PROCEEDINGS OF SPIE

SPIEDigitalLibrary.org/conference-proceedings-of-spie

A multi-source transfer learning-based weighted network for small sample defect inspection

Yue Hong, Chen Li, Haoxin Yan, Erik Valle , Yi Dou, et al.

Yue Hong, Chen Li, Haoxin Yan, Erik Valle , Yi Dou, Chengwei Liao, Peiyuan Zhu, Shidong Zhu, Xiu Li, Xiang Qian, Xiaohao Wang, Xinghui Li, "A multi-source transfer learning-based weighted network for small sample defect inspection," Proc. SPIE 12317, Optoelectronic Imaging and Multimedia Technology IX, 123170F (4 January 2023); doi: 10.1117/12.2644054

SPIE.

Event: SPIE/COS Photonics Asia, 2022, Online Only

A multi-source transfer learning-based weighted network for small sample defect inspection

Yue Hong[a], Chen Li[b], Haoxin Yan[a], Erik Valle[b], Yi Dou[c], Chengwei Liao[d], Peiyuan Zhu[e], Shidong Zhu[e], Xiu Li[a], Xiang Qian[a], Xiaohao Wang[a], and Xinghui Li[a,b,*]

[a]Shenzhen International Graduate School, Tsinghua University, Shenzhen 518055, China
[b]Tsinghua-Berkerley Shenzhen Institute, Tsinghua University, Shenzhen 518055, China
[c]Guangxi Guihua Intelligent Manufacturing Co., Ltd., Nanning 530022, China
[d]Guilin Han Smart Instrument Co.,Ltd., Guilin 541004, China
[e]Shenzhen Han Industrial Technologies Co., Ltd., Shenzhen 518055, China
[*]Corresponding author: li.xinghui@sz.tsinghua.edu.cn

ABSTRACT

Defect inspection is indispensable process in manufacturing, and automatic optical inspection (AOI) has been rapidly applied to various areas. In AOI, artificial intelligence (AI) based deep learning methods are more and more advantageous in many fields. However, obtainment of high-performance deep learning algorithms always requires a large amount of training data, while defect samples are often scarce. So small sample has become one of the key problems in the industrial application of deep learning algorithms. Transfer learning enable us to utilize the knowledge of source domains to improve performance on target domain, which could be used to tackle the small sample problem intuitively. Therefore, this paper proposes a defect inspection network which is based on one of the transfer learning techniques: domain adaptation. We name the network as multi-source and multi-scale weighted domain adaptation network which is based on adversarial learning. Firstly, three adversarial domain adaptation modules are proposed to align feature distributions between multi-source domains and target domain under three scales, which make the backbone extract domain-invariant features. Simultaneously, the weights of domain adaptation module under each scale are set reasonably. Secondly, in order to reduce the effect of negative transfer, a novel similarity weight is proposed, which is applied on domain adaptation modules. Finally, experiments are carried out to prove the effectiveness of our method. The results show that our method can improve the mean average precision(mAP) from 62.3 to 78.5 in the case of 40 samples available for 4 defect categories, which surpasses other counterparts.

Keywords: Automatic optical inspection (AOI), Artificial Intelligence (AI), machine vision, transfer learning, multi-source domain adaptation, adversarial training

1. INTRODUCTION

Surface defects are usually areas of uneven local properties on the product surface. Any abnormality in the production process will lead to product defects. Surface defect is one of the vital factors which affect product quality. It not only affects the appearance of products, but also reduces the service performance. Meanwhile, no matter how the production process develops, defects is inevitable. So surface defect inspection become indispensable process in industrial product manufacturing.

In the beginning, manual inspection is the only choice. But it has many drawbacks such as low speed, low accuracy, relying on the experience of workers and high physical burden on workers, etc. In recent years, automatic optical inspection (AOI) has been rapidly applied to various areas, especially defect inspection. AOI could be divided into two categories: methods based on traditional image processing and methods based on deep learning. The former classifies preprocessed images based on handcraft features. For instance, Zhu *et al.*[1] detect yarn-dyed fabric defect based on autocorrelation function and GLCM. However, handcraft features are designed for specific defect and condition, which means it fails when light condition, background and defect vary. As the research on deep learning further develops, methods based on deep learning show great potential in AOI.

