$t=0$ (active sampling 전)

(randomly sampled. fully annotated w bounding boxes)

$I \rightarrow L_0$

$\quad \searrow U_0$

$\mathbb{M}(\theta_0)$: generic object detector

$t \geq 1$ (active sampling 후)

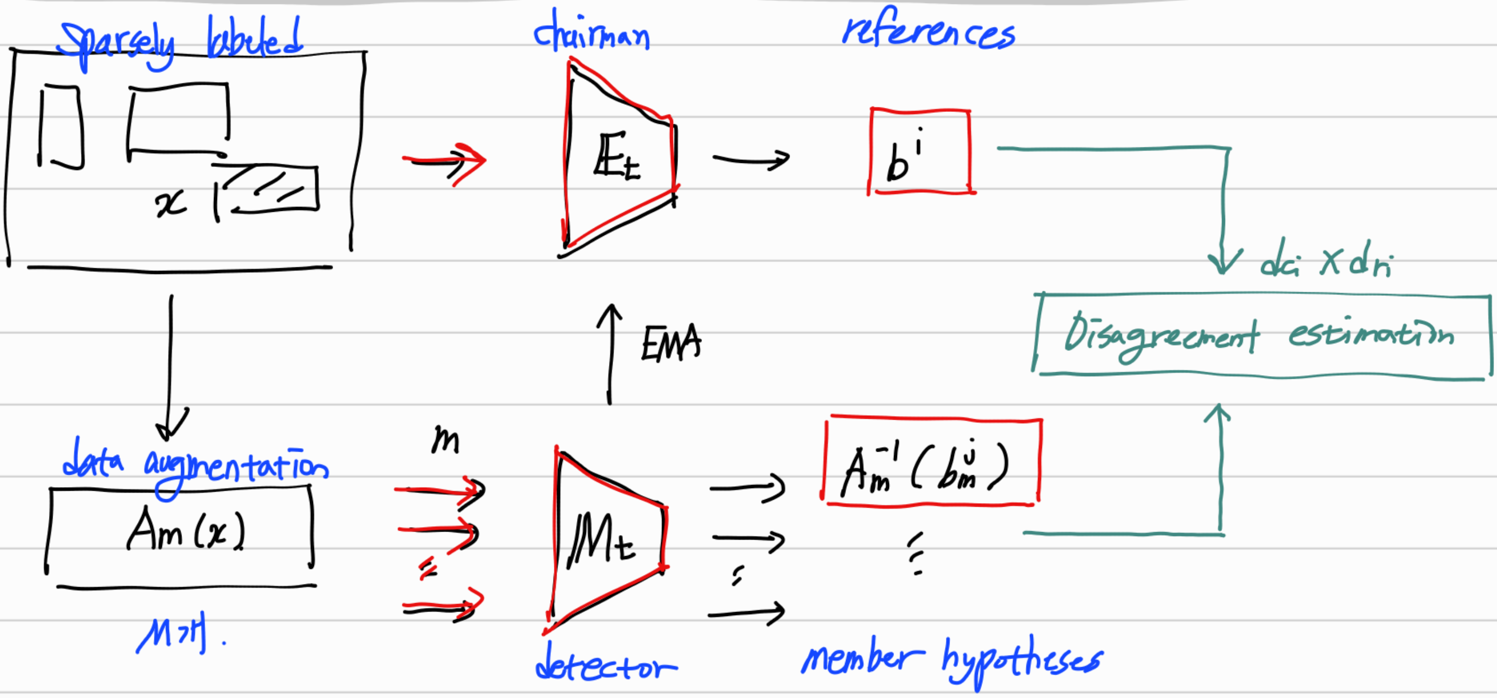
$I \rightarrow L_1 \rightarrow S_1$ (: sparsely labeled images)

$\quad \searrow U_1$

$\mathbb{M}(\theta_t)$ updated w labeled images $L_t \cup S_t$

or all images $L_t \cup S_t \cup U_t$

S_t : sparsely labeled images



chairman : $\mathbb{E}(\theta')$ of $\mathbb{M}(\theta)$: generate box references.

EMA detector

$$\theta_{tr}' = \alpha\, \theta_{tr-1}' + (1-\alpha)\cdot \theta_{tr}$$

(tr : training step within one cycle)

① - $\boxed{\text{Disagreement on classification.}}$

1. Given chairman이 예측한 box candidates : $\{b^i\}$,

member boxes $\{b_m^{ij}\}$ are assigned to each reference box in $\{b^i\}$

using detector - defined assignment strategy, such as max - IoU assigner.

2. Given matched pair of boxes $\boxed{\{b^i, b_m^{ij}\}}$,

classification disagreement 측정.

: CE b/w one - hot chairman prediction g^i & posterior predictive member distribution g^{ij}

$$\boxed{d_c^{ij} = -\mathbb{E}_{g^i}[\log g^{ij}]}$$

3. Disagreement about box b^i is aggregated among M committee members.

$$\boxed{d_c^i = \frac{1}{M}\cdot\sum_m^M \left(\frac{1}{k_{mi}}\sum_j^{k_{mi}} d_c^{ij}\right)}$$

(k_{mi} : # of positively matched member predictions in m-th stochastic view.)

(2)

– Disagreement on localization

1. $\{b^i, b_m^j\}$: matched pair of boxes.

2. $\{A_m^{-1}(b_m^j)\}$: inverse transformations on those boxes.

   fed into localization branch of chairman model E^{reg}.

3. Disagreement over the location of b^i is measured based on chairman re-calibrated boxes.

$$\boxed{d_r^i = \frac{1}{4} \cdot \sum_k^4 \hat{\sigma}_k\left(\{E^{reg}(A_m^{-1}(b_m^j))\}\right)}$$

localization task ⇒ 4개 regression task based on coordinates.

– Our scoring function for box-level detection task

$$d^i = d_c^i \times d_r^i$$

- Supervised loss for labeled images

· fully labeled images $\{x_\ell^i\}$ from L_t ; N_ℓ 개

classification loss

gt class label

localization loss function

$$\boxed{\mathcal{L}_\ell = \frac{1}{N_\ell} \sum_i \ell_{cls}(x_\ell^i, y_\ell^i) + \ell_{loc}(x_\ell^i, t_\ell^i)}$$

fully labeled image 개수. fully labeled image

corresponding box location

- Pseudo-active synergy for sparse images

sparse gt label

pseudo label

$$G(y_s, \hat{y}_{sc})$$

corresponding box

$$= y_s \cup \{ \hat{y}_{sc}^i \mid IOU(\hat{b}_{sc}^i, b_s^j) \leq \lambda_g, \forall b_s^j \in b_s \}$$

jaccard overlap

$G(t_s, \hat{t}_{sr})$ 도 마찬가지.

$\therefore$ supervision quality for sparse images.

$$\boxed{\mathcal{L}_s = \frac{1}{N_s} \cdot \sum_i^{N_s} \mathcal{L}_{cls}(A(x_s^i), G(y_s^i, \hat{y}_{sc}^i)) + \mathcal{L}_{loc}(A(x_s^i), G(t_s^i, \hat{t}_{sr}^i))}$$

sparsely labeled image 개수.