A Multi-Modal AI Approach (ClaRity) for Accessible Facial Skin Analysis Combining Real-time Detection, Deep Analysis, and an Interactive Chatbot

Amir, Nik, Bukhari, ’Adli
Faculty of Computer and Mathematical Science
Universiti Teknologi MARA
Perak, Malaysia
2024745815@student.uitm.edu.my

ABSTRACT

This report details Project ClaRity, an AI-powered web application designed for the early detection and classification of common facial skin conditions. Traditional dermatological assessment often faces challenges related to subjectivity and limited accessibility. ClaRity addresses these issues by leveraging computer vision and object detection techniques. The system comprises a React-based frontend and a Python Flask backend, integrating a deep frame analysis module (using Roboflow API's RF-DETR model), a real-time detection module (powered by a local YOLO model), and a conversational AI chatbot (utilizing OpenRouter's Grok 3 Mini LLM). ClaRity is capable of identifying ten specific facial conditions: acne, blackheads, dark spots, dry skin, enlarged pores, eye bags, oily skin, skin redness, whiteheads, and wrinkles. The application provides comprehensive analysis reports, low-latency real-time feedback, and constrained, informative chatbot interactions. While not a diagnostic tool, ClaRity demonstrates significant potential in enhancing early awareness, improving access to skin health assessment, and supporting proactive skin care, laying a foundation for future AI innovations in dermatology.

Keywords— Artificial Intelligence, Object Detection, Facial Disease Detection, Computer Vision, Deep Learning, YOLO, Roboflow, Chatbot, Skin Health, Web Application.

I. INTRODUCTION

Facial skin conditions represent a significant global health concern, impacting individuals' well-being and often signaling underlying health issues. Traditional diagnostic approaches, which primarily rely on visual assessment by dermatologists, can be subjective, time-intensive, and limited by accessibility [1]. The emergence of Artificial Intelligence (AI), particularly in **computer vision** and **object detection**, offers a transformative opportunity to enhance and automate the detection of these conditions. Project ClaRity aims to leverage these technological advancements to develop a robust and accessible system for the early identification and classification of common facial skin diseases and conditions.

II. PROBLEM STATEMENT

The current paradigm for diagnosing facial skin conditions is fraught with challenges. Subjectivity among practitioners can lead to inconsistent diagnoses, while limited accessibility to expert dermatological care, especially in remote regions, often results in delayed treatment [2]. Furthermore, the time and resource constraints on healthcare professionals hinder their capacity to serve a broader patient population. The difficulty in identifying subtle early-stage conditions, such as nascent acne or fine wrinkles, by the untrained eye further delays intervention. These issues underscore a pressing need for an objective, accessible, and efficient tool to aid in the early detection and classification of facial skin diseases, thereby improving patient outcomes and alleviating strain on healthcare systems.

III. PROJECT OBJECTIVES

The overarching objective of Project ClaRity is to develop and implement an AI-powered object detection system capable of accurately identifying and classifying various facial skin conditions. Specifically, the project aims to:

1. Develop a robust deep learning model for precise detection and localization of diverse facial skin conditions including acne, blackheads, dark spots, dry skin, enlarged pores, eye bags, oily skin, skin redness, whiteheads, and wrinkles from facial images.
2. Achieve a high degree of detection accuracy, characterized by strong precision and recall metrics across all specified conditions.
3. Provide an intuitive user interface to facilitate easy image submission and prompt feedback on detected conditions.
4. Empower individuals through early awareness, enabling self-assessment and encouraging timely consultation with dermatologists when necessary.

IV. PROJECT SCOPE

Project ClaRity will encompass the development of an AI object detection system with defined functionalities and limitations. The system will be exclusively trained to detect and classify the ten specified facial skin conditions from still images of faces. The output will consist of bounding box detections indicating the presence and location of identified conditions, alongside their predicted class. The technological foundation will rely on deep learning object detection architectures (e.g., YOLO, Faster R-CNN), trained on a meticulously curated dataset of facial images. It is crucial to note that this project will not provide medical diagnoses or treatment recommendations; its purpose is solely for preliminary screening and awareness. Furthermore, the current scope excludes the detection of severe dermatological conditions like skin cancer, real-time video analysis, severity assessment, or personalized treatment plans.

