Applications

JDMS

Journal of Defense Modeling and Simulation: Applications, Methodology, Technology
1–12

DOI: 10.1177/15485129241233299
journals.sagepub.com/home/dms
Sage

Military camouflaged object detection with deep learning using dataset development and combination

Kyo-Seong Hwang and Jungmok Ma

Abstract

Camouflaged object detection (COD) is one of the emerging artificial intelligence technologies. COD identifies objects that require attention and time to detect with human eyes due to the similarity in texture or color to the surrounding environment. Despite the importance of camouflage and its detection in military, there is a lack of military camouflaged object detection research. Previous studies point out that the general COD has not been well studied due to the lack of camouflaged datasets, and the situation is worse in the military domain. This study aims at tackling the challenge in two directions. First, we carefully assemble the military camouflaged object (MCAM) dataset, including camouflaged soldiers and people as well as camouflaged military supplies for military COD. The experiment shows that MCAM can generate better performance results than the other benchmark datasets (CAMO, COD10K). Second, military (MCAM) and nonmilitary camouflage datasets (benchmark datasets) are combined and tested to overcome data scarcity. The experiment shows that the nonmilitary camouflage datasets are effective for military COD at a certain level, and a proper combination of military and nonmilitary camouflage datasets can improve the detection performance.

Keywords

Camouflaged object detection, military camouflage, deep learning, data combination

1. Introduction

With the rapid development of science and technology, the utilization of artificial intelligence (AI) technology is expanding beyond diverse fields.[1] Among these diverse fields, the military domain is currently receiving significant attention due to the current international security situation. The battlefield in modern warfare is becoming more precise, unmanned, automated, and networked. AI technology is one of the most important technologies that lead to the change of modern battlefield, and many countries are attempting to adopt AI technology in their future defense systems.[2] Many experts anticipate a great impact on the global economy, including the military sector, in the next decade with the application, development, and adoption of AI. According to reasonable estimates, the economic impact of AI during this period is expected to range from \$1.49 trillion to \$2.95 trillion.[3] Moreover, NATO (North Atlantic Treaty Organization) has designated AI as one of the emerging and disruptive technologies (EDTs) from 2020 to 2040. NATO acknowledges the potential of AI to have a revolutionary impact on NATO operations and capabilities, serving as a key element in converting big data into actionable knowledge and ultimately providing NATO with a decision advantage.[4]

For the application of AI technology to the military sector, the complicated military surroundings make it difficult to use commercial AI directly, such as smoke and flames

Department of Defense Science, Korea National Defense University, Republic of Korea

Corresponding author:
Jungmok Ma, Department of Defense Science, Korea National Defense University, Hwangsanbeol-ro 1040, Nonsan, 33021, Republic of Korea.
Email: jxm1023@gmail.com

from weapons, camouflaged and occluded objects, and motion blur.[5] To overcome the difficulty, this study focuses on the camouflage detection technology in the military domain which can neutralize camouflage technology.

Camouflaged object detection (COD) aims at identifying camouflaged objects that "are seamlessly embedded in their surroundings"[6] and is a type of image segmentation. Figure 1 shows the types of objects in image segmentation. The first one (Figure 1(a)) is an original image. The second one (Figure 1(b)) includes objects that correspond to existing real objects in the original image such as a plate, oranges, a peach, kiwi fruits, nuts, a worm, and a moth. The third one (Figure 1(c)) shows salient objects that are visible and prominent among the generic objects with their large size or vivid color such as kiwi fruits and a peach. The last one (Figure 1(d)) indicates camouflaged objects that require attention and time to detect with human eyes due to the similarity in texture or color to the surrounding environment. The camouflaged objects are a worm and a moth.

The COD technology is expected to be popularly utilized in the field of military Intelligence, Surveillance, Target Acquisition, and Reconnaissance (ISTAR). The authors view that the COD technology can be used along with a general object detector so that the performance of automatic object detection will be improved. Research on COD using deep learning technology first emerged in 2017.[7] Therefore, not much research has been conducted on COD for military purposes. The research on COD in military fields is summarized in Table 1. Zheng et al.[8] assembled a camouflaged people dataset with 1000 images and developed the Dense Deconvolution Network (DDCN) for camouflaged people detection. Fang et al.[9] constructed a camouflaged people dataset with 2600 images and proposed the Strong Semantic Dilation Network (SSDN). Liu et al.[10] made various annotated datasets of camouflaged soldiers which contain 1600 images (600 for labeled, 1000 for unlabeled) and proposed a semisupervised learning strategy to improve the existing COD network. Yi et al.[5] proposed a military object detection dataset (6000 images) since there was no object detection benchmark in military. However, camouflage was just one class out of 20 categories. Zhang et al.[11] built a military camouflaged people dataset captured from 60 videos on the Internet and tested YOLOv5 with the coordinate attention module. Liu and Di[12] made the MHCH2022 dataset and proposed military high-level COD network (MHNet), characterized by four innovative modules. From the previous studies, it can be found that the lack of military camouflaged object data is the main obstacle. It is hard to get sufficient military datasets such as camouflaged soldiers and military supplies due to military security and regulations.[13]

To address this issue, we propose the military camouflaged object (MCAM) dataset consisting of camouflaged soldiers, people, and military supplies. Also, on the basis of the fact that animal camouflage has the same function as military camouflage, we test a combination of benchmark datasets (CAMO, COD10K) consisting mainly of camouflaged animals and the proposed MCAM dataset.

In short, there are two contributions to our paper:

1. We construct MCAM including camouflaged soldiers and people as well as camouflaged military supplies. MCAM should contain the following challenging attributes to reflect a complex military environment: multiobjects, big object, small object, out of view, occlusion, and indefinable boundary. MCAM is analyzed and compared with other camouflaged datasets.
2. It is the first attempt to combine military (MCAM) and nonmilitary datasets (benchmark datasets) to overcome the lack of military datasets in the COD field. The effect of nonmilitary camouflaged object datasets is tested for military camouflaged detection as a single training dataset and augmented dataset to MCAM.