Optoelectronic Imaging and Multimedia Technology IX, edited by Qionghai Dai,
Tsutomu Shimura, Zhenrong Zheng, Proc. of SPIE Vol. 12317, 123170F

Deep neural network can automatically extract features and useful information, avoiding explicit feature design. Nakazawa, Kulkarni *et al.*[2] and Cheon *et al.*[3] use CNN for wafer defects.

At present, most of the deep learning algorithms are based on supervised learning, which generally require a large amount of labeled data for training. And with such big data, supervised learning methods usually surpass unsupervised learning and semi-supervised learning algorithms in performance. Large number of samples are easy to obtain in some fields, while not in most cases. Surface defect is a typical case. Samples with defect are much scarcer than normal samples, so it takes much more time to collect sufficient samples for training, which is called small sample problem. There are some methods to solve the problem. Data augmentation technic can directly increase the number of defect samples. Haselmann *et al.*[4] inject synthetized defects into fault-free surface images for training CNN. Unsupervised or semi-supervised learning algorithms can reduce the reliance on data. He *et al.*[5] designed a steel surface defect inspection system composed of sample generation and semi-supervised learning. Some works[6,7] focus on optimizing the structure of network. Nowadays transfer learning become one of the promising methods to solve the small sample problem.

Transfer learning can utilize knowledge learned from one task to boost the other different but related task, which inspires researchers to use related samples or models to solve the small sample problem. Zyout *et al.*[8] combines transfer learning and Alex Net to inspect PV solar panel surface defects. Zhu *et al.*[9] and Volkau *et al.*[10] applied transfer learning and VGG network to the defect inspection of emulsion pump body and printed circuit board respectively. In this paper, our task is to improve the inspection precision on scarce target samples with adequate similar samples. At present, the exploration of transfer learning in the field of defect inspection is still little, and most of them stay in the pretrain-finetune method and single-source transfer learning method.

Except pretrain-finetune, we use feature-based domain adaptation techniques to solve the small sample problem. This article proposes a multi-source and multi-scale weighted domain adaptation network. In detail, we use YOLOX[11] as main framework, and design a novel domain adaptation module, which is inserted into different part of backbone, to make the backbone extract domain invariant feature. Besides, the module allows multi-source domain data to input, which could take full advantage of all data and hence more meaningful for real-world application. Moreover, to reduce the effect of negative transfer, we design a novel similarity weight as a part of the domain adaptation module. At last, experiments are carried out to prove the effectiveness of our method.

2. PROPOSED METHOD

Multi-source domain adaptation (MSDA) assumes that data is collected from more than one domain which under different distributions. In MSDA, N source domain S_1,S_2,...,S_N and one target domain T are given.$S_j = \left\{(x_i^{S_j}, y_i^{S_j})\right\}_{i=1}^{N_{S_j}}$, N_{S_j} is the number of domain S_j samples. $T = \{(x_i, y_i)\}_{i=1}^{N_T}$. N_T is the number of domain T samples. The distribution of domains are different, referred as domain shift. In common domain adaptation task, the label spaces of all domains are identical. The ultimate goal of MSDA is to train the model with samples from muti-source domains and target domain, and achieve fairly good performance on target domain samples.

2.1 Network Architecture

As mentioned before, we take the YOLOX as the baseline in this work. YOLO series are representative single-stage detectors, and YOLOX shows broad compatibility in practical scenes. Consider the actual demand of speed and complexity of features, we choose YOLOX-S as baseline.

The network architecture of our proposed model is shown in Fig. 1. The Feature extraction is mainly done in the backbone. Therefore following the principle of feature-based domain adaptation, we design a domain adaptation module which could make the backbone extract domain-invariant feature. To better align the source domains and target domain, we embed the domain adaptation modules in three branches of backbone which output three scales of feature. Since there could be significant difference between samples from different domains, the negative transfer should not be neglected. We design similarity weight in each domain adaptation module which measure the similarity of one source domain sample with target domain.

