Optimization Techniques in YOLOv8 for Acne Detection

Aqilah Fedura Ilahi, Putu Harry Gunawan

Telkom University
Bandung, Indonesia

Advancing Human Potential through Technology: Designing Beneficial and Respectful Digital Systems

IEEE Indonesia Section

DIKTISAINTEK BERDAMPAK

ICITACEE 2025

Introduction

Definition - What is Acne Vulgaris ?

- Chronic dermatological condition that mainly affects adolescents and young adults.
- Pathogenesis: hyperkeratinization, excess sebum, Cutibacterium acnes, and inflammation.
- Influencing factors: hormones, diet, genetics.
- Clinically, acne is categorized into:

1. Non-inflammatory lesions: blackheads and whiteheads
2. Inflammatory lesions: papules and pustules

Problem & Context

High Prevalence

- **Global prevalence:** According to Saurat et al. (2024), acne affects approximately 20.5% of individuals aged 16 to 24 years.
- **Indonesia prevalence:** Latifah et al. (2025) reported that in Indonesia, acne is highly prevalent, particularly among females aged 14 to 17 and males aged 16 to 19.

Psychological Impact

In Indonesia, a study by Matthew et al. (2021) reported that over 82% of students with acne experience psychological distress, including low self-esteem and reduced social confidence.

Clinical Challenge

- Lesion types (blackheads, whiteheads, papules, pustules) show visual similarity, making classification difficult.
- Manual detection is slow and subjective.

Need for Automated Solutions

Accurate and timely automated detection is crucial to support both dermatological and mental health care.

Related Works

YOLO Study

- **Gan et al. (2024) –** YOLOv5 for facial image detection involving acne-related categories; mAP@0.5 0.424, precision 0.540, recall 0.446. Performance was limited by data imbalance and variations in image quality.

- **AlSadhan et al. (2024)** – YOLOv7 for skin cancer detection; mAP 0.754, F1-score 0.779. Strong dermatology performance.

- **Azhar et al. (2024)** – YOLOv8 for acne lesion classification and counting; precision 0.800, recall 0.810, F1-score 0.780. Shows YOLOv8's capability for acne detection.

Optimizer Study

Supatman et al. (2023) – Compared six optimizers on YOLOv8m for knee landmark detection.

1. Adam achieved highest mAP@0.5 (0.696).
2. AdamW had fastest training (0.385 h).
3. RAdam showed strong precision and recall.
4. NAdam gave stable results.

Highlighted optimizer's impact on training stability and performance.

Novelty

Previous Research

- Mostly focused on YOLOv5 or YOLOv7 for dermatology tasks, but did not address acne detection with balanced datasets and did not involve optimizer comparison.
- Supatman et al. (2023) compared optimizers on YOLOv8m for knee landmark detection, but the study was not applied to acne datasets and lacked augmentation and balancing strategies.

Our Approach

- Designed specifically for acne detection using YOLOv8s.
- Focuses on four acne types with manual re-annotation, mild augmentation, and class balancing to address imbalance issues.
- Provides the first comprehensive comparison of five optimizers (SGD, Adam, AdamW, NAdam, RAdam) under identical settings.
- Reveals dataset-specific optimizer behaviors not previously reported in dermatology research, providing practical insights for future AI-based dermatology applications.

Materials and Method

Dataset

The dataset was constructed from two sources:

Skin90 (Kaggle)

DermNet (DermNet NZ)

and contains acne lesion images classified into four categories:

Blackheads

Whiteheads

Papules

Pustules

Dataset Specifications:

- Total images: 530 annotated samples
- Annotation performed manually using Roboflow, adopting the YOLOv8 format.
- Each lesion labeled with Class ID and bounding box values (normalized coordinates and dimensions).

