IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁒ Peer-reviewed & Refereed journal ⁒ Vol. 15, Issue 5, May 2026

DOI: 10.17148/IJARCCE.2026.155122

DermAI: A Vision Transformer-Based Web System for Multi-Class Skin Disease Detection Utilizing DINOv2

Priyanka Dnyaneshwar Muley

Department of Computer Science Engineering, International Center of Excellence In Engineering and Management (ICEEM), Maharashtra, India

Abstract: Dermatological conditions represent a significant global health burden, often requiring specialized expertise for accurate diagnosis. Early detection of skin diseases, particularly malignancies like melanoma, is critical for improving patient outcomes. This paper presents **DermAI**, an advanced, production-ready skin disease detection system powered by a fine-tuned **DINOv2-Base** Vision Transformer (ViT) architecture. By leveraging self-supervised features distilled from large-scale image data, DermAI achieves a validation accuracy of **95.57%** across **31 distinct skin disease classes**. The proposed system integrates a high-performance deep learning backend with a user-friendly web interface, facilitating real-time inference and ranking predictions by confidence. Our results demonstrate that foundational vision models can be effectively repurposed for specialized medical diagnostic tasks with high precision and reliability. The global prevalence of dermatological conditions, coupled with a shortage of specialized dermatologists, necessitates the development of accessible, highly accurate automated diagnostic tools. This paper presents **DermAI**, an end-to-end web-based system designed for the classification of 31 distinct skin diseases. By integrating a fine-tuned **DINOv2-Base** Vision Transformer (ViT-B/14) with a robust full-stack web architecture (Python, Flask, HTML, CSS), the system delivers real-time, confidence-ranked predictions to the end-user. The underlying deep learning model -images, achieving a remarkable validation accuracy of **95.57%** over 10 epochs. Through detailed case studies—including the high-confidence identification of malignant melanoma and benign fungal infections—this paper demonstrates that foundational self-supervised models can be effectively deployed via lightweight web frameworks to serve as reliable clinical decision-support systems.

I. Introduction

The rising prevalence of skin diseases necessitates the development of objective, computer-assisted diagnostic tools to assist clinicians and provide preliminary screenings in underserved areas. Traditional dermatological evaluation relies on visual inspection, which is subject to inter-observer variability. Artificial Intelligence (AI) has emerged as a paradigm-shifting tool in healthcare, capable of simulating human cognitive functions for lesion classification. **DermAI** addresses the need for a comprehensive, multi-class diagnostic platform. Unlike models limited to a few common conditions, DermAI classifies 31 categories, providing severity indicators (e.g., "Critical" or "Low") to guide user urgency.

Skin diseases are among the most common human illnesses, ranging from benign fungal infections to highly aggressive malignancies like melanoma. The traditional diagnostic workflow relies heavily on visual inspection, which can be subjective and prone to human error, especially in primary care settings where specialized dermatological expertise may be unavailable. Artificial Intelligence (AI) and deep learning have provided substantial breakthroughs in medical imaging.

However, a highly accurate model is clinically useless if it cannot be accessed easily by practitioners or patients. **DermAI** bridges this gap by encapsulating state-of-the-art AI within a seamless, user-friendly web interface. By leveraging a modern tech stack (HTML/CSS for the frontend, Python/Flask for the backend), DermAI allows users to securely upload localized skin images and receive immediate, categorized feedback regarding the nature and severity of the condition.

II. Related Work

Historically, automated skin lesion analysis relied on Convolutional Neural Networks (CNNs) such as ResNet, VGG-16, and EfficientNet. While effective, CNNs are inherently limited by their localized receptive fields, sometimes struggling to capture the global context of complex, irregularly shaped skin lesions.

Furthermore, many existing solutions in the literature focus purely on the deep learning architecture, neglecting the deployment and integration phases. Studies showcasing deployed models often rely on heavy, platform-specific desktop

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 5, May 2026
DOI: 10.17148/IJARCCE.2026.155122

applications. The shift towards lightweight web frameworks (like Flask or FastAPI) coupled with powerful backend inference engines has become a focal point for modern medical informatics, allowing for broader accessibility across diverse devices.

III. LITERATURE SURVEY

The introduction of Vision Transformers (ViTs) disrupted the computer vision landscape by applying the self-attention mechanisms of Natural Language Processing to image patches. Recent literature emphasizes the superiority of self-supervised foundation models, particularly Meta's **DINOv2**. Unlike models trained purely on supervised datasets, DINOv2 learns robust, generalized visual features through self-distillation, which translates exceptionally well to fine-grained medical classification tasks. Current research indicates that ViT architectures consistently outperform CNNs in multi-class dermatological datasets, providing sharper feature embeddings for visually similar conditions (e.g., distinguishing between different types of carcinomas).

IV. PROPOSED SYSTEM

The DermAI platform is a cohesive integration of frontend user experience and backend deep learning processing. The architecture is divided into three primary layers:

A. Presentation Layer (Frontend)

The user interface is built utilizing standard **HTML** and modern **CSS**, resulting in a sleek, dark-themed, and responsive web application. As observed in the system interface, the design prioritizes user flow with a prominent drag-and-drop zone for image uploads. The interface includes dynamic controls, allowing the user to select the number of top predictions to display (e.g., 3, 5, 7, or 10).

