AUGTARGET DATA AUGMENTATION FOR INFRARED SMALL TARGET DETECTION

Shengjia Chen[1], Jiewen Zhu[1], Luping Ji[1,], Hongjun Pan[2], Yuhao Xu[2]*

School of Computer Science and Engineering,
[1]University of Electronic Science and Technology of China, Chengdu 611731, China
[2]Sichuan University, Chengdu 610065, China

ABSTRACT

Sample shortage has always been a frequently-faced problem for the machine-learning models in infrared small target detection. As one of main limitations, it is hampering the further promotion of target detection performance. In this paper, we propose a simple and effective data augmentation scheme, *AugTarget*, to address this shortage issue of small target samples. Our scheme mainly consists of two crucial algorithms: target augmentation and batch augmentation. The former is designed to generate sufficient targets, by random target representation. The latter is devised to diversify training samples. Moreover, the initially-generated image samples of small targets are further enriched by randomly aggregating the feature representation of different images. The experiments on public datasets demonstrate that our AugTarget could bring an obvious improvement to mean intersection over union (mIoU). Cooperated by the augmentation of our AugTarget, the state-of-the-art (SOTA) method, AGPC could even achieve a distinct performance promotion by 3.03%, 2.17% and 2.63% on MDFA, SIRST-Aug and Merged datasets, respectively. In addition, the experimental results on three baseline models also show the universality & adaptivity of AugTarget to different dataset augmentation. Our source codes are available at https://github.com/UESTC-nnLab/AugTarget.

***Index Terms*—** Infrared small target detection, Data augmentation, Target & batch augmentation

1. INTRODUCTION

Infrared small target detection aims at separating small targets from cluttered backgrounds. It is a challenging task, requiring the accurate prediction of target location and shapes. Benefiting from the advances of deep convolutional neural networks (CNNs) [1], infrared small target detectors [2–4] have achieved further performance promotion on traditional detection models. Many previous works, such as the ones in [2,4–9], mainly focus on designing better CNN-based feature
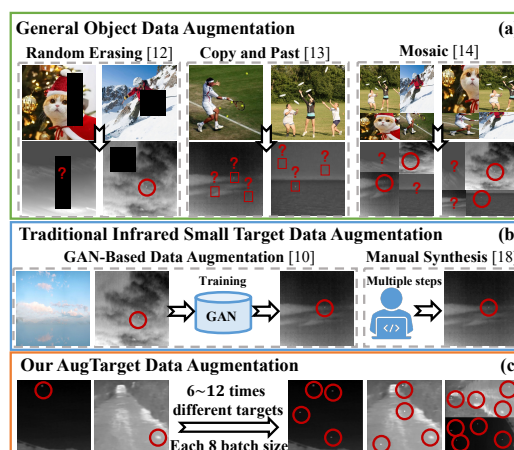
Fig. 1. A visual comparison of data augmentation methods.

extraction models to improve detector performance. In fact, training dataset is also one of significant issues for performance improvement. Currently, inadequate training datasets have become one of the largest obstacles to promote the performance of CNN-based target detectors.

There are some typical characteristics in currently existing datasets, such as i) **limited data size**, ii) **high generating cost** and iii) **tiny target quantity**. Training a CNN-based model with a large number of parameters on limited-size datasets could often cause over-fitting problems [1]. Moreover, cost is often very high to manually generate large-size datasets [10]. Besides these, datasets possibly contain very small number of labeled targets, making detector not to learn sufficient target features in training phase [11]. As a result, how to make full use of existing limited-size data, without increasing labor and time cost, is usually an important issue in infrared small target detection.

In general, data augmentation is an efficient way to augment limited-size datasets, so as to improve the comprehensive performance of deep-learning tasks. Many empirical studies [12–17] have suggested that general object detection models could be improved by augmenting datasets with special algorithms. However, the general data augmentation methods cannot be directly used for infrared small target

detection, due to the small size of targets.

