

迷彩伪装目标检测的视觉特征偏好研究

韩彤^{1,2}, 曹铁勇¹, 郑云飞³, 王杨¹, 陈雷¹, 王焱奎⁴, 付炳阳¹

(1. 陆军工程大学 指挥控制工程学院, 江苏 南京 210007;

2. 95911 部队, 甘肃 酒泉 735000;

3. 陆军炮兵防空兵学院, 江苏 南京 211100;

4. 31401 部队, 吉林 长春 130000)

摘要:迷彩伪装通过设计颜色和纹理图案来破坏目标的固有形状,其检测依赖的视觉特征应与常规目标不同。然而卷积神经网络的黑盒性质使得不同视觉特征对模型识别的贡献程度无法获知。为解决上述问题,借鉴人类视觉系统设计了一种适用于伪装场景的视觉特征解耦方法,解耦并分析目标检测模型在颜色、纹理和形状特征上的偏好程度。具体来说,使用消除单一特征并保留其余特征的解耦框架,以模型的性能下降情况作为偏向性的衡量标准,通过灰度化处理消除图像的颜色特征,使用区域置乱破坏目标的纹理特征,对目标轮廓取内接形状以改变目标的形状特征。在公开的迷彩伪装人员数据集和常规人员检测数据集上分别进行实验,结果显示,迷彩伪装目标的检测主要依赖纹理,常规目标的检测主要依赖形状。

关键词:目标检测;迷彩伪装;特征解耦;人类视觉系统;卷积神经网络

中图分类号:TP391

文献标识码:A

文章编号:1673-629X(2023)12-0193-07

doi:10.3969/j.issn.1673-629X.2023.12.027

Research on Visual Feature Bias of Camouflaged Object Detection

HAN Tong^{1,2}, CAO Tie-yong¹, ZHENG Yun-fei³, WANG Yang¹,

CHEN Lei¹, WANG Ye-kui⁴, FU Bing-yang¹

(1. School of Command & Control Engineering, Army Engineering University of PLA, Nanjing 210007, China;

2. Unit 95911, Jiuquan 735000, China;

3. The Army Artillery and Defense Academy of PLA, Nanjing 211100, China;

4. Unit 31401, Changchun 130000, China)

Abstract: Camouflage uses designed color and texture patterns to disrupt the inherent shape of the target, so the visual features that its detection relies on should be different from those of conventional targets. However, the black box nature of convolutional neural networks makes it impossible to know the contribution of different visual features to model recognition. To solve this problem, a new visual feature decoupling method was designed based on the human visual system, which is suitable for camouflage scenes. This method decouples and analyzes the preference degree of object detection models on color, texture, and shape features. Specifically, an analysis architecture was used to eliminate a single feature while retaining the remaining features, and the performance degradation of the model was used as a measure of bias. Grayscale processing was used to eliminate the color features of images, region scrambling was used to disrupt the texture features of targets, and the inner shapes of targets were extracted to change their shape features. Experiments were conducted on publicly available datasets of camouflaged personnel and conventional personnel detection, respectively, and the results showed that the detection of camouflaged object mainly relies on texture, while the detection of conventional object mainly relies on shape.

Key words: object detection; camouflage; feature decoupling; human visual system; convolutional neural networks

0 引言

迷彩伪装是最基本的军事伪装技术之一,其设计旨在模仿背景的颜色及纹理等特征来降低目标的显著

性,以此规避人眼及机器侦察。近年来,基于卷积神经网络(Convolutional Neural Networks, CNN)^[1]来检测迷彩伪装目标的研究取得了良好的进展^[2-6]。如何进

收稿日期:2023-03-27

修回日期:2023-07-27

基金项目:国家自然科学基金青年科学基金(61801512);国家自然科学基金(62071484);江苏省优秀青年基金(BK20180080)

作者简介:韩彤(1995-),女,硕士,研究方向为目标检测;通信作者:曹铁勇(1971-),男,博士,教授,博导,研究方向为语音和图像处理。

一步提升其检测模型的迁移性和有效性,关键之一在于需要深入分析迷彩伪装目标检测模型的内在机理。然而,神经网络的黑盒性质使得模型的运行机制难以被准确理解。

近来的研究表明,CNN 与人类视觉系统(Human Visual System,HVS)的决策机制有相似之处^[7-8]。研究者尝试分析颜色、形状、纹理特征在常规对象识别中的作用,但相关研究^[9-10]主要针对的是显著性目标分类,其结果无法直接用于指导迷彩伪装目标检测模型的优化。

