

# 迷彩伪装目标图像语义分割数据集的构建

梁新宇<sup>1</sup>, 林浩坤<sup>2</sup>, 杨辉<sup>1</sup>, 肖铠鸿<sup>3</sup>, 权冀川<sup>1\*</sup>

<sup>1</sup>陆军工程大学指挥控制工程学院, 江苏 南京 210007;

<sup>2</sup>华中科技大学软件学院, 湖北 武汉 430070;

<sup>3</sup>中国人民解放军 73676 部队, 江苏 无锡 214400

**摘要** 数据集是基于深度学习语义分割技术的重要组成部分。为了将语义分割技术应用于野外战场环境, 构建一个符合实战场景的数据集至关重要。针对迷彩伪装目标侦察识别的作战保障需求, 分析了野外战场环境及战场侦察图像的特点, 设计了特定场景数据集的构建流程与方法, 构建了具有精细化语义标注的语义分割数据集 CSS, 并通过实验验证了该数据集在语义分割任务上的有效性。

**关键词** 图像处理; 数据集; 语义分割; 迷彩伪装目标

中图分类号 TP183

文献标志码 A

doi: 10.3788/LOP202158.0410015

## Construction of Semantic Segmentation Dataset of Camouflage Target Image

Liang Xinyu<sup>1</sup>, Lin Haokun<sup>2</sup>, Yang Hui<sup>1</sup>, Xiao Kaihong<sup>3</sup>, Quan Jichuan<sup>1\*</sup>

<sup>1</sup>Institute of Command and Control Engineering, Army Engineering University, Nanjing, Jiangsu 210007, China;

<sup>2</sup>College of Software Engineering, Huazhong University of Science & Technology, Wuhan, Hubei 430070, China;

<sup>3</sup>Unit 73676 of The Chinese People's Liberation Army, Wuxi, Jiangsu 214400, China

**Abstract** Dataset is an important part of semantic segmentation technology based on deep learning. In order to apply semantic segmentation technology to the field battlefield environment, it is very important to construct a dataset that conforms to the actual combat scene. In this work, aiming at the operational support requirements for the detection and identification of camouflage targets, the characteristics of the field battlefield environment and battlefield reconnaissance images are analyzed, the construction process and method of the specific scene dataset are designed, and the semantic segmentation dataset CSS with refined semantic annotation is constructed. The effectiveness of the dataset on semantic segmentation tasks is verified by experiments.

**Key words** image processing; dataset; semantic segmentation; camouflage target

**OCIS codes** 100.2960; 100.2980; 100.4996

## 1 引言

语义分割是计算机视觉的高级任务之一,也是解决场景理解和图义推断问题的关键技术。近年来,深度学习的引入使语义分割技术得到了显著的发展。语义分割技术在特定领域的发展逐渐成熟,这不仅得

益于各类深度学习算法的研究,也离不开用于模型和算法训练的数据集支撑,如适用于驾驶场景的KITTI (Karlsruhe institute of technology and Toyota technological institute)数据集<sup>[1]</sup>、适用于城市街景的Cityscapes数据集<sup>[2]</sup>、适用于地物遥感场景的SEN12MS数据集<sup>[3]</sup>。这些场景数据集很好地满足了

收稿日期: 2020-07-03; 修回日期: 2020-07-29; 录用日期: 2020-08-07

\* E-mail: qjch\_cn@sina.com

民用特定领域的图像语义分割需求,但在其他领域,如军事领域,采集难度较大且涉及军事敏感信息,目前没有公开数据集,相关应用数据集的构建和研究较少。针对该问题,本文结合现代迷彩伪装技术的特点和实际战场环境,详细阐述了数据集的构建目的、构建原则、构建过程,构建了一个用于野外场景迷彩伪装目标图像语义分割的 CSS 数据集,并对 CSS 数据集的数据特点进行了全面介绍与分析,最后通过实验验证了 CSS 数据集在语义分割任务上的有效性。

## 2 数据集的构建

基于深度学习的语义分割技术主要研究数

据、算法和算力三大要素。数据在实现分割任务时占据重要地位。一个完备的分割数据集,需要有足够大的数据规模,且能准确表示需要研究的场景信息,以便设计的算法可以对图像进行正确的理解和学习。数据集的好坏决定了分割模型的训练结果,因此,收集并构建一个完备的数据集对基于深度学习的分割技术都是至关重要的。图 1 为构建数据集的一般步骤,包括数据采集、数据处理和实验评估三部分。本节主要从数据采集和数据处理部分出发,介绍了 CSS 数据集的构建目的、采集与构建原则,并详细讨论了 CSS 数据集的构建过程。