A group of visual comparisons for augmentation methods is shown in Fig.1. In (a), to improve generalization ability and performance, the Random Erasing scheme in [12] randomly deletes a continuous region of given image. When it is used for infrared small target, original target could often be lost. The copy-and-paste algorithm in [13] augments small objects by a copying and pasting operation, but it requires bounding boxes. However, a dataset of infrared small targets does not contain expected bounding boxes at the most time.

Another one, Mosaic [14], mixes cropped training images for data augmentation, however it is easy to be cropped out of small targets from original images. In addition, as shown in Fig.1(b), some traditional enhancement methods [10, 18, 19] for small targets use adversarial generative networks (GANs) or manual methods to generate synthetic samples.

In a whole, there are five primary issues to be resolved in data augmentation, we believe. They contain: avoiding target losing, unexpected bounding boxes, without extra training, low cost and generating enough targets. To address them, we propose the AugTarget, a simple and effective sample augmentation scheme, as illustrated in Fig. 1(c). Different from existing other enhancement methods, AugTarget does not require a pre-trained generating model and manual intervention. It could generate new training data exactly in the training of target detector. Furthermore, our AugTarget does not lose targets, instead, it could generate more targets without bounding boxes. In summary, the primary contributions of this paper include three points: i) a target augmentation algorithm is proposed, ii) a batch augmentation approach is designed, and iii) the integrated experiments with other detection models are done to verify the effects of our AugTarget.

2. PROPOSED AUGTARGET

Our AugTarget is a simple, general and efficient data augmentation method. Its goal is to improve infrared small target detection by augmenting heterogeneous targets and increasing sample diversity. To fulfil this aim, an integrated algorithm pipeline, presented in Fig. 2, is designed on the training paradigm of general infrared small target detection model. This pipeline can serve as an add-on block for a target-detection model to achieve high-quality predictions without extra parameter learning.

2.1. Target Augmentation

Designed target augmentation algorithm aims to generate heterogeneous small targets by taking a random strategy as its guidance. In training, this algorithm performs target augmentation simultaneously on both images and ground-truth masks in each mini-batch. The pseudo-code of our target augmentation is given out in Algorithm 1. It mainly consists of four steps, as follows.
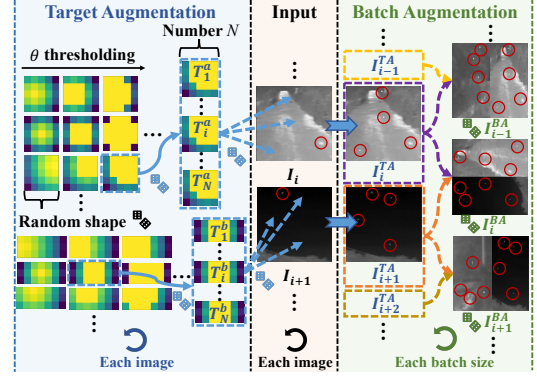

Fig. 2. Overview of proposed AugTarget.

