

基于显著图数据增强的飞行器识别与检测算法



陈鹏鹏¹, 邢晨光¹, 刘波¹, 王若¹, 王思博², 孙舒凡¹, 陈远泽¹

1. 中国航空系统工程研究所, 北京 100012

2. 北京航空航天大学, 北京 100191

摘要: 数据增强技术能够有效解决民用航空等领域机器学习模型训练数据严重不足的问题, 对于提升智能模型的性能至关重要。为此, 针对航空领域飞行器识别和检测的典型应用场景, 本文提出基于显著图的数据增强方法, 在不同种类的飞行器数据集上分别进行分类和检测试验。试验结果表明, 在飞行器目标检测数据集上, 利用本文提出的数据增强方法可以提升飞行器目标检测精度。

关键词: 航空人工智能; 机器学习; 深度学习; 飞行器目标识别; 数据增强

中图分类号: V249

文献标识码: A

DOI: 10.19452/j.issn1007-5453.2023.11.016

民用飞行器识别等任务中, 传统的图像识别检测方法需要通过人工的方式对图片中的特征进行选取, 但随着数据量的增加和飞行器种类的增多, 特征提取变得越发困难。随着深度学习技术的发展, 空中飞行器图像识别也从传统的特征提取开始逐渐转向使用深度神经网络来学习^[1-3]。然而, 空中飞行器的图像资源相对较为短缺, 没有涵盖各种类型飞行器的公开数据集, 数据资源参差不齐, 图像质量难以保障, 可供模型使用的数据量严重不足, 导致该场景下的研究大都集中在飞行器遥感图像方面^[1], 针对单一的飞行器识别的研究较少。而针对飞行器遥感图像识别的任务中, 图像呈现的大多是一定区域内停靠在机场的飞行器, 其数量较多, 且目标很小, 若将该任务下的模型进行预训练后应用于单一飞行器的识别任务中, 会因为网络学习到的飞行器特征并不全面, 模型泛化性较弱。再加上空中飞行器的图像受飞行姿态、天气光照、焦距等客观因素的影响, 导致输入模型中的图像特征并不清晰, 模型没有学习到图像中的关键特征, 影响最终的识别精度。

现有的数据增强方法^[4-10]主要是对图像数据进行预处理, 增强图像中的特征信息, 使模型能够在有限的图像中提

取到更多的语义信息, 数据增强的目的是使模型关注到更深层次更加本质的特征, 从而提升模型的鲁棒性, 增强泛化能力^[11-12]。图像的翻转、旋转、裁剪、添加噪声等都是计算机视觉中常用的数据增强方法, 在具体任务场景下使用这些方法时还需要注意标签数据的变化, 如目标检测中若使用翻转, 则需要将标注框进行相应的调整^[13-19]。剪切法 (Cutout)^[10]通过对图像进行遮挡, 降低模型对数据的敏感性, 从而提升模型的泛化能力。除了对图像进行遮挡的操作外, 研究者也在图像融合操作方面做过尝试, 混合法 (Mixup)^[13-14]便是其中之一。参考文献[20]提出了一种隐式数据增强方法, 该方法通过交换训练样本特征矩信息来生成新的样本, 以达到数据增强的目的。参考文献[21]提出了对抗策略的数据增强优化方法, 它可以为模型生成额外训练数据。参考文献[22]应用两点交叉算法进行数据增强, 提升了机器学习模型的泛化能力。

但是, 以上方法均将图片中的特征均等对待, 事实上, 通过类激活图^[18-19]可视化后可以发现, 图片中不同特征对于深度学习模型训练的影响是不相同的。因此, 基于这些方法进行数据增强会影响深度学习模型在目标检测和分类

收稿日期: 2023-05-24; 退修日期: 2023-08-21; 录用日期: 2023-09-22

基金项目: 航空科学基金(2022Z071020002)

引用格式: Chen Pengpeng, Wang Ruo, Xing Chenguang, et al. Data augmentation method based on saliency map for aircraft identification and detection technology[J]. Aeronautical Science & Technology, 2023, 34(11): 118-124. 陈鹏鹏, 王若, 邢晨光, 等. 基于显著图数据增强的飞行器识别与检测算法[J]. 航空科学技术, 2023, 34(11): 118-124.

