

DOI:10.19651/j.cnki.emt.2210999

一种基于 CenterNet 的多朝向建筑物检测方法

顾东泽 王敬东 姜宜君 廖元晖
(南京航空航天大学自动化学院 南京 211106)

摘要: 航拍图像中的建筑物分布往往是朝向多样的。基于传统卷积神经网络的目标检测算法多以水平锚框作为检测框,在检测多朝向分布的建筑物场景下准确率较低。为此本文提出一种基于 CenterNet 神经网络的目标检测算法,在 CenterNet 模型基础上添加角度预测分支,将朝向角度信息融入网络中。针对 CenterNet 模型在特征提取阶段提取到的建筑物角度特征较少问题,采用非对称卷积替代原先的 3×3 卷积,加强神经网络对于旋转目标角度信息的特征提取能力,并通过改进损失函数降低角度周期性问题对目标检测的影响。改进后的网络可以更加精准的检测出多朝向分布的建筑物。在本文构建的数据集上进行了实验测试,相同环境下网络改进前后建筑物整体检测准确率提升 5.2%,其中 $10^\circ \sim 80^\circ$ 以及 $100^\circ \sim 170^\circ$ 范围内大朝向变化建筑物检测准确率提升了 7.4%。 $0^\circ \sim 10^\circ$ 、 $80^\circ \sim 100^\circ$ 以及 $170^\circ \sim 180^\circ$ 范围内小朝向变化建筑物检测准确率提升了 3.1%,有效提高了多朝向建筑物检测的准确度。

关键词: 关键点检测;多朝向建筑物;非对称卷积;角度周期性

中图分类号: TP391 **文献标识码:** A **国家标准学科分类代码:** 510.4050

A multi-orientation building detection method based on CenterNet

Gu Dongze Wang Jingdong Jiang Yijun Liao Yuanhui
(College of Automation Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China)

Abstract: Buildings in aerial images often have multiple orientations. The target detection algorithm based on the traditional convolutional neural network mostly uses the horizontal anchor frame as the detection frame, which has a low accuracy in detecting the building scene with multi orientation distribution. Therefore, this paper proposes a target detection algorithm based on CenterNet neural network, adds angle prediction branch on the basis of CenterNet model, and integrates the orientation angle information into the network. Aiming at the problem that few building angle features are extracted in the feature extraction stage of CenterNet model, asymmetric convolution is used to replace the original 3×3 convolution to enhance the feature extraction ability of neural network for rotating target angle information, and reduce the impact of angle periodicity on target detection by improving the loss function. The improved network can more accurately detect buildings with multi orientation distribution. In this paper, experimental tests are carried out on the data set built by ourselves. Under the same environment, the overall average precision is improved by 5.2% before and after the network improvement, including 7.4% for buildings with large orientation changes within the range of $10^\circ \sim 80^\circ$ and $100^\circ \sim 170^\circ$. The average precision of buildings with small orientation changes within the range of $0^\circ \sim 10^\circ$, $80^\circ \sim 100^\circ$ and $170^\circ \sim 180^\circ$ has increased by 3.1%, effectively improving the average precision of buildings with multiple orientations.

Keywords: key point detection; multi facing buildings; asymmetric convolution; angle periodicity

0 引言

在城市化加速的大背景下,违章违建现象也层出不穷,违章建筑存在着极大的安全隐患,影响着城市未来的发展。基于人工方式查找违章建筑的方法不仅人力投入巨大,且难以做到对违章信息的及时掌握和查处。近几年基于深度

学习的自动提取建筑物特征的方法有效提高了目标检测的精度^[1-3]。

目前国内外针对建筑物的目标检测方法主要分为基于有锚框和无锚框两个方向。有锚框方向仲伟峰等^[4]提出一种旋转矩形区域的目标检测算法,针对目标尺度变化大的特点,融合了特征金字塔池化模块,但超参数繁多,不适用