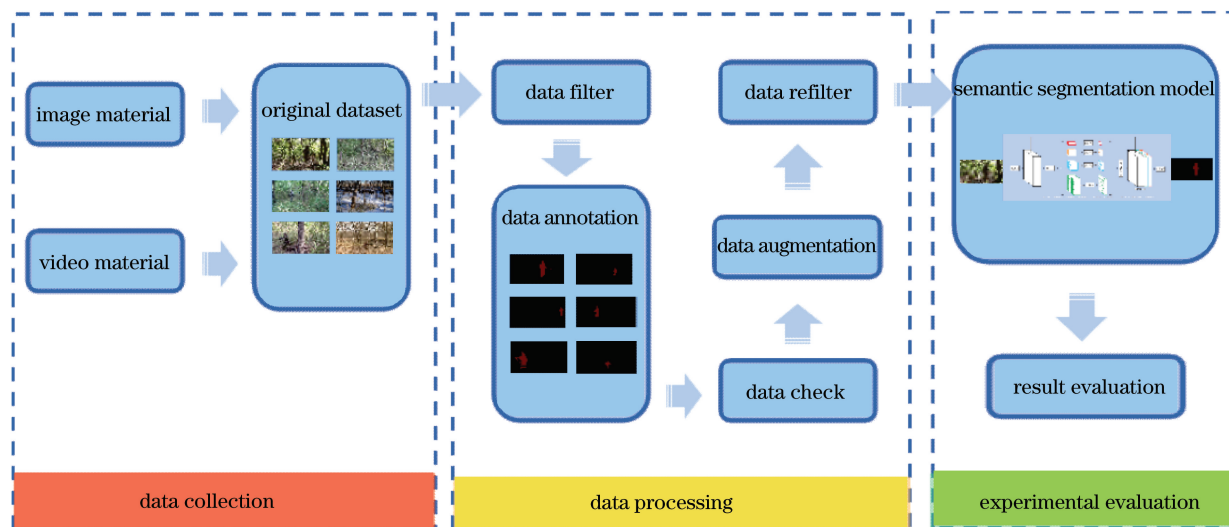


图 1 数据集的构建框架

Fig. 1 Construction framework of the dataset

### 2.1 数据集的构建目的

军事迷彩伪装技术是利用染料、涂料等材料改变目标光学反射特征,实现模仿、融合背景环境的功能,以达到隐蔽自身的目的<sup>[4]</sup>。军事迷彩伪装技术使目标与背景环境在颜色、纹理、形状等特征上非常相似,对战场环境下可见光侦察、目标探测等技术造成了巨大挑战。基于深度学习的语义分割技术可在像素级别对图像进行有效的分类,弥补人类视觉及可见光成像等相关技术难以侦察、识别伪装目标的不足。为了支持分割算法在包含迷彩伪装目标战场侦察场景中的应用,构建一个数据量大、场景丰富、标签质量高的数据集尤为重要。这类数据集可为理解战场环境、高效分割战场侦察图像或识别迷彩伪装目标提供数据支持。

### 2.2 数据集采集与构建原则

目前,没有严格定义图像中的大目标、小目标等术语,也没有统一规定目标的尺寸大小。为便于阐

述,借鉴 MS COCO (Microsoft common objects in context) 数据集<sup>[5]</sup>对目标尺寸(目标在图像中的像素块大小)的定义,结合构建数据集的实际情况和应用需求,根据目标像素点占全图像素点的百分比(下文简称为目标像素点占比)对数据集中的目标尺度进行规定。将目标像素点占比为 1% 及以下的像素区域目标称为小目标;将目标像素点占比在 1%~3% (大于 1%、小于等于 3%) 的像素区域目标称为中目标;将目标像素点占比大于 3% 的像素区域目标称为大目标。

在战场上快速发现敌方军事目标具有重要意义<sup>[6-7]</sup>,但包含迷彩伪装目标的战场图像在场景、设备、保密等方面的采集要求远高于民用领域,其样本的采集难度较大。其次,随着野外迷彩伪装技术的发展,目标隐蔽性较高,对侦察车辆采集的图像进行标注的成本和难度也比民用领域高,其样本数量不能像民用领域包含几十万张或几百万张的规模<sup>[8]</sup>。