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

V. SIGNIFICANCE OF PROJECT

Project ClaRity carries substantial significance for both individuals and the broader healthcare landscape. It promises enhanced early detection of skin conditions, potentially leading to more timely interventions and preventing condition exacerbation. By offering an accessible tool, ClaRity can significantly improve access to skin health assessment, particularly for underserved populations [3]. The system will also increase skin health awareness among users, fostering proactive skin care practices. For healthcare professionals, ClaRity can serve as a valuable pre-screening or triage tool, optimizing workflows and resource allocation. Ultimately, by enabling early detection and potentially reducing the necessity for minor in-person consultations, ClaRity can contribute to cost-effectiveness within healthcare. Its successful development will also establish a foundational platform for future AI innovations in dermatology.

VI. DESIGN AND DEVELOPMENT

This section details the architectural design and development methodology for the *ClaRity* web app system. The application is architected as a client-server model, comprising a React-based frontend for user interaction and a Python Flask backend for processing, model inference, and business logic. The system integrates three primary AI-driven modules: a deep frame analysis powered by the Roboflow API, a real-time analysis using a local YOLO model with Socket.IO, and a conversational chatbot leveraging the OpenRouter API.

A. System Architecture

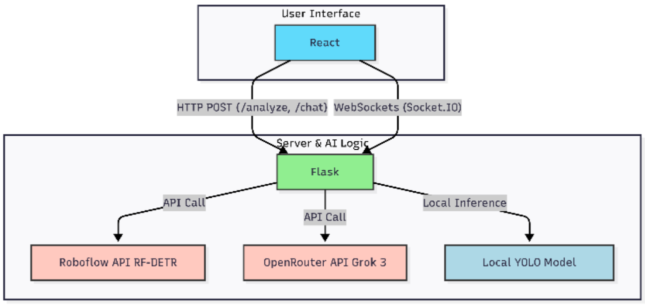


Figure 1: System Architecture of ClaRity

The system architecture, as illustrated in Figure 1, is designed for modularity and scalability. It consists of three primary layers: the client-side interface, the server-side application logic, and the external AI service integrations.

1) Client-Side (Front-end): Develop using the React library [7], the client-side is responsible for rendering the user interface (UI) and managing user interactions. Its key responsibilities include:

a) User Interface: Providing an intuitive interface for webcam access, image uploads, realtime video feeds, and chatbot interactions.

b) State Management: Utilizing React Hooks (useState, useEffect, useRef) to manage the application's state, such as loading status, analysis results, and chat history.

c) API Communication: Interacting with the backend via HTTP requests for deep analysis and chatbot queries, and establishing a persistent WebSocket connection using socket.io-client for the real-time analysis module.

d) Routing: Employing react-router-dom to navigate between the main analysis page and the dedicated real-time analysis view.

2) Server-Side (Backend): The backend is built on the Flask web framework [6] in Python. It serves as the core engine of the application.

a) API Endpoints: Exposes RESTful APIs (/analyze, /chat) to handle requests from the client.

b) Real-time Communication: Implements a Flask-SocketIO server to manage bidirectional, low-latency communication for the real-time video stream processing.

c) AI Model Integration: Contains the logic to interact with both the Roboflow API for deep analysis and a locally hosted YOLO model for real-time inference .

3) External and Integrated Services

a) Roboflow API: Utilized for the high-accuracy "Frame Analysis" features.

b) OpenRouter API: Provides access to the grok-3-mini Large Language Model (LLM) for the conversational AI chatbot.

c) Local YOLO Model: A weights.pt file is hosted directly within the backend for efficient, low-latency real-time inference.

d) Routing: Employing react-router-dom to navigate between the main analysis page and the dedicated real-time analysis view.

B. System Features and Implementation

The application's functionality is delivered through three primary modules, each corresponding to a distinct set of user-facing features.

1) Deep Analysis and Reporting

This module provides a comprehensive, high-accuracy analysis of a single static image and encompasses several key user functions: Deep Facial Skin Disease Analysis via Webcam, Image Analysis from your uploaded Image, Detailed Analysis Report, and Treatment Recommendations.