B. Application Layer (Backend)

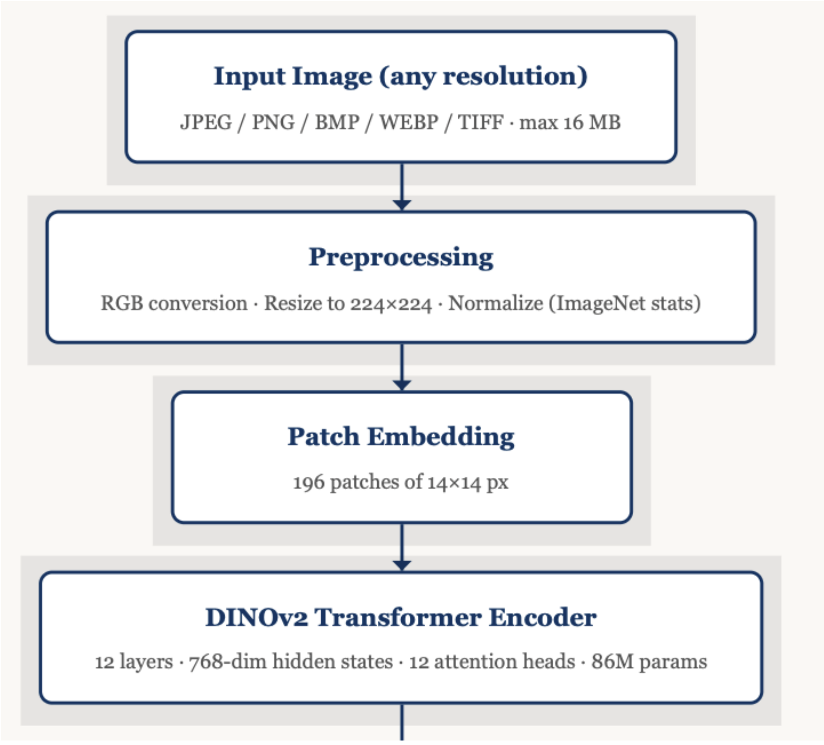
The backend logic and routing are powered by **Python** and the **Flask v3.x** micro-framework. Flask acts as the bridge between the user's web browser and the machine learning model. It handles the asynchronous reception of image files (supporting JPG, PNG, BMP, WEBP up to 16 MB), secures the file payload, and passes it to the data processing pipeline before returning the JSON-formatted predictions back to the frontend for rendering.

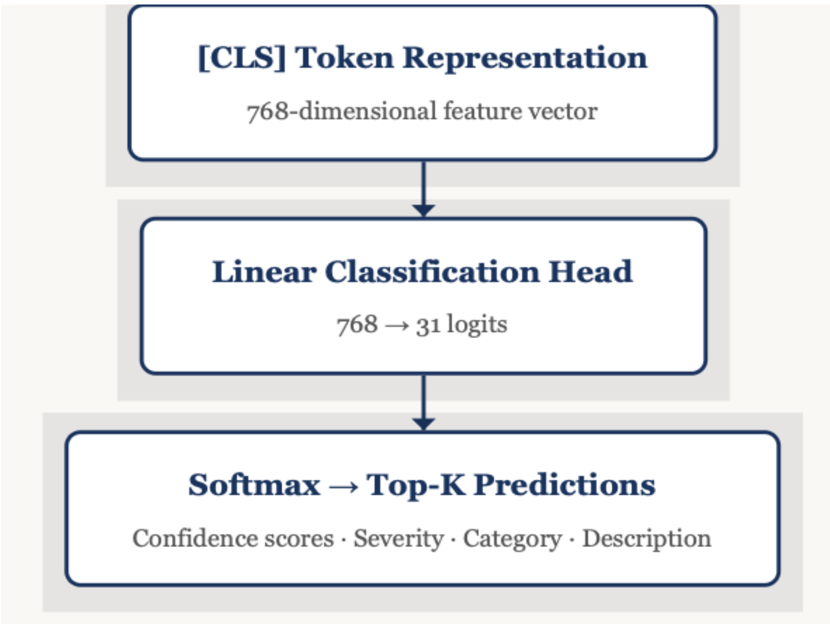
C. AI/Inference Layer

The core intelligence of DermAI is a fine-tuned **DINOv2-Base (ViT-B/14)** architecture. The pipeline consists of:

1. **Image Input:** The uploaded image is resized and normalized to 224×224 pixels.
2. **Patch Embedding:** The image is divided into 196 distinct patches.
3. **Transformer Backbone:** The patches are processed through 12 Transformer Layers, mapping 768-dimensional features.
4. **Classification Head:** A linear layer processes the features into 31 distinct skin disease classes, including *Basal Cell Carcinoma*, *Impetigo*, *Lupus Erythematosus*, and *Psoriasis*.

ISSN (O) 2278-1021, ISSN (P) 2319-5940
International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ∺ Peer-reviewed & Refereed journal ∺ Vol. 15, Issue 5, May 2026
DOI: 10.17148/IJARCCE.2026.155122





V. METHODOLOGY AND VALIDATION

A. Data Pipeline and Tech Stack

To ensure high-performance inference, the system utilizes **PyTorch v2.10.0** and **Transformers v5.3.0**. Image tensor manipulation is handled by **Pillow v12.x** and **NumPy v2.x**, while evaluation metrics are calculated using **scikit-learn v1.8.0**.

B. Training Hyperparameters

The model was fine-tuned for 10 epochs using the following configuration to ensure stable convergence without overfitting:

- **Optimizer:** AdamW (with $\beta_1=0.9$, $\beta_2=0.999$, $\epsilon=1 \times 10^{-8}$)

ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 5, May 2026

DOI: 10.17148/IJARCCE.2026.155122

- **Learning Rate:** 5×10^{-5} with a Linear LR Scheduler and a 10% Warmup Ratio.
- **Batch Sizing:** Train Batch Size of 32, with 4 Gradient Accumulation Steps, yielding an Effective Batch Size of 128.

C. Validation Results

The training phase demonstrated exceptional learning capability. Over the course of 2820 steps, the validation loss steadily decreased from 0.6866 (Epoch 1) to 0.1321 (Epoch 10). Concurrently, the validation accuracy rose from 78.11% to a final, production-ready accuracy of **95.57%**.

VI. RESULTS

The practical efficacy of the Flask-based DermAI system is demonstrated through its real-time inference capabilities on user-uploaded imagery.

