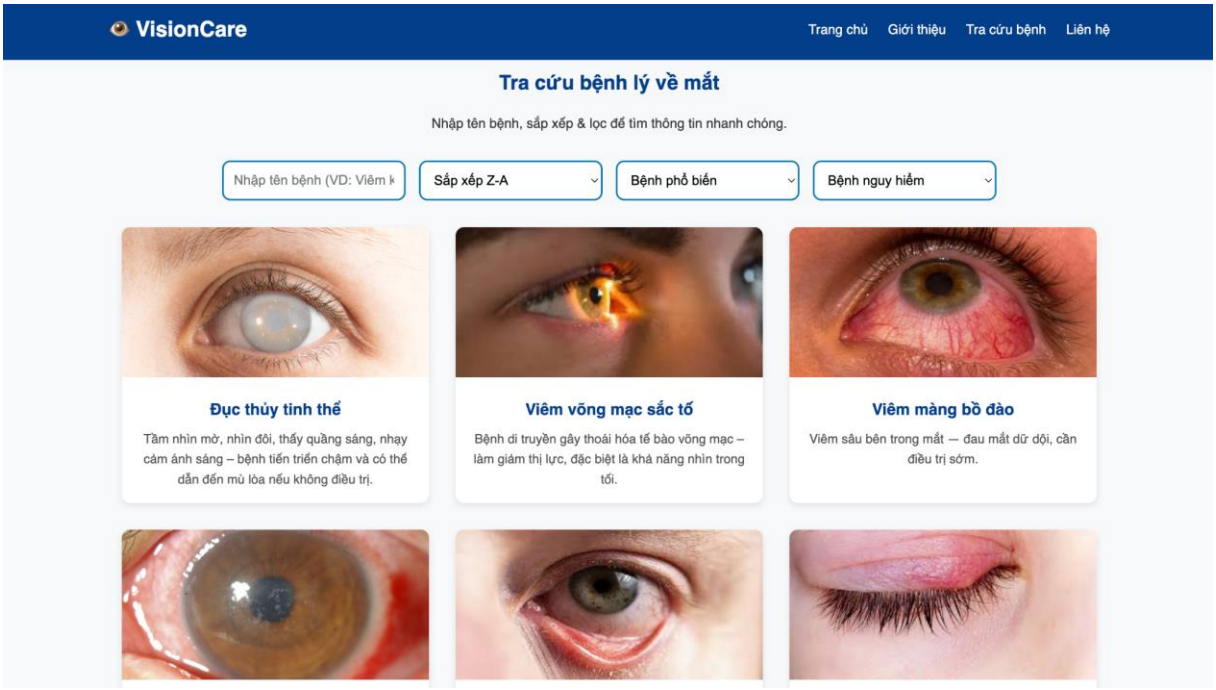
**Figure 14: Grad-cam image and Top-3 diagram similar to the doctor's interface*

1. **Select the model** (11 layers) and upload the photo:
   - Select the model on the sidebar.
   - Upload one or more photos.
2. **Get Results**
   - H ệ th ố NG CH ạ y classification → hi ể n th ị **AI result cards** with labels and probabilities. Displays **Grad-CAM overlay** so that patients can easily visualize important retinal areas.
   - Displays the predicted Top-3 chart (bar chart).
   - Optional view of the LIME demo to see "which zone is the most focused model".,mmm;