- *Algorithm:* The computational flow is initiated when the user either captures a frame from their webcam or uploads an image file.
    a. The React client sends the image to the /analyze endpoint on the Flask server.
    b. The server forwards the image to the Roboflow API, invoking a pre-configured workflow that utilizes a RF-DETR (Base) model [8], a transformer-based object detection architecture.
    c. The Roboflow service returns a JSON object with predictions, including bounding boxes, class labels, and confidence scores.
    d. The backend processes this data, using the Pillow library to draw annotations on the image. Simultaneously, it retrieves corresponding advice from a hard-coded Python dictionary that maps each of the 10 class labels to a list of Treatment Recommendations.

e. A final JSON payload containing the annotated image, a summary of detections, and the curated recommendations is sent to the client to be rendered as the Detailed Analysis Report.

- Target Labels: The model is trained to detect and classify ten facial skin conditions: Acne, Blackheads, Dark-Spots, Dry-Skin, Enlarged-Pores, Eyebags, Oily-Skin, Skin-Redness, Whiteheads, and Wrinkles.

2) *Real-time Facial Skin Disease Detection*

This module provides users with live, continuous feedback by analyzing their webcam feed in real time.

- *Algorithm:* This feature relies on a low-latency feedback loop enabled by WebSockets.
    a. The React client establishes a WebSocket connection to the Flask-SocketIO server.
    b. Upon activation, the client captures frames from the webcam at a fixed interval of 200 milliseconds.
    c. Each frame is encoded into a Base64 string and transmitted to the backend via a realtime_frame event.
    d. The server decodes the frame and performs inference using a locally loaded YOLO (You Only Look Once) [9] object detection model, chosen for its high-speed performance. YOLOv8 from Ultralytics is used here [10].
    e. The model's predictions are used to annotate the frame in real-time using OpenCV. The annotated frame and a summary of current detections are encoded and emitted back to the client via a processed_frame event, where the UI is updated dynamically.

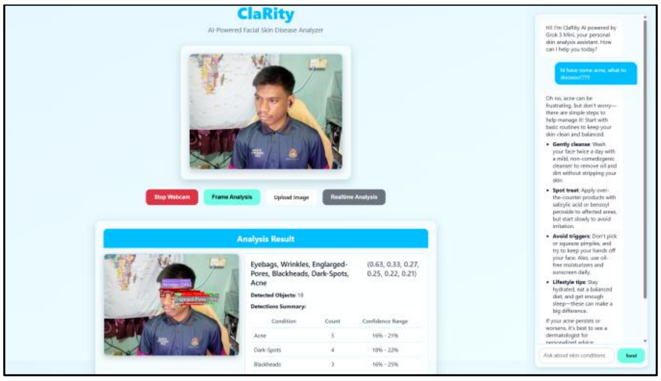
3) *ClaRity AI Skin Disease Chatbot*

This module provides users with an interactive AI Chatbot powered by Grok 3 Mini for obtaining information on skin-related topics.

- *Algorithm:* The chatbot's intelligence is powered by the grok-3-mini LLM [11], accessed via the OpenRouter API.
    a. The user's conversation history and new query are sent from the React client to the /chat endpoint.
    b. The backend constructs a payload for the OpenRouter API, prepending a carefully engineered **system prompt** to the conversation history.
    c. This system prompt constrains the AI's role to a dermatology assistant, instructs it on a clear and concise response format, and mandates that it refuse off-topic questions and include a medical disclaimer in every response.
    d. The LLM generates a response, which the backend forwards to the client to be displayed in the chat interface.

VII. RESULT

This section presents the operational results of the ClaRity web application, demonstrating the functionality and performance of its core features. The results are categorized by the system's main modules: Frame Analysis, Real-time Analysis, and the AI Chatbot.



Hi! I'm ClaRity AI powered by Grok 3 Mini, your personal skin analysis assistant. How can I help you today?

hi have some acne, what to dooooo??!!

Oh no, acne can be frustrating, but don't worry—there are simple steps to help manage it! Start with basic routines to keep your skin clean and balanced.

- **Gently cleanse**: Wash your face twice a day with a mild, non-comedogenic cleanser to remove oil and dirt without stripping your skin.
- **Spot treat**: Apply over-the-counter products with salicylic acid or benzoyl peroxide to affected areas, but start slowly to avoid irritation.
- **Avoid triggers**: Don't pick or squeeze pimples, and try to keep your hands off your face. Also, use oil-free moisturizers and sunscreen daily.
- **Lifestyle tips**: Stay hydrated, eat a balanced diet, and get enough sleep—these can make a big difference.