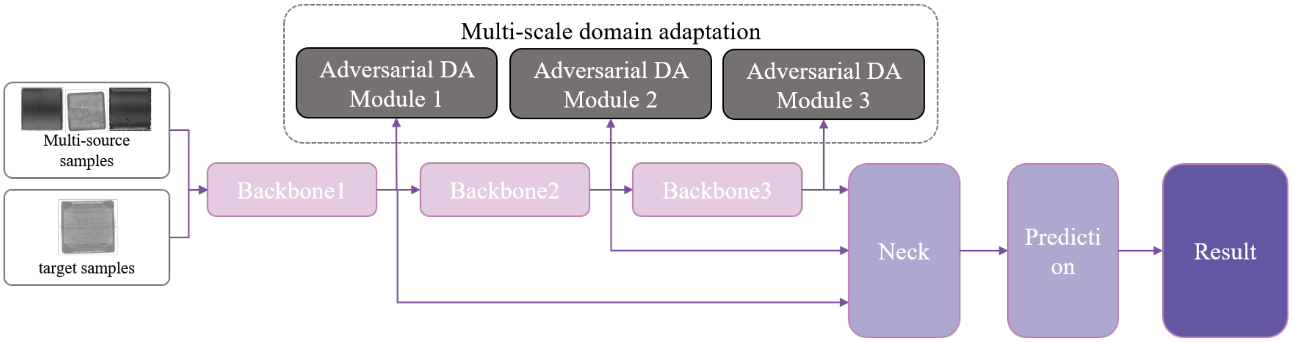


Figure 1. The network architecture of our proposed domain adaptation model.

2.2 Domain Adaptation Module

There are many ways to reduce the domain shift. Adversarial nets become one of the promising methods these years. Further inspired by Ref. 12 which study partial domain adaptation, we propose an adversarial domain adaptation module to achieve weighted domain adaptation. The main structure is of the adversarial domain adaptation module is shown in Fig. 2. Following the sets in Ref. 12, the adversarial domain adaptation module contains one gradient reversal layer (GRL) and two domain classifiers. Domain classifier D measures the similarity weight of each sample while Domain classifier D0 is the one which plays the minimax game and make backbone learn invariant features. GRL changes the sign of gradient in backward propagation which enables the minimax game work simultaneously.

A common adversarial loss is calculated by:

$$\min_F \max_D L(D,F) = \mathbb{E}_{x\sim p_s(x)}[\log D(F(x))] + \mathbb{E}_{x\sim p_t(x)}[\log(1-D(F(x)))]\,, \tag{1}$$

In our multi-source domain adaptation scenario, the loss L_s is used to reduce the domain shift under single-scale:

$$\min_F \max_{D0} L_s(D0,F) = -\sum_i^N \mathrm{CEL}(D0(F(x_i)), y_{d_i})\,, \tag{2}$$

where y_d is the domain label of sample. CEL denotes cross entropy loss.

The loss L_m of the multi-scale domain adaptation are calculated by:

$$L_m = w_1 \cdot L_{s1} + w_2 \cdot L_{s2} + w_3 \cdot L_{s3} \tag{3}$$

The loss L_d is similar with L_s, but L_d only optimizes the parameters of domain classifier D:

$$\max_D L_d(D,F) = -\sum_i^N \mathrm{CEL}(D(F(x_i)), y_{d_i}) \tag{4}$$

The domain classifier D gives the similarity weight of each domain, which will be detailed behind. Thus, the objective function of weighted L_s can be reformulated as:

$$\min_F \max_{D0} L_s(D0,F) = -\sum_i^N w_D \cdot \mathrm{CEL}(D0(F(x_i)), y_{d_i})\,, \tag{5}$$

where w_D is the similarity weight of each source domain , which will be described in next section.

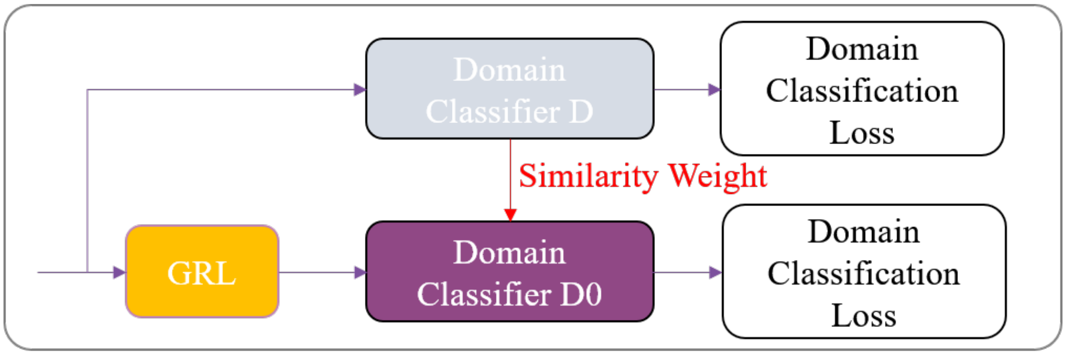


Figure 2. The main structure is of adversarial domain adaptation module.