1.1.2. Eye disease information lookup library feature

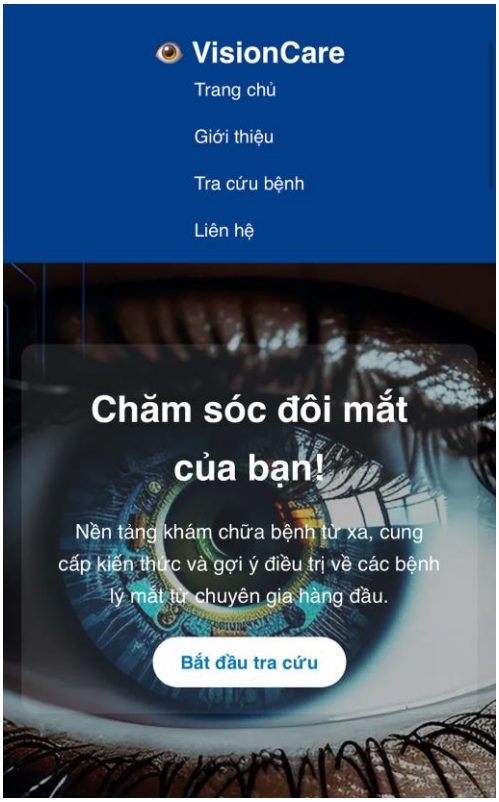


** Figure 15: System Home*



** Figure 16: Eye pathology lookup with all the information you need to know*

When accessing VisionCare, users are greeted by a vivid homepage interface with dynamic typing effects in the header, combined with a prominent hero-section with beautiful background images and a light matte overlay. Right on the main interface, users are quickly oriented through the "Start Lookup" call-to-action button.



** Figure 17: System Home Page on your phone*

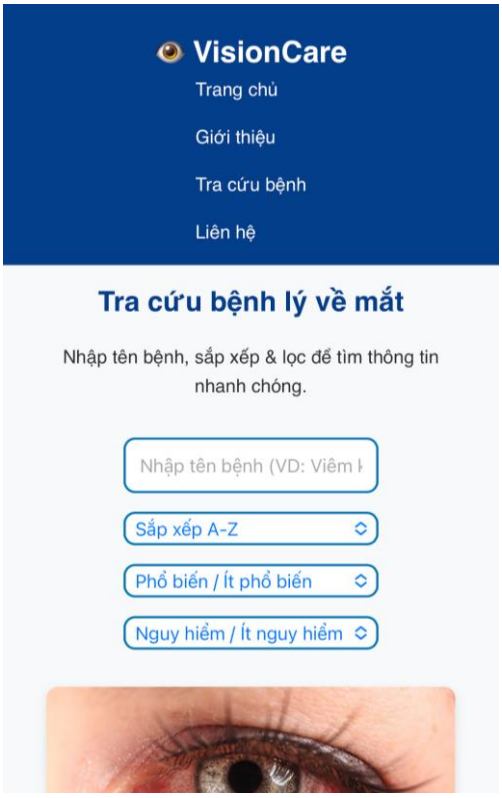
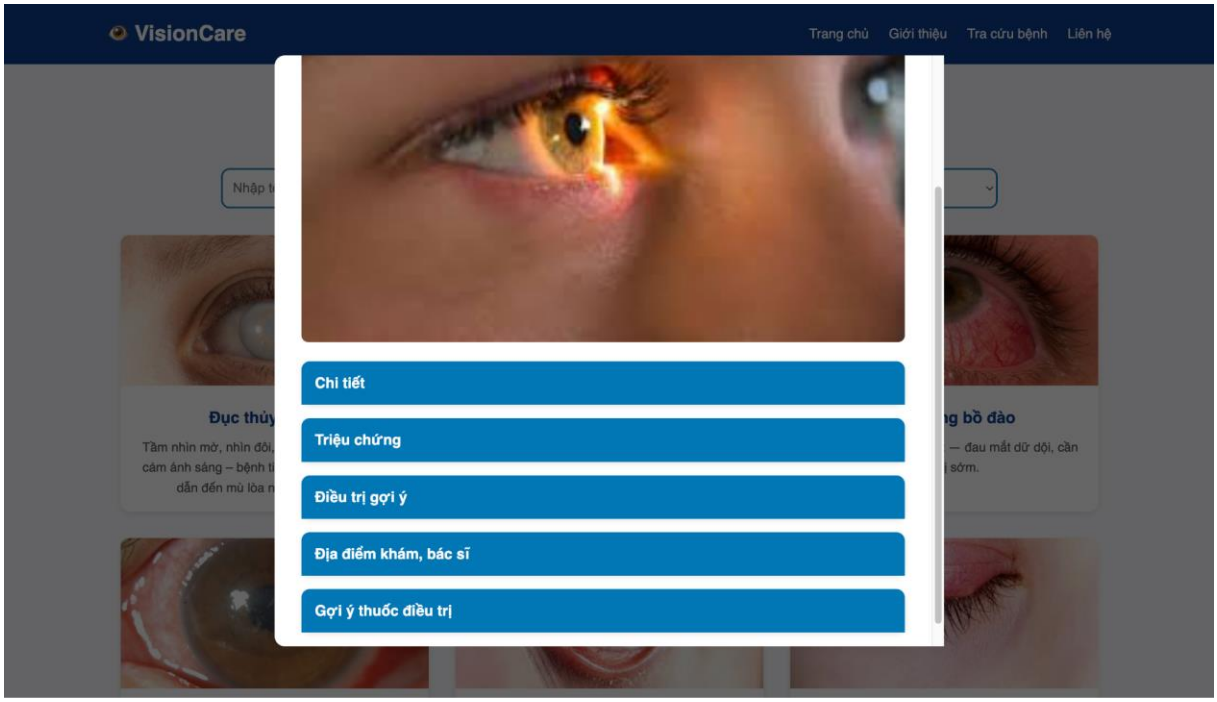


