CenterNet Model Performance Evaluation

Introduction

This report provides a comparative analysis of the CentreNet model's performance on two datasets of PCB components. The task involves detecting various components and performing multi-class classification

Hardware and Environment Config

Hardware Configuration:

- CPU:
- GPU:
- RAM:
- Storage:

Environment Details:

- Operating System:
- Python Version:
- Deep Learning Framework:
- CUDA Version:
- cuDNN Version:

Comparative Analysis for PCB Components Dataset

Metrics Comparison

Metric	1st Dataset	2nd Dataset
Total Inference Time	0.118 seconds	0.144 seconds

FPS	8.509	6.968
mAP (Mean Average Precision)	0.460	0.454
Average Precision	0.460	0.454
Average Recall	0.389	0.500
Average F1 Score	0.502	0.579
Average Mean IoU	0.335	0.408
Overall Precision	0.705	0.688

Key Findings

1. Overall Performance

- **First Dataset:**
    - The CentreNet model achieved an mAP of 0.460 with a total inference time of 0.118 seconds and FPS of 8.509.
    - Inference time and FPS metrics suggest the model is more efficient on this dataset.
    - The Intersection over Union (IoU) and recall values for our model are currently low, indicating that there is significant room for improvement.

- **Second Dataset:**
    - The model showed similar performance with an mAP of 0.454 with a total inference time of 0.144 seconds and FPS of 6.968.
    - It has higher average reacall and mean IoU as compared to the first dataset.

    - The overall precision is 68.8% and needs improvements.

2. F1 Score and Precision

- **First Dataset:**
  - The F1 score of 0.502 reflects a moderate balance between precision and recall, suggesting that while the model performs reasonably well, there is still a need to improve both metrics for better overall performance.
  - With an overall precision of 70.5% , the model demonstrates a strong ability to make accurate predictions, though there is potential to enhance its performance further
- **Second Dataset:**
  - The F1 Score of 0.579 reflects a well-balanced model with good precision and recall.
  - With an overall precision of 0.688, the model demonstrates a good level of accuracy in its predictions, meaning that when the model predicts a positive class, it is correct 68.8% of the time. However, there is still potential to enhance the model's precision for even more reliable results.

3. Recall

- The recall score of 0.389 for the first dataset indicates that the model is identifying only 38.9% of the actual positive cases, suggesting it is missing a significant number of positive instances. For the second dataset, a recall score of 0.500 shows that the model correctly identifies 50% of the positive cases. These relatively low recall scores imply that the model is struggling to capture all relevant instances, which may be due to insufficient training data, imbalanced datasets, or the need for further optimization of the model's parameters.

4. Mean IoU

- The average mean Intersection over Union (IoU) of 0.335 for the first dataset and 0.408 for the second dataset indicates that the model's predictions overlap with the ground truth by only 33.5% and 40.8%, respectively. These low IoU values suggest that the model is not accurately capturing the boundaries of the objects. The performance may be hindered by factors such as insufficient training data, poor quality of annotations, or inadequate model architecture.

Key Findings:

Key observations from these metrics:

1. **Recall Improvement**: The second dataset has a higher recall (0.500) compared to the first (0.389), indicating that the model detects a larger proportion of actual positives in the second dataset.
2. **IoU and Precision**: The average mean IoU and precision are lower for the second dataset, suggesting that while more objects are detected, the accuracy of these detections is slightly reduced.

3. **F1 Score**: The F1 score is higher for the second dataset (0.579 vs. 0.502), showing a better balance between precision and recall.

- The pruned model shows a significant drop in performance. This suggests that pruning may have overly reduced the model’s capacity or introduced issues that led to poor predictions. Despite similar inference time and FPS compared to the raw model, the quality of predictions has deteriorated drastically.
- The quantized model also shows a drop in performance similar to the pruned model. While it has improved inference time and FPS, its predictive quality is poor.

By addressing these areas, the model's performance on object detection tasks can be significantly enhanced, resulting in more accurate and reliable predictions.