2. Related works

2.1. Datasets

Data is critical in deep learning–based object detection models. Generally, the bigger the size of a dataset, the better the object detection models will perform. Also, when a dataset is assembled in accordance to the objective of detection, the object detection models perform better with the dataset.[5] There are some popular open datasets for

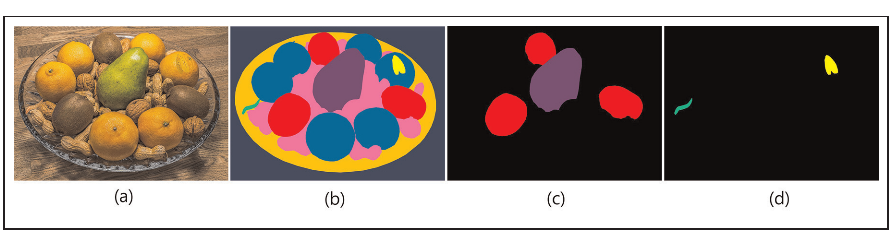

Figure 1. Types of objects in image segmentation: (a) image, (b) generic object, (c) salient object, and (d) camouflaged object.

Table 1. Research on COD in military field.

Researcher	Data quantity	Algorithm	Contents
Zheng et al.[8]	1000	DDCN	First military research on camouflaged people data
Fang et al.[9]	2600	SSDN	Adding convolutions expanding the receptive field in CNN to detect camouflage patterns
Liu et al.[10]	1600	Semi-SINet	Improving performance through semi-supervised learning
Yi et al.[5]	6000	LGA-RCNN	Creating benchmark data in the military object detection
Zhang et al.[11]	1000	CM-YOLOv5s	Adaptation of the YOLOv5s algorithm for COD
Liu and Di[12]	3000	MHNet	Construction of the MHCD2022 Dataset and proposal of MHNet

CNN: convolutional neural network; COD: camouflaged object detection; DDCN: dense deconvolution network; LGA-RCNN: loss-guided attention region-based CNN; SSDN: strong semantic dilation network.

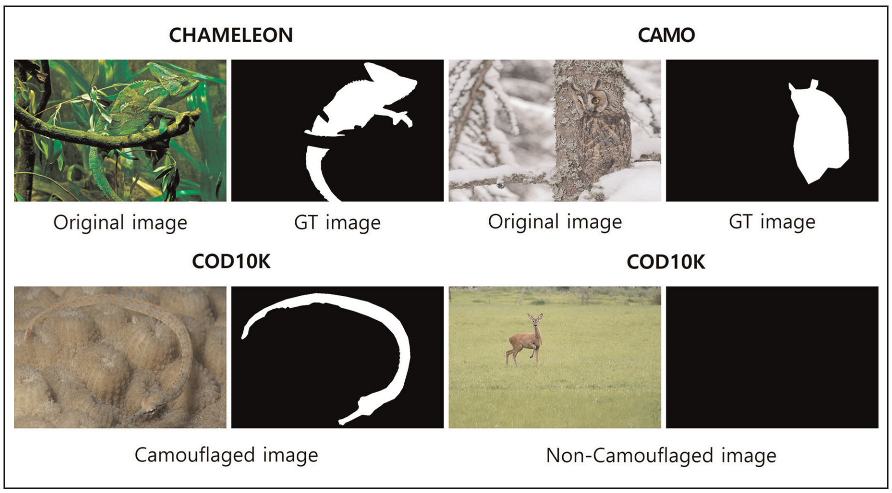

Figure 2. Examples of benchmark datasets.

object detection such as MS COCO,[14] Pascal VOC,[15] and so on.

In this paper, we focus on camouflaged object datasets, and there are three benchmark datasets (CHAMELEON,[16] CAMO,[17] COD10K[6]) which are widely used in the related research. Note that COD has not been actively studied, and COD10K was proposed as the first complete dataset from a camouflage perspective. The benchmark dataset consists of an original image and a Ground Truth (GT) image as shown in Figure 2. Unlike the original image that expresses all color values, the GT image is divided into black background pixels and white camouflaged object pixels since COD algorithm only detects camouflaged objects without classification. The benchmark datasets are mainly composed of camouflaged and noncamouflaged animals in nature. The numbers of images and classes of the benchmark datasets are provided in Table 2. Since the CHAMELEON dataset has only 76 images, the other two datasets are used as benchmark datasets.

2.2. SINet-V2 COD algorithm

There have been some attempts to use deep learning for COD. Le et al. [17] proposed the Anabranch Network which consists of a convolutional neural network (CNN) for classifying camouflaged and noncamouflaged images and a fully connected network for segmentation. Yi et al.[5] used the two-staged object detection method, RCNN (region-based CNN) with an additional module, Loss-Guided Attention, to focus on more discriminative region of objects. Zhang et al.[11] used the one-stage object detection network, YOLOv5 (You only Look Once) with an additional module, Coordinate Attention, to improve the network's attention to camouflaged objects. Jiang et al.[18] proposed the cascade perception network (CPNet) with a ternary cascade perception module which splits into a feature perception module, a spatial perception module, and a key point perception module for improving COD. Fan et al.[6] developed the search identification network

Table 2. Numbers of images and classes of benchmark datasets.

Name	Total number of images		Class
	Camouflage	Noncamouflage	
CHAMELEON[16]	76		Land animal (67), Marine animal (9)
	76	—	
CAMO[17]	2500		Land animal (1700), Marine animal (350), People (450)
	1250	1250	
COD10K[6]	10,000		Land animal (8767), Marine animal (1233)
	5066	4934	

CAMO: camouflaged object; CHAMELEON: cryptic hidden animals masked in environment labelled and evaluated; COD10K: camouflaged object detection 10k images.