1. Malignant Melanoma Detection workflow: The system's file handling capability seamlessly interfaces with local directories, allowing the user to select local files (e.g., Melanoma.jpg). Upon analyzing the lesion, the model successfully identified the image as **Melanoma** with a high confidence score of **79.51%**. Crucially, the HTML/CSS frontend dynamically rendered the severity tags, immediately flagging the prediction as **Malignant** and **Critical**. Secondary differential predictions included *Impetigo* (13.52%, Bacterial, Low Risk) and *Basal Cell Carcinoma* (6.19%, Malignant, High Risk).

2. Benign Fungal Infection Identification: To test the model's certainty on clear-cut non-malignant conditions, an image of *Tinea Nigra* was uploaded (Tinea Nigra Fungal.jpg). The ViT model classified the image with an absolute confidence of **100.0%** for Tinea Nigra. The Flask backend accurately relayed the metadata to the frontend, which categorized it with the tags **Fungal** and **Low** risk, effectively ruling out parasitic or viral infections (e.g., Larva Migrans or Herpes Simplex scored at 0.0%).

The practical efficacy of the Flask-based DermAI system is demonstrated through its real-time inference capabilities on user-uploaded imagery.

1. Malignant Melanoma Detection workflow: The system's file handling capability seamlessly interfaces with local directories, allowing the user to select local files (e.g., Melanoma.jpg). Upon analyzing the lesion, the model successfully identified the image as **Melanoma** with a high confidence score of **79.51%**. Crucially, the HTML/CSS frontend dynamically rendered the severity tags, immediately flagging the prediction as **Malignant** and **Critical**. Secondary differential predictions included *Impetigo* (13.52%, Bacterial, Low Risk) and *Basal Cell Carcinoma* (6.19%, Malignant, High Risk).

2. Benign Fungal Infection Identification: To test the model's certainty on clear-cut non-malignant conditions, an image of *Tinea Nigra* was uploaded (Tinea Nigra Fungal.jpg). The ViT model classified the image with an absolute confidence of **100.0%** for Tinea Nigra. The Flask backend accurately relayed the metadata to the frontend, which categorized it with the tags **Fungal** and **Low** risk, effectively ruling out parasitic or viral infections (e.g., Larva Migrans or Herpes Simplex scored at 0.0%).

The practical efficacy of the Flask-based DermAI system is demonstrated through its real-time inference capabilities on user-uploaded imagery.

1. Malignant Melanoma Detection workflow: The system's file handling capability seamlessly interfaces with local directories, allowing the user to select local files (e.g., Melanoma.jpg). Upon analyzing the lesion, the model successfully identified the image as **Melanoma** with a high confidence score of **79.51%**. Crucially, the HTML/CSS frontend dynamically rendered the severity tags, immediately flagging the prediction as **Malignant** and **Critical**. Secondary differential predictions included *Impetigo* (13.52%, Bacterial, Low Risk) and *Basal Cell Carcinoma* (6.19%, Malignant, High Risk).

2. Benign Fungal Infection Identification: To test the model's certainty on clear-cut non-malignant conditions, an image of *Tinea Nigra* was uploaded (Tinea Nigra Fungal.jpg). The ViT model classified the image with an absolute confidence of **100.0%** for Tinea Nigra. The Flask backend accurately relayed the metadata to the frontend, which categorized it with

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

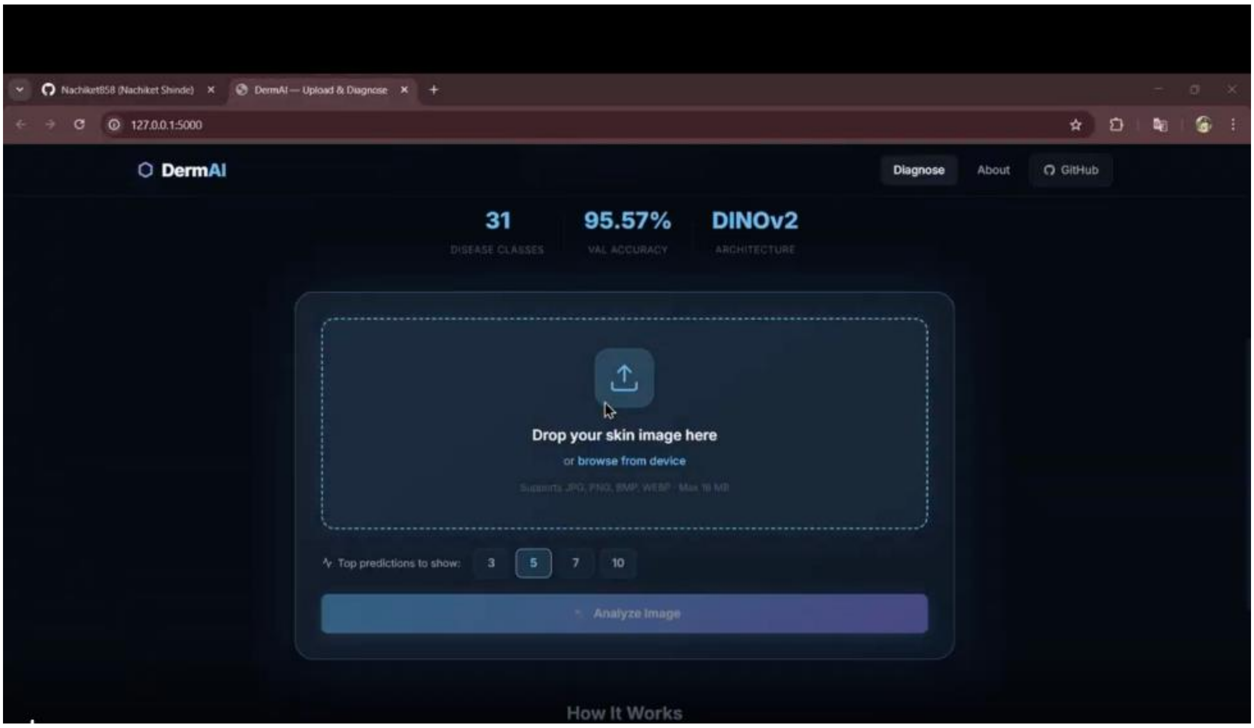
International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁙ Peer-reviewed & Refereed journal ⁙ Vol. 15, Issue 5, May 2026

DOI: 10.17148/IJARCCE.2026.155122

the tags **Fungal** and **Low** risk, effectively ruling out parasitic or viral infections (e.g., Larva Migrans or Herpes Simplex scored at 0.0%).

