IFSCC 2025 full paper (ABSTRACT Nº IFSCC2025-1782)

*“*AI-driven tools for non-invasive skin analysis: A study in detecting lentigines and nevi in human skin*”*

Silva, L[1]; Vieira, P[1]; Pinto, P[2] and Amores da Silva, P[3,4]

[1] Physics Department, NOVA Faculty of Science and Technology, Campus da Caparica, Portugal
[2]PhD Trials, R&D Department, Lisboa, Portugal
[3]ADAT, Lisboa, Portugal
[4]Department of Medical and Health Sciences, School of Health and Human Development, University of Évora, Évora, Portugal

1. Introduction

Skin pigmentation plays a critical role in human appearance and health, reflecting underlying biological and environmental influences. Abnormal pigmentation, particularly hyperpigmentation, manifests as darkened areas of skin caused by excess melanin. Conditions such as lentigines, nevi, melasma, and post-inflammatory hyperpigmentation are among the most common examples and can arise due to aging, sun exposure, inflammation, hormonal factors, or genetic predisposition.

Hyperpigmented lesions on the face are a frequent concern in both dermatological and cosmetic contexts. Although most lesions are benign, they can have substantial psychological and social impact, and in certain cases, may require monitoring for malignant transformation. Reliable identification, classification, and measurement of these lesions are therefore crucial for early diagnosis, treatment planning, and evaluation of therapeutic efficacy.

Conventionally, dermatologists rely on clinical visual inspection and dermatoscopy to assess skin lesions. However, these methods are time-consuming, operator-dependent, and subject to human error. Moreover, the use of standard photographic images (non-dermatoscopic) taken under naturalistic or semi-controlled conditions has increased significantly due to the rise of tele dermatology, remote patient monitoring, and over-the-counter cosmetic product evaluations.

Given these trends, there is growing demand for automated, scalable, and objective tools that can detect and quantify hyperpigmented lesions from facial photographs. The integration of computer vision and artificial intelligence (AI), particularly deep learning, has emerged as a promising strategy to meet this need. Object detection algorithms, especially those based on convolutional neural networks (CNNs), have demonstrated substantial success across various fields, including medical imaging.

This paper presents a comprehensive AI pipeline for the detection and segmentation of facial hyperpigmentation using deep learning object detectors (YOLOv4, Faster R-CNN) coupled with classical segmentation (Otsu's method). Our approach focuses on real-world facial imagery, diverging from the prevalent reliance on dermatoscopic images, and highlights its practical application for both clinical and consumer-oriented dermatological analysis.

2. State of the Art

Automated detection of skin lesions using artificial intelligence has progressed rapidly over the last decade. Traditional approaches in computer vision relied on handcrafted features such as color histograms, edge maps, and texture descriptors combined with classical classifiers like support vector machines (SVMs) or k-nearest neighbors (KNN). However, these techniques generally underperform on heterogeneous and noisy datasets due to limited generalization capabilities.

Deep learning has revolutionized the field by enabling end-to-end training and automatic feature. In particular, convolutional neural networks (CNNs) have demonstrated state-of-the-art performance in image classification, object detection, and semantic segmentation. Object detection methods are especially relevant to dermatology, as they can identify and localize multiple lesions in a single image.

YOLO (You Only Look Once) is a single-stage object detector that frames detection as a regression problem. It divides the image into a grid and predicts bounding boxes and class probabilities directly. Its balance of speed and accuracy makes it particularly attractive for real-world deployment scenarios such as mobile dermatology apps or large-scale clinical screening Faster R-CNN is a two-stage detector that first generates region proposals and then performs classification and bounding box regression on these proposals. Although slower than YOLO, it tends to be more accurate, especially for small or overlapping objects. Its modularity and flexibility make it suitable for high-precision applications such as medical diagnostics.

In dermatological imaging, most AI systems have focused on dermatoscopic datasets to classify melanoma and other skin cancers. Public datasets such as ISIC have driven innovation in segmentation and classification tasks. However, fewer works address the detection of multiple types of hyperpigmented lesions in photographic images of the face, which pose unique challenges related to lighting variability, skin tone differences, and cosmetic artefacts.

In addition to detection, segmentation techniques are often employed to delineate lesion boundaries. Classical thresholding algorithms, such as Otsu's method, are simple yet effective when contrast between lesion and surrounding skin is sufficient. Otsu's method computes an optimal global threshold by maximizing the inter-class variance. Despite its simplicity, it remains a valuable tool for preprocessing and initial analysis in many biomedical applications.