由于真实的野外战场侦察图像涉及军事敏感信息,无法直接使用;同时,在野外作战环境中,为了保证行动的隐蔽性,战术要求一般采取单人或小团体作战模式。因此,结合真实作战场景需求,重点阐述了包含迷彩伪装人像目标图像的采集过程。

综合考虑多种因素,以互联网上关于迷彩伪装测试的 12 个影视资料为数据集的原始素材,该视频资料由定点摄像机拍摄,采集的图像以单类别目标为主,以此构建一个完备、有代表性的迷彩伪装目标图像分割数据集。为了使数据集既满足一般分割算法的实验要求,又尽可能地符合真实战场环境,构建数据集需遵循以下原则。

1) 定点摄像机采集,模拟实际战场环境中侦察车辆定点实施侦察的特点。

2) 数据集需要包含雨林、丛林、雪地和山地 4 种野外环境下多类型的迷彩伪装目标。图 2 为



图 2 数据集包含的 4 类场景示例图。(a)雨林;(b)丛林;(c)雪地;(d)山地

Fig. 2 Sample images for 4 types of scenarios.

(a) Rainforest; (b) jungle; (c) snowfield; (d) mountain



图 3 多尺度目标的示例图像

Fig. 3 Sample images of multi-scale target

### 2.3 数据集的标注

根据数据集的构建原则,提出了一种数据集标注的一般流程,对采集的图像进行精细化标注和流程化的筛选、审核,以保证构建的数据集具有领域代表性和应用稳定性,具体的标注处理流程如图 4 所示。数据集的具体标注流程如下。

1) 人工标注阶段。利用标注工具 Labelme 对目标进行精细化语义标注。

数据集中 4 种不同场景的示例图像,表 1 为数据集中 4 种不同野外场景下 10 种伪装目标的迷彩样式。

表 1 数据集迷彩类型的信息

Camouflage model	Camouflage pattern
German WWII SS camouflage	mountain camouflage
A-TACS AU	mountain camouflage
Chinese type 07 arid camouflage	desert camouflage
Chinese type 07 universal camouflage	universal camouflage
Germansnow camouflage	snowfield camouflage
Russian VSR-93 camouflage	jungle camouflage
Swedish M90 camouflage	jungle camouflage
Greeklizard camouflage	jungle camouflage
Belgianbrush stroke camouflage	rainforest camouflage
Finnish M05 camouflage	rainforest camouflage

3) 采集图像需涵盖多尺度的目标,考虑到真实战场环境下,定点侦察车辆越早发现目标,越有利于己方战局。因此,需要探测和发现尽可能远的目标。距离越远,目标在侦察图像中占据像素数的比例越小。这就意味着采集的目标大多为中、小目标。同时,采集得到的伪装目标应该是多角度、多姿态的。图 3 为数据集中包含的多尺度目标在不同角度、不同姿态下的示例图像。