为此,该文从人类视觉特征角度出发,针对迷彩伪装目标检测的特点,设计了一种新的视觉特征解耦方法,在此基础上研究了 CNN 迷彩伪装目标检测模型对不同视觉特征的偏向性。在迷彩伪装人员数据集与常规人员检测数据集上进行的对比实验表明,CNN 目标检测模型对于迷彩伪装目标偏向于学习其纹理,对于常规目标偏向于学习其形状,颜色特征在二者的检测中均不占主导地位。

1 相关工作

1.1 迷彩伪装目标检测

20 世纪中期,迷彩伪装技术逐渐成熟并广泛用于军事领域。早期研究人员通过设计特定的纹理提取算子来检测迷彩图案^[11-14]。随着深度学习技术的发展,研究人员开始使用基于卷积神经网络的模型来解决迷彩发现问题。Zheng 等人^[15-16]构建了迷彩伪装人员数据集并对迷彩伪装目标进行分割。文献[2-6]使用基于 CNN 的目标检测模型实现了迷彩伪装目标的精准分类与定位。文献[17]使用相同的对抗样本对不同迷彩伪装目标检测模型进行攻击,产生了不同程度的性能下降,然而该文献缺乏对模型机理的进一步研究。文中工作一方面有助于检测模型的改进,另一方面对提升迷彩伪装目标检测的对抗攻击效果有指导意义。

1.2 CNN 模型的视觉特征研究

神经科学研究证明,颜色、形状和纹理是人类视觉系统中最重要的三个特征且相互独立^[18-23]。由于人类视觉系统是许多计算机视觉算法的黄金标准,研究^[7-8]将 HVS 与 CNN 进行比较。2018 年,Geirhos 等人^[9]使用风格迁移构建具有冲突线索的数据集,通过实验揭示了在 ImageNet 上训练的分类模型具有纹理偏重并且提高形状偏重可以提升模型的鲁棒性。此后,研究人员致力于设计提高模型形状偏重的方法^[24-25],对任务本身的特征偏重研究较少。直至 2022 年,文献[10]指出,模型的偏向性是任务相关的。例如,识别同样形状但不同类别的鸟主要依靠颜色,而区分斑马和印有斑马条纹的汽车主要依靠形状。

为了对模型的视觉特征偏重进一步研究,Ge 等人^[10]提出了常规目标分类模型的特征解耦框架,分别提取数据集的颜色、形状、纹理特征。在提取颜色时将图片变换到频域进行相位加扰再逆变换;在提取形状时,通过分割图像的显著区域得到形状掩膜;在提取纹理时,先将图像的显著区域进行灰度化处理,然后将其切割成多个方块,取其中四个方块拼接成一个新的图像,作为该图像的纹理特征。

解耦并分析迷彩伪装目标检测模型的视觉特征有助于解释模型工作机理、提高模型性能。然而目前还没有针对该问题的研究。现有研究旨在提取单一视觉特征,无法用于迷彩伪装目标检测,具体表现在:如果对全图提取单一视觉特征会导致目标位置信息丢失,如果只对目标提取单一视觉特征会破坏目标的伪装性。

2 针对迷彩伪装目标检测的视觉特征分析

2.1 分析框架

与以往框架不同,所提框架旨在分别消除目标某单一特征并保留其余特征。基于此框架,分别在颜色、纹理、形状方面设计解耦方法。

分析框架如图 1 所示。首先,在数据集的训练集上训练好模型;其次,对数据集的原始测试集分别解耦颜色、纹理、形状特征得到三个特征解耦测试集;最后,在训练好的模型上分别验证特征解耦测试集。分别计算模型 mAP 的变化率,并进行归一化处理,用不同特征的占比表示模型的视觉特征偏好。

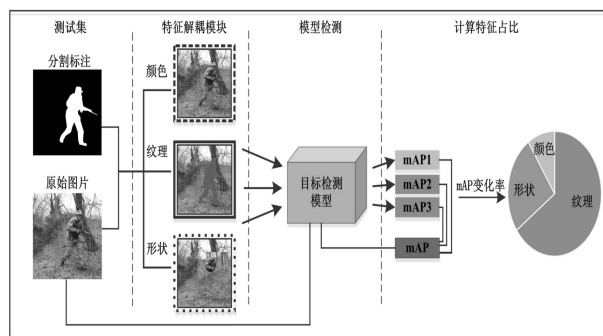


图 1 视觉特征偏好分析框架

2.2 解耦方法

在对不同属性特征进行解耦时应遵循以下原则:一是最大化改变一种属性特征的同时其余属性应尽量保持不变;二是纹理的变化应遵循迷彩的特点和规律,不应破坏其伪装性;三是在迷彩伪装场景下,目标自身结构被迷彩纹理破坏,目标的形状特征体现在目标的外围轮廓。解耦方法将目标分割轮廓视作目标形状,将分割轮廓内的灰度图像视作目标纹理,使用 RGB 空间描述图像的颜色。

