# RT-DETR Performance Report on WaRP-D

## Objective

To analyze the detection performance of RT-DETR on the WaRP-D dataset using Ultralytics, with a focus on high-confidence (conf=0.6) validation for deployment-ready results.

## Experiment Settings

Training configuration:  
• Model: RT-DETR-l (COCO-pretrained)  
• Epochs: 150  
• Batch size: 16  
• Image size: 640x640  
• Augmentations: mosaic=0.5, mixup=0.1, hsv\_h=0.02, hsv\_s=0.8, hsv\_v=0.5  
• Learning rate: 0.001  
• Label smoothing: 0.1  
• Early stopping: Disabled (patience=0)  
• TTA (Test Time Augmentation): Enabled at validation

## Validation with High Confidence Threshold (conf=0.6)

Validation command:  
results = model.val(  
 data='/kaggle/working/data.yaml',  
 imgsz=640,  
 conf=0.6, # Only keep predictions with 60%+ confidence  
 augment=True # Test Time Augmentation for extra robustness  
)  
  
Result snapshot:  
- mAP50: 0.648  
- mAP50–95: 0.548

## Why Use conf=0.6?

What it means:  
The confidence threshold filters model predictions so that only detections with a confidence score of 0.6 (60%) or higher are kept and evaluated.  
  
Why do this?  
- In real-world deployment, we want to reduce false positives—only show detections that the model is very certain about.  
- This makes the output cleaner, with higher precision and fewer “junk” detections.  
- End-users or automated systems can trust these boxes more (important in recycling, quality control, etc.).  
How does it help?  
- Raises precision (percentage of detected objects that are actually correct).  
- Typically, recall may drop slightly (some true objects with lower confidence are missed), but for practical/industrial tasks, high-precision is preferred.  
- mAP50 and mAP50–95 values become a measure of trusted detection performance, not just “best possible” under all conditions.

## When to Use This Approach

For deployment, production, or field use:  
- You want to avoid mistakes more than you want to detect absolutely everything.  
- For waste sorting robots, smart cameras, or operator dashboards, this reduces distraction and error-handling workload.

## Future work

- RT-DETR-l with conf=0.6 delivers strong, reliable, high-precision detection suitable for real-world applications.  
- To further increase accuracy:  
 • Continue using advanced augmentation.  
 • Experiment with different confidence levels (e.g., 0.5 or 0.7) to find the best tradeoff for your use-case.  
 • Enable TTA for further improvement in robustness.

## Conclusion

Evaluating RT-DETR-l at conf=0.6 focuses on the model’s “most certain” predictions—giving you practical, high-precision detection results for deployment, with excellent mAP50 and mAP50–95 on the WaRP-D dataset.