收稿日期:2022-08-08

于其他场景,泛化能力差。李巍等^[5]设计了一种基于 YOLOv3 旋转矩形检测网络,使用旋转矩形框对带角度的目标物体进行定位,但需要人工设置多个不同角度和不同长宽比的阈值,无法自适应检测目标。唐建宇等^[6]针对遥感场景下任意方向排布的目标,在 YOLOv5^[7]的基础上采用了旋转框并融合注意力机制,使用 CSL 标签^[8]表示法实现目标框的旋转,但由于 CSL 标签分类类别过多,分类误差较大,使得预测层过重,最终的目标检测效果不理想。朱煜等^[9]提出了一种旋转框检测网络 R2-FRCNN,将检测过程分为两个阶段,在粗调阶段将水平框转为旋转框,在细调阶段进一步优化旋转框的定位,但存在着检测速度慢, GPU 资源消耗大等问题。曲优等^[10]在 YOLOv3 基础上提出 AAC 对称卷积,将水平框转换为旋转框,提高了旋转目标的检测速度和精度。Zhu 等^[11]在 RetinaNet^[12]基础上采用旋转锚框,并利用特征金字塔网络实现对任意方向的舰船检测。李明阳等^[13]提出一种多尺度可变性卷积神经网络用以检测遥感图像中的飞机目标。无锚框方向蒋光峰等^[14]针对基于 anchor 算法在检测遥感目标时参数过多,运算效率低下的问题,采用类似关键点检测的逻辑提出了一种无框网络 RCNet,实现了较高的检测速度。张国生等^[15]在提出了一种单阶段旋转目标检测网络,采用五点参数法表示旋转目标避免了角度周期性问题,但顶点回归存在较大的误差。

虽然基于卷积神经网络的目标建筑物检测效果已经得到了很大提升,但在实际场景中,建筑物的分布往往是无序的,朝向多样的。基于水平锚框的有框检测算法不仅无法较好的套取目标建筑物,还存在着计算量大,检测速度慢等问题。基于无框网络的多朝向建筑物检测算法则存在着检测精度不高等问题。因此,针对现有模型在多朝向建筑物目标检测方面存在的问题进行进一步研究,在无框检测网络 CenterNet 基础上进行改进,添加角度预测分支,采用非对称卷积组替代原先的水平卷积核以便于增强网络对于建筑物角度信息的学习能力,改进损失函数以便于抑制角度周期性问题对模型角度预测的干扰,提高多朝向建筑物检测的准确率。

1 多朝向建筑物检测网络模型

1.1 CenterNet 在多朝向建筑物检测中的应用

目前的建筑物目标检测大多是基于水平检测框实现的,适宜检测南北朝向或东西朝向的建筑物,也即在建筑物朝向范围为 $0^\circ \sim 10^\circ$ 、 $80^\circ \sim 100^\circ$ 以及 $170^\circ \sim 180^\circ$ 等场景下可以取得较好的检测效果,但建筑物的实际朝向往往变化很大,在 $10^\circ \sim 80^\circ$ 以及 $100^\circ \sim 170^\circ$ 朝向范围内水平检测框无法较好的覆盖目标建筑物,引起检测精度下降。

CenterNet^[16]是一种单阶段的无锚框目标检测模型,采用类似于 CornerNet^[17]的网络结构,以检测框的中心为基础进行目标检测,该模型将目标检测问题转换为关键点

估计问题,使用关键点估计找到中心点并在此基础上回归其他目标对象属性,如长宽,朝向等,每一个目标只有一个正样本,避免了多个检测框的后处理过程,降低了计算量,更适合用于多朝向建筑物目标检测。

CenterNet 网络的具体训练流程如图 1 所示主要是将输入的图片经过主干网络的卷积与池化,再通过卷积得到关于目标中心点位置,目标中心点偏置大小以及目标宽高信息。