2.2.1 消除颜色

改变图像的颜色特征,通常可对目标区域、背景或全图进行处理。由于迷彩伪装目标与背景颜色高度相似,如果只改变目标或背景的颜色会增强目标显著性从而影响检测效果,因此选择对全图进行处理。图像灰度化可以在消除颜色特征时保留形状和纹理特征。灰度化处理通常有取最大值、取平均值和取加权值三种方法。由于人眼对不同颜色敏感度不同,取加权值法能得到较合理的灰度图像。图像 \tilde{X} 中第 i 个像素的灰度值计算为:

$$\tilde{X}_i = (w_1 R + w_2 G + w_3 B) / 3 \quad (1)$$

其中, w_1 为 0.299, w_2 为 0.587, w_3 为 0.114 分别表示图像的 R, G, B 分量加权值。式 1 为灰度心理学公式,该公式的权重系数根据心理学上关于人类视觉系统对绿色最敏感等结论得出。

2.2.2 破坏纹理

纹理是人类视觉系统的一种感知形式,迄今还没有文字或公式化定义,但诸多研究认为:局部纹理体现在像素及其周围空间邻域的灰度分布,全局纹理体现在局部纹理不同程度的重复性^[26-29]。改变纹理的方法包括对像素值的操作(如滤波、仿射变换)和对像素空间关系的操作(如交换、置乱)。像素值的改变可能影响颜色信息,为了保持形状和颜色不变,选择在目标分割轮廓内对像素的空间关系进行操作。此处借鉴像素置乱的思想,对区域内的纹理块进行置乱。

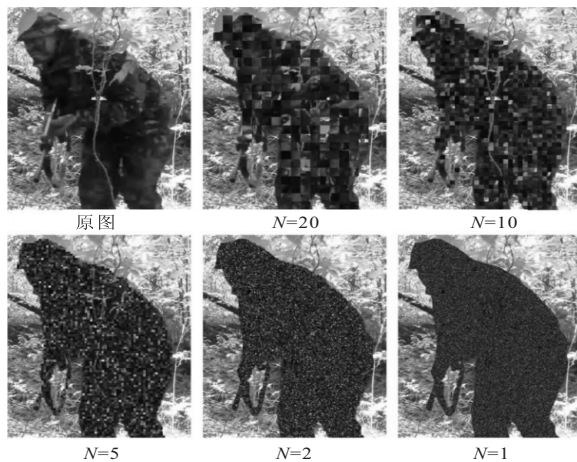


图 2 不同尺度的区域置乱

给定图像 X , 将目标轮廓内的区域切分为若干个 $N \times N$ 像素的纹理块,然后将纹理块的空间位置进行置乱。具体流程如算法 1 所示,其中 n 表示尺寸为 $N \times N$ 的纹理块。由于纹理的破坏程度受置乱区域的大小影响,在此进行了不同尺度的区域置乱实验,效果如图 2 所示。分别设置 N 为 20, 10, 5, 2, 1, 当 $N=1$ 时即为像素置乱。

算法 1: 区域置乱

输入: 图像 X 、图像 X 的二值化 mask、空数组 E

输出: 像素置乱后的图像 \tilde{X}

1. $E = \emptyset, \tilde{X} = X, P = X \odot \text{mask}$
2. while $n \in P$ do
3. $E = E \cup \{X_n\}$
4. end while
5. $E = \text{Shuffle}(E)$
6. while $t \in E$ and $n \in P$ do
7. $\tilde{X}_n = t$
8. end while
9. return \tilde{X}

2.2.3 改变形状

对目标形状的破坏包括消除目标轮廓和改变目标轮廓为其他形状。在消除轮廓时,尝试使用滤波方法将目标轮廓与背景融合,当滤波区域较小时不能达到明显消除轮廓的效果,当滤波区域较大时,背景与前景边界处的纹理产生了较大改变。因此,研究改变目标形状的方法。