任务中的表现。

为此,本文针对飞行器数据的特征加以优化和改进,提出了基于显著图的数据增强方法。该方法提取出飞行器图像中突出的部分,围绕目标对象峰值显著区域选择补丁,而其他部分的像素则设置为0,以增加图像中飞行器对象在整个图片中的语义信息,这些信息可以使模型学习更合适的特征表示,提高深度神经网络最终在分类和检测任务中的准确性和精度。

1 现有的数据增强方法

深度学习模型的成功可以归功于数据的数量和多样性。但收集标记数据是一项烦琐且耗时的任务。因此,现有方法通过应用各种转换(如旋转、翻转等)来增加现有数据的多样性。由于这种简单且廉价的技术显著提高了模型的性能和鲁棒性,因此数据增强已被广泛用于训练深度学习模型。Y. Lecun等^[5]应用数据增强来训练卷积神经网络LeNet进行手写字符识别。他们执行了几种仿射变换,如平移、缩放、剪切等。对于相同的任务,Y. Bengio等^[9]对图像应用了更多样化的变换,如高斯噪声、椒盐噪声、高斯平滑、运动模糊、局部弹性变形和各种遮挡。A. Krizhevsky等^[7]应用基于主成分分析的随机图像补丁裁剪、水平翻转和随机颜色强度变化。除了翻转和裁剪之外,S. Connor等^[8]还应用了色偏、渐晕和镜头失真,以提高深层网络的鲁棒性。D. C. Ekin等^[9]提出了一种称为自动增强法(AutoAugment)的数据增强方法,该方法定义了各种增强技术的搜索空间,并为每个小批量选择最合适的一种。

近年来,数据增强方面的进展可以分为两类:一类是基于区域擦除的方法。区域擦除即在训练中擦除图像的部分区域,旨在鼓励神经网络找到更多的鉴别区域。Zhou Kaiyang等^[10]提出了剪切法,通过随机移除输入训练图像的方形区域,以提高模型的鲁棒性。另一类是基于数据混合的方法,该方法最近在图像分类领域受到广泛关注。与区域擦除增强相比,这类方法通过组合多张图像并相应地融合它们的标签来生成新数据。K. Jang-Hyun等^[13]提出了联合优化两个目标的联合混合法(PuzzleMix)算法,即选择最佳掩码和最佳混合方法。掩码试图揭示两个图像中最显著的数据,最佳传输计划旨在最大化数据显示部分的显著性。Liang Daojun等^[14]首次提出了混合法,通过线性组合图像并使用相同的系数融合它们的标签来生成图像。混合法在处理损坏的目标和提高模型性能方面显示出其优势,但增强后的图像看起来不自然且局部模糊^[10]。C. Summers等^[15]通

过引入一种考虑非线性混合操作的更通用的数据混合形式,进一步改进了混合法。Y. Sangdoon等^[16]提出了剪切混合法,它通过剪切一个图像补丁并粘贴到另一个图像来生成新图像。与混合法类似,标签也是混合的,但与补丁的面积成比例。利用这两种方法的优势,剪切混合法在分类任务和弱监督定位任务中表现出令人印象深刻的性能。

本文提出的方法属于第二类。然而,它与以前的技术有很大不同。当前基于混合数据的方法主要根据混合像素的统计数据来组合标签,如像素数或强度值的比率;而现有的基于补丁的方法中,由于源补丁选择过程中的随机性,它可能会选择一个不包含任何关于源对象的信息像素的区域,并且根据这些无信息补丁的标签混合会误导分类器学习意外的特征表示。

相比之下,本文的方法通过在一些先验信息的帮助下选择源图像补丁来解决上述问题。具体来说,本文首先提取源图像中突出的部分,即需要识别的飞行器,然后围绕飞行器图像峰值显著区域选择补丁,以确保从识别对象部分中进行选择,而其他部分的像素则设置为0,现在选定的补丁包含有关飞行器的相关信息,这些信息可以引导模型学习更合适的特征表示。

