



INTERNATIONAL JOURNAL ON INFORMATICS VISUALIZATION

journal homepage : www.joiv.org/index.php/joiv



Data-Centric Learning Method for Synthetic Data Augmentation and Object Detection

Yeseong Park^a, Hyeongbok Kim^a, Yoon Jung Park^a, Changsin Lee^a, Jinsuk Lee^{a,*}

^a Testworks, AI Research, 42-19, Baekjegobun-ro 41-gil, Songpa-gu, Seoul, 05623, Republic of Korea

Corresponding author: *jslee@testworks.co.kr

Abstract— This paper proposes a deep learning framework for decreasing large-scale domain shift problems in object detection using domain adaptation techniques. We have approached data-centric domain adaptation with Image-to-Image translation models for this problem. It is one of the methodologies that changes source data to target domain's style by reducing domain shift. However, the method cannot be applied directly to the domain adaptation task because the existing Image-to-Image model focuses on style translation. We solved this problem using the data-centric approach simply by reordering the training sequence of the domain adaptation model. We defined the features to be content and style. We hypothesized that object-specific information in images was more closely tied to the content than the style and thus experimented with methods to preserve content information before style was learned. We trained the model separately only by altering the training data. Our experiments confirmed that the proposed method improves the performance of the domain adaptation model and increases the effectiveness of using the generated synthetic data for training object detection models. We compared our approach with the existing single-stage method where content and style were trained simultaneously. We argue that our proposed method is more practical for training object detection models than others. The emphasis in this study is to preserve image content while changing the style of the image. In the future, we plan to conduct additional experiments to apply synthetic data generation technology to various other application areas like indoor scenes and bin picking.

Keywords— Domain adaptation; object detection; image-to-image translation; GAN.

Manuscript received 10 Nov. 2021; revised 28 Jan. 2022; accepted 18 Mar. 2022. Date of publication 31 May 2022.
International Journal on Informatics Visualization is licensed under a Creative Commons Attribution-Share Alike 4.0 International License.



I. INTRODUCTION

Domain adaptation is a way to reuse knowledge obtained from other domains in a new target domain. It is mainly used to reduce the cost of model training in a new domain or when it is not easy to obtain data in that domain. Recently, attempts to utilize domain adaptation are increasing due to the difficulty of constructing datasets in the field of computer vision, such as image classification, object detection, and semantic segmentation. Among image recognition technologies, object detection shows the highest potential and performance, and many attempts have been made to apply it to various fields. We investigated various domain adaptation methods for object detection and found that, in many cases, they are not yet suitable for practical applications. Unlike image classification and semantic segmentation, object detection models perform complex tasks such as classification and localization. Due to these complex tasks, it is not easy to find a direction for improving domain adaptation techniques

that can help object detection models. Thus, research achievements on domain adaptation for object detection remain at the preliminary stage.

To solve the impractical problem of domain adaptation, we used the image-to-image method in our research, which is one of the existing methods in the field. This method directly generates synthetic image data which is used to train the model. Generating synthetic image data is a difficult task, but the advent of Generative Adversarial Nets (GANs) [1] made it possible to overcome some of the existing difficulties. In particular, the method of changing the domain of image data using GAN emerged as one of the useful methods in generating synthetic image data.

Fig. 1 schematically shows how to train an object detection model using synthetic data generated through image-to-image translation. Fig. 1 assumes that there is a model to detect objects in domain B. Domain B dataset x_B is necessary for model training, but it lacks data. If domain A dataset x_A exists, it can be transformed into a domain B dataset x_{AB} through the domain adaptation model G_{AB} .