2.3 Similarity Weight

The domain classifier D identifies the domain of each sample. Both the two domain classifiers output $N+1$ scores for each input sample p_{S1}, p_{S2},..., p_{SN}, p_T. $N+1$ denotes the number of all domains. Each score represents the likelihood of the sample belonging to one specific domain. For example, p_{Sj} denotes the likelihood of this sample coming from the j^{th} source domain distribution.

Transfer learning requires that the source domain and target domain to be related, or it will lead to a worse result on target task, i.e., negative transfer. In realistic scenario, it is quite common that some source domains are similar with target domain, while some aren't. For multi-source domain adaptation, if dissimilar and similar domains are given the same weight, it is conceivable that the result of transfer learning could be terrible. Hence, we propose a novel similarity weight to maximumly mitigate the negative transfer.

Domain similarities p_{S1}, p_{S2},..., p_{SN} measure the similarity of source domain and target domain. p_{Si} denotes the similarity of i^{th} source domain and target domain. As mentioned before, domain classifier D outputs the likelihood of samples coming from each domain distribution. Target samples are no exception. For each target samples, the D gives the $N+1$ scores $p_{S1}, p_{S2}, p_{S3}, ..., p_{SN}, p_T$, among which the first N scores denote likelihood of samples coming from n source domains. We take them to represent the similarity of each source domain and target domain.

And the domain similarity weight is given by the normalization of p_{Sj}:

$$w_{DS_j} = \frac{p_{Sj}}{\sum p_S} \tag{6}$$

If the domain classifier D give target samples a high score on source domain j, it means that D can't distinguish target domain from source domain j. That is source domain j is similar with target domain, and should be assigned high weight. Conversely, if p_{Si} $(i \neq j)$ of target samples is low, the domain i should be assigned a low weight to reduce the negative transfer effect. For target domain, the weight w_{DT} is set to:

$$w_{DT} = 1.0 \tag{7}$$

Then w_D is composed of w_{DS} and w_{DT}.

2.4 Network Optimization

The training loss of proposed network consists of three parts: detection loss of YOLOX, the loss of domain classifier L_d and the weighted multi-scale domain adaptation loss L_m. The overall objective function is written as:

$$L = L_{det} + w_d \cdot L_d \ + w_m \cdot L_m \tag{8}$$

Where w_d and w_m are the hyper-parameter balance the three losses. The stochastic gradient descent (SGD) algorithm is used to optimize the parameters of the proposed network.

3. EXPERIMENTS AND RESULTS

3.1 Implementation Details

Both source domain and target domain are annotated with bonding boxes and corresponding categories. The parameter w_d and w_m are set as 0.8 and 2.4 respectively. We train the network with a learning rate of 0.0004 for 50k iterations. Since our datasets consists of 3 source domains and 1 target domain, the batch size is set to 4.

All experiments are implemented with the Pytorch platform on Nvidia RTX 3090 GPU.

3.2 Dataset and Evaluation

We mainly did our experiment on Magnetic Tile Dataset, which is provided by corporation. Magnetic Tile Dataset is collected from the real production line. And it contains several sub-datasets which are collected from different manufacturers or stations. Sub-datasets are different on appearance, but they are related on distribution. Some crack samples of each domain are shown in Fig. 3. We pick three of the sub-datasets as source domain datasets, and one of them as target domain dataset. It contains four main defects: impurity, crack, fold, broken, some of the typical samples are shown in Fig. 4.

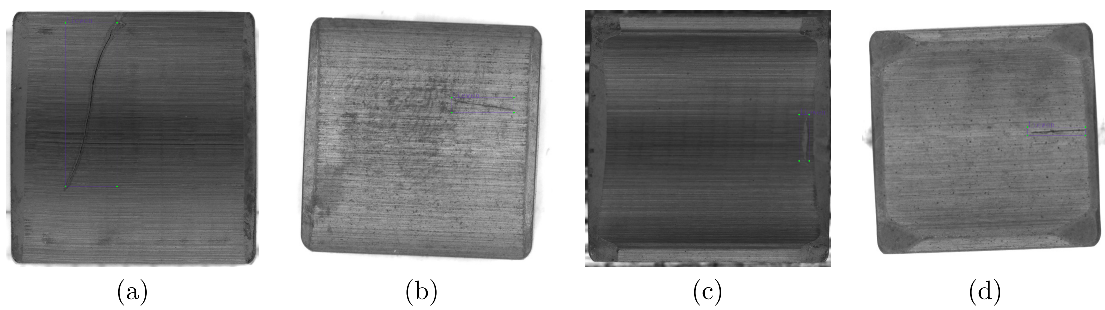


Figure 3. Crack samples of each domain. Images from left to right are from domain 1$\sim$4 respectively.