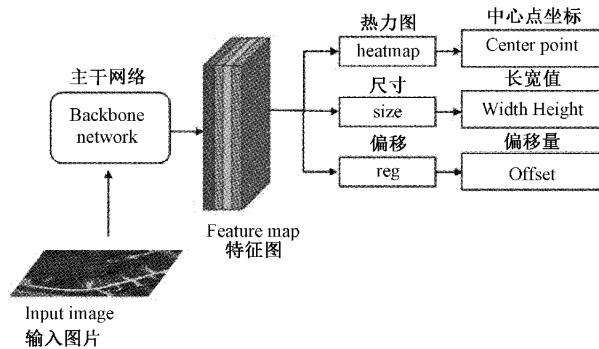


图 1 CenterNet 结构

CenterNet 是以定位目标中心点的方式对建筑物进行检测,受目标建筑物朝向因素的干扰影响较小。采用 CenterNet 网络对多朝向建筑物进行检测,其检测效果如图 2 所示,CenterNet 检测多朝向建筑物时对于建筑物中心点定位比较准确。标记为 1、2、3、4 的近似水平的小朝向变化建筑物套取效果较为精准,但是其他大朝向的建筑物常常会出现水平检测框套取建筑物引入过多背景噪声干扰,检测框互相叠加现象,无法有效分离建筑物与背景。因此本文主要针对偏转范围为 $10^\circ \sim 80^\circ$ 以及 $100^\circ \sim 170^\circ$ 范围内的大朝向变化建筑物检测不准确的问题进行研究。



图 2 CenterNet 水平框检测效果图

1.2 角度预测分支

鉴于 CenterNet 在检测多朝向建筑物时存在的问题,本文在原先的 CenterNet 基础上添加角度预测分支,融入目标建筑物的角度信息,如图 3 所示。其中角度损失函数部分如式(1)所示。

$$L_{ang} = \frac{1}{N} \sum_{k=1}^N |\hat{\theta}_k - \theta_k| \quad (1)$$

式中: L_{ang} 为角度损失, $\hat{\theta}$ 为预测角度, θ 为真值角度值。

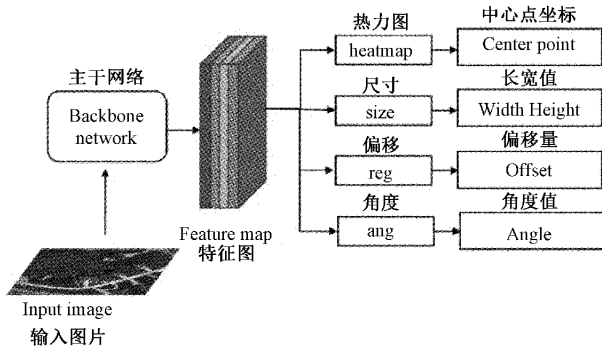


图 3 添加角度预测分支的 CenterNet 结构

为了可以更直观的分析问题,本文将添加角度预测分支后特征提取网络输出的特征图进行了可视化,如图 4 所示。

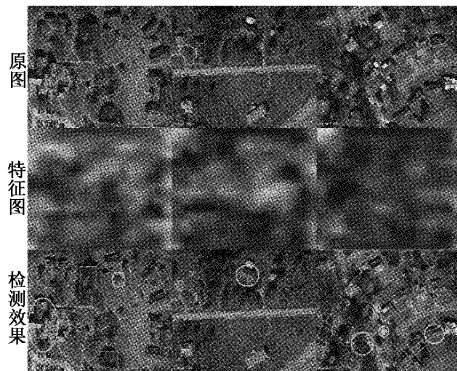


图 4 特征图可视化与输出结果

其中第 1 行为原图,第 2 行为输出特征图,图中颜色越偏亮处表明此处学到的特征信息越多,第 3 行为检测效果图,绿色框为建筑物的检测框,黄色圈圈出的为位置尺寸朝向预测效果均差的建筑物,蓝色圈标记为 1、2 圈出的为位置和尺寸回归正确但角度预测出现偏差的建筑物。结合特征图分析可知,网络模型对黄色圈圈出的建筑物提取的特征信息较少。蓝色圈圈出的建筑物仅角度预测出现偏差是由于多朝向建筑物存在角度周期性问题,模型在回归时损失函数值激增从而影响模型角度预测的精度。因此需要改进特征提取网络增强模型学习角度特征的能力。改进损失函数,抑制角度周期性问题带来的影响。