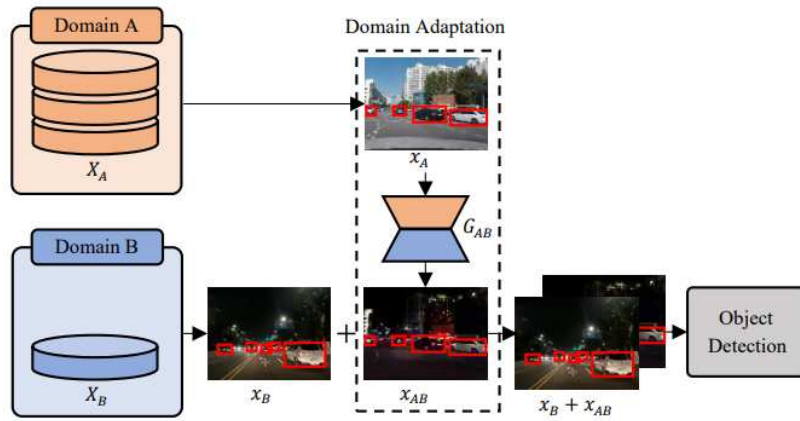


Fig. 1 A conceptual diagram of synthetic data generation and object detection training using image-to-image translation model.

This allows us to obtain additional datasets for model training. In conclusion, if one domain has enough data, image-to-image translation can transform that domain data to generate synthetic data in a new domain with relatively little data. Though promising, image-to-image translation used in generating training data for object detection models is not practical. In image data, a single image often contains multiple objects, and each object has a shape and a location that are critical pieces of information in object detection. As part of translating images, we found that such critical information is also likely to be distorted. The result is that generating good synthetic images for training through image-to-image translation is complicated at best. Building an image-to-image translation model then involves solving the problems of reducing the high cost of training and improving poor performance in inferencing. We want to develop a practical domain adaptation model that generates realistic synthetic images while preserving critical content information.

We provide the following main contributions:

- We propose a data-centric method for domain adaptation. This method is highly efficient as it does not require auxiliary network engineering.
- Our proposed method makes the domain adaptation model good at preserving object-specific information in an image, thus facilitating object detection.
- Experiments demonstrated that the domain adaptation model could be used to train deep learning models as a new data augmentation method.
- We analyzed the appropriate number of synthetic data required for training the deep learning model through experiments.

II. MATERIALS AND METHOD

A. Domain Adaptation for Object Detection

Among several domain adaptation methods, we selected the most practical field of use. Domain adaptation studies for object detection are classified into several main categories [2]: image-to-image translation, adversarial feature learning, pseudo-label-based self-training, domain randomization, and graph reasoning.

- image-to-image translation [3]-[12] converts the target domain image to the source domain or vice versa. This is the most intuitive and easy-to-use methodology

because it visually reduces the differences between domains. This makes it easy to perform object detection training, and most studies and methods use the following approach.

- Adversarial feature learning [13]-[21] performs adversarial training of the object detection model with the help of a domain discriminator. The detector model is trained to fool the domain discriminator, while the domain discriminator learns to classify the domain correctly. This causes the detector to generate domain-independent features. Therefore, the model can detect objects regardless of the domain.
- Pseudo-label-based self-training [22]-[27] learns how to generate a pseudo-label in the target domain using the ground truth label of the source domain. The model predicts pseudo-label in the target domain specified from the source domain. The model gradually learns the object detection model for the target domain.
- Domain randomization [28]-[29] is a method of creating an object detection model regardless of the domain by generating random style data and training it. Thus, it is possible to detect the object in the target domain correctly.
- Graph reasoning [30]-[31] utilizes relationships within or between objects in the detection dataset. By learning the object relationship from the target domain, which is like that of the source domain, the object detection model can also detect the object in the target domain.

We considered practicality, learning difficulty, and performance of the research field. In the case of domain randomization and graph reasoning, there was a problem of either poor performance or very high learning difficulty. In the case of adversarial feature learning and pseudo-label-based self-training, there was a problem that practical application fields were considerably limited. As a result, we conducted a study on the image-to-image translation method.

B. Image-to-image Translation

We classified the image-to-image method into two cases based on the training dataset. One is a model that trains only with the source image, and the other is a model that is guided by labeled data.

Since the advent of GAN, research on the domain adaptation model at the image level using only the source image has been actively conducted. GAN-based image-to-

image translation techniques emerged, such as CycleGAN [3] and UNIT [4] [5]. These models use only the source image as training data, resulting in easy data acquisition and faster model training and application. However, this method does not provide good qualitative results in practice due to the high learning difficulty of the model itself.