2 本文算法流程

为防止模型学习任何不相关的特征表示,本文基于像素点的显著性选择用于数据增强的补丁,以达到增强后的图像包含目标对象相关信息的目的。算法流程如图1(图中为国外飞行器)所示。该方法共包含6个步骤:(1)对图片添加标注信息;(2)通过对图像进行显著图检测,生成可视化的显著性图;(3)通过显著性图选择补丁位置,生成源补丁;(4)保留源补丁图像,将其他位置的像素值置为0,输出修改后的图像;(5)将修改后的图像作为新的训练样本,对模型进行训

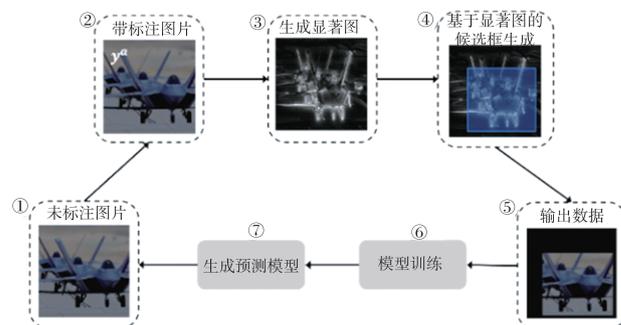


图1 算法流程图

Fig.1 Algorithm flow chart

练;(6)生成新的预测模型。下面重点介绍步骤(3)和步骤(4)。

2.1 补丁的选择

显著性检测的目标是找出有吸引力的像素或区域,并为它们分配更高的强度值^[17]。显著性检测方法可以产生可视化的显著性图,即灰度图,突出感兴趣的对象,将其集中在前景上。令 $I_s \in \mathbb{R}^{W \times H \times C}$ 是一个随机选择的(源)训练图,其标签为 y_s ,从中剪切出一个补丁,那么它的显著性图检测可以表示为

$$I_{vs} = f(I_s) \quad (1)$$

式中, $I_{vs} \in \mathbb{R}^{W \times H \times C}$ 表示源图像 I_s 的视觉显著性图,感兴趣对象具有更高的强度值, $f(I_s)$ 表示显著性检测模型。然后在显著图中搜索具有最大强度值的像素 I_{vs}^i ,其中 i 和 j 分别表示该像素的 x 和 y 坐标,可以表示为

$$i, j = \operatorname{argmax}(I_{vs}) \quad (2)$$

然后选择一个补丁,以第 I_{vs}^i 个像素点为中心,基于均匀分布(0,1)以及图像裁剪百分比参数 λ 确定补丁的大小。

2.2 训练样本生成

随机选择 I_a 作为待增强样本。 I_a 是增强后的样本,样本数据增强的过程可以表示如下

$$I_a = M \odot I_s \quad (3)$$

式中, $M \in \{0,1\}^{W \times H}$ 表示二进制掩码, \odot 表示逐元素乘法。基于参数 λ 将掩码的对应补丁的位置设置为1,其他位置设置为0。与SaliencyMix^[23]方法不同,本文的方法不需要对两张原始图片进行交叉混合操作,本文的方法更加快捷有效。

3 试验与结果分析

为了验证优化后的数据增强方法在对飞行器识别的任务中具有其特定的优势,本文设计了分类和检测两种任务,分别使用不同的数据集进行对比试验。

3.1 试验设计

本文使用的数据集分别来自两种不同类型的飞机,均为国外飞行器。数据集中包含了两种飞机在巡航、爬升、降落、滑行等不同姿态的图片,尽可能地包含了多角度、多方位的飞机状态。试验中均使用Pytorch^[24]深度学习框架实现,并使用GPU加速训练。

令FP表示错误的正例,本为假类但判定为正类。令TP表示正确的正例,一个实例是正类并且也被判定为正类。TN表示正确的反例,一个实例是假类并且被判定为假类。N表示样例总数。本文使用正确率(ACC)作为分类任

