

To execute the code, Follow the following steps

1. Clone the DINO-DETR repository from the link given
2. Follow the steps given the repository to step up the files and execute the script given the repository for MultiScaleDeformableAttention Dependencies.
3. Once all setup, download the dataset from the given link

https://drive.google.com/drive/folders/1FhQ6tug9ti7OHXbt6-4BbmjFaMjyBYNN?usp=drive_link

4. Also download the R50 – scale 4 model from here
https://drive.google.com/file/d/1AwUn5EebmmLBo7njjW_Ng1q9zDrqkNbB/view?usp=drive_link
5. Now execute the cell one by one and you will receive the result

Validation data Analysis

AP is the primary metric for object detection. It calculates precision at various recall thresholds, averaged across Intersection over Union (IoU) thresholds between 0.50 and 0.95 (in 0.05 increments).

AP@[IoU=0.50:0.95] for **all** **areas:** 0.502

This is the most comprehensive measure of performance, combining detections over different object sizes and IoU thresholds. A score of **0.502** indicates the model is fairly accurate in detecting objects across varying IoU thresholds.

AP@[IoU=0.50] for **all** **areas:** 0.833

This measures precision with a more lenient IoU threshold (0.50), sometimes referred to as the "PASCAL VOC" metric. A high value like **0.833** suggests that the model can detect objects accurately with some localization tolerance.

AP@[IoU=0.75] for **all** **areas:** 0.548

At a stricter IoU threshold of 0.75, precision is lower at **0.548**, meaning the model has moderate accuracy in tightly localizing objects.

AP@[IoU=0.50:0.95] for small, medium, and large areas:

- **Small objects (0.402):** The model struggles more with small objects.
- **Medium objects (0.637):** The model performs better here.
- **Large objects (0.849):** The model performs best with large objects, indicating strong localization and detection for bigger targets.

Conclusion: The precision for larger objects is significantly higher (84.9%) than for smaller ones (40.2%).

Validation Result after fine tuning for 5 epochs

Average Precision (AP):

AP@[IoU=0.50:0.95] for all areas: 0.559

- This indicates an overall improvement in the model's ability to detect objects across various IoU thresholds compared to the previous score of **0.502**.

AP@[IoU=0.50] for all areas: 0.890

- This score shows a significant increase from **0.833**, indicating that the model is highly effective in detecting objects with some localization tolerance.

AP@[IoU=0.75] for all areas: 0.637

- This score is also improved from **0.548**, showing better performance in accurately localizing objects.

AP for different object sizes:

- **Small (0.466):** A slight improvement from **0.402**, indicating better handling of small objects.
- **Medium (0.688):** Improved from **0.637**, demonstrating a better detection capability for medium-sized objects.
- **Large (0.874):** This is up from **0.849**, showing continued strength in detecting large objects.

Conclusion: The models performance after fine-tuning, results in a slightly better accuracy in handling small objects on our dataset