Since then, researchers have improved the performance of the domain adaptation model by training it with additional information from existing data. Some models, such as AugGAN [6], [7], were trained with additional segmentation information. Some models, such as GraspGAN [8], were trained with additional behavioral information on top of the segmentation information on existing data. The models were trained not only with raw image data but also were supplied with additional information such as segmentation. As a result, they can learn and recognize objects in images. However, this increases the cost of constructing a dataset of the domain adaptation model and slows down model training and application.

Our research aims to speed up the development of models in various domains. If an image-to-image translation model that requires labeled data is used, the model development cost may exceed that of the existing method depending on the dataset construction difficulty of the model. So, instead of using a model that learns only the original image, such as CycleGAN or UNIT, we studied a method to improve the model's performance.

We performed a qualitative evaluation of each model by generating synthetic data through the image-to-image translation method. We tried changing the daytime driving environment to night using CycleGAN and UNIT.



Fig. 2 Failure cases of CycleGAN and labels for object detection

Fig. 2 shows the output of translating the day environment to night through CycleGAN, and ground truth label information for object detection model training. It can be confirmed that the appearances of objects existing in the image are completely invisible. We interpreted the model's training process as weak in learning the information of individual objects in the image. And we judged that this would be critical for object detection models in which localization is important.

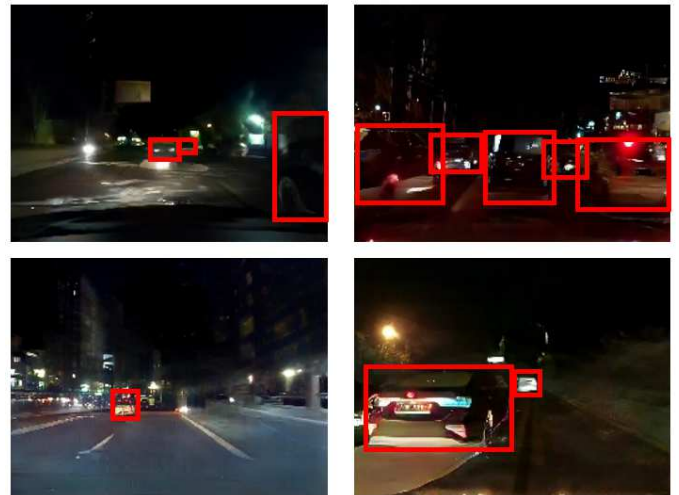


Fig. 3 Failure cases of UNIT and labels for object detection

Fig. 3 shows the output of translating the day environment to night through UNIT and ground truth label information for object detection model training. Compared to CycleGAN, it was confirmed that the object's shape remained, but it was confirmed that the style application was strange.

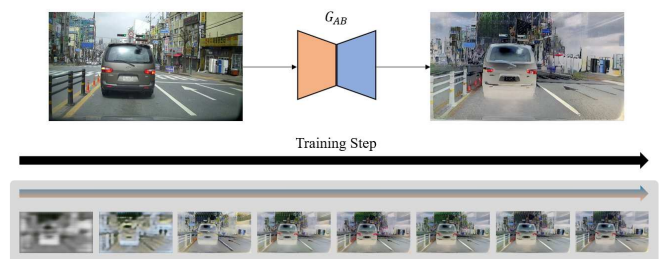


Fig. 4 The final result image sample of the UNIT model and the result by training step.

As shown in Fig. 4, we checked the model's training process by extracting the image translation result of the UNIT model for each training step. We found that once the target domain became complex, generating qualitative data became prohibitively difficult, and training a model suddenly became limited. As a result, we did not find these methods practical for object detection training.

C. Learning Content before Style

We analyzed the results of the previous two models and set two goals for model improvement. One is to maintain the information of the object in the image, and the other is to apply the style of the object in the image well. For this, we referred to studies related to style transfer.