Table I

Class ID	Class Name
0	Blackhead
1	Papule
2	Pustule
3	Whitehead

Figure 1

Table II

Variable	Description
Class ID	Acne lesion category
X-center	Horizontal center coordinate (normalized)
Y-center	Vertical center coordinate (normalized)
Width	Bounding box width (normalized)
Height	Bounding box height (normalized)

Table III

Class ID	X-center	Y-center	Width	Height
0	0.6828	0.8890	0.0189	0.0187
1	0.9531	0.4625	0.0938	0.1031
2	0.9930	0.6813	0.0141	0.0344
3	0.7469	0.5266	0.0656	0.0439

Preprocessing & Augmentation

- Images were processed with Roboflow's pipeline, output formatted for YOLOv8.
- Preprocessing steps:

1. Images resized to 640×640 px
2. Automatic orientation correction

- Augmentation :

1. Horizontal flip, Cutout (3 boxes ≤10%), Mosaic
2. Brightness/exposure ±10%, Light blur (≤5 px)

After Augmentation:

Total images: 1,270 :

Split ratio: 70% training (1,110) – 20% validation (106) – 10% testing (54)

Validation and test sets remained unchanged

Handling Class Imbalance

Figure 2a

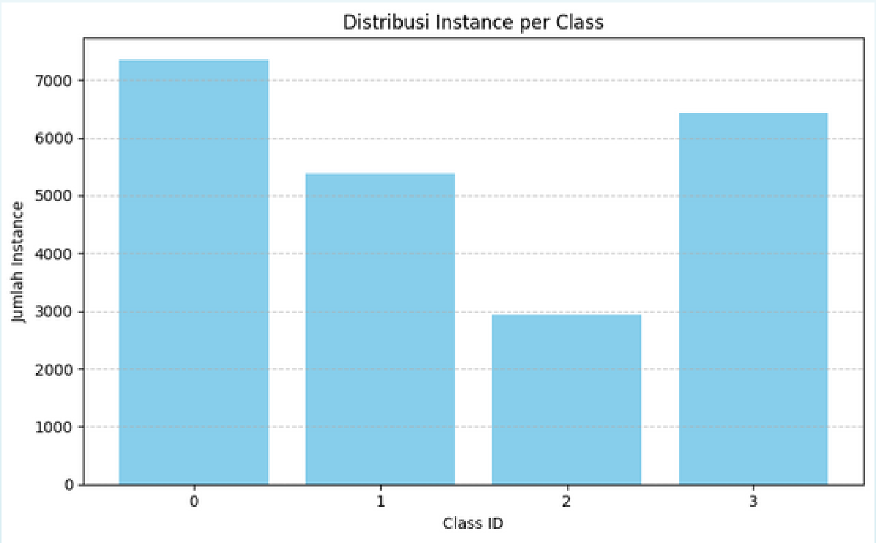
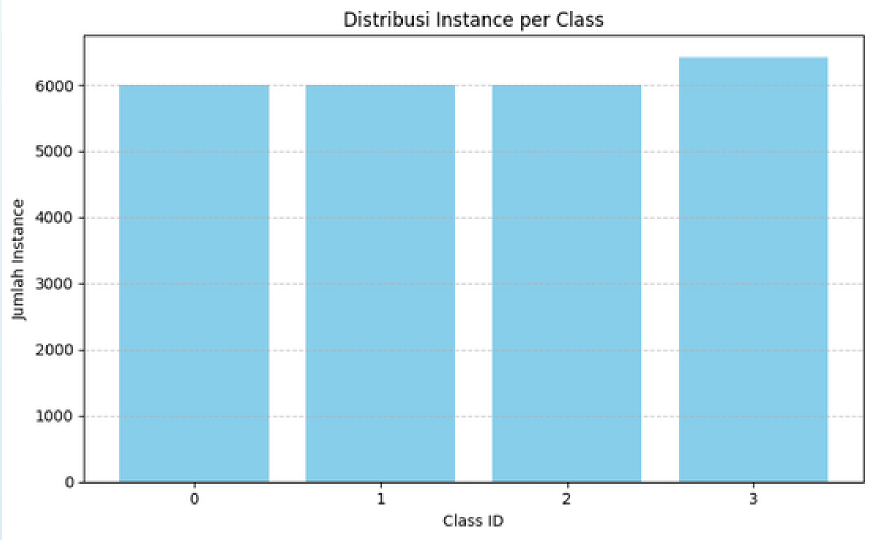