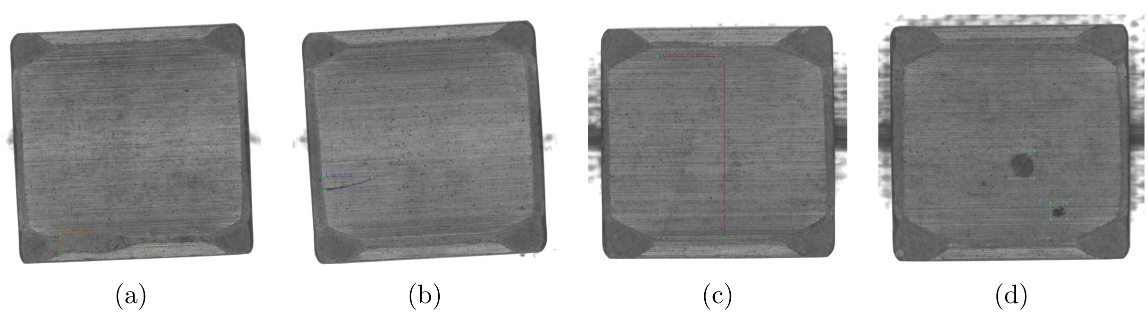


Figure 4. Typical samples of different defect. Images from left to right shows four main defect respectively: broken, crack, fold, impurity.

In our experiment, three source domain datasets have 1230, 1089 and 1263 images respectively, and there are only 40 images for training in target dataset to simulate the small sample problem. In each sub-dataset, the image number of each defect category is balanced.

We use COCO index to evaluate the results and set the IoU threshold to 0.5, so that AP50 is our main evaluation index. AP50 represents the average precision in all categories with the thresholds of 0.5.

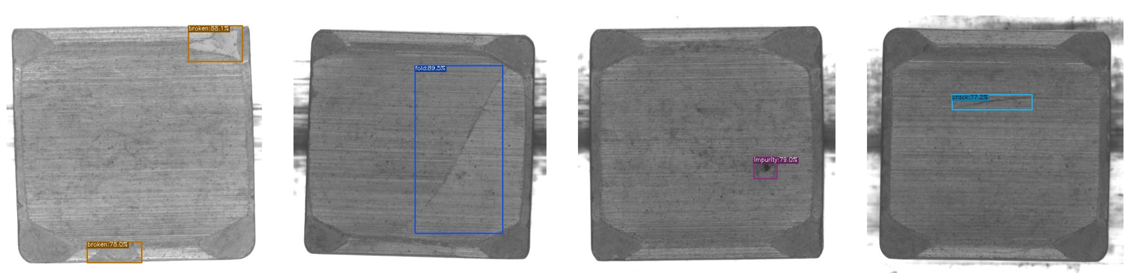
3.3 Main Results

In this section, our proposed method is compared with the baseline and existing methods of domain adaptive object detection. The YOLOX model was trained without domain adaptation. In the experiment by the model of YOLOX (1 src. only), only one source domain images are inputted for training. Since there are three source domains, we take the best result as representative. YOLOX (3 src. only) means all three source domains are combined to a single domain. In the experiment of YOLOX (3 src. + tag. 40), all three source domain and one target domain are combined to one domain for training, which is the same as the proposed method. YOLOX (tag. 40 only) experiment only use 40 target samples to show the effect of small sample problem. Besides, we take two state-of-the-art domain adaptation-based defect inspection methods: sa-da-faster[13] , SW Detection[14] for comparison.

The results are shown in Tab. 1. Our method achieves 78.5 on AP50, outperforming other baselines by a large margin. The AP50 result of the baseline [YOLOX (3 src. + tag. 40)] is 65.1, which is 13.4 points lower than our method. And our method also gets better performance than other domain adaptation method, which shows the power of weighted multi-source domain adaptation. Some of the detection results of our proposed method is shown in Fig. 5.

Table 1. Experiment results of each method.

Method	AP50
YOLOX (1 src. only)	52.7
YOLOX (3 src. only)	63.2
YOLOX (3 src. + tag. 40)	65.1
YOLOX (tag. 40 only)	62.3
sa-da-faster	70.1
SW Detection	71.4
our method	78.5



(a) (b) (c) (d)

Figure 5. Typical detection results of our proposed method.