Studies on the style transfer of images using deep neural networks have been actively conducted [32]-[37]. Leon A. Gatys [32] published a paper on style transfer using convolutional neural networks. Leon A. Gatys attempted to create a new image by separating style and content from features learned through convolutional neural networks, and it was highly successful. Studies related to style transfer were conducted afterward, resulting in excellent research results such as AdaIN [33] and StyleGAN [34] [35].

In some studies, domain adaptation and object detection were attempted using style transfer techniques [36], [37].

Unfortunately, style transfer also had a problem in that it did not significantly reduce the visual gap with the source area. Therefore, it was difficult to apply these techniques for object detection.

As studies related to style transfer were actively conducted and showed good results, researchers agreed that image recognition through deep learning consists of two things. One is content, which represents the structure and shape of an image. The other is style, which refers to the texture and color of an image. We focused on the fact that deep learning models learn them both.

Existing studies on style transfer have been conducted to improve the model for separating style and content. One of these studies is AdaIN [33], where the image is styled through deep networks. However, this model-centric style transfer also had a limit in significantly reducing the visual gap with the source area. This visual gap between domains creates a kind of bias in training deep learning models and causes performance degradation rather than performance improvement. We studied another method to reduce the visual difference between domains to overcome this. We planned data-centric studies rather than models and designed a data training strategy to improve performance using classical models, such as CycleGAN or UNIT.

D. Proposed Method

We hypothesize that training content and style information simultaneously to existing domain adaptation models causes frequent loss of object-specific information in an image during training. Fig. 2 and Fig. 3 show the results of image conversion in which object information is lost. We hypothesized that object-specific information in images was tied closely to the content than the style and thus experimented with methods to preserve content information before style was learned.

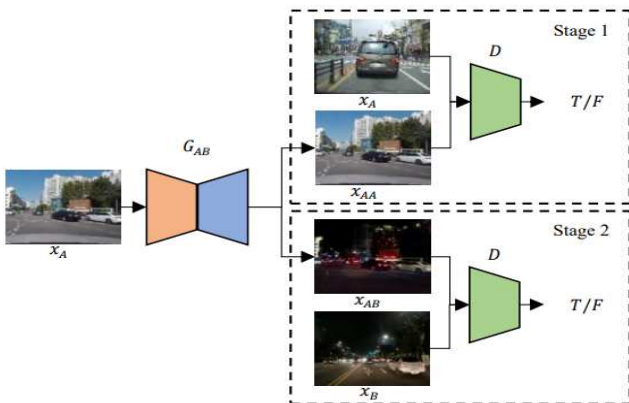


Fig. 5 Diagram of the two-stage training model

We divided the training into two stages, as shown in Fig. 5. The domain adaptation model G_{AB} that changes domain A to B performs training through competition with discriminator D. Then, the training goes through two stages. In the first stage, the source domain data x_A was used as both the input and the output of the model. In the second stage, the target domain data x_B was set as the target output. The model is then trained to generate domain-unchanged data x_{AA} in the first stage and then generates domain-changed data x_{AB} in the

second step. The idea was to let the model focus on training the content first before training the style. For this reason, we named our method, "the two-stage training method." In contrast, we decided to call the existing training domain adaptation method, "a single-stage training method."

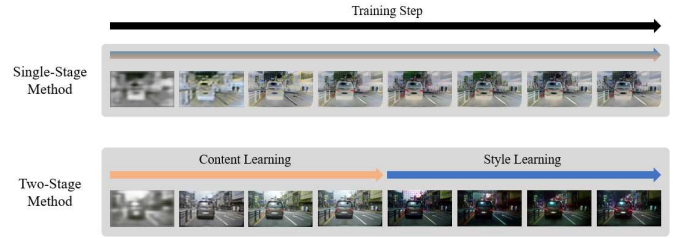


Fig. 6 Comparison of two training methods by training step.

Fig. 6 shows the change in the image output generated by the generator while training is in progress. In the existing single-stage model, the style became fixed first at some point after the beginning of training, and training did not proceed anymore. This quickly learned style produced images with strange styles that had nothing to do with the properties of the objects in the image. This phenomenon occurred in most domains as well as the domains shown in Fig. 6. Based on the following results, we judged that the content and style have different training times and difficulties and concluded that if the training is performed simultaneously, the model is frail to falling into local minimum due to imbalance.