Figure 2b



Balancing strategy (Balancing minimized bias by equalizing class counts) :

- random undersampling : Classes 0 & 3
- oversampling for Classes 1 & 2

Before Balancing

Class 0: 7,356, Class 3: 6,424, Class 1: 5,372, Class 2: 2,933

After Balancing

Class 0: 6,006, Class 1: 6,010, Class 2: 6,014, Class 3: 6,424

Total samples ≈ 24,500

Object Detection

YOLOv8

Why YOLOV8 ?

- YOLOV8 (2023)
- table performance
- Comprehensive documentation
- Proven effectiveness in medical image detection tasks

YOLOv8 Architecture

- Single-stage detection: localization & classification performed simultaneously

- Components:

1. Backbone – feature extraction
2. C2f neck – multiscale feature aggregation
3. Detection head – separates classification & regression for improved output

- Anchor-free design: simplifies training, improves detection for irregular objects

YOLOv8s Variant

- Balanced accuracy & efficiency
- Fast inference for real-time clinical use
- Works well on resource-limited hardware

Optimizers

Five optimizers were compared, each influencing convergence and generalization:

- SGD – Momentum-based; efficient, needs LR tuning
- Adam – Adaptive LR; fast, risk of overfitting
- AdamW – Adam + weight decay; better regularization
- NAdam – Adam + Nesterov momentum; improved convergence
- RAdam – Rectified adaptive LR; stable early training

Evaluation Metrics

Model performance evaluated using four metrics:

1. Precision

$$\text{Precision} = \frac{TP}{TP + FP}$$

2. Recall

$$\text{Recall} = \frac{TP}{TP + FN}$$

3. mAP

$$\text{mAP} = \frac{1}{N}\sum_{i=1}^{N} AP_i$$

4. Training Time

Experimental Setup

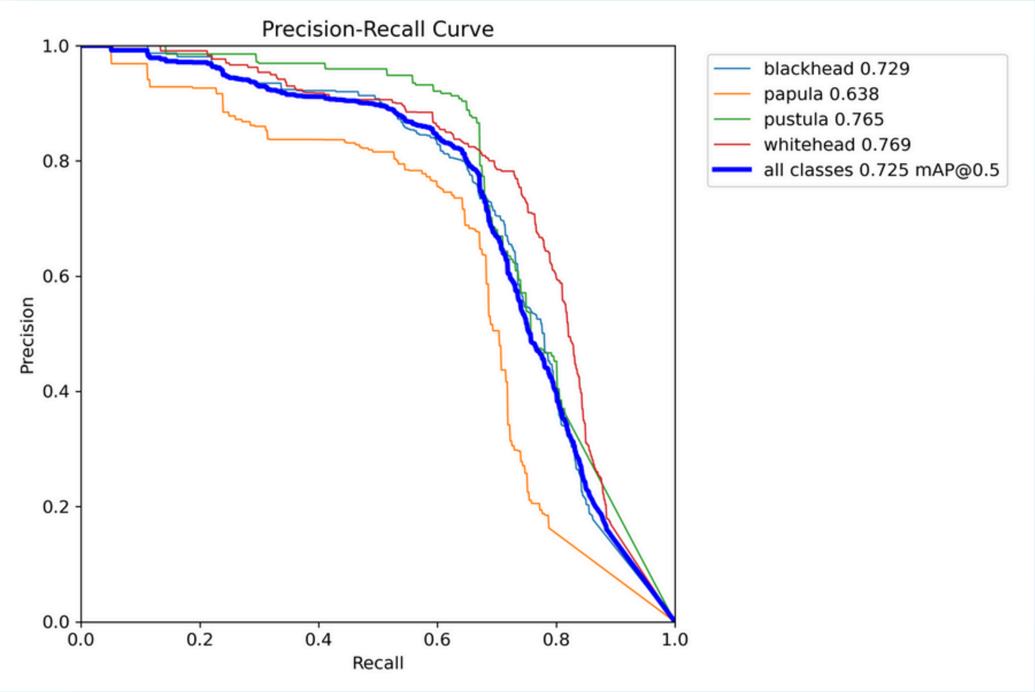
Parameter	Configuration
Model	YOLOv8s (pretrained weights)
Optimizers	SGD, Adam, AdamW, NAdam, RAdam
Batch Size	16
Epochs	150
Learning Rate	Cosine schedule, initial 0.0003, min factor 0.008
Regularization	Weight decay 0.006, Dropout 0.14
Platform	Google Colab, NVIDIA T4 GPU