**Figure 18: Disease lookup on the phone*

The search results are displayed in a grid format with illness-cards that include a profile picture, a brief title and description, ensuring a balance between aesthetics and usability. When tapping on a disease card, the system opens a modal popup, which provides detailed information about the pathology in the form of collapsed/expanded boxes, with the sections "Details", "Symptoms", "Treatment" and "Location", allowing users to freely access information according to their needs. avoid feeling overwhelmed.



**** Figure 19: Details of each disease***

1.1.3. Chatbot for advice and Q&A about eye diseases

One of the outstanding functions inside VisionCare is **EyeVision Chatbot**, an automated consultation system integrated directly into the website. EyeVision Chatbot allows users to interact quickly by selecting suggested topics or entering freehand questions related to eye health. The entire consultation content from the chatbot is focused in-depth on topics such as cataracts, glaucoma, diabetic retinopathy, myopia treatment, eye care after Lasik, treatment of sudden blurred vision, eye nutrition, eye protection habits, and AI technology in the diagnosis of retinal diseases.

EyeVision Chatbot operates by sending user requests to the processing system using the new generation language model Google Gemini (Generative Language API), receiving feedback, then automatically formatting the content (bold, italic, list...) to display smoothly with vivid typing effects. The entire chat history is also stored locally on the browser using localStorage, ensuring convenience and privacy for users.

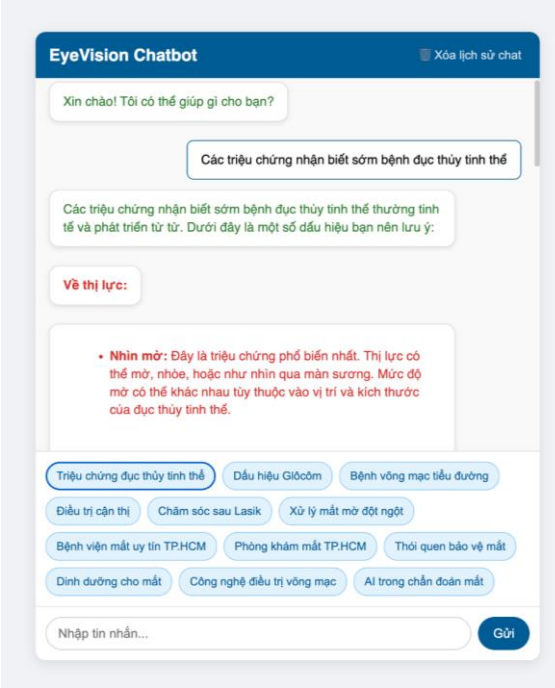
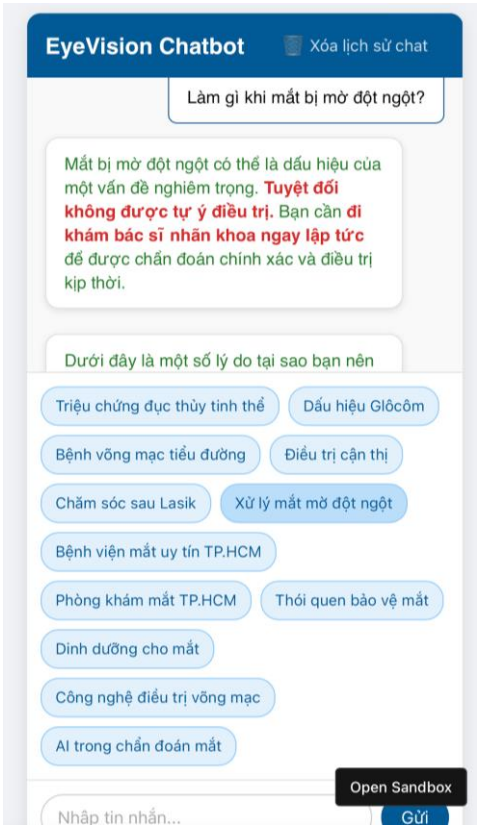