Therefore, we postponed the training of the fast-learned style and performed the training in the proposed two-stage method so that the training of the content would take place first. Changes in the generated image for the proposed two-stage method can also be seen in Fig. 6. We found that the style was naturally applied in the subsequent style training stage.

III. RESULTS AND DISCUSSION

We experimented on domain adaptation to verify that our proposed method is effective. We focused on two aspects of object detection experiments with synthesized data. First, we compared the two datasets' visual quality and quantity metrics generated by the two domain adaptation models. Second, we analyzed the performance metrics of object detection models separately trained with the data generated by the two domain adaptation models. The specifications of the deep learning training server adopted in this work are shown in Table 1.

TABLE I
SPECIFICATIONS OF THE AI DEEP LEARNING TRAINING SERVER

CPU	Intel Xeon Processor (Skylake, IBRS)
GPU	Tesla V100-PCIE-16GB x 2
RAM	16GB RAM x 8
OS	Ubuntu 16.04.6 LTS
AI Framework	Tensorflow v1.15.0

A. The Dataset

We constructed a new model application scenario to validate our experiments. It is an object detection model training in a driving environment. The driving environment was collected by dividing it into day and night, and the nighttime driving environment was collected relatively less than the daytime driving environment. Then, we designed a

model to supplement the insufficient dataset of the nighttime driving environment from the dataset of the day driving environment through domain adaptation.

We collected the driving environment of a car as data for the experiment. We collected 10,000 images of the daytime driving environment and 4000 images of the nighttime driving environment. Fig. 7 and Fig. 8 show samples of each environmental driving image. The data of each environment was labeled for training the object detection model. We consisted of three classes of objects in the image: car, bus, and truck.



Fig. 7 Daytime driving environment dataset.

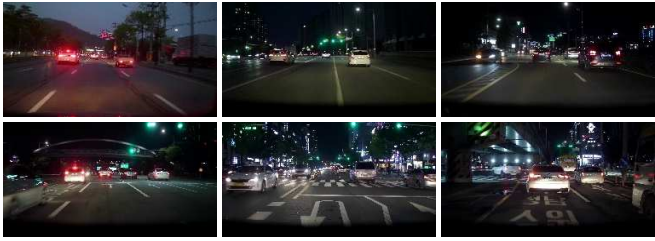


Fig. 8 Nighttime driving environment dataset.

B. Domain Adaptation

The first experiment was the domain adaptation task to transform daytime driving images into nighttime images. The same UNIT model [4] was used for both the single-stage and the two-stage methods. The datasets consisted of 10,000 daytime driving images and 4,000 nighttime driving images.



Fig. 9 Results compare the two domain adaptation methods: the single-stage and the two-stage methods.

As can be seen from the sample images in Fig. 9, the two-stage method produced more plausible nighttime images than those from the single-stage method. The single-stage method frequently generated image data in which some objects became blurred or even disappeared. On the other hand, the two-stage method generated image data where objects in the image were well preserved.

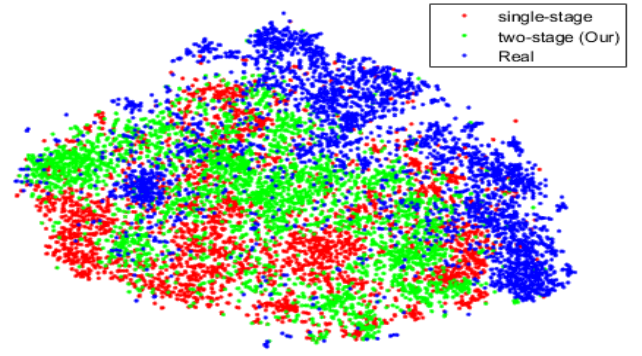


Fig. 10 T-SNE embedding of single-stage synthetic data (red), two-stage synthetic data (green), and real data (blue). Features extracted from ResNet101.