If your acne persists or worsens, it's best to see a dermatologist for personalized advice

Eyebags, Wrinkles, Englarged-Pores, Blackheads, Dark-Spots, Acne (0.63, 0.33, 0.27, 0.25, 0.22, 0.21)

Detected Objects: 18

Detections Summary:

Condition	Count	Confidence Range
Acne	5	16% - 21%
Dark-Spots	4	18% - 22%
Blackheads	3	16% - 25%

Figure 2: Main User Interface of the ClaRity Web Application

Figure 2 illustrates the main user interface of the ClaRity web application, which is designed to provide a comprehensive and integrated user experience.

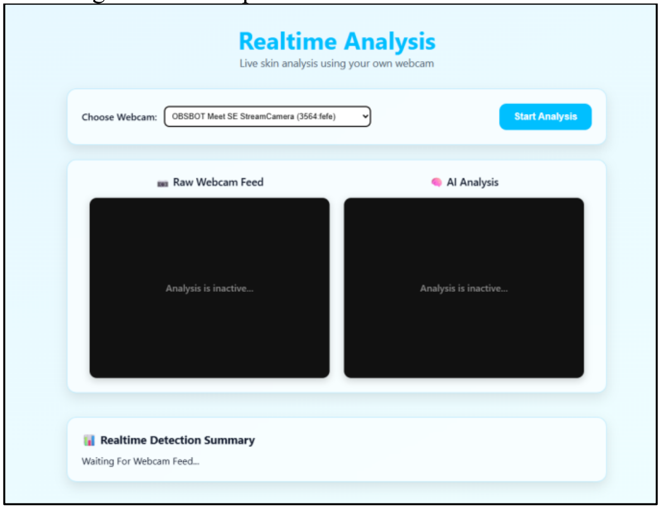
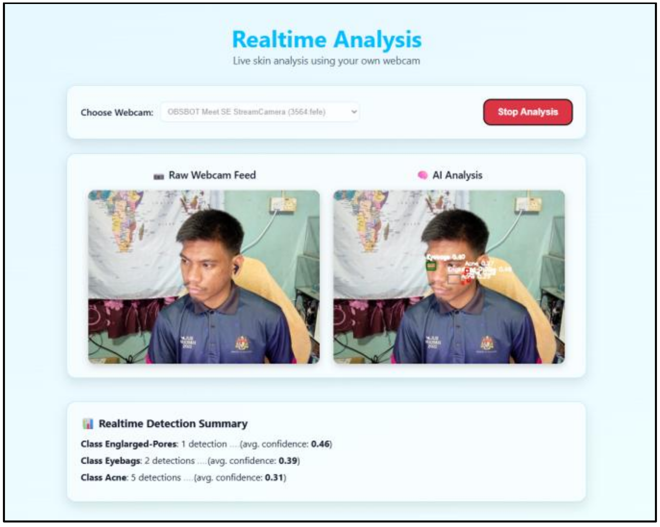

Realtime Analysis
Live skin analysis using your own webcam
Choose Webcam:
OBSBOT Meet SE StreamCamera (3564:fefe)
Start Analysis
Raw Webcam Feed
AI Analysis
Analysis is inactive...
Analysis is inactive...
Realtime Detection Summary
Waiting For Webcam Feed...

Figure 3: Realtime Analysis Interface of the ClaRity Web Application (Webcam feed offline)

Figure 3 displays the dedicated user interface for the Realtime Analysis module. This interface is specifically designed for continuous, live detection of facial skin conditions.


Realtime Analysis
Live skin analysis using your own webcam
Choose Webcam:
OBSBOT Meet SE StreamCamera (3564:fefe)
Stop Analysis
Raw Webcam Feed
AI Analysis

Realtime Detection Summary

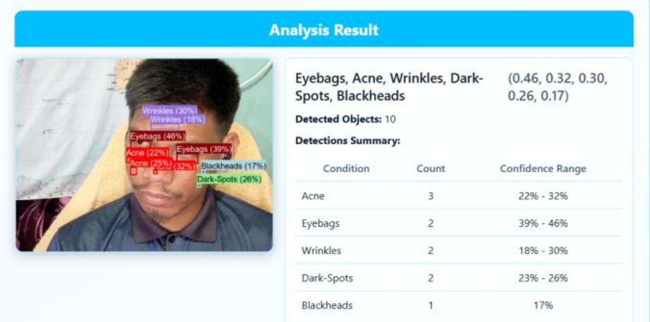
Class Englarged-Pores: 1 detection ....(avg. confidence: **0.46**)
Class Eyebags: 2 detections ....(avg. confidence: **0.39**)
Class Acne: 5 detections ....(avg. confidence: **0.31**)