4. CONCLUSION

In this paper, we present a multi-source transfer learning-based weighted network for defect inspection. With our proposed method, the demand for large-scale dataset is alleviated to a great extent. We take YOLOX as the baseline, while our method could be inserted into other mainstream detection frameworks. Adversarial domain adaptation modules are proposed to align feature distributions between multi-source domains and target domain, which make the backbone extract domain-invariant features. We insert three adaptation modules in three parts of the backbone to addressing the shift of data distribution under multi scales. In order to reduce the effect of negative transfer, a novel similarity weight is proposed. Experimental results on the real-scene defect dataset:

Magnetic Tile Dataset show that our proposed method achieves a good performance in small sample defect inspection.

ACKNOWLEDGMENTS

This work was supported in part by the Shenzhen Stable Supporting Program (Grant No. WDZC20200820200655 001), in part by the Interdisciplinary Foundation of Shenzhen International Graduate School of Tsinghua University (Grant No. JC2021003).

REFERENCES

[1] Zhu, D., Pan, R., Gao, W., and Zhang, J., "Yarn-dyed fabric defect detection based on autocorrelation function and glcm," *Autex research journal* **15**(3), 226–232 (2015).

[2] Nakazawa, T. and Kulkarni, D. V., "Wafer map defect pattern classification and image retrieval using convolutional neural network," *IEEE Transactions on Semiconductor Manufacturing* **31**(2), 309–314 (2018).

[3] Cheon, S., Lee, H., Kim, C. O., and Lee, S. H., "Convolutional neural network for wafer surface defect classification and the detection of unknown defect class," *IEEE Transactions on Semiconductor Manufacturing* **32**(2), 163–170 (2019).

[4] Haselmann, M. and Gruber, D. P., "Pixel-wise defect detection by cnns without manually labeled training data," *Applied Artificial Intelligence* **33**(6), 548–566 (2019).

[5] He, Y., Song, K., Dong, H., and Yan, Y., "Semi-supervised defect classification of steel surface based on multi-training and generative adversarial network," *Optics and Lasers in Engineering* **122**, 294–302 (2019).

[6] Schlegl, T., Seeböck, P., Waldstein, S. M., Schmidt-Erfurth, U., and Langs, G., "Unsupervised anomaly detection with generative adversarial networks to guide marker discovery," in [*International conference on information processing in medical imaging*], 146–157, Springer (2017).

[7] Schlegl, T., Seeböck, P., Waldstein, S. M., Langs, G., and Schmidt-Erfurth, U., "f-anogan: Fast unsupervised anomaly detection with generative adversarial networks," *Medical image analysis* **54**, 30–44 (2019).

[8] Zyout, I. and Oatawneh, A., "Detection of pv solar panel surface defects using transfer learning of the deep convolutional neural networks," in [*2020 Advances in Science and Engineering Technology International Conferences (ASET)*], 1–4, IEEE (2020).

[9] Zhu, C., Zhou, W., Yu, H., and Xiao, S., "Defect detection of emulsion pump body based on improved convolutional neural network," in [*2019 International Conference on Advanced Mechatronic Systems (ICAMechS)*], 349–352, IEEE (2019).

[10] Volkau, I., Mujeeb, A., Wenting, D., Marius, E., and Alexei, S., "Detection defect in printed circuit boards using unsupervised feature extraction upon transfer learning," in [*2019 International Conference on Cyberworlds (CW)*], 101–108, IEEE (2019).

[11] Ge, Z., Liu, S., Wang, F., Li, Z., and Sun, J., "Yolox: Exceeding yolo series in 2021," *arXiv preprint arXiv:2107.08430* (2021).

[12] Zhang, J., Ding, Z., Li, W., and Ogunbona, P., "Importance weighted adversarial nets for partial domain adaptation," in [*Proceedings of the IEEE conference on computer vision and pattern recognition*], 8156–8164 (2018).

[13] Chen, Y., Li, W., Sakaridis, C., Dai, D., and Van Gool, L., "Domain adaptive faster r-cnn for object detection in the wild," in [*Proceedings of the IEEE conference on computer vision and pattern recognition*], 3339–3348 (2018).

[14] Saito, K., Ushiku, Y., Harada, T., and Saenko, K., "Strong-weak distribution alignment for adaptive object detection," in [*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*], 6956–6965 (2019).