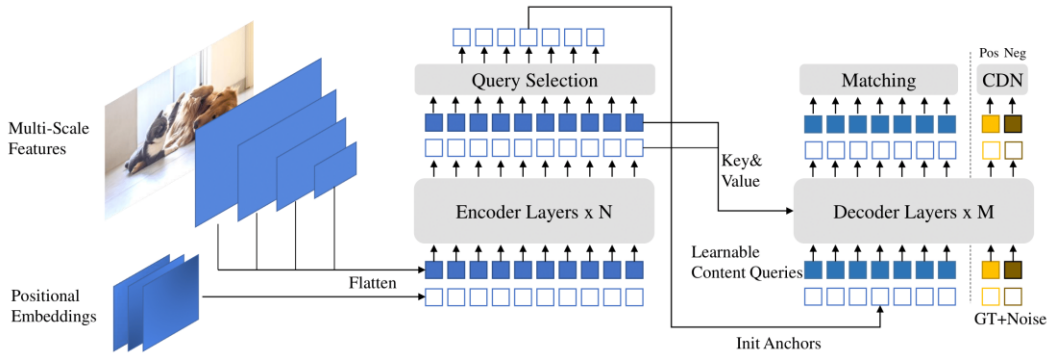
Homework #1 Object Detection

CVPDL 2023 Fall R12725026 秦孝媛

1. Draw the architecture of your object detector
   - Object Detector：DINO



2. Implement details
   - Pretrain-weight：checkpoint0033_4scale.pth
     - 36 epoch setting
     - Backbone: R50
     - Datasets Used for Pre-training: The model was pre-trained on the COCO 2017 dataset
   - Fine-tune Details:

Epochs	300	Weight Decay	1e-4
Optimizer	AdamW	Grad clip	0.1
Learning Rate	1e-4		

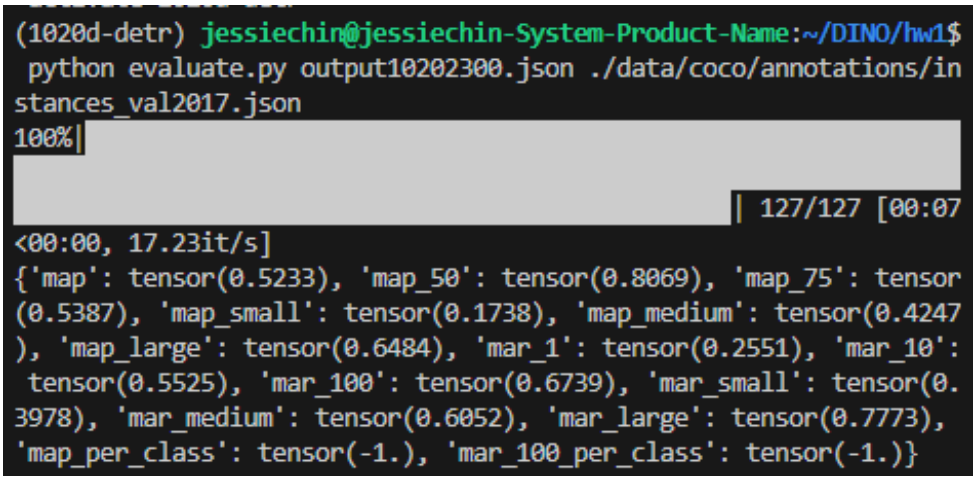
   - Augmentation:
     - Random Horizontal Flip: The images are flipped horizontally with a certain probability.
     - Random Resize and Crop: The images are either resized to certain scales with a max size of 1333 or undergo a sequence of operations: resize to [400, 500, 600], random size cropping

(min size 384, max size 600), and then resized again to scales with a max size of 1333.
    - Normalization: The images are normalized, possibly to a fixed mean and standard deviation typically used for pre-trained models.
- Loss Functions:
    - Auxiliary Decoding Losses: It is hinted that the code supports auxiliary decoding losses, which can be toggled with the --no_aux_loss argument. However, the specifics of this loss aren't detailed in the provided code snippet.
    - Matching Costs: The loss considers class, bounding box (using L1 loss), and Generalized Intersection over Union (GIoU) coefficients.
    - Additional Losses:
    - Mask Loss: Implied by the context but not shown in the code snippet.
    - Dice Loss: Implied by the context but not shown in the code snippet.
    - Bounding Box Loss: Bounding boxes are considered in the matching costs, likely using L1 loss.
    - GIoU Loss: As mentioned in the matching costs.
    - "No-object" Class Coefficient: Implied by the context but not shown in the code snippet.

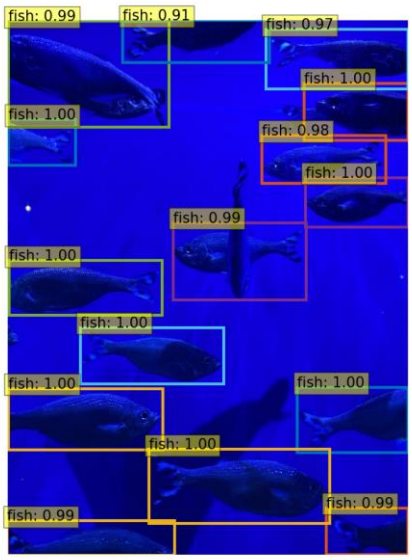
3. Table of your performance for validation set (AP, AP50, AP75)

Validation Index	Value
mAP	0.5233
mAP 50	0.8069
mAP 75	0.5387
mAP small	0.1738
mAP medium	0.4247
mAP large	0.6484

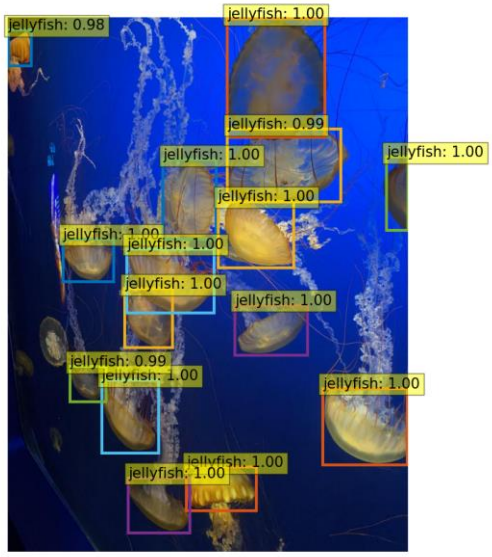
- Screenshot



4. Visualization



Result of `IMG_2277_jpeg_jpg.rf.86c72d6192da48d941ffa957f4780665.jpg`



Result of `IMG_2469_jpeg_jpg.rf.fca5db81cde8b6fe73b8f150e2e16a88.jpg`