# Supplemental Material
# OVANet: One-vs-All Network for Universal Domain Adaptation

In this appendix, we provide details of experiments and additional results.

## 1. Experimental Details

**Implementation.** The batch-size of source and target is 36 for all experiments in UNDA. The initial learning rate of networks is set as 0.01 for new layers and 0.001 for backbone layers in experiments on ResNet50. The learning rate is decayed with inverse learning rate decay scheduling. We follow [1] for setting the learning rate of VGGNet. We use Pytorch [4] to implement our method.

**Category Selection.** After deciding shared categories and source-private categories following default benchmark settings [3, 1], we pick the unknown classes in alphabetical order in Fig. 5 and 6. In Table 2, 3, and 5, we follow default benchmark settings [3, 1]. In Table 6 and Fig. 10(b), which use NAbird, we set the first 300 classes as known and the rest as unknown. In Fig. 9(b) and Fig. 10(a), known and unknown classes of DomainNet are also picked in alphabetical order. For CIFAR10 and CIFAR100 (Table 7), we pick the first $N$ categories as known and rest as unknown, where $N$ denotes the number of known classes. We decide the order of categories of provided label-indexes.

**Semi-superivised Learning.** The implementation of the experiment is the same as UNDA, where we use ResNet50 as a backbone. We simply replace the source with labeled data and the target with unlabeled data.

**Open-Set Recognition.** In this experiment, models are trained from scratch, where we employ WideResNet [8]. We follow the implementation of Fixmatch [6] to train the models. Training is performed similarly to the experiments of OVAN w/o OEM except that we add contrastive loss.

Then, the overall training loss can be computed as follows:

$$\mathcal{L}_{all} = \mathop{\mathbb{E}}_{(\mathbf{x}_i, y_i)\sim\mathcal{D}} \mathcal{L}_{src}(\mathbf{x}_i, y_i) + \lambda \mathop{\mathbb{E}}_{\mathbf{x}_i\sim\mathcal{D}} \mathcal{L}_{simc}(\mathbf{x}_i), \quad (1)$$

$$\mathcal{L}_{src}(\mathbf{x}_i^s, {y_i}^s) = \mathcal{L}_{cls}(\mathbf{x}_i, y_i) + \mathcal{L}_{ova}(\mathbf{x}_i, y_i), \quad (2)$$

where $\mathcal{L}_{simc}$ denotes the instance discrimination loss proposed in SimCLR [2]. We set $\lambda$ as 0.1. We gladly make the code of open-set recognition available when acceptance.

## 2. Additional Results of UNDA

We provide additional results excluded from the main paper due to limited space.

**Open-partial DA on OfficeHome.** We show the result of open-partial DA using OfficeHome in Table A. OVAN outperforms baselines with a large margin in all scenarios.

**Detailed metrics on Office.** In Table B, the accuracy of all samples (ALL), H-score, and accuracy to reject unknown samples (UNK) are described for open-set setting on Office.

**AUROC of Entropy and OVANet.** In experiments of the main paper, we show that OVANet provides good thresholds to decide whether an input sample comes from known or unknown categories. In this analysis, we investigate whether the output of our open-set classifier separates known and unknown samples well. Although our focus is to build a method that determines the threshold well, this analysis will be useful to better understand our method. Here, we compare the entropy of classification output, the value of the predicted category's softmax output, and our one-vs-all classifier's output. To ablate the effect from unlabeled target samples, we train a model without open-set entropy minimization. Table C shows AUROC of each metric. Although the output of OVANet is not always better at calibrating uncertainty of the output than Entropy, OVANet often outperforms the softmax output.