RESULTS AND DISCUSSION

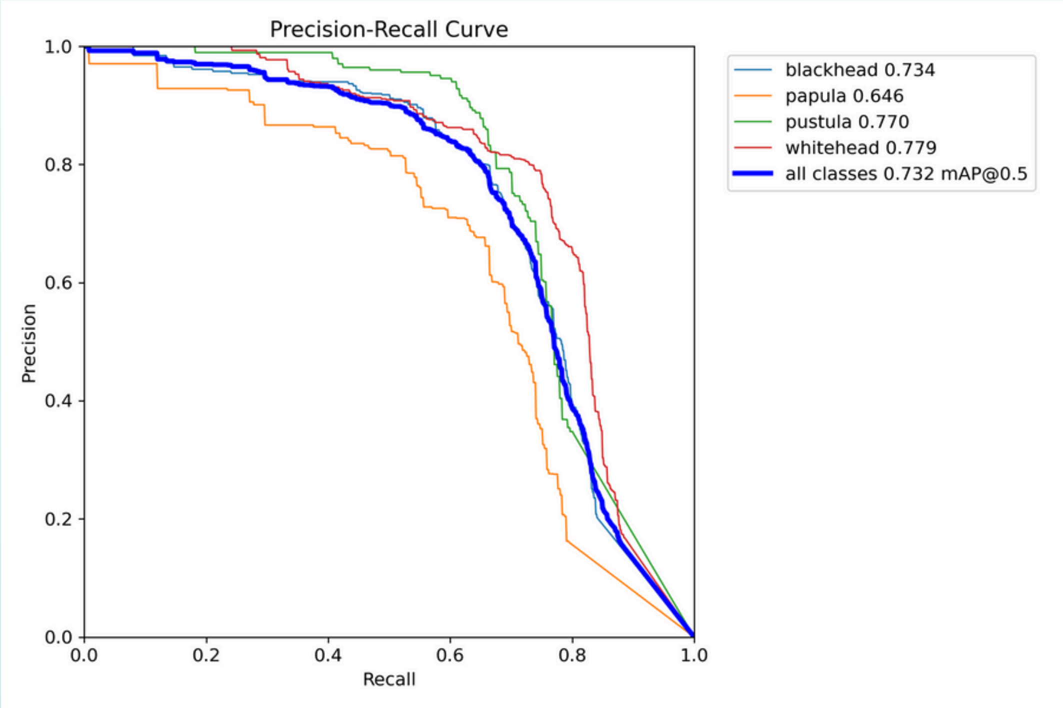
Performance Overview During Training

Optimizer	Precision	Recall	mAP@0.5	Train Time (h)
Adam	0.728	0.671	0.711	0.846
AdamW	0.784	0.668	0.726	0.847
NAdam	**0.823**	0.646	**0.735**	0.847
RAdam	0.788	0.658	0.720	**0.588**
SGD	0.733	**0.676**	0.712	0.826