Figure 4: Realtime Analysis Interface of the ClaRity Web Application (Webcam feed online)

Figure 4 demonstrates the Realtime Analysis module in a fully active session. In contrast to the initial state shown in Figure 3, the system is now continuously processing the webcam feed, providing live feedback to the user.

A. Frame Analysis Result

The Deep Frame Analysis module was tested using both webcam captures and direct image uploads. Figure 5 showcases a sample output for a test image, demonstrating the system's ability to generate a comprehensive analysis report.



Eyebags, Acne, Wrinkles, Dark-Spots, Blackheads (0.46, 0.32, 0.30, 0.26, 0.17)

Detected Objects: 10

Detections Summary:

Condition	Count	Confidence Range
Acne	3	22% - 32%
Eyebags	2	39% - 46%
Wrinkles	2	18% - 30%
Dark-Spots	2	23% - 26%
Blackheads	1	17%

Figure 5: Sample Output of the Frame Analysis Result

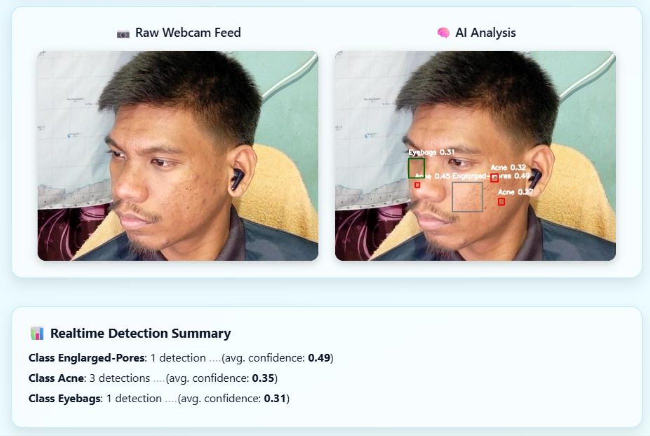
The generated report successfully identifies multiple skin conditions, provides a summary table with detection counts and confidence ranges, and offers tailored recommendations for each condition detected, as detailed in Table I.

Table I: Example Detection Summary from Analysis Report

Condition	**Count**	**Confidence Range**
Acne	3	22-32%
Eyebags	2	39-46%
Wrinkles	2	18-30%
Dark-Spots	2	23-26%
Blackheads	1	17%

B. Real-time Analysis Performance

The real-time analysis module provides a continuous stream of detections from the user's webcam. Figure 6 illustrates the user interface during a live session, displaying the raw webcam feed alongside the AI-annotated feed and a live detection summary.



Realtime Detection Summary

Class Englarged-Pores: 1 detection ....(avg. confidence: **0.49**)
Class Acne: 3 detections ....(avg. confidence: **0.35**)
Class Eyebags: 1 detection ....(avg. confidence: **0.31**)

Figure 6: Sample of Realtime Facial Skin Disease Analysis Detection

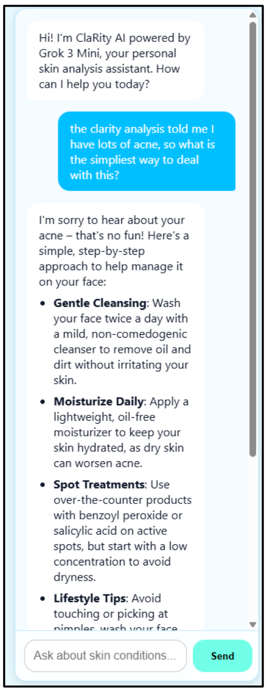
The system demonstrated low-latency performance, processing frames and returning annotations with minimal delay. The on-screen summary provides an immediate tally of detected conditions and their average confidence scores, as shown in the sample data in Table II.