给定一幅宽 W 高 H 的图像 X , 对目标轮廓内区域随机取最大内接圆形或内接矩形,步骤如算法 2 所示。

算法 2: 取目标最大内接圆或内接矩形

输入: 宽 W 高 H 的图像 X , X 的掩膜 mask

输出: 目标最大内接圆或内接矩形的掩膜 mask。

1. 对 mask 先腐蚀后膨胀,消除尖端和噪声,得到 mask
2. 对 mask 进行轮廓提取
3. 随机选择取圆形或取矩形,若取圆形则转步骤 4,若取矩形则转到步骤 5
4. 遍历每个轮廓的所有坐标,取轮廓内点到轮廓的最大值为圆半径,此时的点为圆心。绘制宽 W 高 H 的掩膜 mask_c, 令圆形区域内像素值为 1, 其余为 0
5. 遍历每个轮廓的所有坐标,使用中心扩散法^[30]求四个边界点坐标。绘制宽 W 高 H 的掩膜 mask_r, 令矩形区域内为像素值 1, 其余为 0

为解决取内接形状后,内接形状和原始轮廓之间像素缺失问题,考虑以下两种方案:一是用背景纹理覆盖整个目标真实框,二是用背景纹理覆盖目标轮廓。可视化模型输出后,发现方案一产生的定位偏差更小,方案二仍然能检测到原始目标形状,因此使用方案一更合适。在提取背景纹理时,以目标最小外接矩形为单元,以 8 邻域内的背景单元作为候选区域,如图 3 所示。



图 3 背景候选区域

为了选择与目标颜色相似度最大的背景单元,计算目标区域与背景单元的颜色直方图,使用巴氏系数表示颜色相似度:

$$\rho(p, p_m) = \sum_{i=1}^n \sqrt{p(i)p_m(i)} \quad (2)$$

其中, p, p_m 分别代表目标直方图与候选单元的直方图数据, m 是 1 到 8 之间的整数, ρ 的值域为 0 到 1。取 ρ 值最大的背景单元 M 作为背景纹理。改变形状方法的生成图 \tilde{X} 计算如下:

$$\tilde{X} = X \odot \text{mask}_c + B \odot (1 - \text{mask}_c) \quad (3)$$

其中, mask_c 为内接形状的掩膜, B 为使用背景单元 M 覆盖目标 GT 框区域所得的图像。

3 实验与分析

本节重点对迷彩伪装目标检测模型的视觉特征进行分析,为了比较迷彩伪装目标与常规目标检测任务在视觉特征偏向性的差异,使用第 2 节提出的视觉特征偏好分析框架及解耦方法,选取相同的 CNN 模型在迷彩伪装目标与常规目标两类数据集上进行实验。

3.1 数据集

CAMP^[15] 是一种公开的迷彩伪装数据集,由 2 600 张迷彩伪装人员图片及其分割标注构成。该文对原数据集中的分割标注结果用最小外接矩形框重新标注,以用于检测任务。

由于 CAMP 数据集目标为人,为方便对比,对于常规目标检测的实验,数据集中的目标也应设置为人。SBD 数据集对 PASCAL VOC 数据中没有分割标注的数据重新进行了标注。从 SBD 数据集中选择类别为人的图片,设置为常规人员检测数据集(以下简称为 SBD_PERSON)。实验数据集的类别及样本划分情况见表 1。

表 1 数据集类别及样本划分

数据集名称	样本类别	训练集样本数量	测试集样本数量
CAMP	Person(迷彩伪装)	2 080	520
SBD_PERSON	Person(常规)	2 000	500

3.2 模型

现有 CNN 目标检测模型可分为基于候选区域的 Anchor-based 模型(以 Faster R-CNN^[31] 算法为代表)、基于回归的 Anchor-based 模型(主要有 SSD^[32], RetinaNet^[33], YOLO 系列算法)和 Anchor-free 的检测模型(包括 FCOS^[34], CenterNet2^[35] 等)。实验使用 Faster R-CNN, Cascade R-CNN^[36], Mask R-CNN^[37], SSD, RetinaNet, YOLOv5s, YOLOv7^[38], FCOS, CenterNet2 共 9 种通用的 CNN 目标检测模型。

3.3 评价指标

3.3.1 平均准确度均值

目标检测的结果按是否正确可分为:真正例(True Positive, TP)、真反例(True Negative, TN)、假正例(False Positive, FP)、假反例(False Negative, FN)。由此可计算模型的查准率 p 和查全率 r , 计算公式如下:

$$p = \frac{TP}{TP + FP} \quad (4)$$

$$r = \frac{TP}{TP + FN} \quad (5)$$

平均准确度均值(mean Average Precision, mAP)反映了模型中各类别检测的平均精度的均值,其计算如下:

$$\text{mAP} = \frac{1}{|Q|} \sum_{q \in Q} \text{AP}(q) \quad (6)$$

其中, Q 是数据集中包含的类别数, AP 指某一类别的平均精度, AP 计算如下:

$$\text{AP} = \frac{1}{11} \sum_{r \in \{0, 0.1, \dots, 1\}} \max_p(\tilde{r}) \quad (7)$$