1.3 非对称卷积

建筑物往往呈现出多朝向分布的状态,CenterNet 网络中的 3×3 方核水平卷积在学习多朝向建筑物时感受野往往比较局限,不利于表达模型所学到的角度信息。

ACNet 结构是由 Ding 等^[18]提出,其中作为 ACNet 基本模块的 ACBlock 是由 3×1 纵向卷积、 1×3 横向卷积以及 3×3 方核卷积组成的非对称卷积组。如式(2)所示,利用卷积操作的可叠加性,可以直接替代原始网络中 3×3 的卷积。

$$I * K^{(1)} + I * K^{(2)} = I * (K^{(1)} \oplus K^{(2)}) \quad (2)$$

其中, I 为输入, $K^{(1)}$ 和 $K^{(2)}$ 是两个卷积核, $*$ 为卷积运算。 \oplus 为按位置求和。

相对于原先单一的水平卷积,纵向的卷积可以提取图中建筑物的纵向特征信息,横向的卷积可以提取图中建筑物的横向特征信息,由三者构成的非对称卷积组可以增强网络对于建筑物的朝向特征学习能力。因此本文将原 CenterNet 网络中基本构建单元里的 3×3 卷积替换为 ACNet 中的 ACBlock,提升网络对多朝向建筑物特征的提取能力,补充更多的朝向特征信息。本文在实验中采用的非对称卷积结构块如图 5 所示,在训练阶段,将原来的单个 3×3 的卷积核替换为 ACBlock 的 3 个卷积核训练之后再行叠加操作。

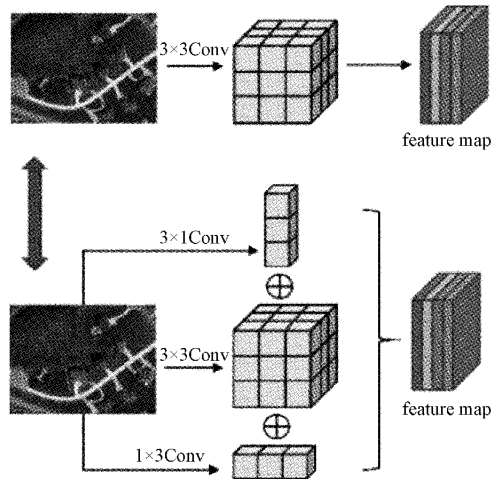


图 5 ACBlock 结构

1.4 损失函数的改进

在将 3×3 卷积替换为非对称卷积之后仍然存在部分建筑物的检测框的朝向与实际真值框的朝向差距较大,但中心点位置和尺寸回归基本正确。图 6 为网络训练损失值随 epoch 变化曲线,查看损失函数曲线发现,损失函数存在突进现象。

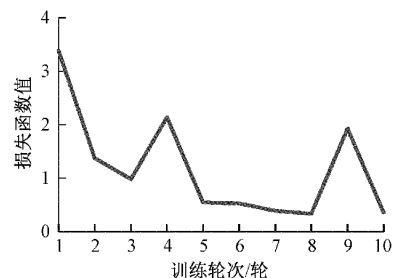


图 6 损失函数曲线

这是由于网络模型在角度回归过程中存在角度周期性问题,如图 7 所示,实线框为真值框,其角度值为 0° ,绿色框为检测框,其角度值为 175° ,尽管虚线框在回归过程中已经接近实际真值框,但二者的角度损失计算值非常大,使得损失函数值激增,影响最终的检测准确率。

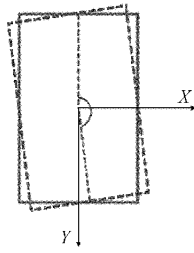


图 7 角度周期性问题

为了抑制角度周期性问题带来的影响,本文对角度损失函数部分做出改进如式(3)所示,在角度损失函数部分前添加系数,用以抑制角度周期性问题带来的影响。当检测框接近真值框时, α 成为一个极小的值,使得角度损失值偏小,抑制了角度周期性问题的影响。

$$L_{ang} = \alpha \times \frac{1}{N} \sum_{k=1}^N |\hat{\theta}_k - \theta_k| \quad (3)$$

$$\alpha = -0.05 \times \ln \frac{int_{area}}{pre_{area} + gt_{area} - int_{area}} \quad (4)$$

其中, int_{area} 为检测框与真值框交集部分面积, pre_{area} 为检测框所覆盖的面积, gt_{area} 为真值框所覆盖的面积。

2 实验及分析

2.1 实验环境

本实验所采用的计算机实验平台配置如下: Intel 的 i7-8700CPU 的和 NVIDIA 的 GTX 1080Ti 的 GPU。本文算法在基于 Ubuntu16.04 系统下的 Pytorch 深度学习框架编写实现。

