Supplementary Material: CLIP the Gap: A Single Domain Generalization Approach for Object Detection

Vidit Vidit[1] Martin Engilberge[1] Mathieu Salzmann[1,2]
CVLab, EPFL[1], ClearSpace SA[2]
firstname.lastname@epfl.ch

A.1. Prompt Curation

As mentioned in our main paper, we curate a set of domain prompts $\mathcal{P}^t$. For this step, we get the words matching the concept *weather* from WordNet [4]. Tab. A.3 lists the hyponyms for the word *weather*. With this list, we filter out words that have cosine distance below 0.5 w.r.t the *weather* in CLIP embedding space. We further improve our wordlist using GloVE [3] to weed out the uncommon words. To this end, we take only the words which are in the 10k most common words of GloVE. Tab. A.1 lists the words after this stage.

draft , easter , elements , depression ,
wave , quiet , snow ,
sunshine , fog , blast , freeze ,
breath , wind , atmosphere ,
rainfall , warming , trade ,
blow , calm , cyclone , rain , air , low , high

Table A.1. Word list after CLIP and GloVE [3] based pruning

We still see there are words that show the same concepts like *draft , easter , elements , depression , air , low , high, wind, atmosphere, air* represent related to the windy or stormy conditions, hence we replace these words with *stormy*. We merge *rainfall* and *rain*, while removing ambiguous words like *blast,warming,breath,quiet*.

A.2. Additional Dataset

We evaluate our approach on (a) Comic and (b) Watercolor [2] while training on VOC [1]. We chose to omit clipart as it has additional categories, while comic and watercolor share the same 6 categories. Our approach on (a) w/-aug. achieves mAP 33.5 vs w/o-aug 32.8, and on (b) w/-aug. mAP 43.4 vs w/o-aug. 42.7. The detector with just CLIP-init attains mAP of 26.2 and 41.9 on (a) and (b), respectively.

For the above experiments, we used the prompts $\mathcal{P}^t$: "*an image in* {*comic, painting, cartoon, digital-art, sketch, watercolor, oil painting*} *style*" and the source prompt p^s "*a realistic image*".

	mAP	
Method	Comic	Watercolor
CLIP-init	26.2	41.9
Ours w/o seg-aug	32.8	42.7
Ours	**33.5**	**43.4**

Table A.2. **VOC results**

A.3. Additional Experiments

The prompts used for the Clip-Random experiment (Sec. 4.5 in the main paper) are arbitrary. For completeness, we give the results of clip-random with another set of prompts: *an image of a cat*, *an image taken in front of a museum*, *a picasso painting*, *an image of the sky* and achieve similar results as in Tab. 7: 51.2(Day), 36.0(Night), 30.4(DuskRainy), 15.1(NightRainy), 37.7(Foggy). We expect similar results for any set of words unrelated to the target domain.

References

[1] M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman. The PASCAL Visual Object Classes Challenge 2012 (VOC2012) Results. http://www.pascal-network.org/challenges/VOC/voc2012/workshop/index.html. 1

[2] Naoto Inoue, Ryosuke Furuta, Toshihiko Yamasaki, and Kiyoharu Aizawa. Cross-domain weakly-supervised object detection through progressive domain adaptation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 5001–5009, 2018. 1

[3] Jeffrey Pennington, Richard Socher, and Christopher D Manning. Glove: Global vectors for word representation. In *Pro-*

scattering , current of air , draft , norther , line storm , doldrums , flurry
roughness , sou'wester , shower , gale , simoon , trade wind , boisterousness
deluge , air mass , southerly , easter , west wind , mizzle , whiff
cold wave , downfall , murk , line squall , elements , bad weather , easterly
sprinkling , balminess , depression , monsoon , cloudburst , chinook wind , good weather
fine spray , harmattan , clemency , tramontane , south wind , gentle breeze , wave
tramontana , draught , windlessness , equinoctial storm , fohn , overcast , cloud cover
quiet , moderate breeze , souther , thundershower , foehn , westerly , stillness
airstream , north wind , whiteout , snow eater , levanter , catabatic wind , clear-air turbulence
fair weather , snow , antitrades , simoom , bize , windiness , fresh gale
drizzle , sunshine , fog , puff of air , bluster , heat wave , light air
blast , zephyr , whole gale , temperateness , thawing , calmness , squall
storminess , prevailing westerly , soaker , snowfall , hot weather , scorcher , fresh breeze
raw weather , puff , rough water , northwest wind , murkiness , cold weather , gentle wind
northwester , freeze , crosswind , gust , mildness , breath , near gale
east wind , wind , northerly , rainstorm , chinook , atmosphere , rain shower
rainfall , high wind , pelter , snow flurry , southeaster , virga , jet stream
breeze , softness , warming , khamsin , atmospheric state , downpour , downdraft
cloudiness , trade , blow , inclementness , antitrade , strong gale , antitrade wind
calm , light breeze , tempestuousness , wester , lull , thaw , choppiness
cyclone , calm air , sleet , sandblast , sea breeze , mistral , air current
hail , southwester , rain , waterspout , sultriness , katabatic wind , torrent
precipitation , headwind , moderate gale , boreas , breeziness , tailwind , air
sou'easter , low , anticyclone , samiel , thermal , strong breeze , Santa Ana
bise , frost , high , sprinkle , inclemency , turbulence , fogginess

Table A.3. Hyponyms of word *weather* in WordNet [4]

ceedings of the 2014 conference on empirical methods in natural language processing (EMNLP), pages 1532–1543, 2014. 1

[4] Princeton University. About wordnet. https://wordnet.princeton.edu, 2010. 1, 2