**Figure 20: EyeVision chatbot interface for eye health consultation on computer*



** Figure 21: EyeVision chatbot interface for eye health consultation on the phone*

1.2. Description of the development platform of the product

1.2.1. Programming Languages

- **Python 3.9+**
  - Serve as the "backbone" for every AI component: training and inferencing CNN/PyTorch models, processing Grad-CAM, LIME, Unet, SQLite storage, generating PDFs with FPDF.
  - Runs on Streamlit for both fast development and an interactive web interface.
- **JavaScript (ES6)**
- **HTML5 & CSS3**

1.2.2. Platform for application development: Web application and software

1.2.3. Description of the operation of the software

1. Startup & Resource Loading

2. Initial Role & Configuration Selection

3. Pathology Lookup & Chatbot (client-side)

4. Photo Upload & AI Inference

- Users click "Select file" to upload 1–n photos of the bottom of the eyes.

- Each image is transferred by Streamlit via PIL → PyTorch tensor → normalize.
- Song song:
  - **Classification** : forward pass via CNN → logits → softmax → label & authentication ấ t.
  - **Segmentation** : forward pass qua U-Net → sigmoid map → threshold → binary mask.
- **Explainable AI**: Grad-CAM & LIME hook deep into layers, generating heatmaps/highlight areas.

5. Displaying Results & Engagement

6. Record Keeping & Report Export

7. Manage & review records

1.3. Conclusion:

The VisionCare **system** represents a comprehensive step forward compared to existing research and products on the market thanks to the synchronous combination of three main feature groups: (1) **AI for multi-disease diagnosis** (classification + segmentation) with the ability to simultaneously identify up to 11 retinal diseases and create detailed vascular masks; (2) **a flexible, intuitive** pathology lookup library; and (3) **EyeVision Medical Chatbot** for instant consultation.

1.3.1 Comparison with previous study:

- SURF/PHOG or CNN models that simply focus on one or two pathologies often ignore the diversity of lesions on the same fundus image. VisionCare uses fine-tuning sequences on EfficientNet, ResNet, and Swin Transformer to achieve > 92% accuracy on ODIR-5K and several other public eye disease datasets, and addresses data imbalances through built-in focal loss and kappa loss.
- Many works only stop at the classification part, while the segmentation of blood vessels helps assess the extent of the extent of the damage, assisting doctors in accurately calculating the "area" of the diseased area. This is a feature that most other products don't have.
- Model **Interpretation features** (Grad-CAM, Grad-CAM++, LIME) improve transparency, overcoming the "black box" disadvantages often found in traditional medical AI systems.

1.3.2. Superiority in deployment and user experience:

- The integration **of SQLite + FPDF** to archive and export PDF reports with full metadata, heatmaps, and mask segmentation forms an end-to-end solution from diagnostics to storage, superior to platforms that only provide static output.