Moreover, we visualized the T-SNE [38] embedding of the real and synthesis data extracted from ResNet101 in Fig. 10. This visual distribution indicates that the synthetic data of the two-stage method were closer to real night images than the results of the single-stage method.

C. Object Detection with Synthetic Data

In the second experiment, we compared the performance of the two object detection models. The models were trained using the data generated by the single-stage and two-stage methods. The target objects for object detection were vehicles in different driving environments. We used Faster RCNN [39] as the object detection model and ResNet101 network. Five different training datasets were created based on the number of real nighttime driving images: namely, 200, 400, 600, 800, and 1000 real nighttime driving images. On top of these real images acting as seeds, we added synthetic data generated by a domain adaptation model. The real and synthetic data ratio was set to 0%, 100%, 200%, 300%, and 400%. We used 2,000 nighttime driving images as a test set to evaluate the performance. We used mAP (mean Average Precision) as a quantitative evaluation indicator of object detection.

TABLE II
RESULTS OF THE MODEL TRAINED WITH THE DATA FROM THE ONE-STAGE TRAINING METHOD

Synthetic Data Ratio	mAP (%)				
	200 data	400 data	600 data	800 data	1000 data
0	51.42	59.79	62.55	66.61	66.62
1	54.88	61.95	64.30	65.56	67.17
2	56.87	62.74	63.67	66.08	67.36
3	56.09	62.66	63.28	65.21	67.38
4	57.32	60.90	63.88	66.86	66.78

TABLE III
RESULTS OF THE MODEL TRAINED WITH THE DATA FROM THE TWO-STAGE TRAINING METHOD

Synthetic Data Ratio	mAP (%)				
	200 data	400 data	600 data	800 data	1000 data
0	51.42	59.79	62.55	66.61	66.62
1	55.69	62.01	64.20	66.86	68.04
2	57.06	63.35	65.29	67.50	68.92
3	58.50	63.24	67.99	68.85	68.89
4	58.20	64.11	67.84	70.22	69.73

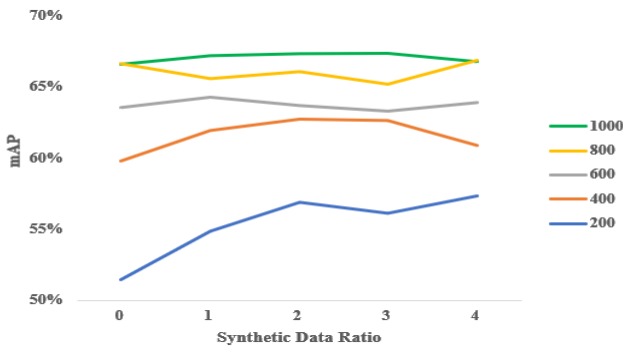


Fig. 11 Result graphs of the single-stage method

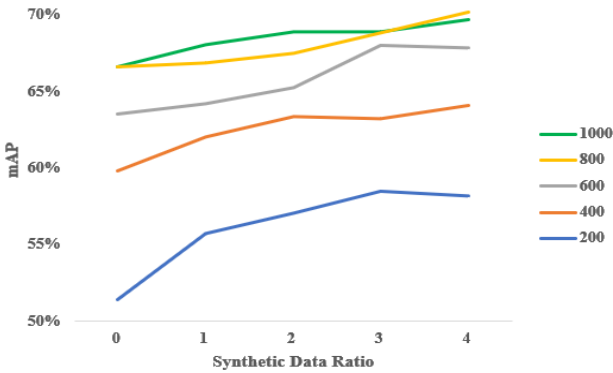


Fig. 12 Result graphs of the two-stage method

Table. 2 and Table. 3 shows the performances of the two object detection models trained with the one-stage training method and the data from the two-stage training method, respectively. The graphs in Fig. 11 and Fig. 12 visually show the performances of the two object detection models. As can be seen from the graphs, mAP obtained from the two-stage model all scored higher than those from the single-stage model. In addition, when the number of real data is more than 400, the mAP score in the single-stage model did not rise well but rather went down. However, the mAP score in the two-stage model steadily increased in all numbers of real data cases. In both models, the mAP score did not go up but instead went down when the ratio of synthetic data exceeded 4. The result shows that the object detection model trained with the two-stage method data demonstrated better performance than that from the single-stage method.