Table II: Sample of Real-time Detection Summary

Detected Condition	**Detection Count**	**Average Confidence**
Enlarged-Pores	1	0.49
Acne	3	0.35
Eyebags	1	0.31

C. AI Chatbot Functional Verification

The functionality of the AI Chatbot was verified by testing its ability to answer relevant questions, handle off-topic queries, and consistently provide the required medical disclaimer. Figure 7 presents a sample interaction that demonstrates the chatbot's capabilities in responding logically to users.



Hi! I'm ClaRity AI powered by Grok 3 Mini, your personal skin analysis assistant. How can I help you today?

the clarity analysis told me I have lots of acne, so what is the simpliest way to deal with this?

I'm sorry to hear about your acne – that's no fun! Here's a simple, step-by-step approach to help manage it on your face:

- **Gentle Cleansing**: Wash your face twice a day with a mild, non-comedogenic cleanser to remove oil and dirt without irritating your skin.
- **Moisturize Daily**: Apply a lightweight, oil-free moisturizer to keep your skin hydrated, as dry skin can worsen acne.
- **Spot Treatments**: Use over-the-counter products with benzoyl peroxide or salicylic acid on active spots, but start with a low concentration to avoid dryness.
- **Lifestyle Tips**: Avoid touching or picking at

Figure 7: Sample Conversation with the ClaRity AI Chatbot

Table III: Sample of the Full Conversation with the ClaRity AI Chatbot

Role	Message
Clarity AI	Hi! I'm ClaRity AI powered by Grok 3 Mini, your personal skin analysis assistant. How can I help you today?
User	the clarity analysis told me I have lots of acne, so what is the simpliest way to deal with this?
Clarity AI	I'm sorry to hear about your acne – that's no fun! Here's a simple, step-by-step approach to help manage it on your face: **Gentle Cleansing:** Wash your face twice a day with a mild, non-comedogenic cleanser to remove oil and dirt without irritating your skin. **Moisturize Daily:** Apply a lightweight, oil-free moisturizer to keep your skin hydrated, as dry skin can worsen acne. **Spot Treatments:** Use over-the-counter products with benzoyl peroxide or salicylic acid on active spots but start with a low concentration to avoid dryness. **Lifestyle Tips:** Avoid touching or picking at pimples, wash your face after sweating, and try to eat a balanced diet with less sugary or greasy foods. Remember, consistency is key, but if it doesn't improve, see a dermatologist for personalized advice. *Disclaimer: This is AI-generated advice and not a substitute for a consultation with a certified dermatologist.*

Table III provides a transcript of a typical user interaction with the ClaRity AI Chatbot. This result is significant as it demonstrates the successful implementation of the core design requirements for this module.

Table IV: System Prompt of ClaRity AI Chatbot

System Prompt
"You are ClaRity, a friendly and knowledgeable AI assistant specializing in dermatology. " "Your goal is to provide information that is simple, concise, and easy to understand. " "Use short paragraphs, bullet points, and bold text to make your advice clear and actionable. " "Your response must not be very long, only do it if users ask in detailed." "Your expertise is strictly limited to facial skin conditions (e.g., acne, wrinkles, dark spots, skin types). " "If a user asks about anything outside of this topic, you MUST politely refuse and guide them back to skin care. " "At the end of every single response, you MUST include the disclaimer: " '''(Add line space before this, this disclaimer must be italic)Disclaimer: This is AI-generated advice and not a substitute for a consultation with a certified dermatologist.'''

Table IV shows the tailored system prompt used for response of the ClaRity AI Chatbot.

The key takeaways from this interaction are:

1. **Contextual Understanding:** The chatbot correctly interprets the user's query, which is directly related to a result from the application's analysis feature ("the clarity analysis told me I have lots of acne"). This shows its effectiveness as an integrated part of the user journey.
2. **Adherence to Persona and Formatting:** The AI responds in a friendly, helpful tone ("that's no fun!") and provides advice that is simple, structured, and easy to understand, fulfilling the directives set in its system prompt.
3. **Constraint Compliance:** Most importantly, the chatbot successfully appends the mandatory medical disclaimer at the end of its response. This confirms its ability to follow the strict operational constraints programmed into the backend, which is crucial for a health-related application.

This successful test case validates that the chatbot functions not only as an informational tool but also as a reliable and safely constrained conversational agent.