OUTPUTS :

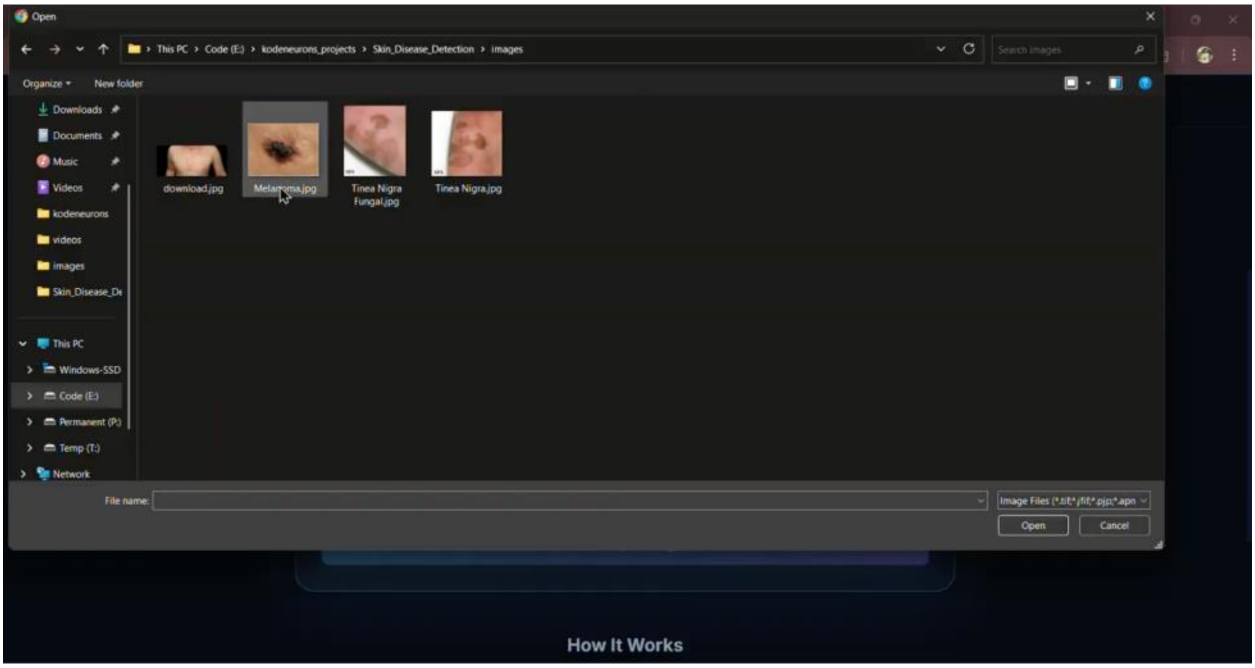


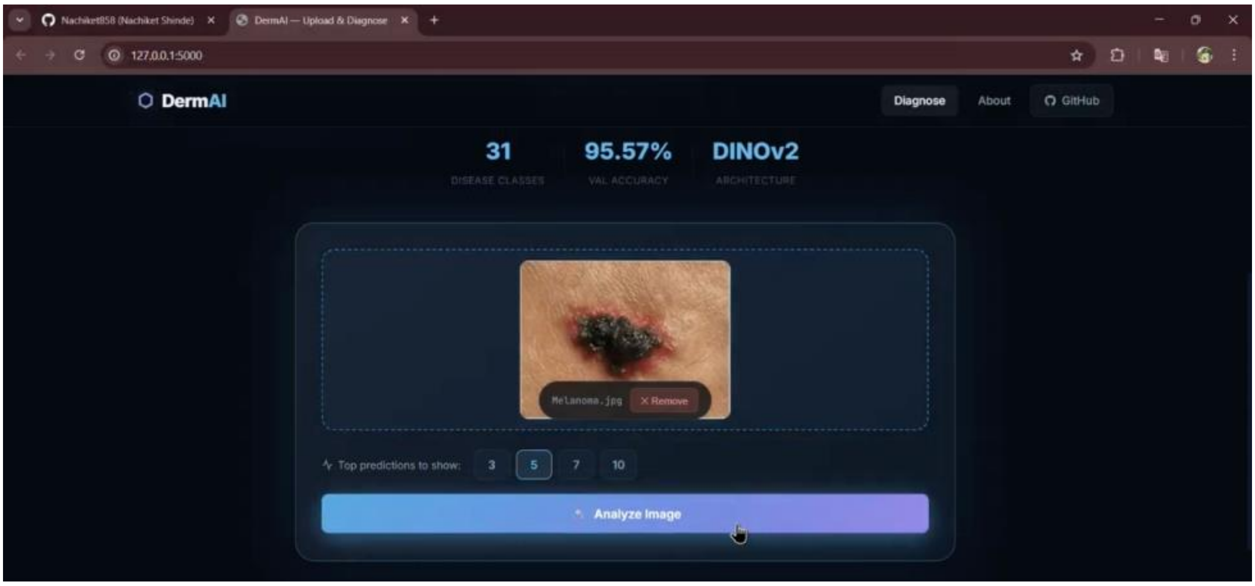
IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