The experiment showed that the existing single-stage model often fell into local minima during training and produced poor-quality images in terms of content and style. On the other hand, the two-stage method generated better images in both categories, as suggested in this paper.

This study solved the problem through a data-centric operation rather than model improvement. This method is easy to do, but the effect of improving the model is clear. This approach allowed us to suggest another direction for deep learning model development.

IV. CONCLUSION

The major objective of this study is to investigate the utilization of domain adaptation to improve the performance of object detection. Considering practicality, we conducted a

study of image-to-image translation among several domain adaptation methodologies. We assumed the existing problem of image-to-image translation models to be information loss of objects in images. By changing the training method, we designed the model to focus on training the content first before training the style. The result images of the proposed method were more plausible and recognizable. Our experiments provide insights into cost-effective and practical methods to solve the lack of data problems.

To prove the feasibility of this method, our results are still encouraging and should be explored in more diverse environments. In the future, we plan to apply synthetic data generation techniques to areas where data is insufficient, and our technique can be applied to a wide range of outdoor applications.

REFERENCES

- [1] Ian J. Goodfellow, Jean Pouget-Abadie, "Generative adversarial nets," *Advances in neural information processing systems* 27, June 2014.
- [2] Poojan Oza, Vishwanath A. Sindagi, "Unsupervised domain adaptation of object detectors: A survey," May 2021.
- [3] Jun-Yan Zhu, Taesung Park, "Unpaired image-to-image translation using cycle-consistent adversarial networks," *Proceedings of the IEEE international conference on computer vision*, March 2017.
- [4] Ming-Yu Liu, Thomas Breuel, and Jan Kautz, "Unsupervised image-to-image translation networks," *Advances in neural information processing systems* 30, March 2017.
- [5] Kyungjune Baek, Yunjey Choi, "Rethinking the truly unsupervised image-to-image translation," *Proceedings of the IEEE/CVF International Conference on Computer Vision*, June 2021.
- [6] Sheng-Wei Huang, Che-Tsung Lin, "Auggan: Cross domain adaptation with gan-based data augmentation," *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018.
- [7] Che-Tsung Lin, Sheng-Wei Huang, "GAN-based day-to-night image style transfer for nighttime vehicle detection," *IEEE Transactions on Intelligent Transportation Systems* 22.2, 2020, pp. 951-963.
- [8] Xue Bin Peng, Marcin Andrychowicz, "Sim-to-real transfer of robotic control with dynamics randomization," 2018 IEEE international conference on robotics and automation (ICRA), IEEE, October 2018.
- [9] Han-Kai Hsu, Chun-Han Yao, "Progressive domain adaptation for object detection," *Proceedings of the IEEE/CVF winter conference on applications of computer vision*, October 2019.
- [10] Zhiqiang Shen, Mingyang Huang, "CDTD: A large-scale cross-domain benchmark for instance-level image-to-image translation and domain adaptive object detection," *International Journal of Computer Vision* 129.3, November 2021, pp. 761-780.
- [11] Chaoqi Chen, Zebiao Zheng, "Harmonizing transferability and discriminability for adapting object detectors," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, March 2020.
- [12] Shan Lin, Fangbo Qin, "Lc-gan: Image-to-image translation based on generative adversarial network for endoscopic images," 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, March 2020.
- [13] Zhenwei He, and Lei Zhang "Multi-adversarial faster-rcnn for unrestricted object detection," *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2019.
- [14] Kuniaki Saito, Yoshitaka Ushiku, "Strong-weak distribution alignment for adaptive object detection," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, April 2019.
- [15] Peng Su, Kun Wang, "Adapting object detectors with conditional domain normalization," *European Conference on Computer Vision*. Springer, Cham, July 2020.
- [16] Yangtao Zheng, Di Huang, "Cross-domain object detection through coarse-to-fine feature adaptation," *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, March 2020.
- [17] Chang-Dong Xu, Xing-Ran Zhao, "Exploring categorical regularization for domain adaptive object detection," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, March 2020.

