# Acne Classification with YOLOv8

Aiden Coffey, Michael Woll

Computer Science Department, Florida Southern College

CSC 3951: AI For Medicine

Dr. Hoan Ngo

May 3, 2024

**Introduction:**

Acne is a prevalent dermatological condition that affects millions globally, manifesting through both non-inflammatory and inflammatory lesions. These lesions predominantly occur on the face, though they can also affect the back and chest. Traditional diagnostic methods rely heavily on visual assessment by dermatologists, which, while effective, can vary in accuracy due to subjective interpretation. The introduction of image segmentation and machine learning technologies has opened new avenues for enhancing the accuracy and consistency of acne lesion identification and classification. This paper explores the application of such advanced AI methodologies, specifically through the use of the YOLOv8 machine learning model, to improve the diagnosis and treatment planning of acne.

The study focuses on leveraging a robust dataset provided by Roboflow, curated by Kritsakorn, which includes diverse acne lesion types categorized into six distinct classes: blackheads, dark spots, nodules, papules, pustules, and whiteheads. By applying image segmentation techniques alongside deep learning algorithms, this research aims to not only classify acne more accurately but also to understand the distribution and frequency of acne types across different demographic groups. The ultimate goal is to pave the way for personalized acne treatment plans tailored to individual patient profiles, thereby enhancing treatment efficacy and patient outcomes. Through the detailed examination of various segmentation methods, the paper seeks to contribute to the broader field of dermatological AI, promoting a shift towards more data-driven, precise, and predictive healthcare solutions.

**Materials & Methods**

This study utilized the YOLOv8 machine learning model, implemented through the Ultralytics framework, to classify acne lesions within a dataset of 929 images sourced from Roboflow. The dataset includes six distinct types of acne: blackheads, dark spots, nodules, papules, pustules, and whiteheads. Each image was annotated with accurate labels corresponding to the lesion type, ensuring a high-quality training set for the model.

Each acne object is represented by 5 numbers. The first number listed is the type of acne being depicted (ranges from 0-5). The next four numbers represent the coordinates and size of the bounding box surrounding the object in the image. These values are normalized to fall between 0 and 1, relative to the dimensions of the image. The format for these values is:

- X_center: The x-coordinate of the center of the bounding box, normalized by the width of the image.
- Y_center: The y-coordinate of the center of the bounding box, normalized by the height of the image.
- Width: The width of the bounding box, normalized by the width of the image.
- Height: The height of the bounding box, normalized by the height of the image.

The training and evaluation of the model were conducted on Google Colab, a cloud-based platform that provides free access to a GPU, enabling efficient handling of deep learning algorithms. This environment was chosen for its accessibility and robust computational resources, allowing for the consistent application of the model without the hardware limitations often encountered with personal computing systems.

**Results, Conclusions, and Discussion**

The training of the YOLOv8 model for acne lesion detection across 929 images revealed significant insights into the model's performance and areas for improvement. Despite aspirations for extensive epoch training, technical constraints limited the process to 2171 epochs, short of the planned 5000. Early stages showed minimal effectiveness, with a mean Average Precision (mAP) as low as 7% at five epochs, predominantly classifying most detections as background. Progress was evident as the model training extended; at 50 epochs, the mAP improved to 14%, showcasing enhanced recognition of papules and dark spots. The introduction of data augmentations at 900 epochs, such as random image rotations and inversions, aimed to enhance model robustness. However, this resulted in a reduced mAP of 12%, suggesting potential overfitting or inadequate training for augmented scenarios. Without augmentations, the mAP at 900 epochs slightly improved to 14%, indicating a delicate balance between augmentation techniques and training efficacy.

The fluctuating mAP values culminating in a peak of 49% at the 2171st epoch highlight the challenges of deep learning applications in medical imaging, particularly in the robust and accurate classification of diverse acne types. This peak performance, while promising, underscores the inconsistency and potential volatility in training deep learning models with limited computational resources and time constraints.

In conclusion, the study demonstrates that while the YOLOv8 model holds substantial potential for dermatological applications, achieving consistent and reliable performance requires further refinement. Future strategies should focus on extending the number of training epochs, optimizing model parameters, and refining augmentation techniques to better suit the complex nature of acne lesion images. Additionally, exploring alternative architectures like YOLOv5 and reassessing the training methodology might yield improved results. These steps are crucial for advancing the model towards clinical utility, where higher precision and reliability are imperative for supporting dermatological diagnostics and treatment planning.