- **EyeVision Chatbot** connects to Google Gemini emulation to increase interoperability, while saving local history to ensure privacy, a highlight compared to medical chatbots that merely provide FAQs.

1.3.3. Conclusion

- **Solving the problem of multiple diseases simultaneously:** VisionCare is the first system at the student project scale to combine the classification of 11 retinal diseases and vascular segmentation, enabling a comprehensive diagnosis based on a single fundus image.
- **High accuracy and reliability:** Fine-tuned transfer learning models achieve > 92% accuracy, supported by FocalLoss and kappa loss, and interpretation of results using Grad-CAM and LIME to enhance physician confidence.
- **Optimal User Experience:** The user-friendly, highly interactive Streamlit interface with a lookup library, chatbot, instantly stores and exports PDF reports, meeting clinical and educational requirements.
- **Scalability and practical application:** The modular architecture allows for the addition of new models (e.g. OCT diagnostics), expansion to hospital systems, integration of tele-medicine, minimizing the shortage of ophthalmologists and shortening diagnosis time.

2. Product reviews

2.1. Application Potential

COMPARISON TABLE OF AI PRODUCTS FOR DIAGNOSING EYE DISEASES IN VIETNAM

Criteria	VisionCare	EyeDr (Ho Chi Minh City Eye Hospital) Ho Chi Minh City)	Optain AI (Australia – implemented in Vietnam)
1. Number of co-diagnosed diseases	11 diseases (multiple retinal disease)	1 disease (glaucoma)	3 diseases (Glaucoma, retinopathy, AMD)
2. Segmentation	✅ Yes (UNet+ResNet)	❌ None	❌ None

3. Explainable AI	✅ Yes (Grad-CAM, LIME)	❌ None	❌ No work b ố
4. Average Accuracy (%)	93–96%	90% Sensitivity, 93% Specialization	95% (as announced)
5. Diagnostic speed	<8 seconds	8–10 seconds	Undisclosed
6. Export reports & keep records	✅ Yes (PDF, ICD, photo)	❌ None	✅ Yes
7. User Interface & Implementation	Web-based, đa nền tảng	Internal Applications	Specialized equipment + software
8. Consultation chatbot functionality	✅ Yes (Gemini AI)	❌ None	❌ None
9. Severity Assessment	✅ Yes (scoring)	❌ None	❌ No work b ố
10. Model scalability	✅ D ễ hu ấ n accumulation ệ n add b ệ nh	❌ None	❌ No work b ố
11. Scope of application	Community, Provincial, Hospital	Provincial and community hospitals	Major hospitals, medical centers

PRODUCT COMPARISON OF AI DIAGNOSIS OF RETINOPATHY ABROAD

Criteria	VisionCare	IDx-DR / LumineticsCore (Digital Diagnostics, Mỹ)	EyeArt (Eyenuk, Mỹ)	AEYE-DS (AEYE Health, Mỹ)

1. Number of co-diagnosed diseases	11 diseases (DR, AMD, Glaucoma...)	1 disease (moderate or higher DR)	1 disease (moderate DR and risk of vision loss)	1 disease (DR)
2. Segmentation	✅ Yes (UNet+ResNet)	❌ None	❌ None	❌ None
3. Explainable AI	✅ Yes (Grad-ORANGE, LIME)	❌ None	❌ None	❌ None
4. Average Accuracy (%)	93–96%	87% Sensitivity, 90% Specialization	96.5% sensitivity, 86% specialization	Undisclosed
5. Diagnostic speed	<8 seconds	~1 minute	~1 minute	~1 minute
6. Export reports & keep records	✅ Yes (PDF, ICD, she nh)	✅ Yes (automatic reporting)	✅ Yes (automatic reporting)	✅ Yes (automatic reporting)
7. User Interface & Implementation	Web-based, đa nền tảng	Specialized equipment with Topcon TRC-NW400 camera	Web-based, integrated with Canon/Topcon cameras	Optomed Aurora Handheld
8. Consultation chatbot functionality	✅ Yes (Gemini AI)	❌ None	❌ None	❌ None

