Supplementary Material for "Transductive Learning for Zero-Shot Object Detection"

Shafin Rahman*†, Salman Khan‡* and Nick Barnes*†
*Australian National University, †Data61-CSIRO, ‡Inception Institute of AI
firstname.lastname@anu.edu.au

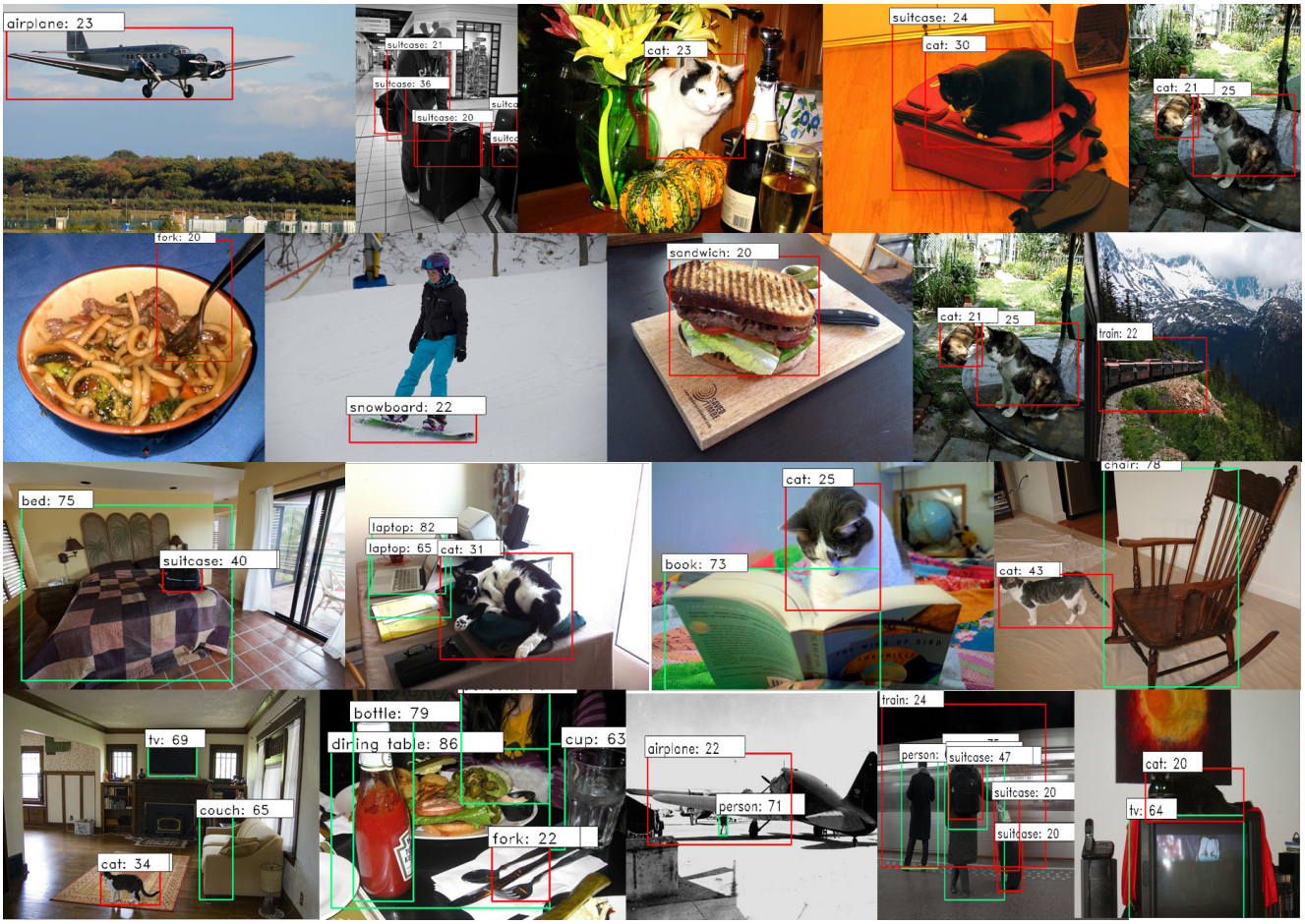

Figure 1: More qualitative results of ZSD (*top two rows*) and GZSD (*bottom two rows*). Red and green bounding boxes represent unseen and seen classes respectively.

1. Qualitative Results

In Fig. 1, we show more qualitative results of our approach.

2. Validation Experiment

To fix λ and t_h, we remove the term $L_d(u)$ and $L'_d(u)$ from Eq. 3 so that the loss becomes independent of η and β. Now, for different values of $t_h = \{.2, .3, .4\}$ and $\lambda = \{.1, .3, .5, .6, .8, 1\}$, we perform a grid search on traditional detection task. In Table 1(a), we report the results of such validation experiment where $t_h = .3$ and $\lambda = .2$ performs the best. The reason $t_h = .3$ works the best is that the same value was used during fixed pseudo-labeling for the seen classes. The hyper-parameter λ controls the balance between fixed and dynamic pseudo-labeling. When $\lambda = 1$, the network is trained with only fixed pseudo-labeling. Thus, traditional detection performs similarly

(38.25, 38.95, 39.48) across different t_h. Then, keeping all the chosen hyper-parameters fixed, we run another grid search for $\beta = \{.1, .3, .5, .7, .9, 1\}$ and $\eta = \{1, 2, 3, 4, 5\}$ on the same traditional detection task. β and η control the object/background imbalance of dynamic pseudo-labeling. Table 1(b) shows that $\beta = 0.1$ and $\eta = 1$ are the recommended values from our validation experiments.

(a) $L_d = L_d(s)$ case: Varying λ and t_h with $L_d(u) = L'_d(u) = 0$

$\lambda(\rightarrow)$	.2	.4	.5	.6	.8	1
t_h=.2	0.0	0.0	0.0	0.0	38.80	38.25
t_h=.3	**39.57**	37.51	24.91	0.7	37.87	38.95
t_h=.4	0.0	0.0	0.0	0.0	0.55	39.48

(b) $L_d = L_d(s) + L_d(u) + L'_d(u)$ case: Varying β and η

$\beta(\rightarrow)$	.1	.3	.5	.7	.9	1
η=1	**43.38**	41.06	39.14	40.98	41.73	38.81
η=2	41.59	39.05	39.40	37.65	41.47	39.98
η=3	40.66	41.13	41.89	40.47	39.65	39.87
η=4	40.64	40.48	40.00	40.15	42.42	39.82
η=5	40.22	36.73	41.73	37.85	32.94	40.94

Table 1: mAP scores of validation experiments. Selected hyper-parameters: $\lambda = .2$, $t_h = 0.3$, $\beta = .1$ and $\eta = 1$.