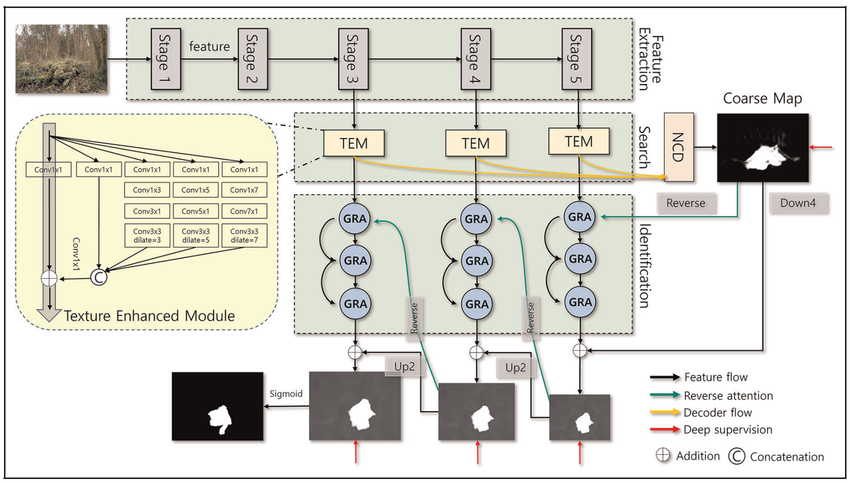

Figure 3. Pipeline of SINet-V2 algorithm (revised from Fan et al.[19]).

(SINet) which includes two modules, search and identification modules inspired by hunting. SINet was compared to 12 other detection models with three benchmark datasets (CHAMELEON, CAMO, COD10K) and obtained the best performance. Liu et al.[10] used the SINet with an additional edge attention module to enhance the detection of military camouflaged people. Since this paper focuses on the military COD using dataset development and combination, the newest version of SINet called SINet-V2 is utilized as a COD baseline algorithm. The SINet-V2 is currently the best performance algorithm in the field of COD.[19]

As shown in Figure 3, SINet-V2 is largely composed of feature extraction, search, and identification modules. The feature extraction module extracts characteristics of input data with Res2Net-50.[20] The search module extracts feature candidates using texture-enhanced modules (TEM) for Stages 3–5, which is related to the image segmentation process, and maintains semantic consistency between layers. Next, the approximate location of camouflaged objects is obtained through a neighbor connection decoder (NCD). In the Identification module, the approximate location of the acquired camouflaged object is subdivided into a deeper layer as it cascades through the group reversal attachment (GRA) to locate the final camouflaged object. We utilized the structure of the SINet-V2 algorithm in our research with different hyperparameters as described in Section 4.1.

3. MCAM dataset

In this paper, we carefully assembled a dataset using 1000 images including camouflaged soldiers and military supplies. It is called the MCAM dataset. The MCAM dataset will be analyzed and compared with other camouflaged datasets in this section.

3.1. Data collection and preprocessing for MCAM

The MCAM dataset was assembled with three steps: data collection, preprocessing, and labeling. Manual web crawling was performed for data collection from websites

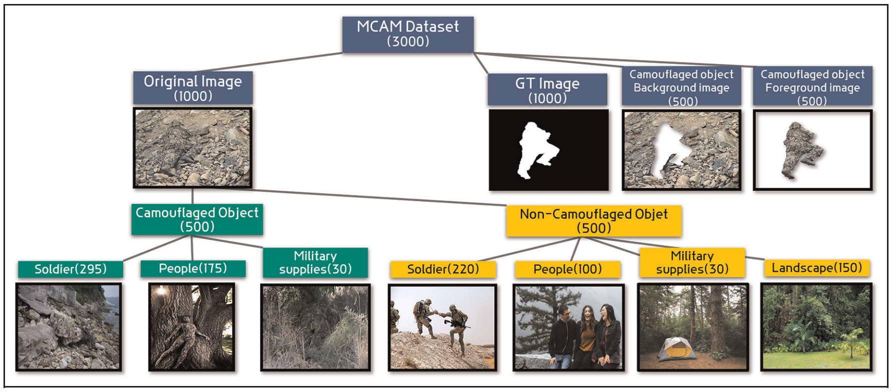

Figure 4. Organization of MCAM dataset.

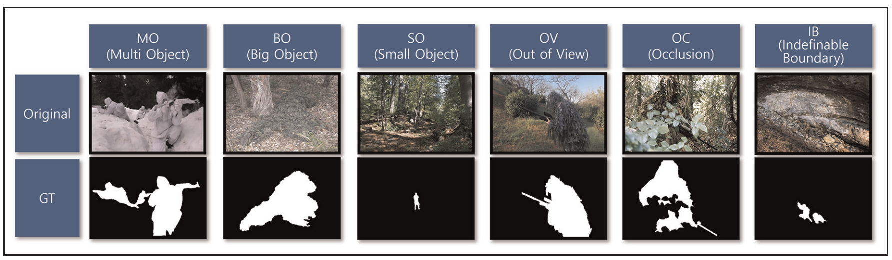

Figure 5. MCAM dataset attributes.

such as Google, Unsplash, Pixabay, Yandex, and FreeImages. The search terms included ''Military camouflage,'' ''Hidden soldier,'' ''Camouflaged soldier,'' ''Camouflaged people,'' ''Camouflaged military supplies,'' and ''Camouflaged Object'' for military camouflaged objects and ''Soldier,''‘People ''Military supplies,'' ''Weapon,'' and ''Background'' for noncamouflaged objects. Among the camouflaged objects, camouflaged people with body paint were considered as military camouflaged objects to reflect the complexity of urban warfare. Note that there are also camouflaged people in CAMO, but the images are different. In the data preprocessing step, the collected images were inspected and removed if they failed the inspection. The Lasso tool in Photoshop was used for data labeling. The tool generated GT, foreground, and background images from the original images as shown in Figure 4.