Figure 1.a: Precision-Recall Curve @ 5 epochs

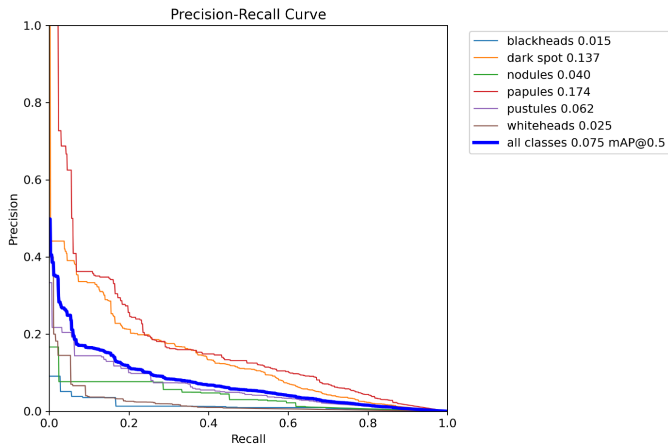


Figure 1.b: Confusion Matrix @ 5 epochs


Figure 1.c: F1-Confidence Curve @ 5 epochs

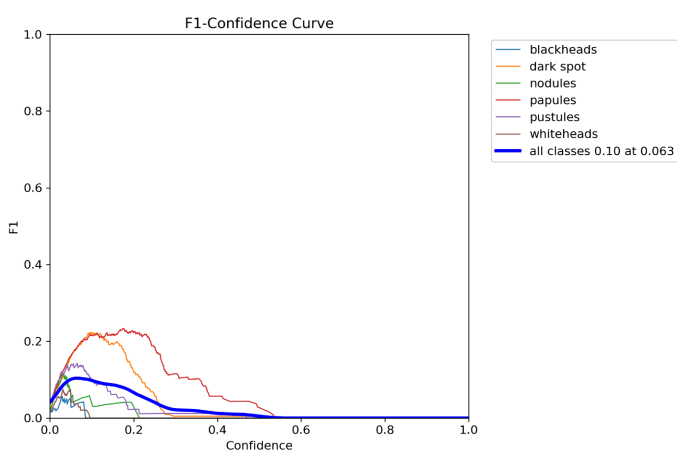

Figure 2.a: Precision-Recall Curve @ 50 epochs

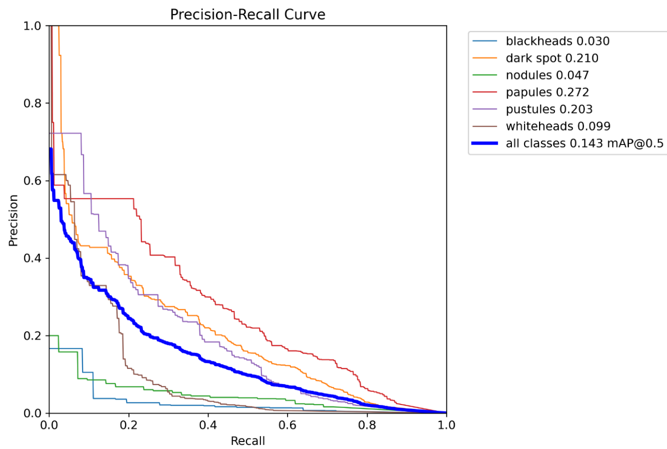


Figure 2.b: Confusion Matrix @ 50 epochs

