2.4 Calculation Results

Model Performance Evaluation:
After completing the training process, the object detection model was thoroughly evaluated using standard metrics commonly adopted in computer vision, particularly those provided by Roboflow. These metrics help quantify how accurately and reliably the model can detect fuselage cracks in aircraft images.

Reported Metrics:

- **Precision:**
  Precision measures the proportion of correctly identified cracks among all detections made by the model. High precision indicates that the model generates few false positives, which is crucial in industrial inspections to avoid unnecessary maintenance or inspection alerts.
- **Recall:**
  Recall represents the proportion of actual cracks that the model successfully detects. A high recall value ensures that few cracks are missed, which is critical for safety and compliance in aircraft maintenance.
- **Mean Average Precision (mAP):**
  The mAP summarizes the model's detection performance across all confidence thresholds and object classes. It provides a single, comprehensive score that balances precision and recall, making it a widely used benchmark for object detection models.
- **Confusion Matrix:**
  The confusion matrix visually presents the number of true positives, false positives, true negatives, and false negatives. This matrix helps identify specific cases where the model may struggle, such as misclassifying textured surfaces or shadows as cracks.

Observations:

- **Successful Detection:**
  The model demonstrates strong performance in detecting visible fuselage cracks in most test images. Cracks of varying lengths and orientations were correctly identified, confirming the model's capability to generalize across different parts of the fuselage.
- **False Positives:**
  Some false positives were observed, particularly on textured, scratched, or dirty surfaces. These areas occasionally resemble cracks in visual appearance, highlighting a limitation that could be addressed through further dataset augmentation or model refinement.
- **Influence of Image Quality:**
  Detection performance was notably improved with higher-resolution images and well-lit conditions. Poor lighting or low-resolution images sometimes led to

missed detections or lower confidence scores, indicating the importance of image acquisition standards in practical inspection scenarios.

Summary:

Overall, the evaluation confirms that the trained model is effective for automated fuselage crack detection. While most cracks are detected accurately, special attention should be paid to reducing false positives and ensuring consistent image quality for optimal deployment in real-world inspection workflows.