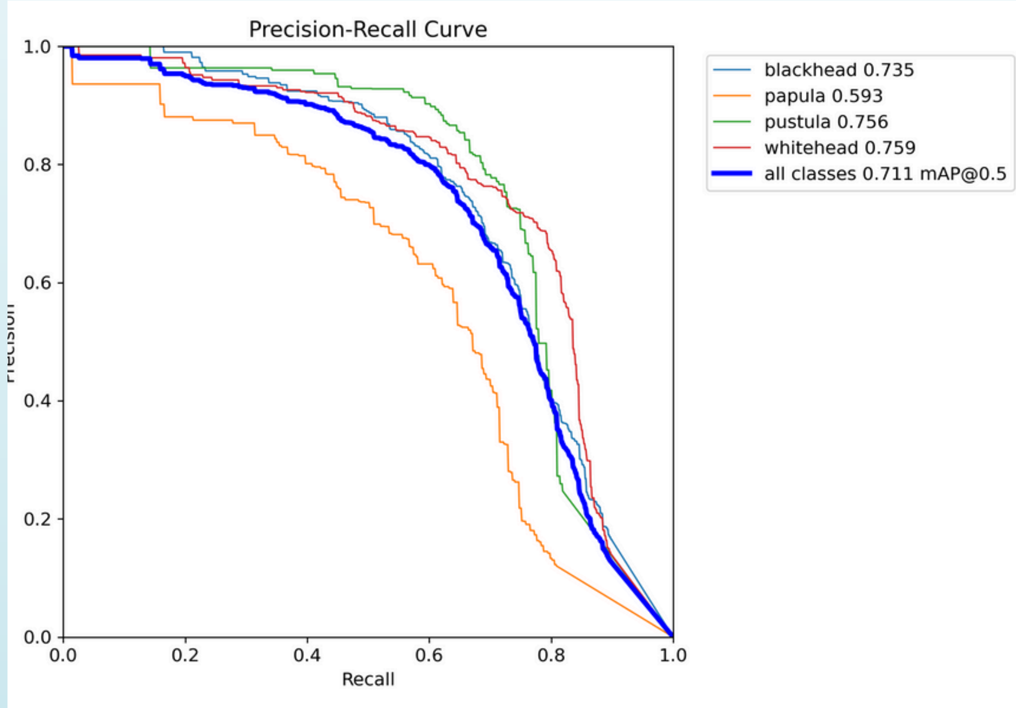
Precision--Recall Curve Analysis



Precision--Recall Curve Analysis AdamW

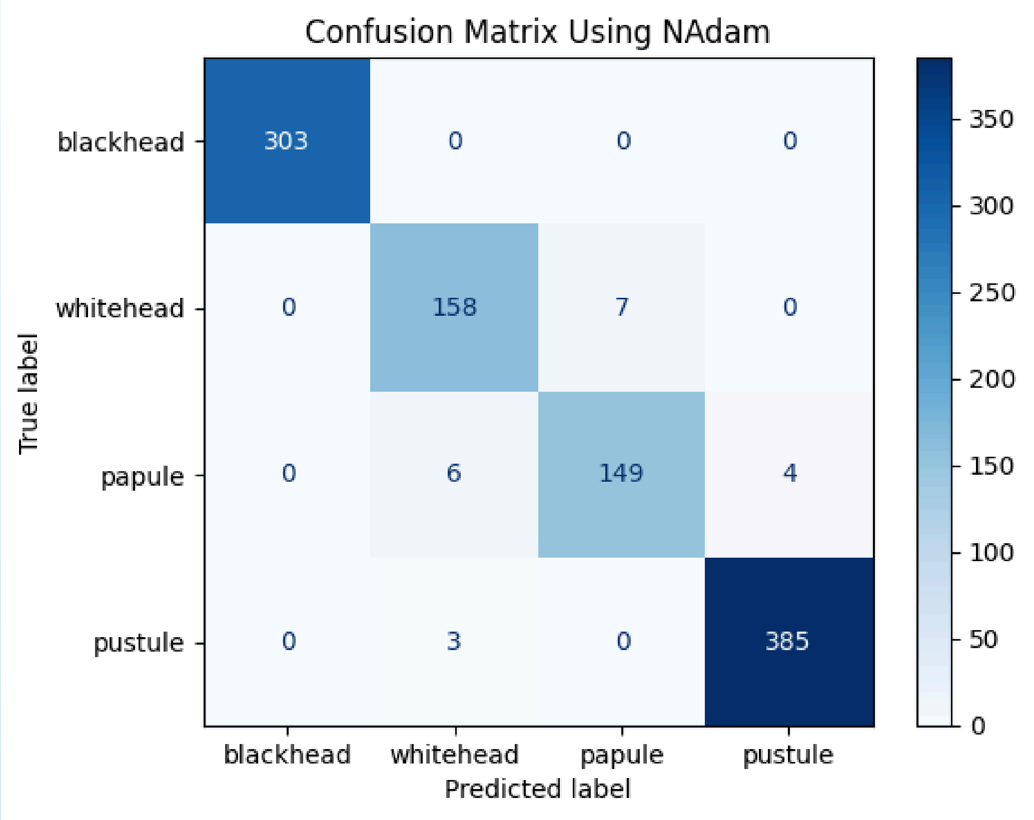


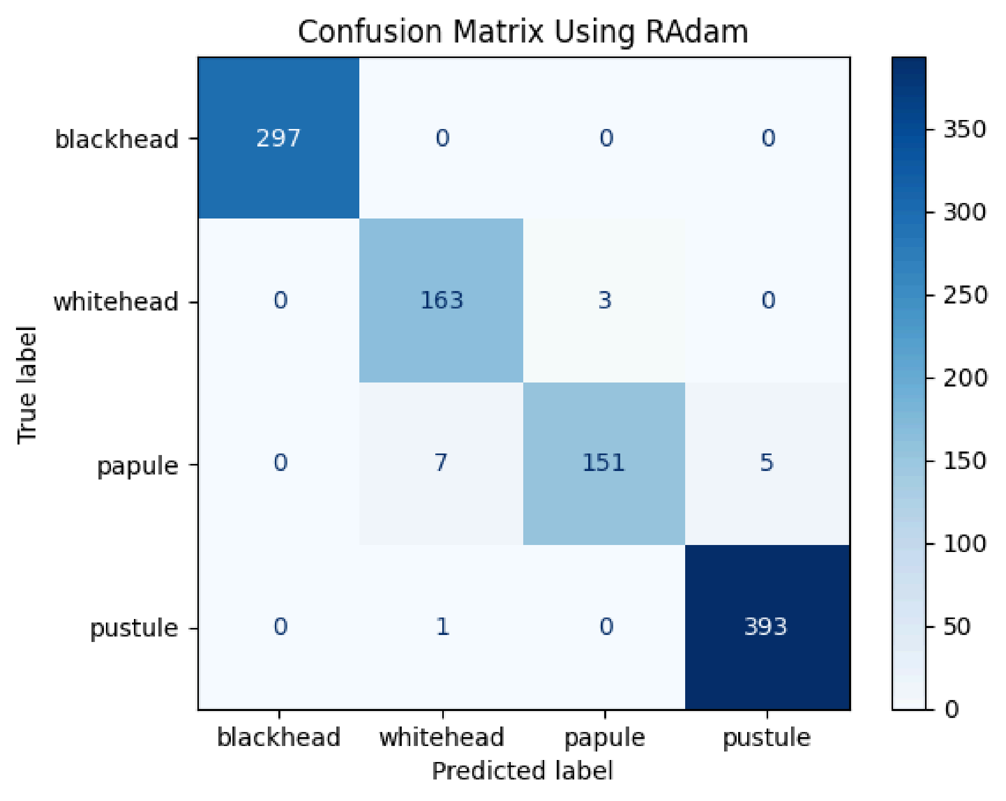
Precision--Recall Curve Analysis NAdam

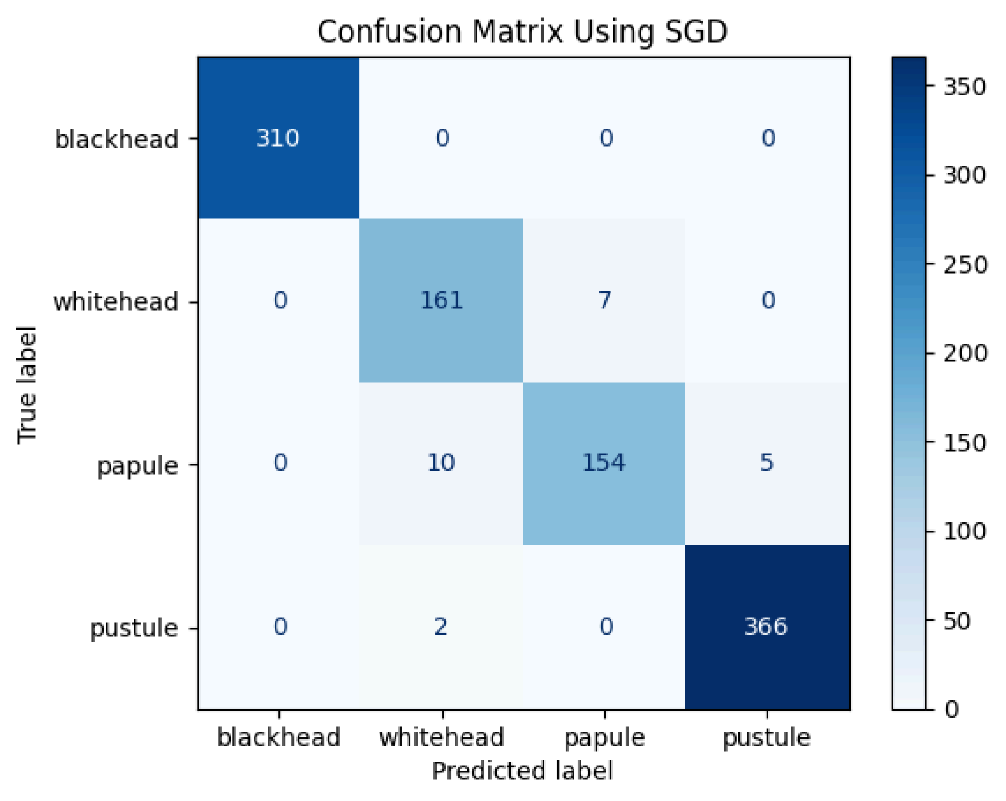