Algorithm 1: Pseudocode of Target Augmentation

Input: image $I \in \mathbb{R}^{W \times H}$, mask $M \in \mathbb{R}^{W \times H}$
1   Initialize maximum area and number of targets A, N;
2   Initialize shape and position of targets h, w and x, y;
3   **for** $i = 1$ **to** N **do**
4      $h \leftarrow \sqrt{A \times r}; w \leftarrow \sqrt{A/r}$;
5      $x \leftarrow Random(0, H - h_i); y \leftarrow Random(0, W - w_i)$;
6   **if** $I[x_i, y_i] > 0$ **then**
7         $\boldsymbol{m}_1, \boldsymbol{m}_2 = \texttt{numpy.mgrid}[: h_i, : w_i]$;
8         $\boldsymbol{T}_i = (\boldsymbol{m}_1 - h_i/2) \times 2 + (\boldsymbol{m}_2 - w_i/2) \times 2$;
9         $\boldsymbol{T}_i = -\boldsymbol{T}_i + \sum_{(H \times W)} \boldsymbol{I}_{(i,j)}/(H \times W)$;
10        $\boldsymbol{T}_i[\boldsymbol{T}_i < \boldsymbol{I}_{\min}] = \boldsymbol{I}_{\min}$, $\boldsymbol{I}_{\min}$ is the minimum pixel value in $\boldsymbol{I}$;
11        $\theta = (\boldsymbol{I}_{\min}/\boldsymbol{I}_{\max}) \times \sum_{(H \times W)} \boldsymbol{I}_{(i,j)}/(H \times W)$;
12        $\boldsymbol{T}_i[\boldsymbol{T}_i > \theta] = \boldsymbol{I}_{\max}$
13   **end**
14   **end**
15   $\boldsymbol{I}^{\text{TA}}_{x_i, y_i} \leftarrow \boldsymbol{T}_i, \boldsymbol{M}^{\text{TA}}_{x_i, y_i} \leftarrow 255$;
Output: image $\boldsymbol{I}^{\text{TA}} \in \mathbb{R}^{W \times H}$, mask $\boldsymbol{M}^{\text{TA}} \in \mathbb{R}^{W \times H}$

Step 1: Initialization. Firstly, the height h and width w of given target are randomly initialized by a random strategy [12] with the pre-defined maximum area A for targets. The shape of target is $h = \sqrt{A \times r}$ and $w = \sqrt{A/r}$, where r is a random aspect ratio. The number of iterations depends on the pre-defined target numbers N, and N pairs of (h, w) are obtained after iteration. The coordinates (x, y) of target are randomly initialized in a given image. The ranges of coordinates (x, y) follow $0 < x < H - h_i$ and $0 < y < W - w_i$, respectively. Symbol i denotes the index of the ith target.

Step 2: Generation of heterogeneous targets. Next, target augmentation is performed on the input image $\boldsymbol{I} \in \mathbb{R}^{W \times H}$ and mask $\boldsymbol{M} \in \mathbb{R}^{W \times H}$. We use *Numpy* operation to generate two grid matrices of dense pixels, $\boldsymbol{m}_1 \in \mathbb{R}^{h_i \times w_i}$ and $\boldsymbol{m}_2 \in \mathbb{R}^{h_i \times w_i}$. We synthesise the infrared small target in a simple way, $\boldsymbol{T}_i = (\boldsymbol{m}_1 - h_i/2) \times 2 + (\boldsymbol{m}_2 - w_i/2) \times 2$. After that, the target matrix $\boldsymbol{T}_i \in \mathbb{R}^{h_i \times w_i}$ with infrared target characteristics is generated, visually shown in Fig. 2.

Step 3: Thresholding. The simple target matrix given above may suffer from two drawbacks. First, we find that it is difficult for detection model to capture the features of insignificant targets in infrared images with complex backgrounds. Second, synthetic target is relatively homogeneous.

Thus, we consider thresholding this target matrix. In details, we use a threshold θ to enhance the saliency of targets and increase the diversity of targets.

Step 4: Adding augmented targets into image. Finally, we set the region $\boldsymbol{I}^{\text{TA}}_{x_i,y_i} \in \mathbb{R}^{h_i \times w_i}$ as a determined target region. Meanwhile, we choose the region of same position with (x_i, y_i) as target mask. Because this mask is a binarized image, we set all the pixel value of target region to 255. In contrast, all background region pixels are set to 0.

2.2. Batch Augmentation

The purpose of our batch augmentation algorithm is to utilize target augmented images to improve the diversity of sample images, and enrich the feature representation of targets. The pseudo-code of batch augmentation is listed in Algorithm 2. Our algorithm consists of two steps, as follows.