## References


[1] Silvia Bucci, Mohammad Reza Loghmani, and Tatiana Tommasi. On the effectiveness of image rotation for open set domain adaptation. In *Eur. Conf. Comput. Vis.*, pages 422–438. Springer, 2020. 1

[2] Ting Chen, Simon Kornblith, Mohammad Norouzi, and Geoffrey Hinton. A simple framework for contrastive learning of visual representations. In *Int. Conf. Mach Learn.*, pages 1597–1607. PMLR, 2020. 1

[3] Bo Fu, Zhangjie Cao, Mingsheng Long, and Jianmin Wang. Learning to detect open classes for universal domain adaptation. In *Eur. Conf. Comput. Vis.* Springer, 2020. 1, 2

[4] Adam Paszke, Sam Gross, Soumith Chintala, Gregory Chanan, Edward Yang, Zachary DeVito, Zeming Lin, Alban Desmaison, Luca Antiga, and Adam Lerer. Automatic differentiation in pytorch. *arXiv*, 2017. 1

| Method | OfficeHome (10 / 5 / 50) A2C | A2P | A2R | C2A | C2P | C2R | P2A | P2C | P2R | R2A | R2C | R2P | Avg |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| OSBP [5] | 39.6 | 45.1 | 46.2 | 45.7 | 45.2 | 46.8 | 45.3 | 40.5 | 45.8 | 45.1 | 41.6 | 46.9 | 44.5 |
| UAN [7] | 51.6 | 51.7 | 54.3 | 61.7 | 57.6 | 61.9 | 50.4 | 47.6 | 61.5 | 62.9 | 52.6 | 65.2 | 56.6 |
| CMU [3] | 56.0 | 56.9 | 59.1 | 66.9 | 64.2 | 67.8 | 54.7 | 51.0 | 66.3 | 68.2 | 57.8 | 69.7 | 61.6 |
| OVANet | **62.8** | **75.6** | **78.6** | **70.7** | **68.8** | **75.0** | **71.3** | **58.6** | **80.5** | **76.1** | **64.1** | **78.9** | **71.8** |

Table A: **Open-partial domain adaptation on OfficeHome (H-score)**.

| A2D | | | A2W | | | D2A | | |
|---|---|---|---|---|---|---|---|---|
| All | H-Score | UNK | All | H-Score | UNK | All | H-Score | UNK |
| 88.3 | 90.5 | 88.0 | 87.4 | 88.3 | 86.6 | 87 | 86.7 | 95.4 |

| W2A | | | W2D | | | D2W | | |
|---|---|---|---|---|---|---|---|---|
| All | H-Score | UNK | All | H-Score | UNK | All | H-Score | UNK |
| 88.5 | 88.3 | 91.5 | 98.3 | 98.4 | 96.9 | 97.5 | 98.2 | 98.1 |

Table B: **Open-set domain adaptation on Office (10 / 0 / 11).**. ResNet50 is used as a backbone.

| Confidence Score | Office (10 / 0 / 11) A2D | D2W | W2A | OfficeHome (10 / 5 / 50) R2A | R2C | A2P | P2R | DNet R2P (150 / 50 / 145) |
|---|---|---|---|---|---|---|---|---|
| Softmax | 87.6 | 97.6 | 85.9 | 80.9 | 67.8 | 81.3 | 88.2 | 65.5 |
| Entropy | 88.3 | 97.1 | 86.2 | 82.1 | 68.5 | 82.4 | 88.7 | 66.8 |
| One-vs-All Classifier | 90.1 | 97.3 | 87.3 | 81.4 | 67.5 | 82.9 | 88.9 | 67.3 |

Table C: **AUROC score.** The score measures how well known and unknown samples are separated.


[5] Kuniaki Saito, Shohei Yamamoto, Yoshitaka Ushiku, and Tatsuya Harada. Open set domain adaptation by backpropagation. In *Eur. Conf. Comput. Vis.*, 2018. 2

[6] Kihyuk Sohn, David Berthelot, Chun-Liang Li, Zizhao Zhang, Nicholas Carlini, Ekin D Cubuk, Alex Kurakin, Han Zhang, and Colin Raffel. Fixmatch: Simplifying semi-supervised learning with consistency and confidence. *arXiv preprint arXiv:2001.07685*, 2020. 1

[7] Kaichao You, Mingsheng Long, Zhangjie Cao, Jianmin Wang, and Michael I. Jordan. Universal domain adaptation. In *IEEE Conf. Comput. Vis. Pattern Recog.*, 2019. 2

[8] Sergey Zagoruyko and Nikos Komodakis. Wide residual networks. *arXiv preprint arXiv:1605.07146*, 2016. 1