9. Severity Assessment	✅ Yes (scoring)	✅ Yes (moderate or higher DR classification)	✅ Yes (moderate DR and risk of vision loss)	✅ Yes (Medium level DR or higher)
10. Model scalability	✅ Dễ dàng accumulation và add bệnh	❌ None	❌ None	❌ None
11. Scope of application	Community, Provincial, Hospital	Clinics and hospitals in the US	Clinics and hospitals in the US	Clinics and hospitals in the US

VisionCare ushers in a new era for the diagnosis and monitoring of retinal diseases in the medical community and the general public:

- **Scale and scope**:
  - **District-level hospitals, commune health stations**: Easy to deploy on common computers, helping to improve the quality of early screening at low cost.
  - **Private Clinic**: Quickly integrate into the ophthalmology examination process, shortening the waiting time for results.
  - **Tele-ophthalmology program**: Connection between the grassroots and city experts, especially useful for remote areas.
- **Practical Application**:
  - At the local private ophthalmology clinic (Hoc Mon), the trial phase in March applied VisionCare to screen for diabetic retinopathy for more than 100 images of the patient's fundus and retina, shortening the average diagnosis time from 15 minutes to 4-5 minutes. the rate of early detection increased.

2.2. Effectiveness when applying the product

Criteria	VisionCare	Traditional methods	Previous single AI applications
Diagnosis time	<5 seconds/time	5–10 minutes every time	~1 minute of single image

Accuracy ()	93–96% (kappa quadratic ≥ 0.85) and 96-98% accompanied by the doctor's expertise	~90%-95% of diagnostic doctors have no external support	85–92% (classification only)
Pathological range	In parallel with multiple diseases, incorporate vascular segmentation to increase professional accuracy and differentiate results	Only individual diseases and it is difficult to see the blood vessels of the retina to identify the disease	Only classification or segmentation

3. Infrastructure Requirements

Ingredients	Minimum configuration
Hardware	4-core CPU, 16 GB RAM, (optional, GTX 1060 GPU (6 GB))
Software	Python ≥3.9, PyTorch ≥1.12, Streamlit ≥1.18
OS	Windows 7,10,11 / Ubuntu 20.04 LTS
Network	Internet bandwidth ≥5 Mbps (for telemedicine)

4. The product is developed over an estimated period of time

- Total: 7 months
- Period: From 11/2024 to 5/2025

5. Product Usage Instructions

https://shorturl.at/Lg4md

6. Self-assessment of the remaining unresolved aspects of the product to overcome.

6.1. Data Diversity and Quality

- *Problem:* The ODIR-5K dataset is primarily focused on cases from a local area, not fully covering variation in race, age, and imaging conditions.

- *Fix:* Establish cooperation with ophthalmology centers at home and abroad to collect diverse fundus images. Implement the process of inspecting image quality (resolution, brightness) and augmenting appropriate data (GAN, image transformation).

6.2. Lack of multimodal integration

- *Problem:* Only fundus images are used, not combined with OCT, eye ultrasound, or clinical information (diabetes, blood pressure).
- *Fix:* Extend the architecture to a multimodal system, incorporating OCT images and electronic data (EHR) inputs. Develop an API to connect to the hospital management system to obtain pathological information and examination history.

6.3. Health & Security Regulatory Compliance

- *Problem:* The system has not been tested according to HIPAA/GDPR standards, has not been end-to-end codified, and has not passed CE/FDA certification.
- *Fix:* Apply HTTPS, OAuth 2.0, AES encryption for sensitive data; coordinate with medical certification bodies for security audits and clinical trials.

CONCLUSION

1. The development direction of the product in the future

In the coming time, the VisionCare platform will continue to be upgraded and improved comprehensively with specific development directions as follows:

- **Database expansion:** Continuously collect and update large and diverse data sources on retinal images, widely covering common and rare eye diseases, in order to improve the system's accurate diagnostic capabilities.
- **Integrating new technologies:** Applying next-generation artificial intelligence models (such as transformers, multitasking learning models, and semi-supervised learning) to improve accuracy, sensitivity, and the ability to detect diseases at the earliest stages.
- **Cross-platform integration:** Develop an app version for mobile devices (iOS and Android), making it easier and more convenient for users.
- **Product Internationalization:** Expand the international market by adding multiple languages, suitable for various countries and regions, and strictly comply with international health data security and safety standards (such as HIPAA and GDPR).

