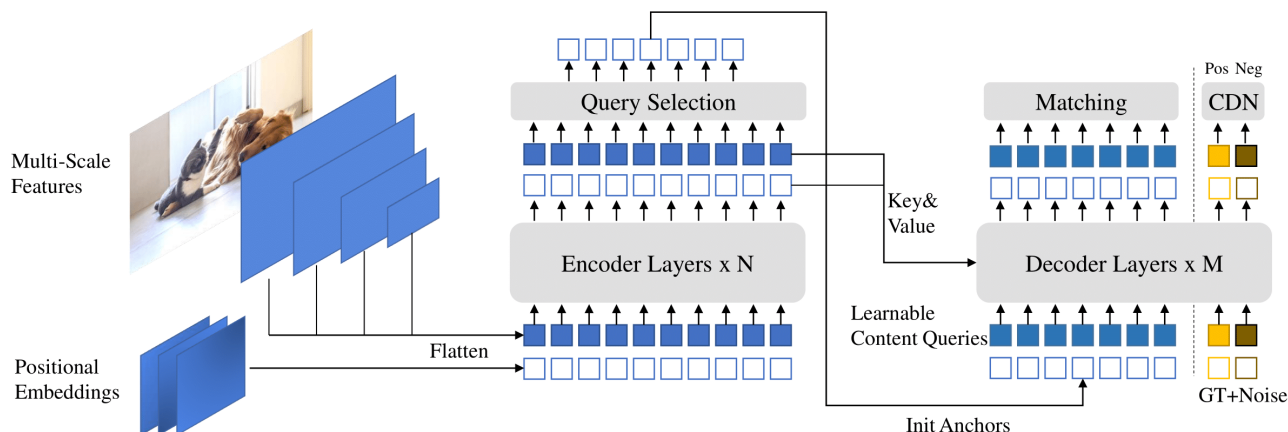


# CVPDL-HW1

## 1. Draw the architecture of your object detector

我使用的 object detector 叫做 "DINO", 來自這篇 [DINO: DETR with Improved DeNoising Anchor Boxes for End-to-End Object Detection](#)。他的架構如下圖所示：



主要可以分成 4 個部分：

1. **Backbone**：提取圖片的特徵。
2. **Transformer**：用來做 object detection 的主要部分，利用提取出來的圖片特徵透過 self-attention 來找出圖片中的物件。
3. **Query Selection**：DINO 使用 mixed query selection strategy，選擇 positions 作為 anchors，並且使用 learnable query embeddings 作為 content queries，選擇出最好的 query，並且利用這些 query 來做 object detection。
4. **Contrastive Denoising Training**：也就是圖中的 CDN，DINO 利用這個方式來提高 bounding box 的準確度，並且降低重複偵測的問題。

## 2. Implement details

1. **Augmentation**：縮放圖片大小、裁切圖片、水平翻轉圖片、將圖片轉成 tensor、將圖片 normalize。
2. **Loss function**：在定界框的預測上使用 L1 loss、GIOU loss；在類別的預測上使用 focal loss。
3. **Parameter settings**：我使用的 backbone 是 SwinL，batch size 是 1，learning rate 是 0.0001，optimizer 是 AdamW，weight\_decay 是 0.0001，訓練 12 個 epoch。其餘的參數都是使用 paper 中的參數。

## 3. Table of your performance for validation set

我用 DINO 本身的 evaluation code 來評估模型的表現，得到的結果如下：

AP	AP <sub>50</sub>	AP <sub>70</sub>
0.574	0.854	0.597

Average Precision (AP) @[ IoU=0.50:0.95   area= all   maxDets=100 ]	= 0.574
Average Precision (AP) @[ IoU=0.50   area= all   maxDets=100 ]	= 0.854
Average Precision (AP) @[ IoU=0.75   area= all   maxDets=100 ]	= 0.597

但我用助教給的 `evaluate.py` 得到的結果如下：

<b>AP</b>	<b>AP<sub>50</sub></b>	<b>AP<sub>70</sub></b>
0.582	0.88649	0.6089

```
device = cuda0 if torch.cuda.is_available() else -1 # device id: cuda0 or -1
100% |██████████████████████████████████████████████████████████████████████████████| 127/127 [00:06<00:00, 20.51it/s]
{'map': tensor(0.5852), 'map_50': tensor(0.8649), 'map_75': tensor(0.6089), 'map_small': tensor(0.2192), 'map_medium': tensor(0.4548), 'map_large': tensor(0.7220), 'mar_1': tensor(0.2676), 'mar_10': tensor(0.5884), 'mar_100': tensor(0.7014), 'mar_small': tensor(0.4108), 'mar_medium': tensor(0.6270), 'mar_large': tensor(0.8013), 'map_per_class': tensor(-1.), 'mar_100_per_class': tensor(-1.)}
```

## 4. Visualization

我是用 `IMG_2514_jpeg.jpg.rf.6ccb3859d75fc5cfe053b1c1474254b2.jpg` 這張圖片來做 visualization，結果如下：

