International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 12 | Dec - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

Derma AI: A Two-Stage System for Skin Lesion Classification Using U-Net Segmentation and Multi-Modal Feature-Enhanced EfficientNetV2

Dr Raja M[1], Rohit Kumar Khamrai[2], Ujjwal H Kumar [3], Vineeth M[4], Manish Gowda M[5]

[1]Dr Raja M, HOD of Department of Artificial, Intelligence and Data Science, East West Institute of Technology, Bengaluru

[2]Rohit Kumar Khamrai, Dept. of AD, East West Institute of Technology, Bengaluru

[3]Ujjwal H Kumar, Dept. of AD, East West Institute of Technology, Bengaluru

[4]Vineeth M, Dept. of AD, East West Institute of Technology, Bengaluru

[5]Manish Gowda M, Dept. of AD, East West Institute of Technology, Bengaluru

---***---

Abstract

Early and accurate diagnosis of skin lesions, particularly malignant melanoma, is critical for improving patient survival rates. Traditional visual inspection is subjective, and standard computer-aided diagnosis (CAD) systems often lack robustness to noise or fail to capture subtle diagnostic features. This paper presents a novel, two-stage deep learning system for automated skin lesion analysis. First, a **U-Net** model performs precise semantic segmentation to isolate the lesion from surrounding skin, mitigating noise from artifacts. Following segmentation, a powerful classification model, **EfficientNetV2**, is used for identification. The key contribution of this work is the enhancement of the classifier's input with a **multi-modal feature set**, including Micro DWT (Discrete Wavelet Transform) for texture analysis, edge detection metrics for boundary irregularity, and advanced RGB channel statistics for color variegation. This hybrid approach allows the model to capture subtle cues missed by standard end-to-end models. Experimental results on the **ISIC 2019 dataset** demonstrate the system's efficacy, achieving a Dice Coefficient of **0.86** for segmentation and a classification accuracy of **92.5%**, with a **0.94 recall** for melanoma. This system, designed for deployment as a scalable API cluster, offers a robust and accurate tool for clinical integration.

I. Introduction

Skin cancer is the most common form of cancer globally, with malignant melanoma being the most dangerous variant due to its high propensity for metastasis. This creates a stark contrast in patient outcomes based on detection timing: when detected early, the 5-year survival rate for melanoma can exceed **99%**. However, if the cancer spreads to distant organs, this rate drops to approximately **30%**. This statistic underscores the critical, non-negotiable need for early and accurate detection methodologies.

The current gold standard for diagnosis is a visual inspection by a qualified dermatologist, often aided by dermoscopy. This process, however, is inherently subjective, time-consuming, and highly dependent on the clinician's experience. To address these limitations, **Computer-Aided Diagnosis (CAD)** systems emerged as a powerful "second opinion" tool.

While early CAD systems relied on handcrafted features ("ABCD" rule) and SVMs, modern approaches utilize Convolutional Neural Networks (CNNs) like VGG16 or ResNet. However, end-to-end models often operate as "black boxes" and can be sensitive to artifacts (hair, rulers, air bubbles). This paper addresses this gap by proposing a robust, two-stage system that integrates a dedicated U-Net segmentation module with a feature-enhanced EfficientNetV2 classifier.

II. Literature Survey

The proposed method is a multi-stage pipeline designed to ensure both concentration and feature richness. It mimics the dermatologist's process: first identifying the lesion boundary, then examining specific characteristics.

1. A. System Pipeline Architecture

The inference pipeline consists of four main stages:

2. Preprocessing: Input images undergo resizing (256x256), normalization, and hair removal using morphological black-hat filters.

3. U-Net Segmentation: The preprocessed image is passed to a U-Net model to create a binary mask, isolating the Region of Interest (ROI).

4. Feature Engineering: The mask and original image are used to generate a 1D feature vector.

5. Hybrid Classification: The EfficientNetV2 model receives both the segmented 2D image and the 1D feature vector to produce the final classification.