The combination of deep learning-based detection with classical segmentation provides a hybrid framework that benefits from the strengths of both paradigms. Detection localized regions of interest, while segmentation enables pixel-level analysis, supporting metrics like lesion area, perimeter, and shape complexity. This dual approach enhances interpretability and usability in clinical environments.

3. Materials and Methods

This study used two datasets containing high-resolution facial images obtained at PhD Trials®, after informed written consent of the subjects. All photographs were taken using professional DSLR cameras under standardized lighting and pose conditions to ensure consistency. Subjects were positioned at a fixed distance and angle relative to the camera, with hair and accessories removed from the frame to ensure unobstructed cheek exposure.

Dataset 1 consisted of 20 manually annotated facial images from female volunteers, focused on the left cheek area. Each image was labelled with bounding boxes using MATLAB's Image Labeler tool to identify two primary types of pigmentation abnormalities: lentigines and nevi. Annotation was performed by trained researchers and validated by dermatologists to ensure accuracy.

Dataset 2 included 943 images with mixed gender representation and greater lesion diversity. These images were annotated semi-automatically using the best-performing model trained on Dataset 1, followed by manual validation and correction by clinical experts. This two-stage annotation approach enabled the creation of a larger dataset with consistent and clinically relevant labels (Figure 1).

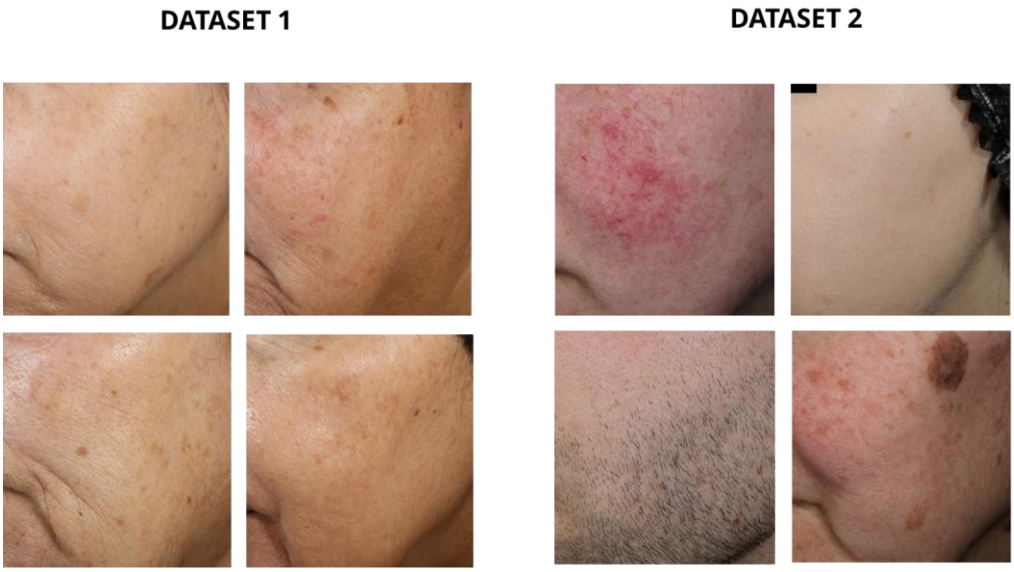

Figure 1. Two diverse datasets were used. Dataset 1 consisted of images from 20 female volunteers while dataset 2 was composed of images from 943 volunteers both genders and greater lesion variability.

All images were preprocessed through cropping to isolate the cheek region. Data augmentation was applied to increase training data variability and improve generalization (Figure 2). The augmentation techniques included horizontal flipping, brightness and contrast adjustments, rotation (±15°), Gaussian blur, and noise addition.

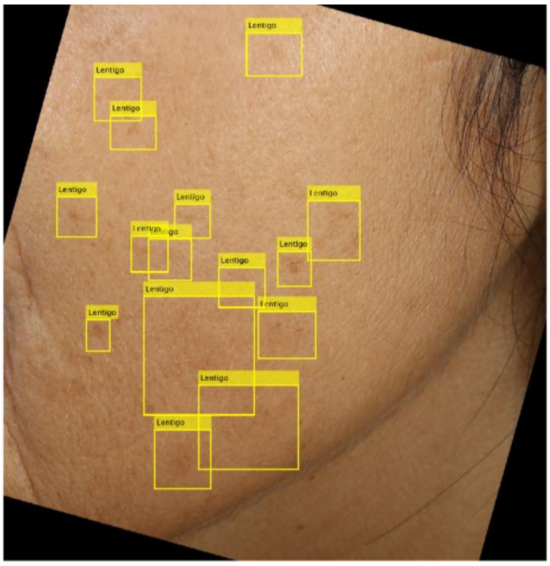

Figure 2. – Data augmentation used Horizontal reflection and Random rotations between -30 and 30 degrees

For object detection, YOLOv2 and YOLOv4 were implemented in MATLAB using transfer learning with pretrained weights on the COCO dataset. Hyperparameters such as batch size (16), learning rate (0.001), and maximum epochs (100) were selected empirically. Faster R-CNN was trained in Detectron2, a PyTorch-based framework, using a ResNet-50 backbone and default configurations optimized for medical imaging tasks.