- [18] Chaoqi Chen, Zebiao Zheng, "I3net: Implicit instance-invariant network for adapting one-stage object detectors," Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, March 2021.
- [19] Vibashan VS, Vikram Gupta, "Mega-cda: Memory guided attention for category-aware unsupervised domain adaptive object detection," Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, March 2021.
- [20] Yuhua Chen, Haoran Wang, "Scale-aware domain adaptive faster r-cnn," International Journal of Computer Vision 129.7, May 2021, pp. 2223-2243.
- [21] Zhenwei He, and Lei Zhang, "Domain adaptive object detection via asymmetric tri-way faster-r-cnn," European conference on computer vision. Springer, Cham, July 2020.
- [22] Mehran Khodabandeh, Arash Vahdat, "A robust learning approach to domain adaptive object detection," Proceedings of the IEEE/CVF International Conference on Computer Vision, November 2019.
- [23] Aruni RoyChowdhury, Prithvijit Chakrabarty, "Automatic adaptation of object detectors to new domains using self-training," Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, April 2019.
- [24] Seunghyeon Kim, Jaehoon Choi, "Self-training and adversarial background regularization for unsupervised domain adaptive one-stage object detection," Proceedings of the IEEE/CVF International Conference on Computer Vision, September 2019.
- [25] Ganlong Zhao, Guanbin Li, "Collaborative training between region proposal localization and classification for domain adaptive object detection," European Conference on Computer Vision. Springer, Cham, September 2020.
- [26] Petru Soviany, Radu Tudor Ionescu, "Curriculum self-paced learning for cross-domain object detection," Computer Vision and Image Understanding 204, 103166, January 2021.
- [27] Jinhong Deng, Wen Li, "Unbiased mean teacher for cross-domain object detection," Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 2021.
- [28] Taekyung Kim, Minki Jeong, "Diversify and match: A domain adaptive representation learning paradigm for object detection," Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, May 2019.
- [29] Adrian Lopez Rodriguez, and Krystian Mikolajczyk, "Domain adaptation for object detection via style consistency," November 2019.
- [30] Qi Cai, Yingwei Pan, "Exploring object relation in mean teacher for cross-domain detection," Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, December 2019.
- [31] Minghao Xu, Hang Wang, "Cross-domain detection via graph-induced prototype alignment," Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, March 2020.
- [32] Leon A. Gatys, Alexander S. Ecker, and Matthias Bethge. "Image style transfer using convolutional neural networks," Proceedings of the IEEE conference on computer vision and pattern recognition, 2016.
- [33] Xun Huang, and Serge Belongie "Arbitrary style transfer in real-time with adaptive instance normalization," Proceedings of the IEEE international conference on computer vision, July 2017.
- [34] Tero Karras, Samuli Laine, and Timo Aila, "A style-based generator architecture for generative adversarial networks," Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, March 2019.
- [35] Tero Karras, Samuli Laine, "Analyzing and improving the image quality of stylegan," Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, March 2020.
- [36] Hyeongkeun Lee, Kyungmin Lee, "Applying FastPhotoStyle to Synthetic Data for Military Vehicle Detection," 2020 20th International Conference on Control, Automation and Systems (ICCAS). IEEE, 2020.
- [37] David Kadish, Sebastian Risi, and Anders Sundnes Løvlie. "Improving object detection in art images using only style transfer," 2021 International Joint Conference on Neural Networks (IJCNN), IEEE, May 2021.
- [38] Laurens van der Maaten, Geoffrey Hinton, "Visualizing Data using t-SNE," Journal of Machine Learning Research 9, November 2008.
- [39] Shaoqing Ren, Kaiming He, "Faster r-cnn: Towards real-time object detection with region proposal networks," Advances in neural information processing systems 28, January 2016, pp. 91-99.