务的评价指标,即预测正确的结果占总样本的百分比,如式(4)所示

$$\text{ACC} = \frac{\text{TP} + \text{TN}}{N} \quad (4)$$

式中,ACC为识别正确率;FN为错误的反例,本为正类但判定为假类;FP为错误的正例,本为假类但判定为正类。本文使用平均精度(mAP)作为目标检测任务的评价指标。与mAP相关的两个评估指标为精确率(P)和召回率(R),如式(5)、式(6)所示

$$P = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (5)$$

$$R = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (6)$$

以召回率R和精确率P分别为横纵坐标建立坐标系,得到P-R曲线图,通过计算P-R曲线下的面积得到平均精度(AP),所有类别的AP做平均,即可得到mAP。

3.2 试验结果

3.2.1 分类试验

本文分别使用残差神经网络ResNet18^[25]、残差神经网络ResNet50^[25]和视觉几何组网络(VGG)^[26]三种模型在不同型号飞行器的数据集上做对比试验。试验中所用的数据集信息见表1,试验结果如图2所示。

表1 数据集划分

Table 1 Data set partition table

数据集	任务	训练数据集	验证数据集
国外飞行器1	分类	1600	160
国外飞行器2	分类	1600	160

图2试验结果表明,原始数据集在经过数据增强后,图像分类的正确率在残差神经网络ResNet18和残差神经网络ResNet50模型上均有一定幅度的提升,数据增强前正确率均未达到或仅是接近90%,数据增强后,其正确率均超过或接近90%;在视觉几何组网络(VGG)模型上指标有较大幅度的提升,数据增强后正确率提升较为明显。

3.2.2 检测试验

在检测任务中,本文使用基于深度神经网络的目标识别模型(YOLOX)^[26]在三种不同类别的数据集上分别试验。试验中所用的数据集信息见表2。

为了进一步验证优化后的数据增强方法在目标检测任务中的表现,本文引入了马赛克(Mosaic)数据增强^[27]做对比,试验中重叠区域交并比(IoU)阈值为0.5,在此阈值下,试验结果见表3。

由表3可以看出,国外某飞机1数据集在原始图像上已经有较高的平均精度mAP,在此基础上,马赛克数据增强

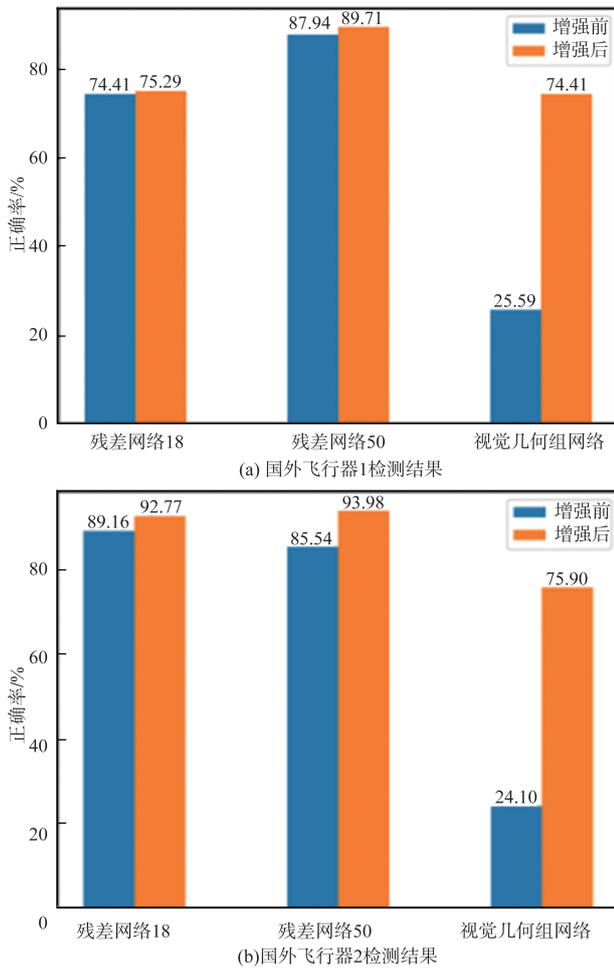


图2 不同数据集上的分类效果对比

Fig.2 Comparison between classification effects on different data sets

表2 数据集划分表

Table 2 Data set partition table

数据集	任务	训练数据集	验证数据集
国外飞行器1	定位检测	1600	160
国外飞行器2	定位检测	1600	160

表3 数据增强方法在目标检测任务中的mAP对比

Table 3 mAP comparison of different data enhancement methods in the target detection task