Algorithm 2: Pseudocode of Batch Augmentation

Input: $\boldsymbol{I}^{\text{TA}} \in \mathbb{R}^{B \times C \times W \times H}$, $\boldsymbol{M}^{\text{TA}} \in \mathbb{R}^{B \times C \times W \times H}$

1 Initialize random coefficient $\eta = Random(0, B-1)$;
2 **for** $i = 1$ **to** B **do**
3 $\quad \boldsymbol{I}^{\text{TA}}_{\text{list}} = \cup_{i=1}^{N} \cup_{j=1}^{B} [\boldsymbol{I}_j : \boldsymbol{I}_{j+1}]$;
4 $\quad \boldsymbol{M}^{\text{TA}}_{\text{list}} = \cup_{i=1}^{N} \cup_{j=1}^{B} [\boldsymbol{M}_j : \boldsymbol{M}_{j+1}]$;
5 $\quad$ **if** $\eta > B/2 - 1$ **then**
6 $\quad\quad \boldsymbol{I}^{\text{BA}}_i \leftarrow \text{concat}((\boldsymbol{I}^{\text{TA}}_{\text{list}}[i], \boldsymbol{I}^{\text{TA}}_{\text{list}}[i+1]), 3)$;
7 $\quad\quad \boldsymbol{M}^{\text{BA}}_i \leftarrow \text{concat}((\boldsymbol{M}^{\text{TA}}_{\text{list}}[i], \boldsymbol{M}^{\text{TA}}_{\text{list}}[i+1]), 3)$;
8 $\quad$ **else**
9 $\quad\quad \boldsymbol{I}^{\text{BA}}_i \leftarrow \text{concat}((\boldsymbol{I}^{\text{TA}}_{\text{list}}[i], \boldsymbol{I}^{\text{TA}}_{\text{list}}[i+1]), 2)$;
10 $\quad\quad \boldsymbol{M}^{\text{BA}}_i \leftarrow \text{concat}((\boldsymbol{M}^{\text{TA}}_{\text{list}}[i], \boldsymbol{M}^{\text{TA}}_{\text{list}}[i+1]), 2)$;
11 $\quad$ **end**
12 **end**
13 $\boldsymbol{I}^{\text{BA}}, \boldsymbol{M}^{\text{BA}} \leftarrow \text{interpolate}((\boldsymbol{I}^{\text{BA}}, \boldsymbol{M}^{\text{BA}}), \text{size} = [W, H])$;
14 $\boldsymbol{I}^{\text{BA}} = \text{concat}((\boldsymbol{I}^{\text{BA}}, \boldsymbol{I}^{\text{TA}}), 0)$;
15 $\boldsymbol{M}^{\text{BA}} = \text{concat}((\boldsymbol{M}^{\text{BA}}, \boldsymbol{M}^{\text{TA}}), 0)$;

Output: image $\boldsymbol{I}^{\text{BA}} \in \mathbb{R}^{2B \times C \times H \times W}$, mask $\boldsymbol{M}^{\text{BA}} \in \mathbb{R}^{2B \times C \times H \times W}$

Step 1: Batch image and mask splitting. We split batch images $\boldsymbol{I}^{\text{TA}} \in \mathbb{R}^{B \times C \times W \times H}$ and masks $\boldsymbol{M}^{\text{TA}} \in \mathbb{R}^{B \times C \times W \times H}$. Then, we obtain an image list, $\boldsymbol{I}^{\text{TA}}_{\text{list}} = \cup_{i=1}^{N} \cup_{j=1}^{B} [\boldsymbol{I}_j : \boldsymbol{I}_{j+1}]$ and a mask list $\boldsymbol{M}^{\text{TA}}_{\text{list}} = \cup_{i=1}^{N} \cup_{j=1}^{B} [\boldsymbol{M}_j : \boldsymbol{M}_{j+1}]$.