4) 构建数据集时应该考虑不同季节、不同光照环境以及目标部分遮挡等多种因素的影响效果。



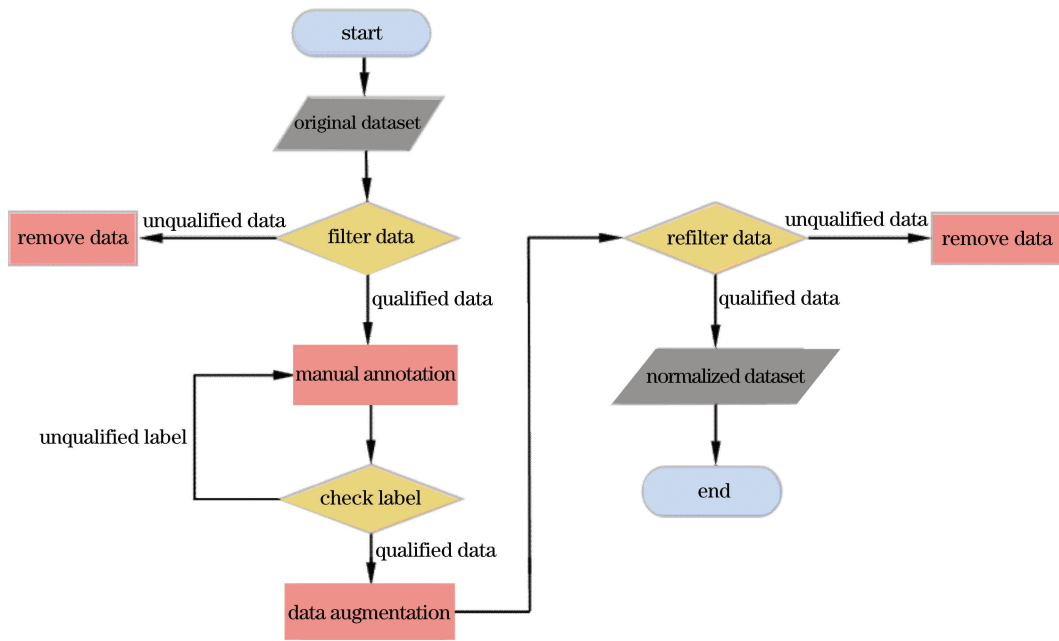


图 4 数据集的标注流程图

Fig. 4 Annotation flow chart of the dataset

一方面通过数据扩增可以有效增加分割模型的泛化能力。

为了保证迷彩伪装目标图像内容的自然化、随机化,数据扩增阶段采用指定变换方法的随机扩增策略,按照一定概率对图像分别进行缩放、裁剪、旋转和镜像操作。

3) 数据筛选阶段。该阶段共进行两次数据筛选。第一次筛选主要剔除不符合构建原则的图像数据,以免进行无效标注,浪费人力、时间;第二次数据筛选是为了保证数据扩增产生的图像符合实际,且满足数据集的构建原则。

### 3 CSS 数据集与实验评估

#### 3.1 CSS 数据集

结合实际战场环境,按照数据集构建原则,经过精细化语义标注和严格的审核筛查,最终构建的 CSS 语义分割数据集包含 10852 张分辨率为  $1280 \times 720$  的图像,包含 4 类战场环境、10 种迷彩类型。数据集的部分标注样例如图 5 所示。

表 2 为 CSS 数据集中多尺度目标的分布情况,图 6 为 CSS 数据集中 3 种尺度的目标在 4 种野外场景下的分布统计图。

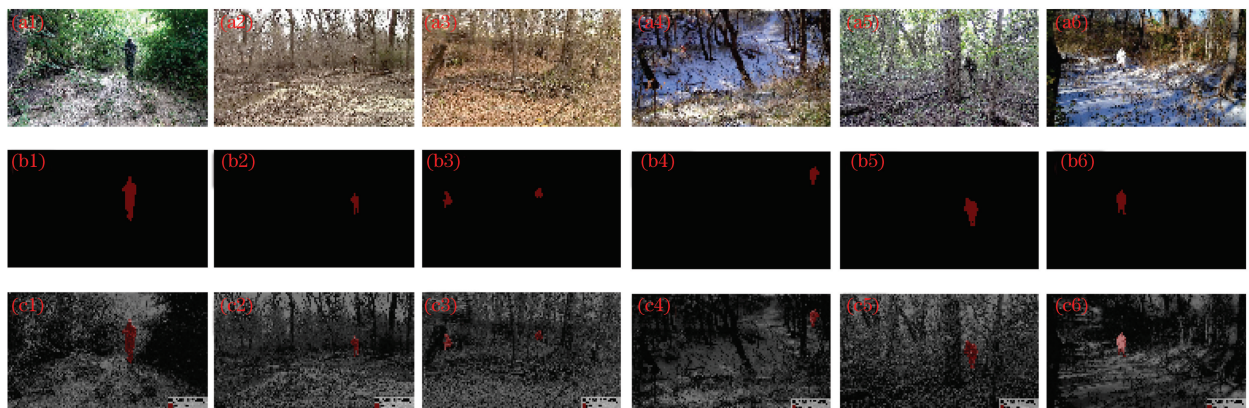


图 5 CSS 数据集的标注样例图像。(a)原始图像;(b)标注图像;(c)带掩码的标注图像

Fig. 5 Annotated sample images of the CSS dataset. (a) Original images; (b) annotated images; (c) annotated images with mask



表 2 数据集多尺度目标的分布信息

Table 2 Distribution information of multi-scale targets in the dataset

Target-scale	Small target	Medium target	Large target
Label number	5388	3419	2045

图 7 为 CSS 数据集中 10 种迷彩类型的分布情况,从实际战场侦察需求上看,CSS 数据集涵盖了不同场景、不同天候、不同迷彩类型的多尺度伪装目标,可以满足不同迷彩伪装目标的检测识别及语义分割的实验需求。