数据集	增强前	马赛克增强	本文的方法
国外飞行器1	98.08	98.85	98.85
国外飞行器2	97.26	99.31	99.52

和优化后的数据增强对目标检测的结果均有小幅度提升。在国外飞行器2数据集上,马赛克数据增强和优化后的数据增强同样是对检测精度有小幅度的提升,但优化后的数据增强提升的百分比相比较马赛克数据增强有更大的幅度。

3.3 试验分析

从3.1节和3.2节的试验结果可以看出,本文提出的优化后的数据增强方法在图像的分类和检测任务中有不错的表现,为了进一步解释该方法对模型训练中图像特征的提取的有效性,本文将检测任务中基于深度神经网络的目标识别模型(YOLOX)训练过程中的图像进行了可视化,如图3所示。

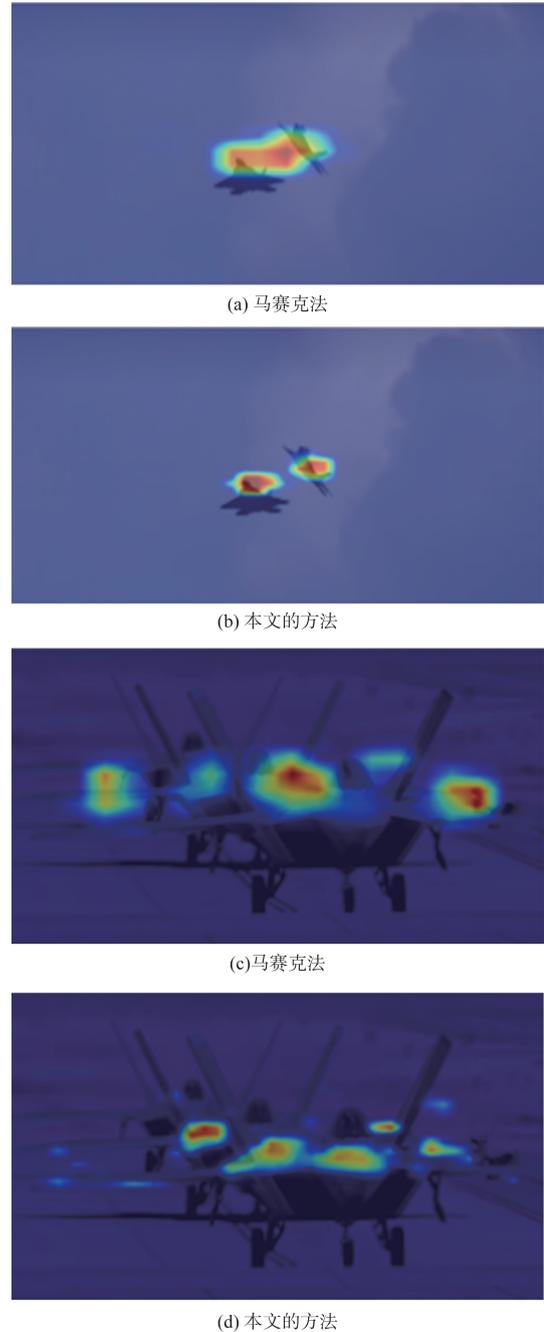


图3 不同数据增强方法的可视化结果对比

Fig.3 Comparison of visual results of different data enhancement methods

图3利用类激活图(CAM)^[27]的方法对国外某飞行器降落姿态的飞行器热力图进行可视化,分别为马赛克数据增强^[27]的可视化结果和本文提出方法的可视化结果。从图3中可以明显地看出,马赛克法对于空中距离较近的两个飞