2.2 数据集介绍

本文在谷歌地球软件中采集了美国各地区的 1 000 张含有建筑物的航拍图像作为数据集,尺寸为 1 024×768 像素。选取其中 100 张作为测试集,其余图片用于模型的训练和验证。数据集的标注软件为 rolabelImg。标注格式为 xml 文件,最后统一转换为 COCO 数据集格式。为了提高模型对于多朝向建筑物检测的鲁棒性,本文对训练数据进行了相应的数据处理,通过对图片进行改变亮度,对比度,增加不同噪声干扰等操作,最终得到 7 200 张图片用于模型训练。处理后效果如图 8 所示。第 1 行从左至右依次为提升亮度,降低亮度以及改变色度后的图片展示。第 2 行从左至右依次为改变锐度,未作改变以及改变对比度后的图片展示。第 3 行从左至右依次为添加高斯噪声,椒盐噪声以及随机噪声后的图片展示。

2.3 实验结果及分析

为了验证上述改进措施的有效性,本文对改进前后的 3 种网络进行消融实验,对改进前后的网络在测试集上的准确率 AP 和平均角度误差 MAE 作对比,如表 1 所示。其中 AP 代表整体建筑物检测准确率, AP_l 代表大朝向变化建筑物的检测准确度, AP_s 代表小朝向变化建筑物的检测准确度。

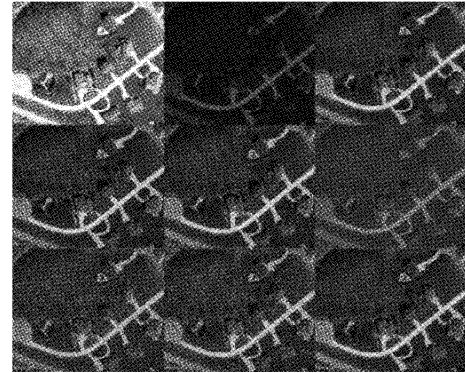


图 8 数据处理后的效果图

表 1 网络改进测试结果前后对比

网络结构	AP	AP_l	AP_s
CenterNet	0.759	0.552	0.802
CenterNet+ACBlock	0.798	0.592	0.828
CenterNet+ACBlock+损失函数改进	0.811	0.626	0.833

由表 1 可以看出,改进后的网络对建筑物整体准确率上升了 5.2%,其中小朝向变化建筑物检测准确率提升了 3.1%,大朝向变化建筑物检测准确率提升了 7.4%。进一步分析发现,ACBlock 这项改进对多朝向建筑物检测准确率提升作用较大,这说明之前的网络对于目标的旋转特征信息提取还不足,增加建筑物目标的旋转特征可以较大提升其检测准确率。改进损失函数之后,检测准确率有所提升,平均角度误差再次降低,表明对损失函数的改进有效抑制了角度周期带来的影响。

网络改进前后的损失函数曲线及检测结果对比如图 9 和 10 所示。图 9 中 1 号曲线(虚线)为改进前的损失函数曲线,由于受角度周期问题的影响,在训练过程中会出现损失值剧增的情况,增加模型的学习难度。2 号曲线(实线)为改进后的损失函数曲线,改进后的损失函数值相比于改进前的损失函数值更小,且曲线更加平缓,说明改进后的损失函数有效抑制了角度周期性问题带来的影响。图 10 中