即给定 11 个召回率值 r , 找出大于 r 的所有 \tilde{r} 中对应的最大精度值 p , 然后计算这些 p 值的平均值。

3.3.2 mAP 变化率

不同模型训练得到的初始 mAP 值不同,为了便于分析,使用 mAP 变化率作为评价指标。将 mAP 变化率定义为视觉特征解耦后 mAP 值的变化量占原始 mAP 值的比率,其计算如下:

$$T = \frac{m - \tilde{m}}{m} \times 100\% \quad (8)$$

其中, m 代表原始测试集 mAP, \tilde{m} 代表特征解耦后的测试集 mAP。

3.4 实验方法

3.4.1 区域置乱的尺度选择

为比较不同置乱尺度下的纹理特征破坏程度,使用颜色直方图衡量颜色相似度,使用 SSIM 指标衡量纹理相似度。给定图像 X 、经纹理破坏后的图像 Y , SSIM 指标计算如下:

$$\text{SSIM}(X, Y) = \frac{(2\mu_X\mu_Y + c_1)(2\sigma_{XY} + c_2)}{(\mu_X^2 + \mu_Y^2 + c_1)(\sigma_X^2 + \sigma_Y^2 + c_2)} \quad (9)$$

其中, μ_X, μ_Y 分别为 X, Y 的平均值, $\sigma_X, \sigma_Y, \sigma_{XY}$ 分别为 X 的标准差、 Y 的标准差、 XY 的协方差。计算不同尺寸 N 下纹理块置乱图与原图在纹理和颜色特征的相似度,结果见表 2。置乱后的测试集与原始测试集的颜色相似度始终较高, SSIM 值随 N 逐渐减小,这说明颜色特征的统计量不受置乱区域的大小影响,纹理特征的破坏程度随 N 的减小而增大。故取 $N=1$ (像素值乱)的区域置乱用于特征解耦方法中纹理特征的破坏。

表 2 不同尺度纹理置乱图与原图相似度

N	SSIM	颜色相似度
20	0.858	0.985
10	0.718	0.974
5	0.536	0.955
2	0.385	0.945
1	0.338	0.94

3.4.2 解耦方法的有效性验证

文献[10]制作了人类视觉上分别具有颜色偏重、形状偏重和纹理偏重的三个分类数据集并且对每个数据集分别训练了颜色、形状和纹理三种特征编码器。为验证所提解耦方法的有效性,使用所提解耦方法对不同属性偏置数据集分别解耦其对应偏置属性的特征。将解耦后的数据输入特征编码器中,与将原始图像输入特征编码器输出的准确率进行比较,结果见表 3。括号外和括号内的数字分别表示不使用解耦方法和使用解耦方法后的准确率。

表 3 特征解耦数据与原始数据在视觉特征编码器的表现

数据集	解耦方法	颜色编码器	纹理编码器	形状编码器
颜色偏置数据集	消除颜色	24% (5.6%)	31% (32%)	14% (13%)
纹理偏置数据集	消除纹理	23% (20.6%)	60% (8.7%)	19% (19%)
形状偏置数据集	消除形状	16% (14.5%)	66% (60%)	43% (7%)

由表 3 知,在特定属性偏置数据集上使用视觉特征解耦方法,可以消除对应属性的特征,导致该属性特征编码器失效,但不影响其他属性特征编码器的性能。因此,提出的特征解耦方法能够消除数据集某一属性特征,同时不改变其他属性特征。

3.4.3 方法对比

文献[10]中提取图像单一特征的解耦方法仅适用于常规目标的分类任务,而该文提出的分析框架及方法在分类任务和检测任务、常规场景和伪装场景中均可使用。为比较两种方法的效果,在常规目标的分类任务上进行分析。使用文献[10]提供的数据集,在数据集原始图像上训练好模型,分别将使用文献[10]中方法所得图像和使用文中方法所得图像输入模型,模型准确率如表 4 所示。解耦方法为表格第一列,方法中是否保留颜色、纹理、形状特征示于表格第二至四列,Ds1,Ds2,Ds3 分别为颜色偏置数据集、纹理偏置数据集、形状偏置数据集。

文献[10]提取单一特征输入模型,模型在特征解耦数据的准确率越高表示在该特征偏好越强。文中方

法消除图像的单一特征,模型准确率越低表示在该特征的偏好越强。由表 4 知,当使用文献[10]中方法时,对数据集 Ds1,模型的准确率均保持较低水平,难以比较模型的偏向性。对 Ds2,模型对纹理特征有较强偏好,但对颜色特征和形状特征的偏向性难以区分;对于 Ds3,模型在颜色和形状特征的偏向性难以区分。使用文中方法,模型在数据集上对颜色、纹理、形状特征的偏向性区分更加明显。