(a) 原图



(b) 显著图



(c) 原图



(d) 显著图

图4 可视化显著图与原始图像对比

Fig.4 Comparison of visual saliency and original image

行,模型的关注点将其混淆在一起。而使用本文优化后,距离较近的两个飞行器模型的聚焦点也能分别关注。

同时,本文提出的数据增强方法的有效性也可以通过显著图可视化的方式验证。特征图可以呈现图像中的像素点对图像分类结果的影响,即图4中亮色部分为神经网络感兴趣的部分。从图4中的特征图中可以看出,经过数据增强后,飞机机体的像素点亮色更加明显,这表明神经网络的关注点更加聚焦于飞机本身而不是图像的其他地方,使得模型的分类和检测结果更加精确。

4 结束语

针对在数据集有限甚至缺少的情况下使用深度学习方法对空中飞行器图像识别的任务,本文提出了基于显著图概念的数据增强方法。在此方法的基础上,对飞行器数据的特征加以优化和改进。通过试验可以看出,本文所提出的方法对不同类别飞行器的识别和检测精度都有所提高,通过数据增强的方法较好地解决了因样本资源稀缺和样本质量较差等客观因素导致的模型性能较差的问题,提高了模型在单一飞行器识别任务中的水平。

AST

参考文献

- [1] Deepan P, Sudha L R. Effective utilization of YOLOv3 model for aircraft detection in remotely sensed images[J]. Materials Today:Proceedings, 2021(3): 214-7853.
- [2] 刘准钊,程咏梅,潘泉,等. 证据冲突下自适应融合目标识别算法[J]. 航空学报, 2010 (7): 1426-1432.
Liu Zhunga, Cheng Yongmei, Pan Quan, et al. Target identification by adaptive combination of conflicting evidence [J]. Acta Aeronautica et Astronautica Sinica, 2010 (7): 1426-1432. (in Chinese)
- [3] 施立瑞,王帅帅,肖昊. 面向航空目标检测的神经网络加速器设计[J]. 航空科学技术, 2022, 33(5):89-96.
Shi Lirui, Wang Shuaishuai, Xiao Hao, et al. Design of neural network accelerator for aeronautical target detection[J]. Aeronautical Science & Technology, 2022, 33(5): 89-96. (in Chinese)
- [4] Nuñez-Ramirez D, Mata-Mendoza D, Cedillo-Hernandez M. Improving preprocessing in reversible data hiding based on contrast enhancement[J]. Journal of King Saud University: Computer and Information Sciences, 2022, 34(8): 5468-5477.
- [5] Lecun Y, Bottou L, Bengio Y, et al. Gradient-based learning

- applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11): 2278-2324.
- [6] Bengio Y, Bastien F, Bergeron A, et al. Deep learners benefit more from out-of-distribution examples[C]//Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics, 2011: 164-172.
- [7] Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deep convolutional neural networks[J]. Communications of the ACM, 2017, 60(6): 84-90.
- [8] Connor S, Taghi M K. A survey on image data augmentation for deep learning[J]. Journal of Big Data, 2019, 6(1): 1-48.
- [9] Ekin D C, Barret Z, Dandelion M, et al. Autoaugment: Learning augmentation strategies from data[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019: 113-123.
- [10] Zhou Kaiyang, Liu Ziwei, Qiao Yu, et al. Domain generalization: A survey[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, 45(4): 4396-4415.
- [11] 宋婷, 贺丰收, 程宇峰. 深度学习技术在雷达目标检测中的研究进展[J]. 航空科学技术, 2020, 31(10):12-20.
Song Ting, He Fengshou, Cheng Yufeng. Research progress of deep learning technology in radar target detection[J]. Aeronautical Science & Technology, 2020, 31(10): 12-20. (in Chinese)
- [12] 李胜男, 王景霖, 沈勇, 等. 基于预分类机制及GRNN的直升机飞行状态识别[J]. 航空科学技术, 2021, 32(8):18-23.
Li Shengnan, Wang Jinglin, Shen Yong, et al. Helicopter flight status recognition based on pre-classification mechanism and GRNN[J]. Aeronautical Science & Technology, 2021, 32(8):18-23. (in Chinese)
- [13] Jang-Hyun K, Wonho C, Hyun O S. Puzzle mix: Exploiting saliency and local statistics for optimal mixup[C]. International Conference on Machine Learning, 2020: 5275-5285.
- [14] Liang Daojun, Yang Feng, Zhang Tian, et al. Understanding mixup training methods[J]. IEEE Access, 2018, 6: 58774-58783.
- [15] Summers C, Dinneen M J. Improved mixed-example data augmentation[C]. 2019 IEEE Winter Conference on Applications of Computer Vision, 2019: 1262-1270.
- [16] Sangdoon Y, Dongyoon H, Sanghyuk C, et al. CutMix: Regularization strategy to train strong classifiers with localizable features[C]. International Conference on Computer Vision, 2019.
- [17] Zhou Tao, Fang Dengping, Cheng Mingming, et al. RGB-D salient object detection: A survey[J]. Computational Visual Media, 2021, 7: 37-69.
- [18] Ramprasaath R S, Michael C, Abhishek D, et al. Grad-cam: Visual explanations from deep networks via gradient-based localization[C]. Proceedings of the IEEE International Conference on Computer Vision, 2017: 618-626.
- [19] Mitchell W, Gabriel I, Mike L, et al. Robust fine-tuning of zero-shot models[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 7959-7971.
- [20] Suzuki T. Techaugment: Data augmentation optimization using teacher knowledge[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 10904-10914.
- [21] Boyi L, Felix W, SerNam L, et al. On feature normalization and data augmentation[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2021:12383-12392.
- [22] Rishi R, Jimson M, Santhosh K, et al. Crossover based technique for data augmentation[J]. Computer Methods and Programs in Biomedicine, 2022, 218: 106716.
- [23] Dominik L, Jacek M. An overview of mixing augmentation methods and augmentation strategies[J]. Artificial Intelligence Review, 2023, 56(3): 2111-2169.
- [24] Dimitrios D, Georgios Z, Kostas S, et al. AdaPT: Fast emulation of approximate DNN accelerators in PyTorch[J]. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 2023,42(6): 2074-2078.
- [25] Tammina S. Transfer learning using vgg-16 with deep convolutional neural network for classifying images[J]. International Journal of Scientific and Research Publications, 2019, 9(10): 143-150.
- [26] Zhang Mingjiang, Wang Chengyuan, Yang Jungang, et al. Research on engineering vehicle target detection in aerial photography environment based on YOLOX[C]. 2021 14th International Symposium on Computational Intelligence and Design (ISCID). IEEE, 2021: 254-256.

- [27] Ramprasaath R S, Michael C, Abhishek D, et al. Grad-cam: Visual explanations from deep networks via gradient based localization[C]//Proceedings of the IEEE International Conference on Computer Vision, 2017: 618-626.

Data Augmentation Method Based on Saliency Map for Aircraft Identification and Detection Technology

Chen Pengpeng¹, Xing Chenguang¹, Liu Bo¹, Wang Ruo¹ Wang Sib², Sun Shufan¹, Chen Yuanze¹

1. *China's Aviation System Engineering Research Institute, Beijing 100012, China*

2. *Beihang University, Beijing 100191, China*

Abstract: With the development of artificial intelligence application in aviation domain, artificial intelligence technology also plays an important role in classification and detection tasks. This paper focuses on the data augmentation methods in the field of computer vision, which proposes a data augmentation algorithm based on improved saliency map and conducts classification and detection experiments on different kinds of aircraft data sets. Experiments show that the data augmentation algorithm based on the improved saliency map has better performance in the aircraft identification and detection tasks.

Key Words: aviation artificial intelligence; machine learning; deep learning; aircraft target recognition; data augmentation

Received: 2023-05-24; **Revised:** 2023-08-21; **Accepted:** 2023-09-22

Foundation item: Aeronautical Science Foundation of China (2022Z071020002)