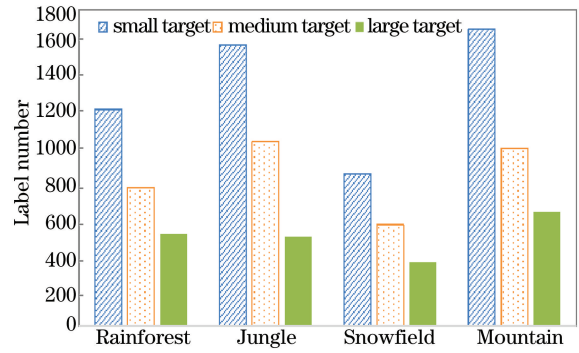


图 6 多尺度目标在不同场景下的分布信息

Fig. 6 Distribution information of multi-scale targets in different scenarios

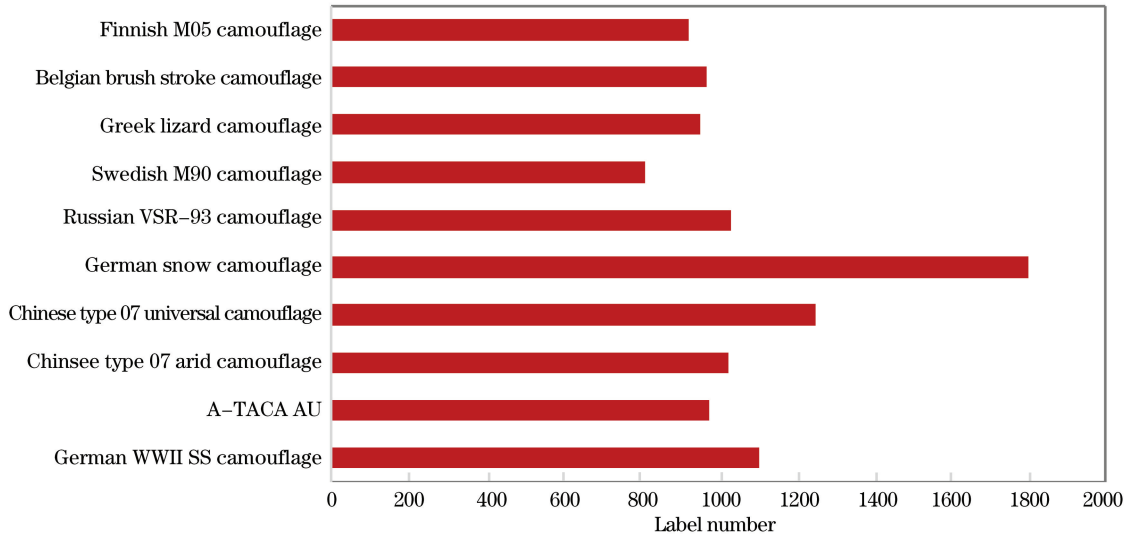


图 7 CSS 数据集中的迷彩类型分布

Fig. 7 Distribution of camouflage types in the CSS dataset

### 3.2 实验评估

为了验证 CSS 数据集在语义分割任务中的有效性,挑选了 SegNet<sup>[9]</sup>、PSPNet<sup>[10]</sup>、DeepLab v3+<sup>[11]</sup> 三类有代表性的语义分割模型,以 CSS 数据集为输入数据,选用交并比 (IoU)<sup>[12]</sup>、平均交并比 (mIoU)<sup>[12]</sup> 为评价指标,根据模型的实验结果对 CSS 数据集进行分析评估。IoU 表示分割结果与原始图像真值的重合程度,取值范围为 [0, 1],可表示为

$$P_{IoU} = \frac{\sum_{i=1}^K n_{ii}}{\sum_{j=1}^K n_{ij} + \sum_{j=1}^K n_{ji} - n_{ii}}, \quad (1)$$