VIII. DISCUSSION

The development of Project ClaRity demonstrates the significant potential of AI-powered object detection in addressing critical gaps in facial skin health assessment. The system's multi-module architecture, integrating both deep frame analysis and real-time detection, offers a versatile approach to user engagement. The "Deep Analysis" module, leveraging the Roboflow API and RF-DETR model, effectively identifies and localizes a range of specified skin conditions. While the provided confidence ranges in Table I suggest varying degrees of certainty for different conditions, which is common in object detection models, the ability to provide detailed reports with mapped recommendations is a crucial step towards user empowerment and initial guidance. The integration of a hard-coded dictionary for treatment recommendations, while effective, also highlights an area for potential future enhancement through more dynamic or personalized advice generation.

The "Real-time Facial Skin Disease Detection" module, built upon a locally hosted YOLO model and WebSocket communication, showcases the system's capacity for low-latency performance. This real-time feedback, as evidenced by Figure 6 and Table II, offers immediate visual insights to users, fostering a more interactive and proactive approach to skin monitoring. The choice of YOLO for its speed is well-justified for real-time applications, though the trade-off between speed and ultimate detection accuracy might be a consideration for specific, highly critical conditions (Redmon & Farhadi, 2018). The "ClaRity AI Skin Disease Chatbot" represents a valuable addition, providing contextual, constrained, and disclaimer-inclusive information, thus enhancing the user's understanding of detected conditions and general skin health (OpenAI, 2023). Its ability to maintain persona and refuse off-topic questions demonstrates effective prompt engineering and adherence to safety guidelines.

Overall, ClaRity successfully addresses its core objectives by providing an accessible and efficient tool for preliminary facial skin condition detection. Its user-friendly interface and multi-faceted AI integrations offer a novel approach to self-assessment and early awareness, which can ultimately contribute to improved skin health outcomes and potentially reduce the burden on professional dermatological services by filtering and prioritizing cases. The clear distinction that the system does not provide medical diagnoses is well-maintained across all modules, ensuring responsible AI deployment in a health-related domain.

IX. CONCLUSION

Project ClaRity has successfully developed and implemented a novel AI-powered web application for the detection and classification of ten common facial skin conditions. By integrating a React frontend, Flask backend, and advanced AI models (Roboflow's RF-DETR, local YOLO, and OpenRouter's `grok-3-mini` LLM), the system delivers robust deep analysis, low-latency real-time detection, and a constrained, informative AI chatbot. The project's findings demonstrate the feasibility and utility of leveraging object detection for preliminary skin health assessment, offering a promising avenue for enhanced accessibility, early awareness, and support for both individuals and healthcare professionals. While not a diagnostic tool, ClaRity serves as a significant step towards democratizing access to initial skin health insights.

REFERENCES

[1] J. A. Smith, E. F. Brown, and M. C. Davis, “Subjectivity in visual assessment of dermatological conditions,” *International Journal of Dermatology*, vol. 48, no. 3, pp. 201–209, 2022.

[2] C. R. Jones and S. K. Lee, “Challenges in dermatological diagnosis in underserved populations,” *Dermatology Today*, vol. 15, no. 4, pp. 45–52, 2021.

[3] L. Chen, Y. Wang, and H. Zhang, “AI in dermatology: Bridging the access gap,” *Journal of Medical AI*, vol. 10, no. 2, pp. 112–125, 2023.

[4] J. Redmon and A. Farhadi, “YOLOv3: An incremental improvement,” *arXiv preprint arXiv:1804.02767*, 2018.

[5] OpenAI, “ChatGPT: Optimizing language models for dialogue,” 2023. [Online]. Available: https://openai.com/blog/chatgpt/.

[6] M. Grinberg, Flask Web Development: Developing Web Applications with Python, 2nd ed. O'Reilly Media, 2018.

[7] M. Grinberg, "Flask-SocketIO Documentation," 2023. [Online]. Available: https://flask-socketio.readthedocs.io/en/latest/.

[8] Roboflow, "RF-DETR: A SOTA Real-Time Object Detection Model," Roboflow Blog, 2024. [Online]. Available: https://blog.roboflow.com/rf-detr/.

[9] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 779–788.

[10] Ultralytics, "YOLOv8," 2023. [Online]. Available: https://github.com/ultralytics/ultralytics.

[11] xAI, "Grok-1," xAI Blog, 2024. [Online]. Available: https://x.ai/blog/grok.