表 4 不同解耦方法对模型准确率的影响比较

方法	颜色	纹理	形状	Ds1	Ds2	Ds3
原图	✓	✓	✓	0.95	0.98	0.97
文献 [10]	✓	×	×	0.213	0.167	0.122
	×	✓	×	0.261	0.565	0.523
	×	×	✓	0.138	0.153	0.166
	×	✓	✓	0.591	0.877	0.94
文中	✓	×	✓	0.217	0.13	0.324
	✓	✓	×	0.625	0.77	0.68

3.5 实验结果

在 CAMP 数据集上应用所提视觉特征解耦方法,效果如图 4 所示。将原始测试集与特征解耦后的测试集输入训练良好的目标检测模型,检测结果见表 5,括号外数据为 mAP 值,括号内数据为 mAP 变化率。

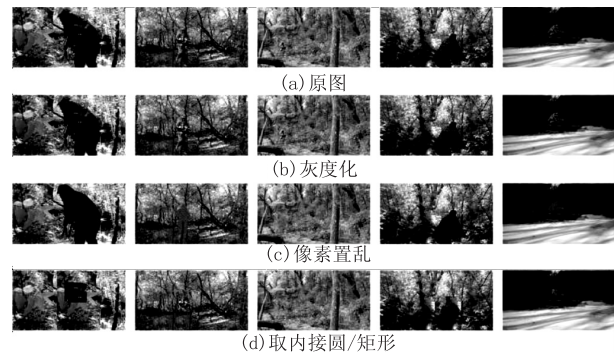


图 4 特征解耦效果

由表 5 知,对于 CAMP 数据集,表中 9 个 CNN 目标检测模型皆对纹理特征更为敏感。当消除颜色特征时,模型的 mAP 变化率范围为 2.1%~9.8%,均值为 4.5%;破坏纹理特征后,模型的 mAP 变化率范围为 55.9%~86.6%,均值为 74.1%;改变形状特征后,模型的 mAP 变化率范围为 26.1%~59.6%,均值为 40.2%。取三种特征解耦数据上的 mAP 变化率均值,归一化处理后,可得模型在 CAMP 数据集上的视觉特征偏向性为:纹理(62%)>形状(34%)>颜色(4%)。

在 SBD_PERSON 上进行同样的实验,实验结果见表 6。由表 6 知,对于 SBD_PERSON 数据集,模型皆对形状特征更为敏感。消除颜色后,模型的 mAP 变化率均值为 8.5%;破坏纹理后,模型的 mAP 变化率

均值为42%;改变形状后,模型的mAP变化率均值为68.4%。取三种特征解耦数据上的mAP变化率均值,归一化处理后,可得模型在SBD_PERSON数据集上

表5 CAMP数据集特征解耦后模型mAP

颜色	纹理	形状	Faster R-CNN	Cascade R-CNN	Mask R-CNN	SSD300	RetinaNet	YOLOv5s	YOLOv7	FCOS	CenterNet2
✓	✓	✓	0.921	0.901	0.90	0.855	0.926	0.915	0.92	0.80	0.897
×	✓	✓	0.888 (↓3.6%)	0.813 (↓9.8%)	0.881 (↓2.1%)	0.787 (↓8%)	0.888 (↓4.1%)	0.883 (↓3.5%)	0.895 (↓2.7%)	0.776 (↓3%)	0.864 (↓3.7%)
✓	×	✓	0.406 (↓55.9%)	0.202 (↓77.6%)	0.307 (↓66.2%)	0.235 (↓72.5%)	0.35 (↓61.1%)	0.15 (↓83.6%)	0.14 (↓84.8%)	0.17 (↓78.8%)	0.12 (↓86.6%)
✓	✓	×	0.601 (↓34.7%)	0.666 (↓26.1%)	0.626 (↓30.4%)	0.468 (↓45.3%)	0.67 (↓27.6%)	0.537 (↓41.3%)	0.372 (↓59.6%)	0.382 (↓52.3%)	0.495 (↓44.8%)