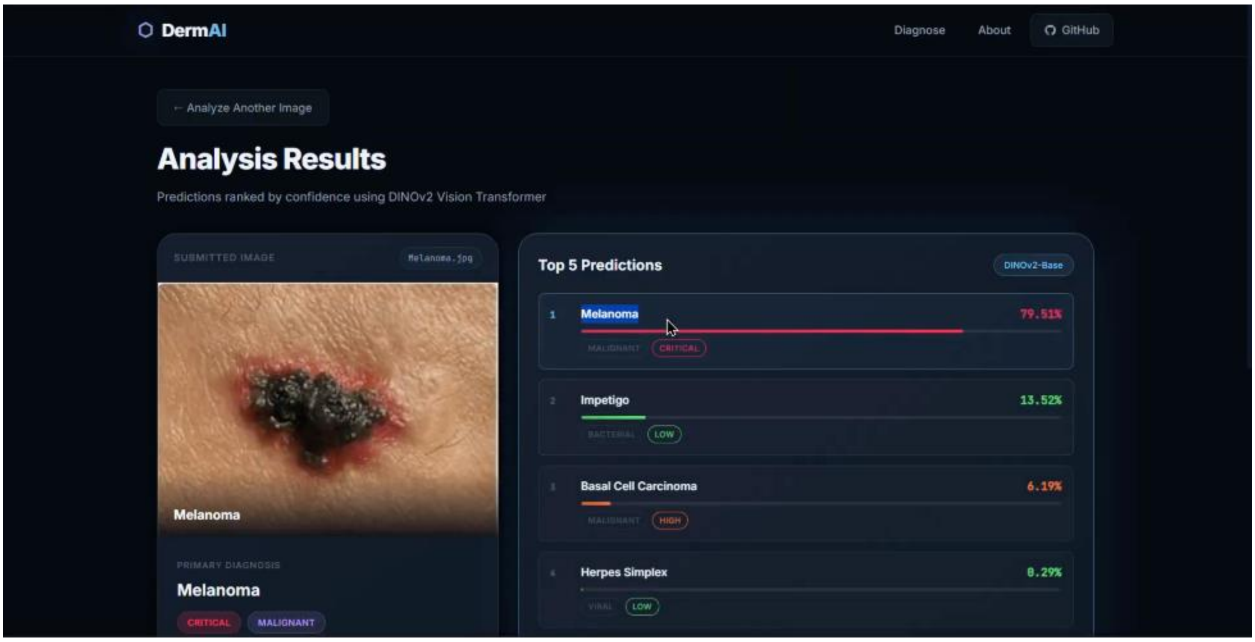
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 5, May 2026

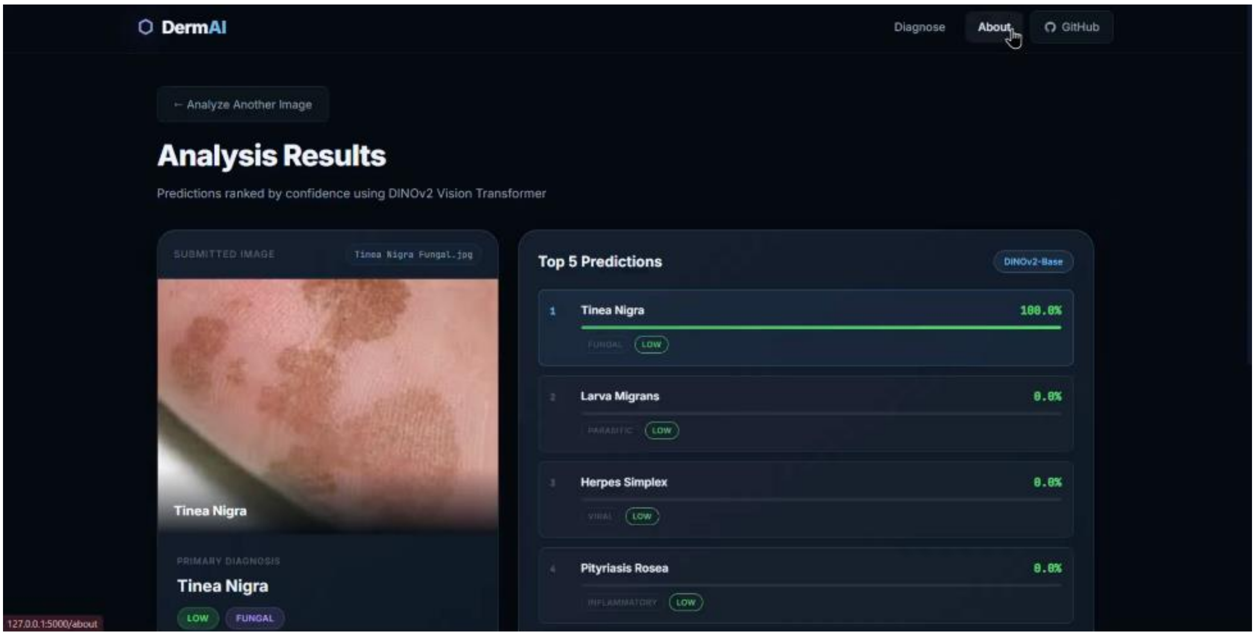
DOI: 10.17148/IJARCCE.2026.155122





International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁑ Peer-reviewed & Refereed journal ⁑ Vol. 15, Issue 5, May 2026
DOI: 10.17148/IJARCCE.2026.155122