Precision--Recall Curve Analysis SGD

Confusion Matrix Analysis



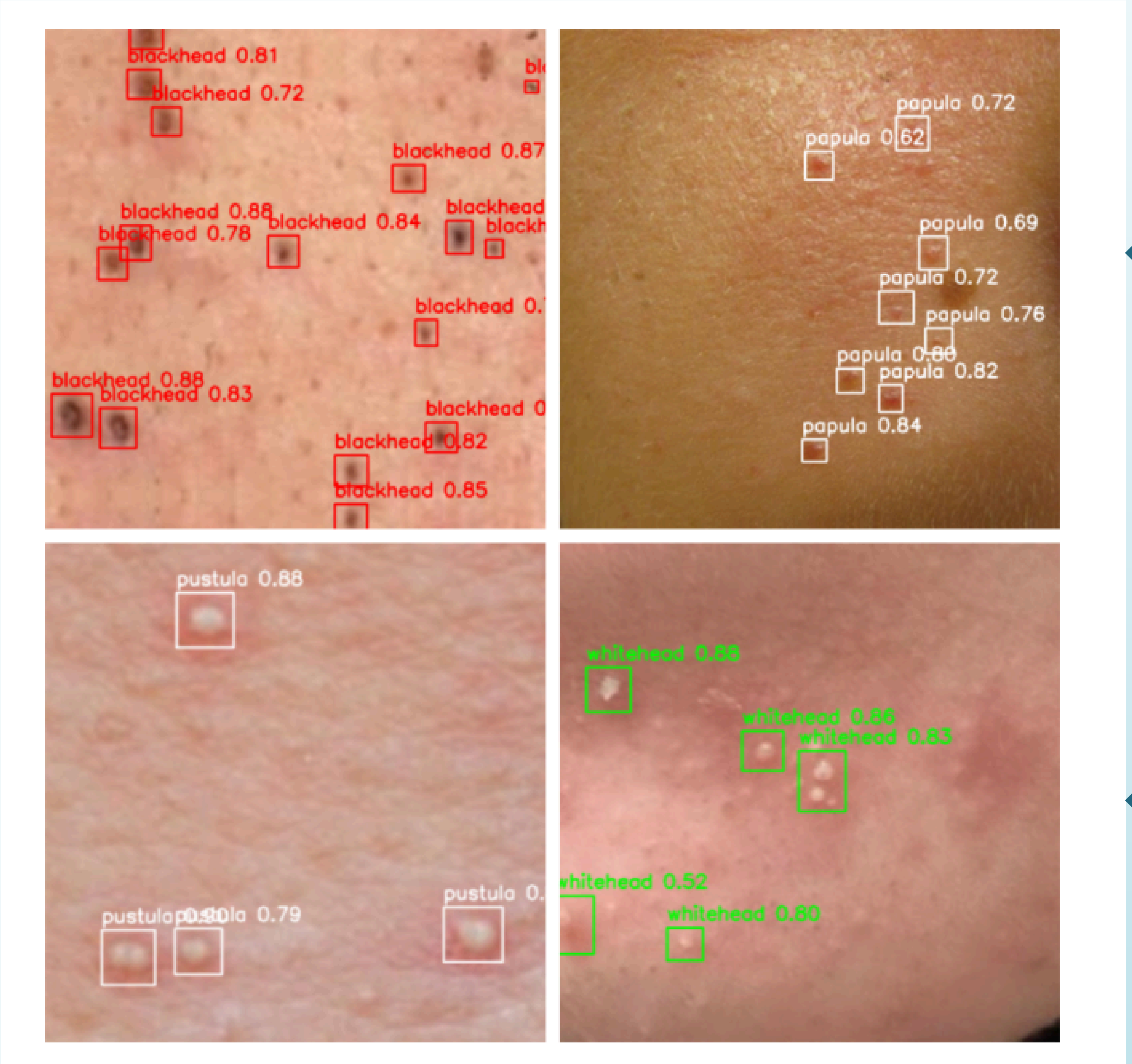




Detection Result Using NAdam

Blackheads

Papules



pustules

Whiteheads

CONCLUSION

This study evaluated five optimizers: Adam, AdamW, NAdam, RAdam, and SGD on YOLOv8s for multi-class acne detection. Using a balanced dataset from Skin90 and DermNet, experiments showed that NAdam achieved the highest precision (0.823) and mAP@0.5 (0.735), SGD obtained the highest recall (0.676), and RAdam trained the fastest (0.588 h) with competitive accuracy. Papules remained the most difficult class, often misclassified as whiteheads or pustules. Tests on unseen images confirmed the robustness of the NAdam-optimized model, highlighting the importance of optimizer selection in enhancing YOLOv8-based acne detection

FUTURE WORK

- External Validation
- Advanced Augmentation
- Evaluation of New Architectures

Thank You