表6 SBD_PERSON数据集特征解耦后模型mAP

颜色	纹理	形状	Faster R-CNN	Cascade R-CNN	Mask R-CNN	SSD300	RetinaNet	YOLOv5s	YOLOv7	FCOS	CenterNet2
✓	✓	✓	0.925	0.899	0.904	0.897	0.923	0.902	0.9	0.835	0.914
×	✓	✓	0.848 (↓8.3%)	0.80 (↓11%)	0.839 (↓7.2%)	0.83 (↓7.5%)	0.856 (↓7.3%)	0.823 (↓8.8%)	0.816 (↓9.3%)	0.748 (↓10.4%)	0.852 (↓6.8%)
✓	×	✓	0.676 (↓26.9%)	0.496 (↓44.8%)	0.59 (↓34.7%)	0.593 (↓33.9%)	0.68 (↓26.3%)	0.51 (↓43.5%)	0.325 (↓63.9%)	0.45 (↓46.1%)	0.385 (↓57.9%)
✓	✓	×	0.262 (↓71.7%)	0.362 (↓59.7%)	0.27 (↓70.1%)	0.309 (↓65.6%)	0.31 (↓66.4%)	0.26 (↓64.2%)	0.12 (↓86.7%)	0.29 (↓65.3%)	0.31 (↓66%)

基于上述研究结果,得出如下结论:尽管不同模型对不同视觉特征的敏感程度略有不同,总体来讲,现有基于CNN的通用目标检测模型在学习迷彩伪装目标的特征时具有较强的纹理偏好,在学习常规目标的特征时具有较强的形状偏好。

同时,相对于形状和纹理特征,颜色特征对二者的检测影响较小。因此,迷彩伪装目标的检测任务不应与常规目标的检测任务一概而论,针对迷彩伪装目标检测的模型改进可以从设计特征网络使之捕捉更精细的纹理特征、使用纹理增强的方法进行数据处理等方面入手。

4 结束语

通过实验对比了迷彩伪装目标与常规目标的视觉特征在目标检测任务中的敏感程度。实验证明:对于CNN目标检测模型,迷彩伪装目标的检测主要依赖其纹理,常规目标的检测主要依赖其形状。同时,颜色特征在二者的检测中不占主导地位。

实验结论可用于指导下一步迷彩伪装技术的发展和迷彩目标检测模型的改进。实验中运用的视觉特征解耦方法和思路可用于验证不同数据集在计算机视觉任务中的特征偏向性,指导设计具有特定偏向性的数据集和网络模型。

参考文献:

- [1] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[J]. Communications of the ACM, 2017, 60(6): 84-90.
- [2] 邓小桐, 曹铁勇, 方正, 等. 改进 RetinaNet 的伪装人员检测方法研究[J]. 计算机工程与应用, 2021, 57(5): 190-196.
- [3] 王 杨, 曹铁勇, 杨吉斌, 等. 基于 YOLOv5 算法的迷彩伪装目标检测技术研究[J]. 计算机科学, 2021, 48(10): 226-232.
- [4] 王焯奎, 曹铁勇, 王 杨, 等. CAMOU-YOLO: 一种迷彩伪装目标检测模型[J]. 计算机技术与发展, 2022, 32(12): 29-36.
- [5] 吴 涛, 王伦文, 朱敬成. 改进的 YOLOv3 算法对伪装目标检测[J]. 火力与指挥控制, 2022, 47(2): 114-120.
- [6] 张 伟, 周旗开, 李睿智, 等. 基于注意力机制的迷彩伪装人员检测算法研究[J]. 医疗卫生装备, 2022, 43(9): 1-7.
- [7] CADIEU C F, HONG H, YAMINS D L K, et al. Deep neural networks rival the representation of primate IT cortex for core visual object recognition[J]. PLoS Computational Biology, 2014, 10(12): e1003963.
- [8] YAMINS D L K, HONG H, CADIEU C F, et al. Performance-optimized hierarchical models predict neural responses in higher visual cortex[J]. Proceedings of the National A-