The MCAM dataset consists of 1000 original images, 1000 GT images (500 camouflaged object images and 500 black background images), 500 foreground images corresponding to camouflaged objects, and 500 background images excluding camouflaged objects. The original and GT images are used as training, validation, and test data for SINet-V2. The foreground and background images of camouflaged objects are used for the comparison of other benchmark datasets. The original images consist of 500 camouflaged and 500 noncamouflaged object images. The camouflaged object images are divided into camouflaged soldiers wearing ghillie suits or military uniforms, camouflaged people with body paint, and camouflaged military supplies such as tanks, military vehicles, rifles, and tents. The noncamouflaged object images include distinct objects such as soldiers, people, and military supplies. Also, landscape images without camouflaged objects are included in the noncamouflaged object images. For clarification, camouflaged objects are determined only by the harmony between the objects and the backgrounds. For example, while a soldier with a green combat uniform in the forest can be a camouflaged person, the same soldier with a green combat uniform on the sand can be a noncamouflaged person. The reason why the MCAM dataset is composed of camouflaged images and noncamouflaged images is that the algorithm can boost the performance of COD by simultaneously training correct and noncorrect data.[17]

3.2. Analysis of MCAM

To understand the MCAM dataset, the 6 challenging attributes were utilized from the previous studies:[6,19] MO (multiobjects), BO (big object), SO (small object), OV (out of view), OC (occlusion), and IB (indefinable boundary) as shown in Figure 5. SC (shape complexity) was

Table 3. Explanation of dataset attributes.

Attribute	Explanation
MO	Image contains at least 2 objects
BO	Ratio of pixel between object and image $\geqslant$ 0.5
SO	Ratio of pixel between object and image $\leq$ 0.1
OV	Object is clipped by image boundaries
OC	Object is partially occluded by other objects
IB	Foreground and background of image have similar colors (Bhattacharyya distance of RGB histogram between foreground and background $\leq$ 0.9)

BO: big object; IB: indefinable boundary; MO: multiobjects; OC: occlusion; OV: out of view; RGB: red, green, and blue; SO: small object.

excluded since the attribute was difficult to be objectively evaluated. The description for each corresponding attribute is provided in Table 3.

Based on the 6 attributes, the co-occurrence matrix of the MCAM dataset is provided in Figure 6. The horizontal and vertical axes correspond to each attribute, and the numbers in the distribution diagram refer to the number of images with both horizontal and vertical attributes. For example, the number of images with both MO and OV is 17.

The resolution distribution of the MCAM dataset is shown in Figure 7. The horizontal axis is the number of pixels in the horizontal direction of the image, and the vertical axis is the number of pixels in the vertical direction of the image. The MCAM dataset is built to contain more high-resolution images than the other benchmark datasets.

3.3. Comparison of datasets

Comparisons between the MCAM dataset and benchmark datasets (CAMO, COD10K) were conducted in terms of three aspects: generalization, photography, and color contrast.

3.3.1. Generalization. Cross-dataset generalization[21] is a method of showing the bias of a dataset by cross-applying multiple datasets to train and test with the same object detection model. That is, the method provides the performance comparison (percent drop) between one dataset (self) and other comparing datasets (others) as test data. In this paper, the number of training and test data was set to 400 and 100, respectively, with SINet-V2. A Structure Measure or S-measure (Equation (1) in Section 4.2) was used as a performance result of COD, and Table 4 shows the average of the S-measures from 5 experiments with a random extraction of training and test data. Although the percent drops of MCAM and COD10K are similar, the mean others value of MCAM (0.645) is similar to that of CAMO (0.677), which indicates good generalization. Furthermore, it can be seen that when MCAM dataset was used as test data, the performance of CAMO (0.606) and COD10K (0.558) showed the lowest among other cases. In other words, military camouflaged data are more challenging to detect.

3.3.2. Photography. Since the most of image data are photographed by humans, there can be an issue of object size and center bias (tendency to place a photographing object in the middle of the image).[22] The point is that the issue of photography should not be a major factor in generating different performance of COD algorithm when datasets are compared. The first graph in Figure 8 shows the statistical distributions representing the camouflaged object size of the MCAM, CAMO, and COD10K datasets. The X value (object size) represents the ratio of the number of camouflaged object pixels to the number of entire image pixels. The object size ratio of MCAM dataset is ranged from 0.01 to 0.66, with an average value of 0.14. The size of

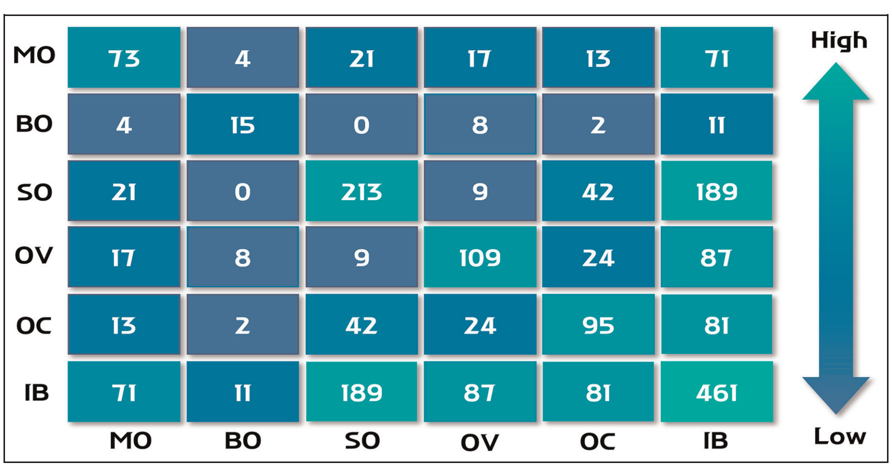

Figure 6. Co-occurrence matrix of MCAM dataset.

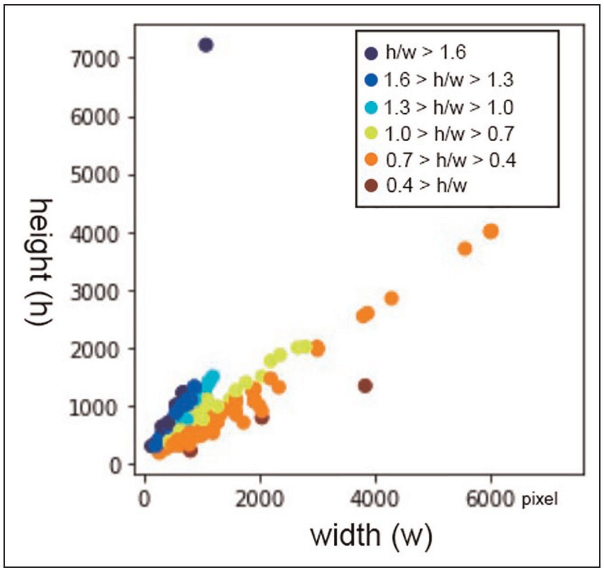

Figure 7. Resolution distribution of MCAM dataset.

camouflaged object in MCAM tends to be smaller than CAMO dataset and larger than COD10K dataset.