2. Future aspirations

We want VisionCare to become a free, shared platform with the Vietnamese community and trusted in the field of artificial intelligence eye disease diagnosis in Vietnam and reach out to the world. With that vision, we hope that this platform will make an important contribution to the prevention and reduction of blindness rates,

improving the quality of life for the community. At the same time, we are always ready to cooperate with domestic and foreign experts, hospitals, and research institutions to constantly improve technology, expand practical applications and spread the humanistic medical value that the product aims at.

REFERENCES

Bhati, A., Gour, N., Khanna, P., & Ojha, A. (2022). Discriminative Kernel Convolution Network for Multi-Label Ophthalmic Disease Detection on Imbalanced Fundus Image Dataset. *arXiv preprint arXiv:2207.07918*. https://arxiv.org/abs/2207.07918

Cleveland Clinic. (2024). *Fundus photography: What it is & when it is needed*. Cleveland Clinic. https://my.clevelandclinic.org/health/diagnostics/fundus-photography

Kaur, P., & Kaur, R. (2023). Evaluating the efficacy of teleophthalmology in delivering ophthalmic care to underserved populations: A literature review. *Healthcare*, 12(8), 818. https://doi.org/10.3390/healthcare12080818

Li, N., Li, T., Hu, C., Wang, K., & Kang, H. (2021). A benchmark of ocular disease intelligent recognition: One shot for multi-disease detection. *arXiv preprint arXiv:2102.07978*. https://arxiv.org/abs/2102.07978

Singh, D., Agarwal, S., & Mishra, S. (2025). Retinal Fundus Multi-Disease Image Classification using Hybrid CNN-Transformer-Ensemble Architectures. *arXiv preprint arXiv:2503.21465*. https://arxiv.org/abs/2503.21465

World Health Organization. (2019). *World report on vision*. World Health Organization. https://www.who.int/publications-detail-redirect/world-report-on-vision

Cleveland Clinic. (2023). *Retinal Imaging: Purpose & Procedure Details*. https://my.clevelandclinic.org/health/diagnostics/25046-retinal-imagingCleveland Clinic

Bhati, A., Gour, N., Khanna, P., & Ojha, A. (2023). Discriminative kernel convolution network for multi-label ophthalmic disease detection on imbalanced fundus image dataset. *Computers in Biology and Medicine*, 153, 106519. https://doi.org/10.1016/j.compbiomed.2022.106519

Li, N., Li, T., Hu, C., Wang, K., & Kang, H. (2021). A benchmark of ocular disease intelligent recognition: One shot for multi-disease detection. In *Benchmarking, Measuring, and Optimizing* (pp. 177–193). Springer. https://doi.org/10.1007/978-3-030-71058-3_11ACM Digital Library+2OUCI+2PMC+2

Singh, D., Agarwal, S., & Mishra, S. (2025). Retinal fundus multi-disease image classification using hybrid CNN-transformer-ensemble architectures. *arXiv preprint arXiv:2503.21465*. https://arxiv.org/abs/2503.21465

Dolar-Szczasny, J., Barańska, A., & Rejdak, R. (2023). Evaluating the efficacy of teleophthalmology in delivering ophthalmic care to underserved populations: A literature review. *Journal of Clinical Medicine*, 12(9), 3161. https://doi.org/10.3390/jcm12093161

PyTorch. (2024). *PyTorch documentation (Version 2.7)*. https://pytorch.org/docs/stable/index.htmlWikipedia+3PyTorch+3PyTorch+3

Google AI. (2024). *Gemini API documentation*. Google AI. https://ai.google.dev/gemini-api/docs

Ho Chi Minh City, May 28, 2025

Signature of the author/group of authors

Le Truong Minh Dang Nguyen Le Quynh Chau