**Skin Disease Detection at the Edge Using OpenVINO**

**Author: Your Name**

Affiliation: Your Institution, City, Country

Date: September 03, 2025

**Abstract**

This paper presents a practical system for skin disease classification that runs entirely on CPU using Intel's OpenVINO toolkit for efficient edge inference. The application ingests a dermatoscopic image, performs preprocessing, executes a lightweight convolutional network compiled to OpenVINO, and returns the predicted disease along with an interpretable occlusion-sensitivity heatmap and a clinical severity assessment. A classical K-Nearest Neighbors (KNN) fallback provides resilience when the IR model is missing or incompatible. The system emphasizes latency and deployability over dataset-level accuracy, achieving sub-second inference on commodity hardware.

**Index Terms—**

OpenVINO, edge AI, dermatology, skin lesion analysis, explainable AI, KNN.

**I. Introduction**

Skin cancer and other dermatological conditions represent a major global health burden. Early triage and risk assessment can improve outcomes via timely specialist referral. Recent advances in deep learning have enabled computer-assisted diagnosis; however, compute and memory constraints often limit deployment at the edge. We implement an end-to-end system that performs low-latency inference on CPU via OpenVINO [1], deliverable as a Flask web application. The solution also integrates interpretable heatmaps based on occlusion sensitivity [2] and a clinically informed severity score to aid decision support.

**II. Related Work**

OpenVINO has been widely adopted to optimize and deploy DNNs for CPU and heterogeneous devices [1]. Explainability in medical imaging often leverages saliency methods such as occlusion maps [2] and Grad-CAM [3]. Dermatology datasets such as ISIC facilitate benchmarking skin lesion classifiers [4]. Classical methods, including KNN with color-texture features, remain useful baselines and fallbacks [5].

**III. System Overview**

The system is implemented as a Flask application that exposes a simple web UI for image upload. On submit, the server reads the image and invokes the OpenVINO runtime to execute the IR model (XML+BIN). If model loading fails due to legacy IR versions, a KNN fallback computes per-class scores from color histograms and simple gradient features. For interpretability, an occlusion-sensitivity heatmap perturbs local regions to estimate their contribution to the predicted class. Additionally, a rule-based clinical severity score summarizes lesion characteristics (area ratio, border irregularity, color heterogeneity, asymmetry, edge density, texture complexity).

**IV. Methods**

**A. Data and Classes**

The deployment includes 9 classes, e.g., Actinic keratosis, Atopic Dermatitis, Benign keratosis, Dermatofibroma, Melanocytic nevus, Melanoma....

**B. Preprocessing**

Images are resized to the network's input resolution, channels are ordered to match the model, and a batch dimension is added prior to inference. For KNN, features include per-channel color histograms and a simple gradient-magnitude histogram with L2 normalization.

**C. OpenVINO Inference**

The Intel OpenVINO runtime reads the IR graph and compiles it for CPU execution. Inference outputs are converted to probabilities using a numerically stable softmax. The top-1 class and its probability (0–1) are reported as the predicted label and accuracy, respectively.

**D. KNN Fallback**

If the IR model cannot be loaded (e.g., legacy IR version), a KNN fallback computes per-class scores as inverse-distance accumulations over the k nearest neighbors. The scores are transformed with softmax to obtain a proper probability distribution, and the top-1 probability is presented as accuracy (0–1).

**E. Explainability via Occlusion Sensitivity**

We adopt occlusion sensitivity [2]: sliding a patch over the image while re-scoring the model, quantifying how occluding a region lowers the target-class score. The resulting map is normalized to [0, 1] and overlaid on the image for transparency.

**F. Clinical Severity Scoring**

We define a lightweight severity score inspired by ABCD criteria, aggregating area ratio, border irregularity, color heterogeneity, asymmetry, edge density, and texture complexity into a 0–1 score with heuristic thresholds. The score is mapped to Low/Moderate/High/Very High levels accompanied by actionable advice.

**V. Experiments**

Latency was measured on a commodity CPU using the OpenVINO path, yielding sub-second end-to-end response time (upload + preprocess + inference + rendering). As the repository prioritizes deployment, we report per-image confidence rather than dataset-level accuracy; the README explicitly notes the model is not tuned for high accuracy. Qualitative examples demonstrate that the occlusion heatmaps highlight salient lesion regions.

**VI. Discussion**

The system balances deployability, interpretability, and responsiveness. OpenVINO ensures efficient CPU inference, while the KNN fallback provides robustness. The clinical severity score and heatmaps can support triage, but they do not replace clinical judgment. Future work includes stronger backbones, better calibration, and external validation on diverse datasets.

**VII. Limitations and Ethical Considerations**

This system is intended for educational and assistive purposes, not as a diagnostic device. Dataset bias, image quality, and domain shift can affect predictions. Any deployment must ensure privacy, informed consent, and appropriate regulatory review.

**VIII. Conclusion**

We presented an edge-ready skin lesion analysis tool leveraging OpenVINO for fast CPU inference, with a classical KNN fallback, interpretable occlusion heatmaps, and a clinical severity score. The design demonstrates how to deliver responsive, explainable, and practical AI capabilities within constrained environments.

**References**

[1] OpenVINO Toolkit Documentation, Intel, 2025. [Online]. Available: https://docs.openvino.ai/

[2] M. D. Zeiler and R. Fergus, 'Visualizing and Understanding Convolutional Networks,' in ECCV, 2014, pp. 818–833.

[3] R. R. Selvaraju et al., 'Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization,' in ICCV, 2017, pp. 618–626.

[4] No Authors, 'ISIC Archive: International Skin Imaging Collaboration,' [Online]. Available: https://www.isic-archive.com/

[5] G. Bradski, 'The OpenCV Library,' Dr. Dobb’s Journal of Software Tools, 2000.