The second graph in Figure 8 represents the center bias distributions of the datasets. The X value (center bias) is the distance from the center of the image to the center of the camouflaged object divided by the distance from the center to the edge of the image. The center of the camouflaged object is the center of the bounding box around the object. MCAM dataset shows a similar range of center bias ratio compared with the benchmark datasets.

3.3.3. Color contrast of foreground and background. The third graph in Figure 8 represents the color contrast distributions of the datasets. The X value (color contrast) is the Bhattacharyya distance between the RGB distribution of the foreground corresponding to camouflaged objects and the RGB distribution of the background where camouflaged objects were removed from the image.[22] The lower X value means that the difference between the RGB distribution of the foreground and background is smaller, which indicates that the camouflage degree of camouflaged objects is higher. The MCAM dataset has a broader and various camouflage degree compared to the benchmark datasets, but its overall camouflage degree is lower than that of the benchmark datasets. The analysis shows that military camouflage objects such as soldiers, people, and supplies have a lower and wider camouflage degree than that of animals in the photographs in these databases. A camouflage degree of a dataset is crucial in learning and testing of COD algorithm, and a wide range of a camouflage degree also can decrease the performance of COD due to the versatility.

4. Experiments

The flowchart of the research is depicted in Figure 9. In Section 3, data collection, preprocessing, labeling, and analysis have been conducted. In the experimental phase, the MCAM dataset created in the study and the benchmark datasets (CAMO and COD10K) were combined in various ways for the purpose of this research (i.e., the two contributions described in Section 1). Subsequently, the

Table 4. Cross-dataset generalization with structure measure scores.

Test on:Train on:	CAMO	COD10K	MCAM	Self	MeanOthers	PercentDrop
CAMO	0.800	0.747	0.606	0.800	0.677	15.3%
COD10K	0.615	0.728	0.558	0.728	0.587	19.3%
MCAM	0.638	0.652	0.806	0.806	0.645	19.9%

MCAM: military camouflaged object.

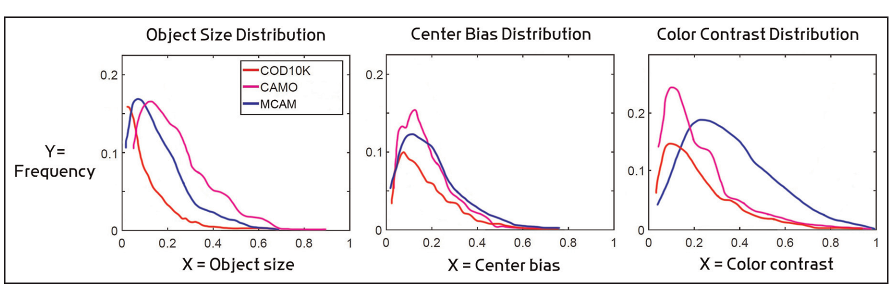

Figure 8. Object size, center bias, and color contrast distributions for 3 datasets.

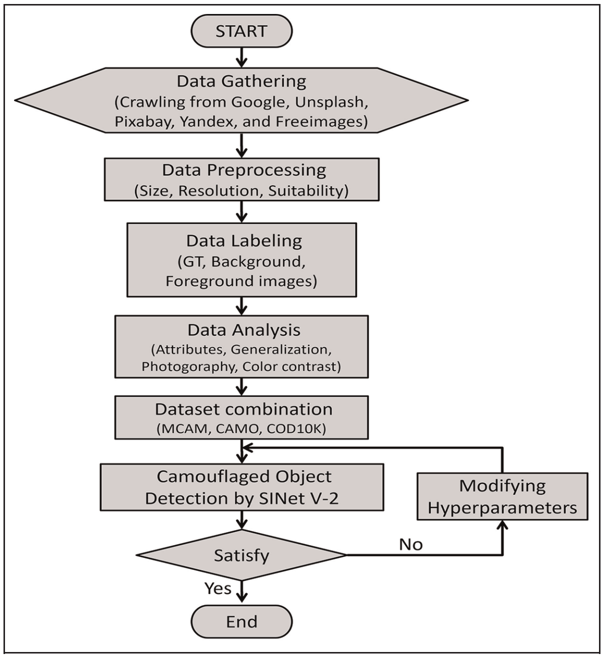

Figure 9. Research flow chart.

performance of COD was measured and analyzed using the SINet-V2 algorithm for each dataset created.

4.1. Experimental objective and setup

The experimental objectives are threefold: 1) performance test with proposed MCAM dataset, 2) performance test with nonmilitary benchmark datasets (CAMO, COD10K), and 3) combination effect of MCAM and two benchmark datasets for military COD. SINet-V2 is used as COD baseline algorithm.

The experiment was conducted using Google Colaboratory, which provides high-performance GPUs based on cloud environments with CPU of the Intel Xeon® CPU at 2.30 GHz, 25.46 GB RAM, and GPU of the Tesla P100 (16 GB). The hyperparameters in the training process of SINet-V2 are as follows: epoch of 100, learning rate of 0.001, and batch size of 36. For data splitting, the ratio of training and validation data was set to 3:1 by class which is often used to prevent overfitting as shown in Table 5. For example, MCAM has 3 classes of camouflaged objects (soldier, people, and military supplies) and the camouflaged training set has 221 data of soldiers (3/4 of 295). For performance testing, a total of 200 data (100 camouflaged and 100 noncamouflaged) was randomly extracted from MCAM and called Military Test. The Military Test data will be used as test data for all the datasets.

