Algorithm 1 Optimal Transport Assignment (OTA)

Input:

I is an input image
A is a set of anchors
G is the gt annotations for objects in image I
γ is the regularization intensity in Sinkhorn-Knopp Iter.
T is the number of iterations in Sinkhorn-Knopp Iter.
α is the balanced coefficient in Eq. 2

Output:

π^* is the optimal assigning plan

$m \leftarrow |G|, n \leftarrow |A|$
$P^{\text{cls}}, P^{\text{box}} \leftarrow$ Forward(I,A)
$s_i(i = 1, 2, ..., m) \leftarrow$ Dynamic k Estimation
$s_{m+1} \leftarrow n - \sum_{i=1}^{m} s_i$
$d_j(j = 1, 2, ..., n) \leftarrow$ OnesInit
pairwise cls cost: $c_{\text{cls}}^{ij} = \text{FocalLoss}(P_j^{\text{cls}}, G_i^{\text{cls}})$
pairwise reg cost: $c_{\text{reg}}^{ij} = \text{IoULoss}(P_j^{\text{box}}, G_i^{\text{box}})$
pairwise Center Prior cost: $c_{ij}^{\text{cp}} \leftarrow (A_j, G_i^{\text{box}})$
$bg\ cls$ cost: $c_{\text{cls}}^{\text{bg}} = \text{FocalLoss}(P_j^{\text{cls}}, \varnothing)$
fg cost: $c^{\text{fg}} = c_{\text{cls}} + \alpha c_{\text{reg}} + c_{\text{cp}}$
compute final cost matrix c via concatenating $c_{\text{cls}}^{\text{bg}}$ to the last row of c^{fg}
$v^0, u^0 \leftarrow$ OnesInit
for t=0 **to** T **do**:
    $u^{t+1}, v^{t+1} \leftarrow \text{SinkhornIter}(c, u^t, v^t, s, d)$
compute optimal assigning plan π^* according to Eq. 11
return π^*

$$c^{fg} = c_{cls} + \alpha c_{reg} + c_{cp}$$

$$c_{ij} = L_{ij}^{cls} + \lambda L_{ij}^{reg}$$

Q1. OTA와 SimOTA의 차이가 정확하게 뭔지, Cost function의 차이인 건가?

Q2. Anchor free detectors은 object의 중심위치를 예측하여 경계선까지의 거리를 regression한다. 논문의 Multi positives section에서 anchor free는 object의 중심 위치만 선택하고 다른 양질의 예측들은 무시를 한다고했는데, 구체적으로 어떤 양질의 예측들을 무시하는지 궁금하다.

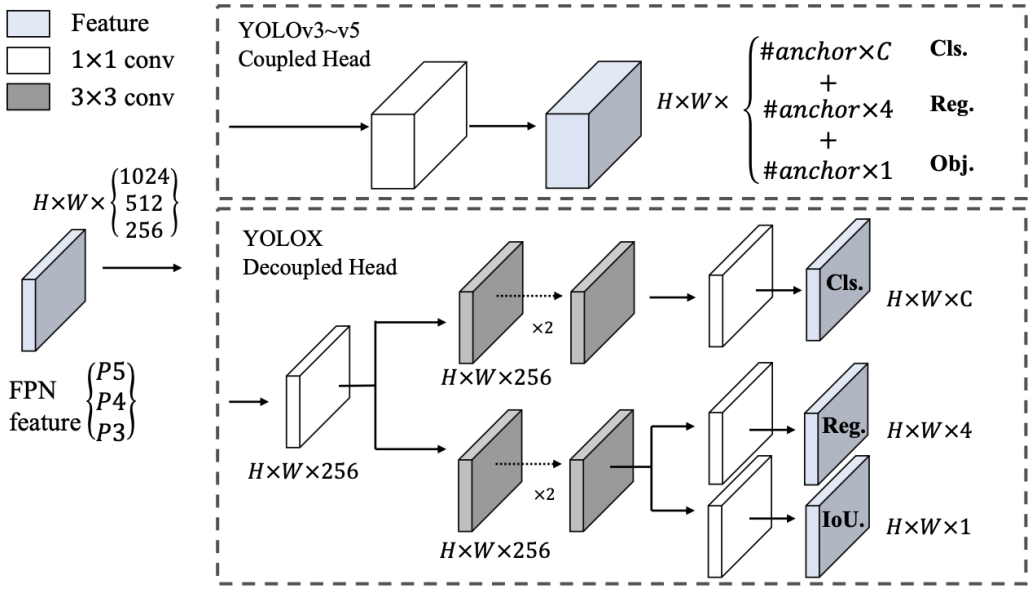

Figure 2: Illustration of the difference between YOLOv3 head and the proposed decoupled head. For each level of FPN feature, we first adopt a 1×1 conv layer to reduce the feature channel to 256 and then add two parallel branches with two 3×3 conv layers each for classification and regression tasks respectively. IoU branch is added on the regression branch.

Q3. 기존 Coupled Head 에서 발생한 문제점이 Regression 과 Classification 의 충돌이라고 하는데, 어떤 부분에서 충돌이 발생하는건지 궁금합니다.

Q4. Backbone은 어떤 방식으로 레이어가 결정되고 개발이 되는지?