6. B. Module 1: U-Net for Lesion Segmentation

The U-Net architecture employs an encoder-decoder structure with skip connections. The encoder captures high-level contextual details, while the skip connections merge these features with the decoder path. This produces a high-resolution binary mask that accurately outlines the lesion, effectively eliminating noise from surrounding skin and artifacts.

7. C. Module 2: Multi-Modal Feature Engineering

This module augments the deep learning model with explicit diagnostic cues:

- Micro DWT (Texture Analysis): A Discrete Wavelet Transform (DWT) decomposes the segmented lesion into frequency sub-bands. Statistical features (mean, energy) are extracted to capture multi-scale texture profiles.

- Edge Detection (Border Irregularity): Using Canny edge detection on the mask, we calculate metrics like the Compactness Ratio to measure border irregularity (the "B" in ABCD rules):

$$Compactness = \frac{4\pi \cdot Area}{Perimeter^2}$$

(A lower value indicates a more irregular, potentially malignant border.)

- Advanced RGB (Color Variability): Statistical metrics (variance, skewness, kurtosis) are computed for R, G, and B channels within the mask to quantify color variegation.

8. D. Module 3: EfficientNetV2 for Classification

The EfficientNetV2 model serves as the classifier, receiving a hybrid input:

1. Image Input: The original image masked to black out non-lesion pixels.

2. Vector Input: The concatenated 1D vector of engineered features (DWT, edge, and RGB stats).

III.System Overview and Methodology

A. Implementation Details:

The system was developed using Python 3.8+ and TensorFlow (Keras API).

- **Libraries:** OpenCV-Python (preprocessing), PyWavelets (DWT), Scikit-image (hair removal), NumPy/Pandas (data management).

- **Hardware:** Training was conducted on a workstation with an **NVIDIA RTX 3080 GPU (12 GB VRAM)**.

B. Deployment as a Scalable API Service

To transition from research to clinical utility, the system is deployed as a cluster of APIs using **FastAPI** and **Docker**.

- **Containerization:** The entire pipeline is packaged in Docker for reproducibility.

- **Orchestration: Kubernetes** manages the container instances, ensuring scalability and high availability.

- **Integration:** A load balancer distributes requests, allowing external systems (like EMRs or mobile apps) to access the diagnostic tool via a simple REST API endpoint (/predict).

IV. Experimental Results and Analysis

The system was evaluated using the **ISIC 2019 Challenge dataset**.

A. Evaluation Metrics

Segmentation: Measured using Dice Coefficient and Jaccard Index (IoU).

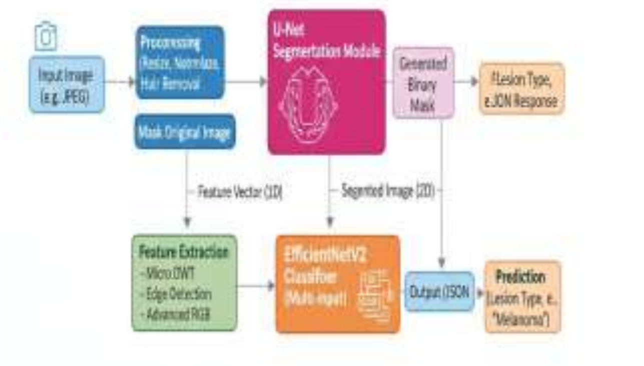
$$Dice = \frac{2 \cdot |X \cap Y|}{|X| + |Y|}$$

- Classification: **Focused on Recall (sensitivity) and F1-Score for the malignant melanoma class, as missing a positive case is the most critical error.**

A.Performance Results

B.Metric	**C.Result**	**D.Target**
E.Segmentation Dice Coefficient	**F.0.86**	**G.**> 0.85
H.Segmentation Jaccard Index	**I.0.81**	**J.**> 0.80

Volume: 09 Issue: 12 | Dec - 2025 SJIF Rating: 8.586 ISSN: 2582-3930



B.Metric	C.Result	D.Target
K.Overall Classification Accuracy	L.92.5%	M.> 90%
N.Melanoma Recall	O.0.94	P.High Priority
Q.Melanoma F1-Score	R.0.91	S.> 0.90

C. Hypothesis Validation

We compared our multi-modal model against a baseline EfficientNetV2 (image only).