The experimental design is shown in Table 6. Experiment #1 was designed to check the performance of SINet-V2 trained with MCAM for military COD. Experiments #2–#3 were designed to check the performance of SINet-V2 trained with nonmilitary benchmark datasets mainly composed of camouflaged animals. Experiments #4–#6 were designed to see the effect of the combination among datasets. Note that the different sample sizes of training and validation sets are intended to emphasize the necessity of military camouflage data after conducting preliminary experiments with the same training samples.

4.2. Evaluation indexes

This paper uses four evaluation indexes in the experiment: S-measure (structure measure), E-measure (enhanced-alignment measure), wF-measure (weighted F-measure), and MAE (mean absolute error). The evaluation indexes have been usually used in the performance test of COD.[8,9,23] The first three indexes are similar to human visual functions to meet the purpose of COD, and higher values them represent better COD performance. The three indexes were calculated with the camouflaged Military Test set in Table 5.

S-measure[24] is defined as

$$S_\alpha = \alpha s_0 + (1-\alpha)s_r \tag{1}$$

where s_0 is the object-aware structural similarity between an SM (nonbinary saliency map) and a GT map, s_r is the region-aware structural similarity between an SM and a GT map, and α is a constant set as 0.5.

Table 5. Dataset split.

	MCAM		CAMO		COD10K		Military test
	Training set	Validation set	Training set	Validation set	Training set	Validation set	Test set
Camouflaged	300	100	938	313	3800	1266	100
Noncamouflaged	300	100	937	312	3700	1234	100
Total	600	200	1875	625	7500	2500	200

MCAM: military camouflaged object.

Table 6. Experimental design.

Experiment	Dataset for training and validation
#1(M)	MCAM
#2(C1)	CAMO
#3(C2)	COD10K
#4(MC1)	MCAM + CAMO
#5(MC2)	MCAM + COD10K
#6(MCC)	MCAM + CAMO + COD10K

MCAM: military camouflaged object.

Table 7. Quantitative result of experiment.

	$S_\alpha\uparrow$	$E_\varnothing\uparrow$	$F_\beta^w\uparrow$	$MAE\downarrow$
#1(M)	0.810	0.855	0.703	0.062
#2(C1)	0.775	0.815	0.647	0.075
#3(C2)	0.690	0.687	0.507	0.094
#4(MC1)	**0.835**	**0.887**	**0.738**	0.025
#5(MC2)	0.771	0.802	0.634	**0.015**
#6(MCC)	0.793	0.822	0.668	0.037

The best result is highlighted in bold.

E-measure[25] is defined as

$$E_\varnothing = \frac{1}{\omega \times h}\sum_{x=1}^{\omega}\sum_{y=1}^{h}\theta(\varnothing_{FM}) \tag{2}$$

where ω is the width of map, h is the height of map, and $\theta(\varnothing_{FM})$ is the enhanced alignment matrix.

wF-measure[26] is defined as

$$F_\beta^\omega = \frac{(1+\beta^2)precision^\omega \times \ Recall^\omega}{\beta^2 Precision^\omega + Recall^\omega} \tag{3}$$

where ω and β are the parameters that control the flaws and different weights, respectively. $Precision = \frac{TP}{TP+FP}$, $Recall = \frac{TP}{TP+FN}$ (TP: true-positive, FP: false-positive, FN: false-negative in confusion matrix).

While the three indexes focus on the detection of camouflaged objects when camouflaged objects are present, MAE focuses on the false detection of camouflaged objects when camouflaged objects are not present. If the trained SINet-V2 generates many false-positive results, that can affect the efficiency of military operations. MAE is capable of pixel-by-pixel evaluating background as well as foreground of images and calculated with the noncamouflaged Military Test set in Table 5.

MAE[27] is defined as

$$MAE = \frac{1}{w \times h}\sum_{x=1}^{\omega}\sum_{y=1}^{h}|P(x,y) - \ G(x,y)| \tag{4}$$

where w is the width of images, h is the height of images, $P(x,y)$ is the prediction value of (x,y), and $G(x,y)$ is the GT value of (x,y).

5. Results

The quantitative results of the experiment are provided in Tables 7 and 8. Overall, the trained SINet-V2s based on the experimental design in Table 6 show the performance as follows: S_α of 0.690–0.835, $E_\varnothing$ of 0.687–0.887, F_β^w of 0.507–0.738, and MAE of 0.015–0.094 as shown in Table 7. From Experiment #1–#3, it can be seen that MCAM outperformed the other benchmark datasets over the four indexes even though it has the smallest training and validation sample sizes. Also, the benchmark datasets showed the usefulness for military COD (S_α is greater than 0.6) although they were mainly composed of animals in nature. CAMO showed better results than COD10K even though COD10K had more data. The lower performance results of COD10K for camouflaged soldiers and people in Table 8 explained the performance difference. Note that CAMO has a class of people as shown in Table 2.

Experiments #4–#6 are the cases of combination of the datasets, and MC1 or MCAM + CAMO showed the best result over the first three evaluation indexes, which is (S_α,$E_\varnothing$,F_β^w) = (0.835, 0.887, 0.738). In other words, MC1 is the best for detecting military camouflaged objects when military camouflaged objects are present. MC2 showed the best performance for MAE, which is related to the false detection of camouflaged objects when camouflaged objects are not present. By comparing the single cases and the combination cases, it can be seen that the increase of data size did not simply enhance the performance (e.g., Experiment #1 vs Experiments #5 and #6). However, the experiment indicates that a proper augmentation can be helpful for improving the performance of military COD (e.g., Experiment #1 vs Experiment #4).

Therefore, MCAM showed the best result as a single dataset and MC1 provided the best detection as a combined dataset for military COD. Figure 10 shows the qualitative results of the experiment, and the yellow boxes represent the best results. Figure 11 provides two examples of military COD with MC1. From the original image, the GT image and the COD image with MC1 can be visually compared.