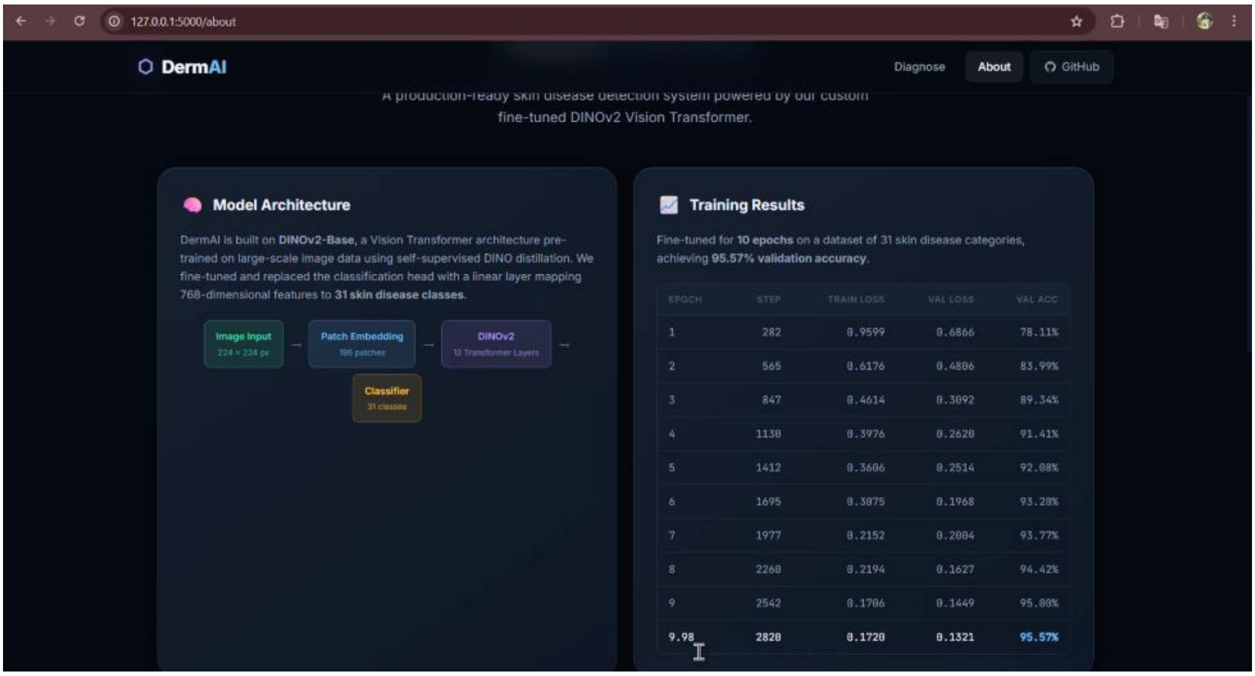


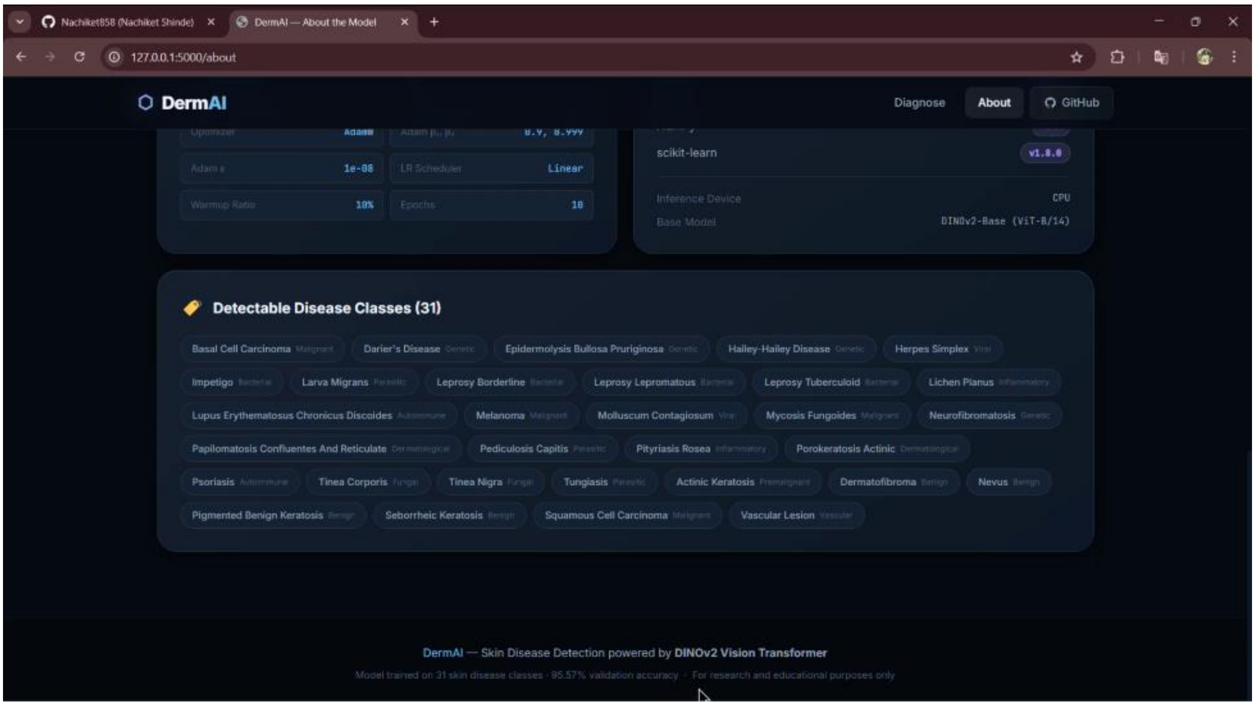
IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