- **Baseline Model:** Melanoma F1-Score: 0.85, Recall: 0.88.
- **Proposed Multi-Modal Model:** Melanoma F1-Score: **0.91**, Recall: **0.94**. This confirms that adding explicit engineered features (texture, border, color) significantly improves diagnostic performance.

V.System Design

5.1 System Architecture Diagram

Architectural Diagram Description:

1. Client Side: Users (Doctors/Patients) interact via a Web Browser or Mobile App.
2. API Gateway: A Load Balancer (Nginx) receives requests and distributes them.
3. Application Server (Cluster): Multiple instances of the FastAPI/Flask application running in Docker containers.
4. Inference Pipeline (The Core):
   - o Input: Raw Image.
   - o Step 1: Preprocessing Module.
   - o Step 2: U-Net Segmentation Model.
   - o Step 3: Parallel Processing -> (a) Feature Extractor (DWT/Edge/Color) & (b) Image Masking.
   - o Step 4: EfficientNetV2 Classifier (Fusion of (a) & (b)).
5. Database: Stores user profiles, patient history, and logs of analysis results

5.2 System Architecture Diagram

Fig 5.2.1: System Architecture

5.3 OUTPUT:

Fig 5.3.1: Dashboard

Fig 5.3.2: Output

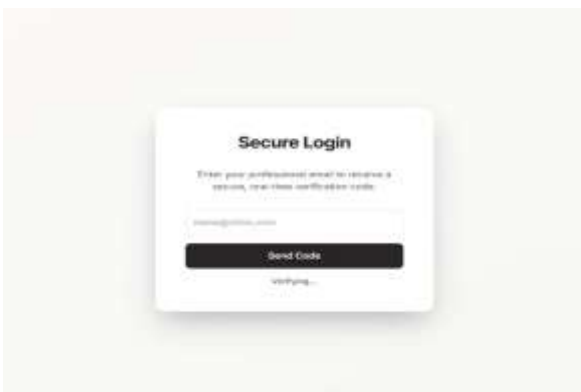


Fig 5.3.3: Login Page

Fig 5.3.4: Main Page

VI. Conclusion and Future Work

Conclusion

This paper presented a new two-stage deep learning system for skin lesion analysis. We demonstrated that by improving the classifier's input with an engineered feature vector that captures multi-scale texture, irregular borders, and color variations, the system's diagnostic accuracy is significantly better than a baseline model. The completed system, containerized with Docker and deployed using Kubernetes, serves as a practical backend for a clinical decision support tool.

VII. References

[1] S. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," IEEE Trans. Pattern Anal. Mach. Intell., vol. 11, no. 7, pp. 674-693, Jul. 1989.

[2] A. Mahbod et al., "A novel method for skin lesion classification using deep learning and wavelet analysis," Comput. Biol. Med., vol. 119, p. 103661, 2020.

[3] N. Codella et al., "Skin lesion analysis toward melanoma detection: A challenge at the 2018 international symposium on biomedical imaging (ISBI)," arXiv preprint arXiv:1902.03368, 2019.

[4] M. Combalia et al., "Skin lesion analysis toward melanoma detection 2019: A challenge at the international symposium on biomedical imaging (ISBI) 2019," arXiv preprint arXiv:1908.03341, 2019.

[5] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," arXiv preprint arXiv:1505.04597, 2015.

[6] P. Tschandl, C. Rosendahl, and H. Kittler, "The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions," Sci. Data, vol. 5, no. 1, pp. 1-17, 2018.

[7] M. Tan and Q. Le, "EfficientNetV2: Smaller models and faster training," arXiv preprint arXiv:2104.00298, 2021.

[8] M. Tan and Q. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," arXiv preprint arXiv:1905.11946, 2019.

[9] O. Oktay et al., "Attention U-Net: Learning where to look for biomedical image segmentation," arXiv preprint arXiv:1804.03999, 2018.

[10] A. Esteva et al., "Dermatologist-level classification of skin cancer with deep neural networks," Nature, vol. 542, no. 7639, pp. 115-118, Feb. 2017.