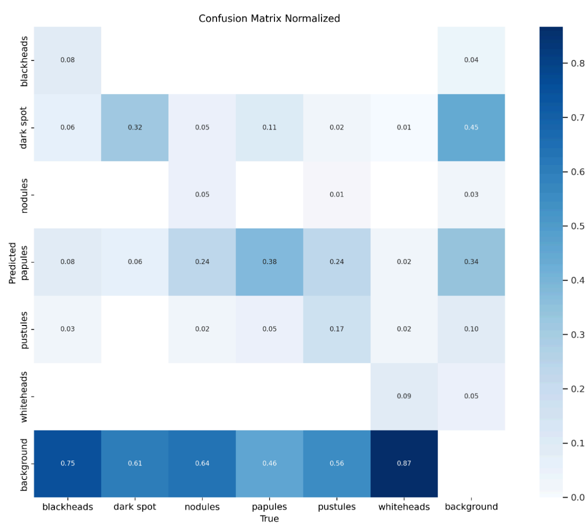


Figure 2.c: F1-Confidence Score @ 50 epochs

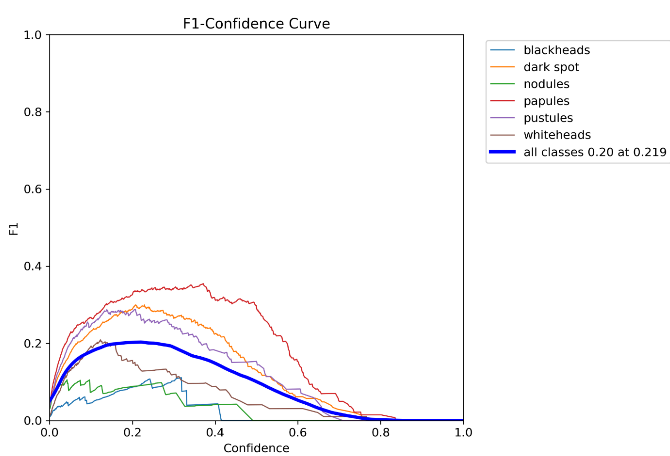

Figure 3.a: Precision-Recall Curve @ 300 epochs

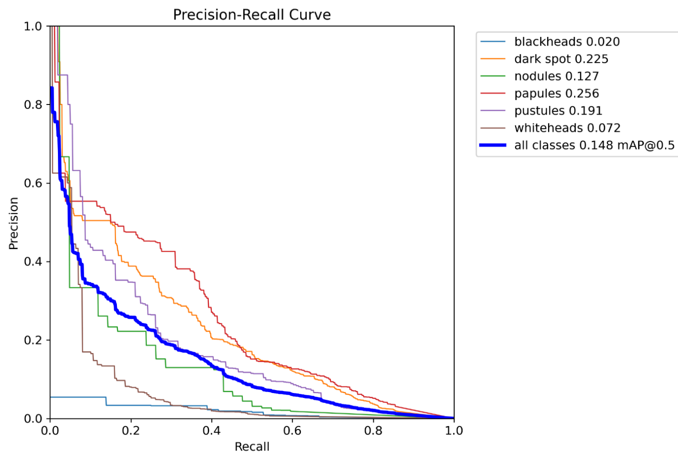


Figure 3.b: Confusion Matrix @ 300 epochs

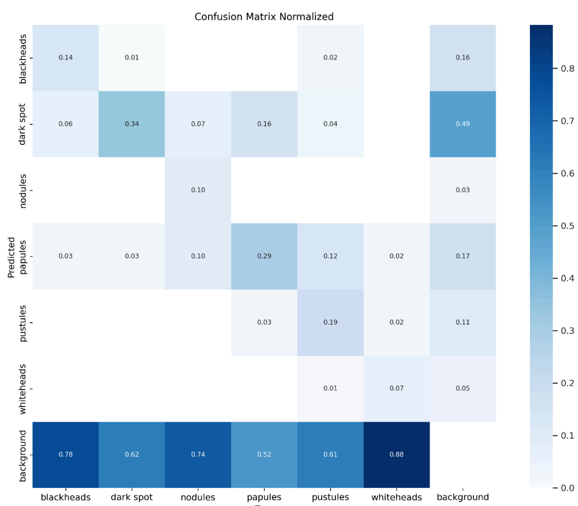


Figure 3.c: F1-Confidence Score @ 300 epochs

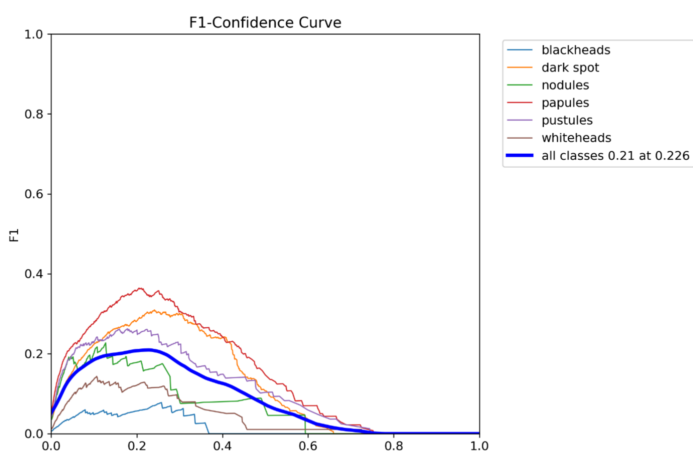