6. Conclusion

Despite the importance of camouflage and its detection in military, there is a lack of military COD research. The previous studies pointed out that the general COD has not been well studied due to the lack of camouflaged datasets and the situation is worse in the military domain. This

Table 8. S_α and *MAE* for each class in experiment.

Evaluation index	Camouflaged object detection ($S_\alpha\uparrow$)			False camouflaged object detection ($MAE\downarrow$)		
Class	Soldier	People	Military supplies	Soldier	People	Military supplies
#1(M)	0.826	0.800	0.757	0.052	0.077	0.080
#2(C1)	0.779	0.764	0.770	0.060	0.086	0.073
#3(C2)	0.671	0.581	0.761	0.084	0.044	0.115
#4(MC1)	**0.842**	**0.826**	**0.811**	0.022	0.023	0.047
#5(MC2)	0.752	0.783	0.769	**0.013**	0.019	**0.011**
#6(MCC)	0.808	0.790	0.730	0.028	**0.012**	0.083

MAE: mean absolute error.
The best result is highlighted in bold.

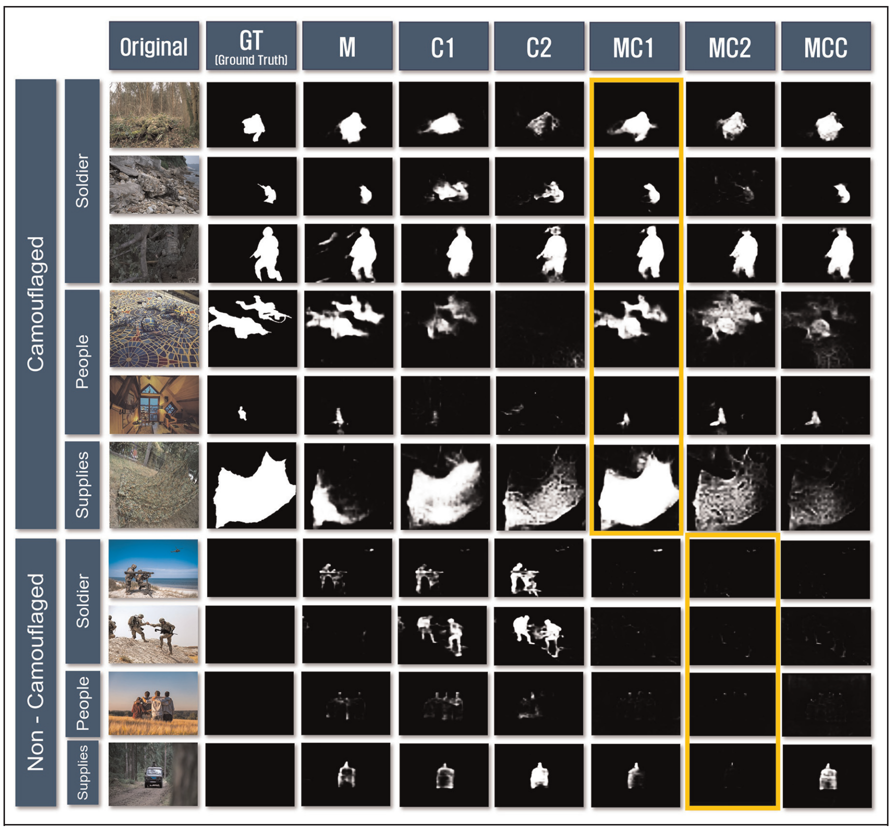

Figure 10. Qualitative result of experiment.

study aims at tackling the challenge in two directions. First, we carefully assembled the MCAM dataset, including camouflaged soldiers and people as well as camouflaged military supplies for military COD. The experiment showed that MCAM generated better performance result than the other benchmark datasets (CAMO, COD10K). Second, military (MCAM) and nonmilitary camouflage datasets (benchmark datasets) were combined and tested to overcome data scarcity. The experiment showed that the nonmilitary camouflage datasets were effective for military COD in a certain level and a proper combination of military and nonmilitary camouflage datasets could improve the detection performance.

In the future, the MCAM dataset needs to be enriched by including more various military supplies. Considering the difficulty of military data collection, synthetic data generation from simulated environments such as game engines and simulation models can be explored. It will be interesting to utilize the deep learning-based military COD for improving camouflage. Furthermore, to enhance COD performance, it will be essential to explore research on the fusion of various technologies such as robotics, sensors,

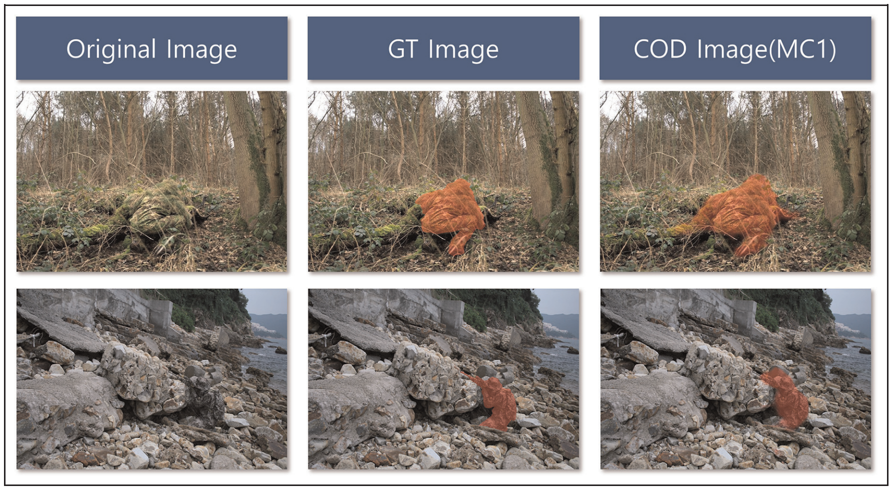

Figure 11. Qualitative result of COD with MC1.

and generative AI. Also, for military applications, there is a need for the enhancement of computing power to detect camouflaged objects in videos.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iDs

Kyo-Seong Hwang https://orcid.org/0000-0002-0985-0808
Jungmok Ma https://orcid.org/0000-0001-5390-8483