- cademy of Sciences, 2014, 111(23):8619–8624.
- [9] GEIRHOS R, RUBISCH P, MICHAELIS C, et al. ImageNet-trained CNNs are biased towards texture; increasing shape bias improves accuracy and robustness[EB/OL]. [2018–11–29]. <https://arxiv.org/pdf/1811.12231.pdf>.
 - [10] GE Y H, XIAO Y, XU Z, et al. Contributions of shape, texture, and color in visual recognition[C]//European conference on computer vision. Berlin: Springer, 2022: 369–386.
 - [11] TANKUS A, YESHURUN Y. Convexity-based visual camouflage breaking[J]. Computer Vision & Image Understanding, 2001, 82(3): 208–237.
 - [12] BHAJANTRI N U, NAGABHUSHAN P. Camouflage defect identification: a novel approach[C]//International conference on information technology. Bhubaneswar: IEEE, 2007.
 - [13] 武国晶, 吕绪良, 邢海宁, 等. 三维凸面分析法在迷彩伪装检测中的应用[J]. 解放军理工大学学报: 自然科学版, 2015, 16(6): 582–586.
 - [14] SENGOTTUVELAN P, WAHI A, SHANMUGAM A. Performance of decamouflaging through exploratory image analysis[C]//First international conference on emerging trends in engineering and technology. Nagpu: IEEE, 2008: 6–10.
 - [15] ZHENG Y F, ZHANG X W, WANG F, et al. Detection of people with camouflage pattern via dense deconvolution network[J]. IEEE Signal Processing Letters, 2019, 26(1): 29–33.
 - [16] FANG Z, ZHANG X W, DENG X T, et al. Camouflage people detection via strong semantic dilation network[C]//Proceedings of the ACM turing celebration conference. Chengdu: ACM, 2019: 134.
 - [17] WANG Y, FANG Z, ZHENG Y F, et al. Dual attribute adversarial camouflage toward camouflaged object detection[J]. Defence Technology, 2023, 22(2): 166–179.
 - [18] HERRERAS E B. Cognitive neuroscience: the biology of the mind[J]. Cuadernos de Neuropsicología/Panamerican Journal of Neuropsychology, 2010, 4(1): 87–90.
 - [19] GEGENFURTNER K R, RIEGER J. Sensory and cognitive contributions of color to the recognition of natural scenes[J]. Current Biology, 2000, 10(13): 805–808.
 - [20] PEUSKENS H, CLAEYS K G, TODD J T, et al. Attention to 3-d shape, 3-d motion, and texture in 3-d structure from motion displays[J]. Journal of Cognitive Neuroscience, 2004, 16(4): 665–682.
 - [21] PUCE A, ALLISON T, ASGARI M, et al. Differential sensitivity of human visual cortex to faces, letterstrings, and textures: a functional magnetic resonance imaging study[J]. Journal of Neuroscience, 1996, 16(16): 5205–5215.
 - [22] KUBILIUS J, BRACCI S, OP DE BEECK H P. Deep neural networks as a computational model for human shape sensitivity[J]. PLoS Computational Biology, 2016, 12(4): e1004896.
 - [23] GATYS L A, ECKER A S, BETHGE M. Texture and art with deep neural networks[J]. Current Opinion in Neurobiology, 2017, 46: 178–186.
 - [24] LI Y W, YU Q H, TAN M X, et al. Shape-texture debiased neural network training[EB/OL]. [2021–03–30]. <https://arxiv.org/pdf/2010.05981v2.pdf>.
 - [25] LEE S J, HWANG I, KANG G C, et al. Improving robustness to texture bias via shape-focused augmentation[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. New Orleans: IEEE, 2022: 4323–4331.
 - [26] HAWKINS J K. Textural properties for pattern recognition[C]//Picture processing and psychopictorics. New York: Academic Press, 1970: 347–370.
 - [27] 王润生. 图像理解[M]. 长沙: 国防科学技术大学出版社, 1995: 145–146.
 - [28] HARALICK R M. Statistical and structural approaches to texture[J]. Proceedings of the IEEE, 1979, 67(5): 786–804.
 - [29] TUCERYAN M, JAIN A K. Texture analysis[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. New York: IEEE, 1993: 235–276.
 - [30] 谢新华, 梁 栋, 张香倩, 等. 基于图像处理的目标物体最大内接矩形面积的检测[J]. 科学技术与工程, 2015, 15(17): 193–197.
 - [31] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, 39(6): 1137–1149.
 - [32] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector[C]//European conference on computer vision. Amsterdam: Springer, 2016: 21–37.
 - [33] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2020, 42(2): 318–327.
 - [34] TIAN Z, SHEN C, CHEN H, et al. FCOS: fully convolutional one-stage object detection[C]//Proceedings of the IEEE/CVF international conference on computer vision. Seoul: IEEE, 2019: 9627–9636.
 - [35] ZHOU X Y, KOLTUN V, KRÄHENBÜHL P. Probabilistic two-stage detection[EB/OL]. [2021–03–12]. <https://arxiv.org/pdf/2103.07461.pdf>.
 - [36] CAI Z W, VASCONCELOS N. Cascade R-CNN: delving into high quality object detection[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. Salt Lake City: IEEE, 2018: 6154–6162.
 - [37] HE K M, GKIOXARI G, DOLLÁR P, et al. Mask rcnn[C]//Proceedings of the IEEE international conference on computer vision. Venice: IEEE, 2017: 2961–2969.
 - [38] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real time object detectors[EB/OL]. [2022–07–06]. <https://arxiv.org/pdf/2207.02696.pdf>.