Model performance was evaluated using mean Average Precision (mAP) at an IoU threshold of 0.5, along with precision, recall, and F1-score per class. Five-fold cross-validation was conducted to ensure robust evaluation, and results were averaged across all folds.

Post-detection, the segmented regions of interest (ROIs) were processed using Otsu's thresholding method to extract binary masks for lesion analysis (Figure 3). Each cropped ROI was converted to grayscale, histogram equalized and filtered using morphological operations such as erosion and dilation. The segmented masks were then analyzed to compute quantitative lesion metrics, including area (in pixels), circularity, and eccentricity (Figure 4).

Figure 3. – Once the Region of interest is defined procedures will be applied to each Region of Interest (RoI) detected by the detection algorithm

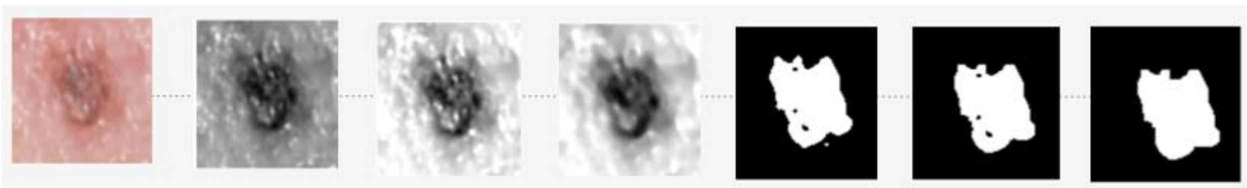

Figure 4. – Procedures are applied to each skin lesion: conversion to greyscale, increase contrast, increase brightness, noise reduction and threshold analysis.

Finally, a graphical user interface (GUI) was developed using the open-source Python framework Streamlit, to enable easy image loading, automatic inference, and visualisation of bounding boxes and segmented masks (Figure 1). The tool was tested with five clinical users for usability, accuracy, and interpretability of results.

4. Results

All three models trained for this study—YOLOv2, YOLOv4, and Faster R-CNN—demonstrated successful learning with convergence observed within the first 60 epochs. YOLOv4 exhibited the best balance between speed and accuracy, while Faster R-CNN achieved the highest precision in distinguishing lesion types.

The performance metrics for each model are summarized in Table 1. Faster R-CNN attained the highest mean Average Precision (mAP) at 0.88, followed closely by YOLOv4 at 0.84. YOLOv2 showed the lowest mAP but maintained real-time inference capability, making it suitable for embedded applications despite its reduced accuracy.

Segmentation analysis using Otsu's method provided accurate lesion masks in cases with strong contrast between lesion and surrounding skin. Morphological refinement of the masks was particularly useful in smoothing boundaries and removing noise. Quantitative analysis of lesion area revealed consistent measurements across multiple test images, supporting the feasibility of longitudinal monitoring.

A time-series case study was conducted on a subset of 10 patients undergoing treatment with depigmentation cream. The masks generated over 4-week intervals demonstrated measurable reduction in lesion area, which correlated with clinical assessments. This result highlights the utility of the system for objective follow-up in clinical trials and cosmetic research.

Table 1 – Detection performance per model (average across 5 folds):

Model	Precision	Recall	F1-Score	mAP (IoU=0.5)
YOLOv2	0.76	0.78	0.77	0.75
YOLOv4	0.85	0.87	0.86	0.84
Faster R-CNN	0.89	0.85	0.87	0.88

5. Discussion

The performance outcomes of this study indicate that combining deep learning-based object detection with classical segmentation provides a robust framework for analyzing hyperpigmentation in facial skin images. The high precision and recall scores achieved by YOLOv4 and Faster R-CNN suggest these models are capable of generalizing to varied facial features and pigmentation patterns.

Faster R-CNN demonstrated slightly superior precision, which can be attributed to its two-stage detection architecture that prioritizes accurate localization. This makes it particularly suitable for dermatological applications where diagnostic accuracy is critical. However, YOLOv4's near real-time performance and excellent recall make it attractive for possible applications in teledermatology and mobile deployment.