References

1. Oh Y, Busogi M, Ransikarbum K, et al. Real-time quality monitoring and control system using an integrated cost effective support vector machine. *J Mech Sci Technol* 2019; 33: 6009–6020.
2. Ryu J and Ma J. Automatic target recognition of military SAR images using super-resolution based convolutional neural network. *J Inst Control Robot* 2022; 28: 22–27.
3. Chen N, Christensen L, Gallagher K, et al. Global economic impacts associated with artificial intelligence. *Analysis Group*, 2016, https://www.analysisgroup.com/globalassets/content/insights/publishing/ag_full_report_economic_impact_of_ai.pdf
4. Reding DF and Eaton J. *Science & technology trends 2020-2040*. Brussels: NATO Science & Technology Organization, 2020, pp. 71–73, https://securitydelta.nl/media/com_hsd/report/406/document/190422-ST-Tech-Trends-Report-2020-2040.pdf
5. Yi X, Wu J, Ma B, et al. MOD: benchmark for military object detection. arXiv:2104.13763v2, 2021, https://arxiv.org/pdf/2104.13763v1.pdf
6. Fan D-P, Ji G-P, Sun G, et al. Camouflaged object detection. In: *2020 IEEE/CVF conference on computer vision and pattern recognition (CVPR)*, Seattle, WA, 13–19 June 2020, pp. 2777–2787. New York: IEEE.
7. Li Y, Hou X, Koch C, et al. The secrets of salient object segmentation. In: *2014 IEEE conference on computer vision and pattern recognition*, Columbus, OH, 23–28 June 2014, pp. 280–287. New York: IEEE.
8. Zheng Y, Zhang X, Wang F, et al. Detection of people with camouflage pattern via dense deconvolution network. *IEEE Signal Process Lett* 2018; 26: 29–33.
9. Fang Z, Zhang X, Deng X, et al. Camouflage people detection via strong semantic dilation network. In: *Proceedings of the ACM turing celebration conference-China*, Chengdu China, 17–19 May 2019, pp. 1–7. New York: ACM.
10. Liu Y, Wang CQ and Zhou YJ. Camouflaged people detection based on a semi-supervised search identification network. *Def Technol* 2021; 21: 176–183.
11. Zhang W, Zhou Q, Li R, et al. Research on camouflage target detection method based on improved YOLOv5. *J Phys Conf Ser* 2022; 2284: 012018.
12. Liu M and Di X. Extraordinary MHNet: military high-level camouflage object detection network and dataset. *Neurocomputing* 2023; 549: 126466.
13. Yoon JH. Issues and prospects of AI utilization in the defense field: organizational capability, technology maturity, institutional congruence. *Korean J Def Anal* 2021; 33: 453–479.
14. Lin T-Y, Maire M, Belongie S, et al. Microsoft COCO: common objects in context. In: Fleet D, Pajdla T, Schiele B, et al. (eds) *Computer vision—ECCV 2014. ECCV 2014*. Cham: Springer, 2014, pp. 740–755.

15. Everingham M, Van Gool L, Williams CK, et al. The Pascal visual object classes (VOC) challenge. *Int J Comput Vis* 2010; 88: 303–338.
16. Skurowski P, Abdulameer H, Błaszczyk J, et al. Animal camouflage analysis: CHAMELEON database. Unpublished manuscript, 2018, https://www.polsl.pl/rau6/chameleon-database-animal-camouflage-analysis
17. Le T-N, Nguyen TV, Nie Z, et al. Anabranch network for camouflaged object segmentation. *Comput Vis Image Underst* 2019; 184: 45–56.
18. Jiang X, Cai W, Ding Y, et al. Camouflaged object detection based on ternary cascade perception. *Remote Sens* 2023; 15: 1188.
19. Fan D-P, Ji G-P, Cheng M-M, et al. Concealed object detection. *IEEE Trans Pattern Anal Mach Intell* 2022; 44: 6024–6042.
20. Gao S, Cheng M-M, Zhao K, et al. Res2Net: a new multi-scale backbone architecture. *IEEE Trans Pattern Anal Mach Intell* 2019; 43: 652–662.
21. Torralba A and Efros AA. Unbiased look at dataset bias. In: *CVPR 2011*, Colorado Springs, CO, 20–25 June 2011, pp. 1521–1528. New York: IEEE.
22. Li C, Zhao X and Wan Y. The camouflage color target detection with deep networks. In: *Neural information processing: 24th international conference, ICONIP 2017*, Guangzhou, China, 14–18 November 2017.
23. Xu X, Chen S, Lv X, et al. Guided multi-scale refinement network for camouflaged object detection. *Multimed Tools Appl* 2023; 82: 5785–5801.
24. Cheng MM and Fan D-P. Structure-measure: a new way to evaluate foreground maps. *Int J Comput Vis* 2021; 129: 2622–2638.
25. Fan D-P, Gong C, Cao Y, et al. Enhanced-alignment measure for binary foreground map evaluation. In: *Proceedings of the 27th IJCAI*, Stockholm, 13–19 July 2018, pp. 698–704, https://www.ijcai.org/proceedings/2018/0097.pdf
26. Margolin R, Zelnik-Manor L and Tal A. How to evaluate foreground maps? In: *Proceedings of the IEEE Computer Society conference on computer vision and pattern recognition*, Columbus, OH, 23–28 June 2014, pp. 248–255. New York: IEEE.
27. Perazzi F, Krahenbuhl P, Pritch Y, et al. Saliency filters: contrast based filtering for salient region detection. In: *2012 IEEE conference on computer vision and pattern recognition*, Providence, RI, 16–21 June 2012, pp. 733–740. New York: IEEE.

Author biographies

Kyo-Seong Hwang is an ROK Army officer. He has a Master's degree in Defense Science at Korea National Defense University.

Jungmok Ma has a Master's degree in Industrial Engineering from Pennsylvania State University and a PhD from the University of Illinois at Urbana-Champaign. He is a professor in the Department of Defense Science at Korea National Defense University.