式中,  $K$  为图像像素类别的数量,  $n_{ii}$  为实际类型为  $i$ 、预测类型为  $i$  的像素总数,  $n_{ij}$  为实际类型为  $j$ 、预测类型为  $i$  的像素总数,  $n_{ji}$  为实际类型为  $i$ 、预测类型为  $j$  的像素总数<sup>[13]</sup>。mIoU 表示图像像素的 IoU 在所有像素类别上的平均值,可表示为

$$P_{mIoU} = \frac{1}{K} \left( \frac{\sum_{i=1}^K n_{ii}}{\sum_{j=1}^K n_{ij} + \sum_{j=1}^K n_{ji} - n_{ii}} \right). \quad (2)$$

mIoU 的代表性与简单性较强,是目前图像语义分割领域最常用的准确度评价指标<sup>[14]</sup>。实验策略方面,将数据集按照 7 : 2 : 1 的比例划分为训练集、验证集、测试集,分别针对大、中、小目标进行语义分割实验,并通过 mIoU 评估不同模型在 CSS 数据集上的表现,结果如表 3 所示。

表 3 不同模型在 CSS 数据集上的 mIoU

Table 3 mIoU of different models on the CSS dataset unit: %

Method	SegNet	PSPNet	DeepLab v3+
Small target	69.77	73.24	83.91
Medium target	75.16	82.83	86.94
Large target	85.65	88.27	93.20

图 8 为不同模型对多尺度目标的分割效果,可以通过目标的预测标签与真实标签的重合程度判断分割效果。实验结果表明,针对不同类型的目标,选用的模型在 CSS 数据集上的表现良好,这表明 CSS 数据集用于迷彩伪装目标图像语义分割任务时能够满足模型需要。特别地,在大目标语义分割效果上  $P_{mIoU}$  可达到 85% 以上,分割效果如图 8(a) 所示。但中、小目标的语义分割效果仍然有很大的提升空间,分割效果如图 8(b)、图 8(c) 所示,特别是小目标图像出现了不同程度的错分割、漏分割问题。一方面是由于小目标在原始图像中

所占像素区域小、特征少,分割难度大。一般情况下,卷积神经网络在特征提取阶段,会通过若干池化操作提取高级别的语义信息;虽然可以通过上采样等方法恢复空间分辨率,但小目标池化处理过程中会造成细节信息丢失;当池化层达到一定数量时,小目标的部分甚至全部信息都会丢失。另一方面,由于迷彩伪装与背景十分相似,待分割目标与背景的融合度很高,导致分割难度大幅度增加。因此,在包含迷彩伪装目标的 CSS 数据集上,针对小目标图像的语义分割任务,保留低层特征信息是十分必要的。

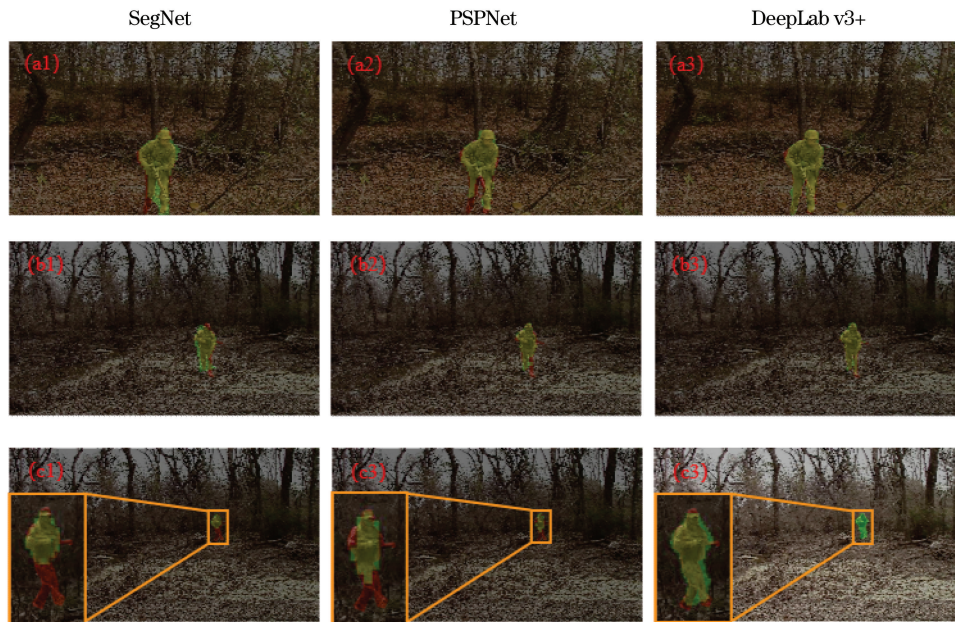


图 8 不同模型对不同目标类型的语义分割效果。(a)大目标;(b)中目标;(c)小目标

Fig. 8 Segmentation effects of different models according to different target types. (a) Large target; (b) medium target; (c) small target