Step 2: Image concatenation. In this part, random coefficient η is used to choose the dimension of image concatenation. If $\eta > B/2 - 1$, we concatenate two split images along width direction, obtaining $\boldsymbol{I}^{\text{BA}} \in \mathbb{R}^{B \times C \times H \times 2W}$. Otherwise, we concatenate them along height direction, achieving $\boldsymbol{I}^{\text{BA}} \in \mathbb{R}^{B \times C \times 2H \times W}$. Moreover, to ensure that the shapes of input and output images are consistent, we perform a down-sampling operation on the images generated by concatenating. By this operation, we obtain the augmented batch images $\boldsymbol{I}^{\text{BA}} \in \mathbb{R}^{B \times C \times H \times W}$ and their masks $\boldsymbol{M}^{\text{BA}} \in \mathbb{R}^{B \times C \times H \times W}$.

After random concatenation, the data distribution of images has been changed. Therefore, to keep consistent with original data distribution, we concatenate all augmented images with their original images in batches. Finally, small target feature representation and image diversity are enriched, while original data distribution is maintained.

3. EXPERIMENTS

Three recent detection models, including the AGPC [2], ACM [3] and IAA [4], are taken as the baselines on the three public datasets: MDFA [18] and SIRST-Aug [2], as well as their merging. For baseline AGPC, we keep evaluation scheme unchanged, except the learning rate. For baselines ACM and IAA, we take their original evaluation strategies. In particular, we modified the IAA data pre-processing process to generate bounding boxes during the training process. The evaluation metrics contain Precision, Recall, mean Intersection over Union (mIoU), F-measure and the Area Under Curve (AUC).

3.1. Comparisons with State-of-the-Art Ones

In experiments, two categories of detection methods are chosen for comparisons. One is model-driven, including the SRWS [20], MPCM [21] and PSTNN [22]. The other is data-driven, containing the DNANet [5], MDvcFA [18], IAA [4], ACM [3] and AGPC [2].

Table 1. Comparisons with the state-of-the-art methods.

Dataset	Methods	Precision	Recall	mIoU	F-measure	AUC
MDFA	SRWS [20]	81.38	3.94	3.90	7.51	46.41
	MPCM [21]	64.72	14.99	13.86	24.34	84.76
	PSTNN [22]	72.36	15.44	14.58	25.45	60.87
	DNANet [5]	-	-	46.73	-	-
	MDvsFA [18]	66.00	54.00	-	60.00	**91.00**
	IAA [4]	60.46	77.08	-	61.83	-
	IAA+ours	54.58	75.25	-	63.27↑1.44	-
	ACM [3]	54.33	70.27	44.17	61.28	85.69
	ACM+ours	60.23	68.45	47.15↑2.98	64.08↑2.80	85.81↑0.12
	AGPC [2]	59.39	72.41	48.43	65.25	86.82
	AGPC+ours	64.82	71.41	**51.46**↑3.03	**67.95**↑2.70	86.99↑0.17
SIRST-Aug	SRWS [20]	84.18	3.60	3.57	6.90	57.06
	MPCM [21]	93.57	15.09	14.93	25.99	78.86
	PSTNN [22]	93.66	9.49	9.43	17.24	35.56
	IAA [4]	82.57	79.01	-	80.75	-
	IAA+ours	77.90	87.81	-	82.56↑1.81	-
	ACM [3]	84.60	79.63	69.55	82.04	90.55
	ACM+ours	86.60	82.45	73.13↑3.58	84.48↑2.44	91.59↑1.04
	AGPC [2]	83.23	85.42	72.88	84.31	93.44
	AGPC+ours	84.49	87.04	**75.05**↑2.17	**85.74**↑1.43	**93.78**↑0.34
Merged	IAA [4]	76.16	68.30	-	72.02	-
	IAA+ours	72.18	74.86	-	73.49↑1.47	-
	ACM [3]	75.23	75.45	60.44	75.34	88.79
	ACM+ours	73.20	81.42	62.72↑2.28	77.09↑1.75	91.18↑2.39
	AGPC [2]	74.53	83.84	65.17	78.91	91.94
	AGPC+ours	75.76	86.58	**67.80**↑2.63	**80.81**↑1.90	**93.95**↑2.01

Experiment comparisons are shown in Table 1. From it, we could observe three obvious findings. One is that our Aug-Target could help all baselines to effectively promote their performance metrics on all three datasets. Another is that different baselines could obtain different metric promotion degrees, although with same AugTarget. Finally, our scheme could also help to refresh a new SOTA peak on baselines.