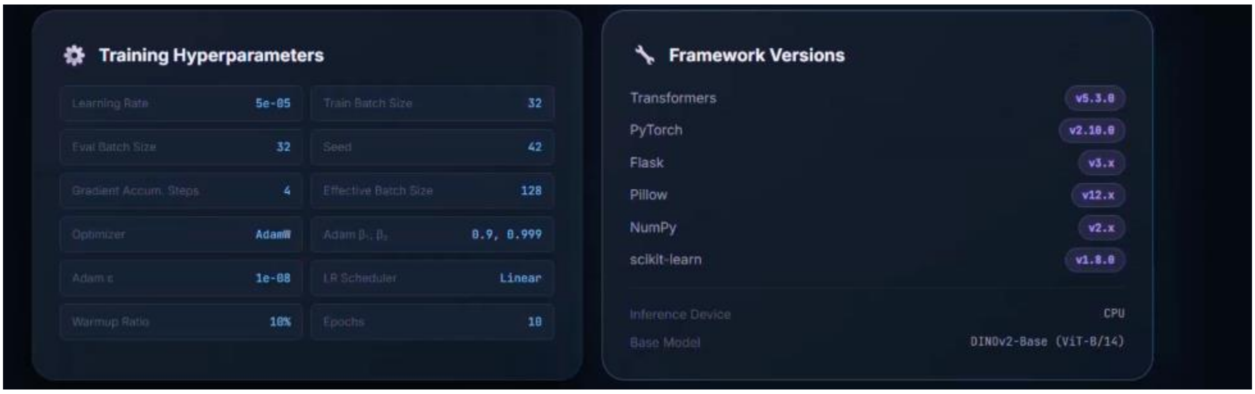
Impact Factor 8.471 ⁒ Peer-reviewed & Refereed journal ⁒ Vol. 15, Issue 5, May 2026

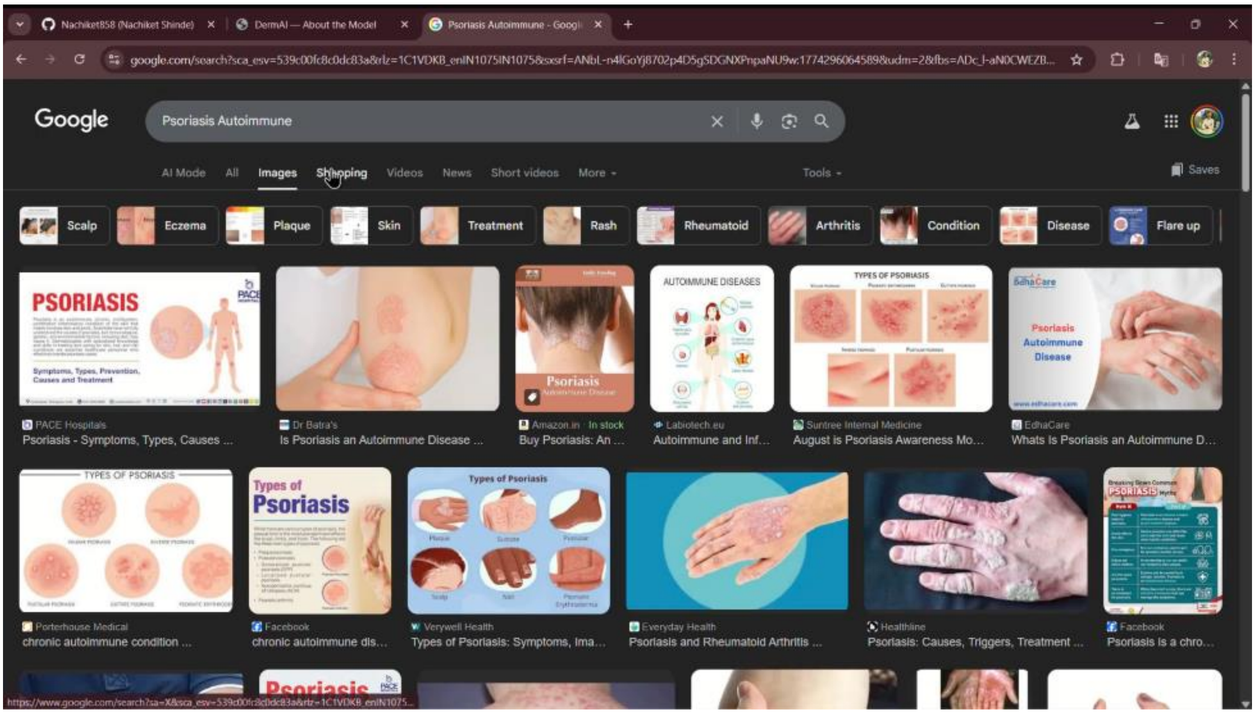
DOI: 10.17148/IJARCCE.2026.155122





IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940
International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 5, May 2026
DOI: 10.17148/IJARCCE.2026.155122