Figure 4.a: Precision-Recall Curve @ 900 epochs

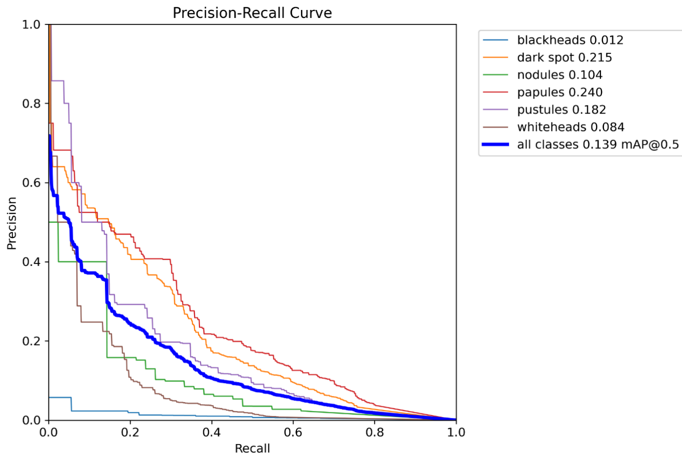


Figure 4.b: Confusion Matrix @ 900 epochs

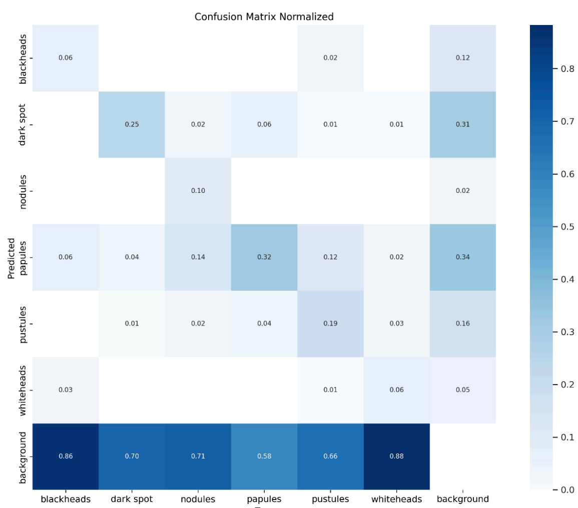


Figure 3.c: F1-Confidence Score @ 300 epochs

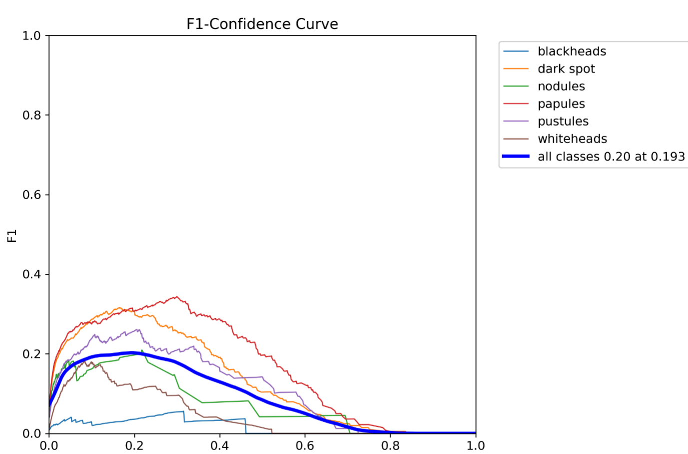

Figure 5.a: Original image

Figure 5.b: Labeled image


Figure 5.c: Model labeled image