## 4 结 论

根据野外战场环境的真实情况,创建了包含迷彩伪装目标的 CSS 语义分割数据集,并通过不同模型的语义分割实验,分析、验证了 CSS 数据集在语义分割任务上的有效性,为将语义分割技术应用于野外战场环境目标侦察提供了有价值、有意义的数据支持。针对实验中发现的问题,下一步计划从迷彩伪装小目标图像的数据特点出发,研究设计有效的语义分割算法,以改善分割性能,增强语义分割的实用性。

### 参 考 文 献

- [1] Geiger A, Lenz P, Stiller C, et al. Vision meets robotics: the KITTI dataset [J]. The International Journal of Robotics Research, 2013, 32 (11): 1231-1237.
- [2] Cordts M, Omran M, Ramos S, et al. The cityscapes dataset for semantic urban scene understanding [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 3213-3223.
- [3] Schmitt M, Hughes L H, Qiu C, et al. SEN12MS: a curated dataset of georeferenced multi-spectral SENTINEL-1/2 imagery for deep learning and data fusion[EB/OL]. [2020-06-23]. <https://arxiv.org/abs/1906.07789>.
- [4] Zhang L X. Digital camouflage design and camouflage effect evaluation based on natural background [J]. Journal of Xi'an Technological University, 2019, 39 (2): 358.

- 张立新. 基于自然背景的数码迷彩设计及伪装效果评价[J]. 西安工业大学学报, 2019, 39(2): 358.
- [5] Lin T Y, Maire M, Belongie S, et al. Microsoft COCO: common objects in context [M]//Fleet D, Pajdla T, Schiele B, et al. Computer Vision - ECCV 2014. Lecture Notes in Computer Science. Cham: Springer, 2014, 8693: 740-755.
- [6] Zhang H R, Li Y B, Xing R K, et al. Evaluation of air defense missile infrared camouflage capability based on set pair analysis [J]. Laser & Optoelectronics Progress, 2018, 55(7): 070402.  
张海瑞, 李彦彬, 邢瑞康, 等. 基于集对分析的防空导弹装备红外伪装能力评估[J]. 激光与光电子学进展, 2018, 55(7): 070402.
- [7] Guo T, Hua W S, Liu X, et al. Comprehensive evaluation of optical camouflage effect based on hyperspectra [J]. Laser & Optoelectronics Progress, 2016, 53(10): 101002.  
郭彤, 华文深, 刘恂, 等. 一种基于高光谱的光学伪装效果综合评价方法[J]. 激光与光电子学进展, 2016, 53(10): 101002.
- [8] Zhuo L, Chen X Q, Xie Z P, et al. Simulation learning method for discovery of camouflage targets based on deep neural networks [J]. Laser & Optoelectronics Progress, 2019, 56(7): 071102.  
卓刘, 陈晓琪, 谢振平, 等. 基于深度神经网络的迷彩目标发现仿真学习方法[J]. 激光与光电子学进展, 2019, 56(7): 071102.
- [9] Badrinarayanan V, Kendall A, Cipolla R. SegNet: a deep convolutional encoder-decoder architecture for image segmentation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(12): 2481-2495.
- [10] Zhao H S, Shi J P, Qi X J, et al. Pyramid scene parsing network [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 6230-6239.
- [11] Chen L C, Zhu Y K, Papandreou G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation [M]//Ferrari V, Hebert M, Sminchisescu C, et al. Computer Vision-ECCV 2018. Lecture Notes in Computer Science. Cham: Springer, 2018, 11211: 833-851.
- [12] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation [EB/OL]. [2020-06-25]. <https://arxiv.org/abs/1411.4038>.
- [13] Tian X, Wang L, Ding Q. Review of image semantic segmentation based on deep learning [J]. Journal of Software, 2019, 30(2): 440-468.  
田萱, 王亮, 丁琪. 基于深度学习的图像语义分割方法综述[J]. 软件学报, 2019, 30(2): 440-468.
- [14] Garcia-Garcia A, Orts-Escolano S, Oprea S, et al. A survey on deep learning techniques for image and video semantic segmentation [J]. Applied Soft Computing, 2018, 70: 41-65.