Segmentation results confirmed the utility of applying Otsu's thresholding after detection. While not as precise as deep learning-based segmentation methods, Otsu’s approach does not needs a training process, is computationally efficient and performs reliably in images with good contrast. When combined with morphological filters, it supports accurate estimation of lesion area—an essential parameter for monitoring treatment outcomes.

The integration of this system into a graphical user interface provides a practical tool for non-experts in clinical settings. Usability testing indicated high satisfaction with the interface's simplicity and interpretability, supporting its potential for wider adoption in clinical trials, dermatological assessments, and cosmetic product evaluation (Figure 5).

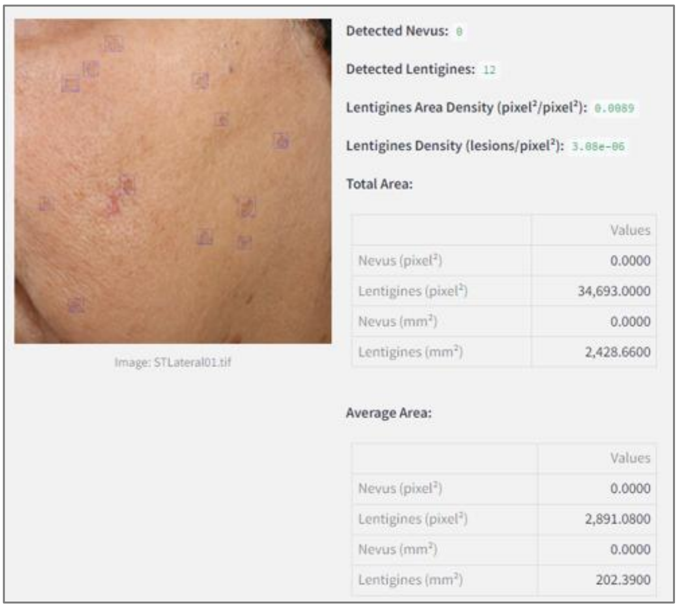

Figure 5. – The user interface allows the definition and extraction of metrics: number of lentigines and nevus detected, lentigines density, total area of each lesion and average area of each lesion in pixels² and mm²

Nonetheless, certain limitations exist. The datasets used in this study were composed primarily of lighter phototypes (Fitzpatrick I–III), and generalization to darker skin tones (Fitzpatrick V–VI) requires further validation. Hyperpigmented lesions on darker skin are often harder to distinguish due to reduced contrast, posing challenges to both detection and segmentation.

Moreover, the current models are limited to two lesion classes: lentigines and nevi. Expanding the system to include other pigmentation disorders such as melasma, post-inflammatory hyperpigmentation, or malignant lesions would significantly increase clinical applicability. This would require a larger, more diverse dataset and expert annotation to maintain label integrity. Lastly, although the segmentation approach was effective in this study, the use of more advanced architecture such as U-Net or Mask R-CNN could enhance boundary precision, particularly in cases with poor lighting or overlapping lesions. These methods also enable instance segmentation, which could support more detailed morphometric analysis in future versions of the system.

6. Conclusion

This paper presents a robust and interpretable approach for the detection and analysis of facial hyperpigmentation using deep learning and classical image processing. By integrating YOLOv4 and Faster R-CNN object detectors with Otsu-based segmentation, the system enables reliable lesion localization and quantitative analysis from standard facial photographs.

The solution demonstrated strong performance across multiple evaluation metrics, and its deployment in a user-friendly graphical interface enhances its clinical usability. Applications include diagnostic support in dermatology, cosmetic treatment evaluation, and automated patient follow-up in telemedicine settings.

Future developments will focus on extending the model's capabilities to a broader range of lesion types, improving segmentation precision with deep learning-based architectures, and validating system performance across diverse skin tones and image acquisition conditions. These steps will be essential for translating the technology into real-world dermatological and commercial environments.

7. References List

[1] Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You Only Look Once: Unified, Real-Time Object Detection. In CVPR.

[2] Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. In NeurIPS.

[3] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. In CVPR.

[4] Otsu, N. (1979). A Threshold Selection Method from Gray-Level Histograms. IEEE Transactions on Systems, Man, and Cybernetics, 9(1), 62–66.

[5] Lin, T. Y., Goyal, P., Girshick, R., He, K., & Dollár, P. (2017). Focal Loss for Dense Object Detection. In ICCV.

[6] Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. In MICCAI.

[7] Pham, T. C., Luong, C. M., Visani, M., & Hoang, V. D. (2021). Skin Lesion Segmentation and Classification: A Review. Journal of Healthcare Engineering.