For examples, the mIoU of origal ACM is 44.17% on MDFA, and it rises to 47.15% by AugTarget, obtaining an obvious promotion of 2.98%. On SIRST-Aug, AugTarget helps ACM to obtain an F-measure improvement of 2.44%, however it makes AGPC only achieve a 1.43% rise. Furthermore, without AugTarget, the current SOTA peak of AUC on Merged, generated by AGPC, is 91.94%. Once cooperated by

AugTarget, new peak could be refreshed to 93.95%, earning 2.01% increment.
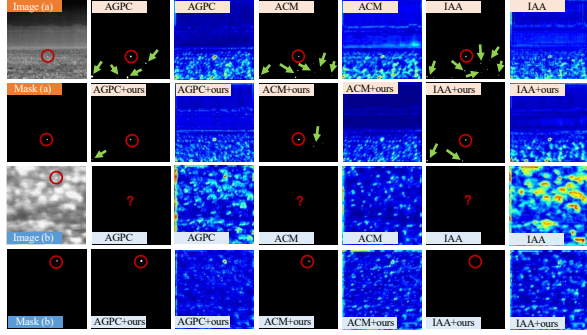

Fig. 3. Visualization comparisons of detection on MDFA.

In Fig.3, we also present two groups of detection comparisons in the complex scenes from MDFA. Red circle denotes that a true target correctly detected, and green arrow means that a false target is wrongly detected. Blue images represent visualized feature maps. In image (a), it's easy to see that three baselines often detect false targets. And in image (b), they don't detect any target. However, they could accurately detect this target via the help of AugTarget.
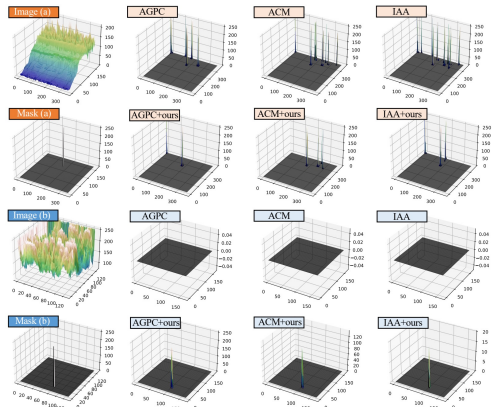

Fig. 4. 3D visualization of complex scenes on MDFA.

Furthermore, corresponding to Fig.3, Fig.4 exhibits two groups of 3D visualization on images (a) and (b), respectively. Each spike indicates a possible target region. It's also believed that our AugTarget could make three baselines more robust to capture the efficient features of true targets in complex background scenes.

3.2. Ablation Study

In order to evaluate the effects of different components, including the TA (Target Augmentation without θ), TA (with θ) and BA (Batch Augmentation) in our AugTarget, we do two groups of ablation experiments. The results on two datasets are listed in Table 2.

From metric comparisons, we could see two finding points. One is that all components in our AugTarget have always effects on helping baseline to improve metrics. The other is that different component assembling schemes could generate different effect degrees on performance metrics.

Table 2. Ablation study on MDFA and SIRST-Aug.