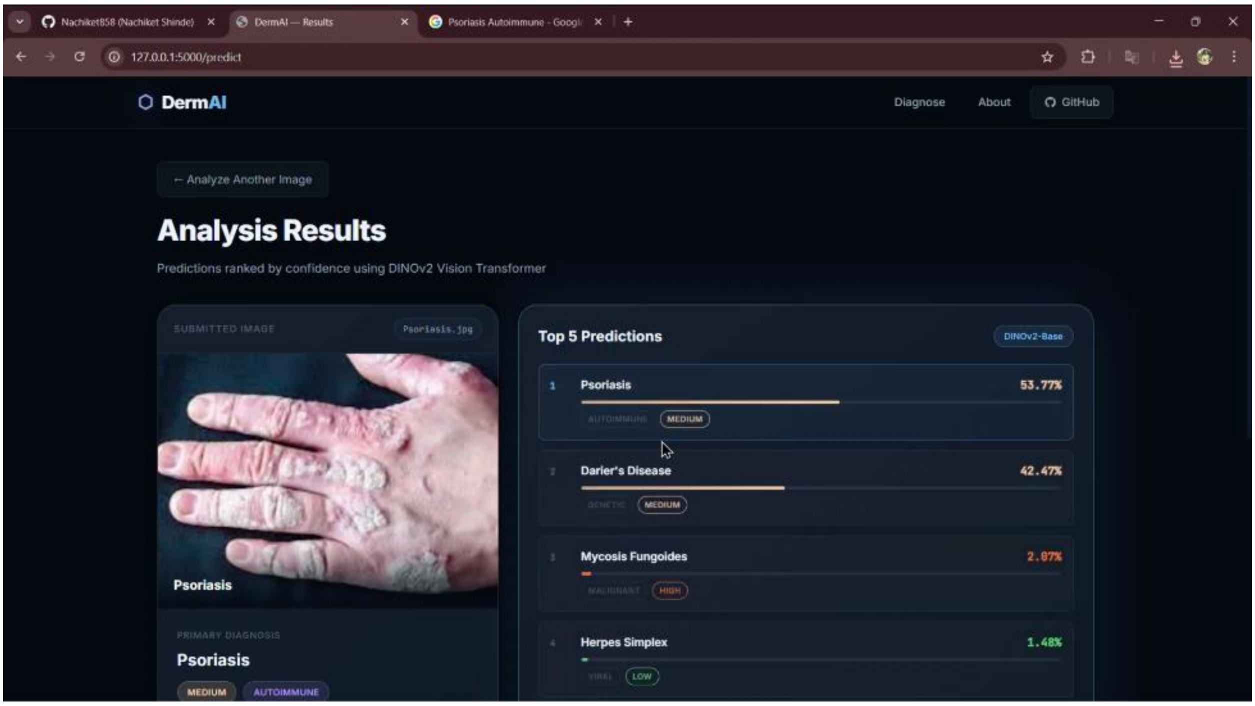


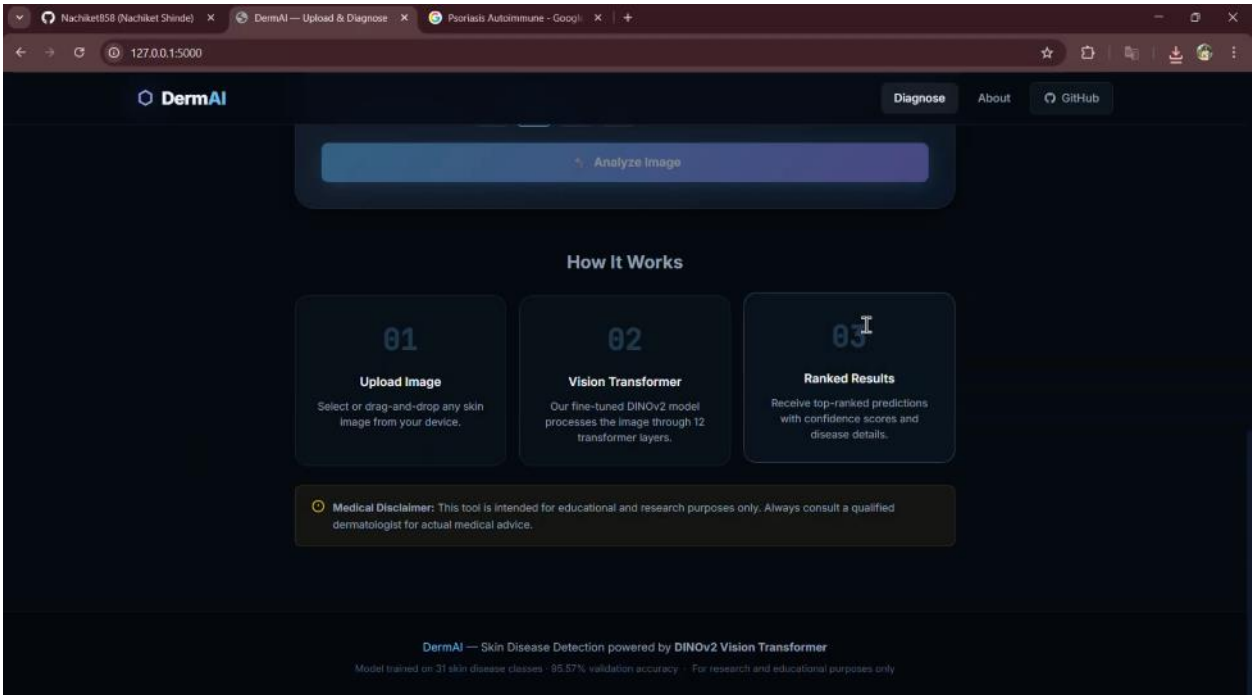
IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁙ Peer-reviewed & Refereed journal ⁙ Vol. 15, Issue 5, May 2026

DOI: 10.17148/IJARCCE.2026.155122





VII. CONCLUSION AND FUTURE SCOPE

The **DermAI** project successfully illustrates the powerful synergy between advanced self-supervised Vision Transformers (DINOv2) and accessible web technologies (Python, Flask, HTML, CSS). By achieving a 95.57% validation accuracy across a diverse set of 31 disease classes, the system proves that foundational models can be fine-tuned for highly specialized medical tasks and deployed via lightweight web frameworks for real-world use on standard CPU inference devices.

Future Scope:

1. **Mobile-First Adaptation:** Transitioning the HTML/CSS frontend into a Progressive Web App (PWA) or Flutter application for offline capabilities and direct smartphone camera integration.

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 5, May 2026
DOI: 10.17148/IJARCCE.2026.155122

2. **Dataset Diversification:** Expanding the training dataset to include a wider spectrum of Fitzpatrick skin types to eliminate diagnostic bias.
3. **Advanced Security:** Implementing rigorous user authentication and VAPT (Vulnerability Assessment and Penetration Testing) protocols to secure sensitive patient image data in compliance with medical data regulations.

REFERENCES

[1]Meta AI Research. (2023). *DINOv2: Learning Robust Visual Features without Supervision*.
[2]Dosovitskiy, A., et al. (2020). *An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale*. ICLR.
[3] Grinberg, M. (2018). *Flask Web Development: Developing Web Applications with Python*. O'Reilly Media.
[4]Paszke, A., et al. (2019). *PyTorch: An Imperative Style, High-Performance Deep Learning Library*. NeurIPS.
[5] A, J., et al. (2024). DEEPSCAN: Integrating Vision Transformers for Advanced Skin Lesion Diagnostics. *The Open Dermatology Journal*.