Method	Dataset	TA	TA (θ)	BA	mIoU (%)	F-measure (%)
Baseline (AGPC)	MDFA				48.43	65.25
+AugTarget	MDFA	✓			50.36 ↑1.93	66.99 ↑1.74
+AugTarget	MDFA			✓	49.33 ↑0.90	66.07 ↑0.82
+AugTarget	MDFA	✓	✓		50.82 ↑2.39	67.39 ↑2.14
+AugTarget	MDFA	✓	✓	✓	**51.46** ↑3.03	**67.95** ↑2.70
Baseline (AGPC)	SIRST-Aug				72.88	84.31
+AugTarget	SIRST-Aug	✓			73.84 ↑0.96	84.95 ↑0.64
+AugTarget	SIRST-Aug			✓	73.60 ↑0.72	84.79 ↑0.48
+AugTarget	SIRST-Aug	✓	✓		74.31 ↑1.43	85.26 ↑0.95
+AugTarget	SIRST-Aug	✓	✓	✓	**75.05** ↑2.17	**85.74** ↑1.43

For example, TA could help AGPC to raise mIoU from 48.43% to 50.36% on MDFA. Similarly, BA could also make AGPC obtain the mIoU gain from 48.43% to 49.33%. Moreover, the TA with θ could improve AGPC's F-measure only from 84.31% to 85.26% on SIRST-Aug. However, the assembling of TA(θ) and BA could help AGPC to achieve an F-measure rise from 84.31% to 85.74%. The rise degree of the former is 0.48% less than the latter.
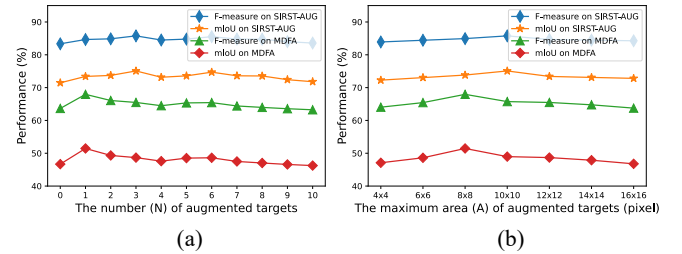

Fig. 5. Performance comparisons with different settings.

We also explore the relation between parameters and performance. Fig. 5 exhibits their relevance. In (a), we could observe that both mIoU and F-measure could reach their peaks when N=1 on MDFA, however they obtain peaks when N=3 on SIRST-Aug. In (b), both metrics obtain their peaks when the maximum size of augmented targets A is $8{\times}8$ on MDFA, while both reach peaks when A is $10{\times}10$ on SIRST-Aug.

4. CONCLUSIONS AND FUTURE WORK

To address the shortage of annotated datasets and overcome the limitations of traditional methods, this paper proposes the AugTarget, a simple and efficient augmentation scheme. It contains target augmentation and batch augmentation. The comparisons and ablation experiments verify that our scheme could effectively boost recently-new models to refresh their SOTA performance peaks. In future work, we will continue to improve our scheme, so as to obtain higher-quality sample and better adaptiveness to complex scenarios.

5. REFERENCES

[1] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems*, 2012, pp. 1106–1114.

[2] Tianfang Zhang, Siying Cao, Tian Pu, and Zhenming Peng, "AGPCNet: Attention-guided pyramid context networks for infrared small target detection," *arXiv preprint arXiv:2111.03580*, 2021.

[3] Yimian Dai, Yiquan Wu, Fei Zhou, and Kobus Barnard, "Asymmetric contextual modulation for infrared small target detection," in *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 2021, pp. 950–959.

[4] Kewei Wang, Shuaiyuan Du, Chengxin Liu, and Zhiguo Cao, "Interior attention-aware network for infrared small target detection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–13, 2022.

[5] Boyang Li, Chao Xiao, Longguang Wang, Yingqian Wang, Zaiping Lin, and et al Li, Miao, "Dense nested attention network for infrared small target detection," *IEEE Transactions on Image Processing*, 2022.

[6] Shengjia Chen, Zhixin Li, and Zhenjun Tang, "Relation R-CNN: A graph based relation-aware network for object detection," *IEEE Signal Processing Letters*, vol. 27, pp. 1680–1684, 2020.

[7] Shuhan Qi, Jianjun Du, Mingyan Wu, Hong Yi, and et al Tang, Linlin, "Underwater small target detection based on deformable convolutional pyramid," in *IEEE International Conference on Acoustics, Speech and Signal Processing*. IEEE, 2022, pp. 2784–2788.

[8] Bin Zhao, Chunping Wang, Qiang Fu, and Zishuo Han, "A novel pattern for infrared small target detection with generative adversarial network," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 5, pp. 4481–4492, 2020.

[9] Qingyu Hou, Zhipeng Wang, Fanjiao Tan, Ye Zhao, and et al Zheng, Haoliang, "RISTDNet: Robust infrared small target detection network," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1–5, 2021.

[10] Jun-Hyung Kim and Youngbae Hwang, "GAN-based synthetic data augmentation for infrared small target detection," *IEEE Transactions on Geoscience and Remote Sensing*, 2022.

[11] Shengjia Chen, Zhixin Li, and Xiwei Yang, "Knowledge reasoning for semantic segmentation," in *IEEE International Conference on Acoustics, Speech and Signal Processing*. IEEE, 2021, pp. 2340–2344.

[12] Zhun Zhong, Liang Zheng, Guoliang Kang, Shaozi Li, and Yi Yang, "Random erasing data augmentation," in *Proceedings of the AAAI conference on artificial intelligence*, 2020, pp. 13001–13008.

[13] Mate Kisantal, Zbigniew Wojna, Jakub Murawski, Jacek Naruniec, and Kyunghyun Cho, "Augmentation for small object detection," *arXiv preprint arXiv:1902.07296*, 2019.

[14] Alexey Bochkovskiy, Chien-Yao Wang, and Hong-Yuan Mark Liao, "YOLOv4: Optimal speed and accuracy of object detection," *arXiv preprint arXiv:2004.10934*, 2020.

[15] Pengguang Chen, Shu Liu, Hengshuang Zhao, and Jiaya Jia, "GridMask data augmentation," *arXiv preprint arXiv:2001.04086*, 2020.

[16] Haotao Wang, Chaowei Xiao, Jean Kossaifi, Zhiding Yu, Anima Anandkumar, and Zhangyang Wang, "AugMax: Adversarial composition of random augmentations for robust training," *Advances in neural information processing systems*, vol. 34, pp. 237–250, 2021.

[17] Dan Hendrycks, Norman Mu, Ekin D Cubuk, Barret Zoph, Justin Gilmer, and Balaji Lakshminarayanan, "AugMix: A simple data processing method to improve robustness and uncertainty," in *International Conference on Learning Representations*, 2020.

[18] Huan Wang, Luping Zhou, and Lei Wang, "Miss detection vs. false alarm: Adversarial learning for small object segmentation in infrared images," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2019, pp. 8509–8518.

[19] Xuerui Dai, Xue Yuan, and Xueye Wei, "Data augmentation for thermal infrared object detection with cascade pyramid generative adversarial network," *Applied Intelligence*, vol. 52, no. 1, pp. 967–981, 2022.

[20] Tianfang Zhang, Zhenming Peng, Hao Wu, Yanmin He, Chaohai Li, and Chunping Yang, "Infrared small target detection via self-regularized weighted sparse model," *Neurocomputing*, vol. 420, pp. 124–148, 2021.

[21] Yantao Wei, Xinge You, and Hong Li, "Multiscale patch-based contrast measure for small infrared target detection," *Pattern Recognition*, vol. 58, pp. 216–226, 2016.

[22] Landan Zhang and Zhenming Peng, "Infrared small target detection based on partial sum of the tensor nuclear norm," *Remote Sensing*, vol. 11, no. 4, pp. 382, 2019.