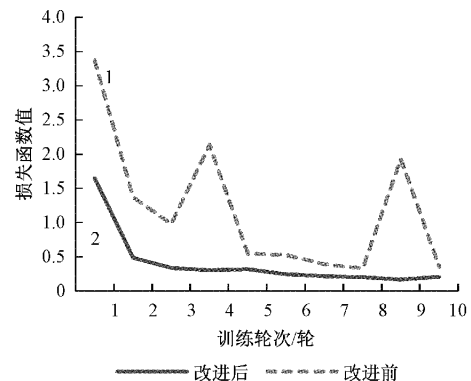


图 9 损失函数曲线对比

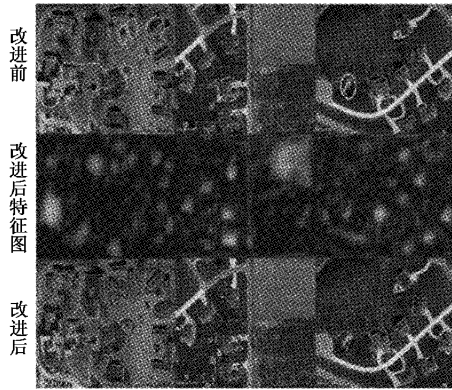


图 10 改进前后检测效果对比图

第 1 行为改进前的检测效果图,其中圆圈圈出的建筑物为漏检、误检建筑物。第 2 行为改进后可视化的特征图,第 3 行为改进后的检测效果图,其中圆圈圈出的建筑物为改进后被正确检测出的建筑物。与改进前的检测结果相比,多朝向建筑物中心位置定位,尺寸回归以及角度预测精度均有所提高。

3 结 论

本文针对水平锚框算法检测多朝向建筑物检测准确率低的问题,在 CenterNet 基础上添加角度预测分支。针对 3×3 卷积感受野较局限导致网络对建筑物朝向特征信息学习不足问题,采用 ACNet 中的非对称卷积模块替代原先的 3×3 卷积。针对角度周期性问题导致部分建筑物角度预测不准确的现象,本文对损失函数作出改进,改善角度损失函数部分,抑制角度损失值突然波动带来的影响。最后通过消融实验证明了两种改进方法都起到了一定作用,建筑物总体检测准确率达到 0.811,提升了 5.2%,其中小朝向变化建筑物检测准确率有所提高,达到了 3.1%,大朝向变化建筑物检测准确率提升较大,达到了 7.4%。

参考文献

- [1] 彭宝钗,郑成. 基于深度学习的建筑物检测方法[J]. 电脑编程技巧与维护,2022(2):136-138.
- [2] 化嫣然,张卓,龙赛,等. 基于改进 YOLO 算法的遥感图像目标检测[J]. 电子测量技术,2020,43(24):87-92.
- [3] 宋廷强,李继旭,张信耶. 基于深度学习的高分辨率遥感图像建筑物识别[J]. 计算机工程与应用,2020(8):26-34.
- [4] 仲伟峰,郭峰,向世明,等. 旋转矩形区域的遥感图像舰船目标检测模型[J]. 计算机辅助设计与图形学学报,2019(11):1935-1945.
- [5] 李巍,戴朝霞,张向东,等. 旋转目标检测算法在卫星影像中的应用[J]. 计算机工程与应用,2021,57(16):134-141.
- [6] 唐建宇,唐春晖. 基于旋转框和注意力机制的遥感图像目标检测算法[J]. 电子测量技术,2021,44(13):

- 114-120.
- [7] 董丽君,曾志高,易胜秋,等. 基于 YOLOv5 的遥感图像目标检测[J]. 湖南工业大学学报,2022,36(3):44-50.
- [8] YANG X, YAN J. Arbitrary-oriented object detection with circular smooth label[C]. European Conference on Computer Vision,2020:677-694.
- [9] 朱煜,方观寿,郑兵兵,等. 基于旋转框精细定位的遥感目标检测方法研究[J]. 自动化学报,2020,45:1-10.
- [10] 曲优,李文辉. 基于锚框变换的单阶段旋转目标检测方法[J]. 吉林大学学报(工学版),2022,52(1):162-173.
- [11] ZHU M, HU G, ZHOU H, et al. Arbitrary-oriented ship detection based on retinanet for remote sensing images[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2021, 14:6694-6706.
- [12] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[C]. Proceeding of the IEEE International Conference on Computer Vision, 2017:2980-2998.
- [13] 李明阳,胡显,雷宏. 基于可变形卷积神经网络的遥感图像飞机目标检测[J]. 国外电子测量技术,2020,39(7):121-126.
- [14] 蒋光峰,胡鹏程,叶桦,等. 基于旋转中心点估计的遥感目标精确检测算法[J]. 计算机应用研究,2021,38(9):2866-2870.
- [15] 张国生,冯广,李东. 基于姿态表示的航空影像旋转目标检测网络[J]. 广东工业大学学报,2021,38(5):40-47.
- [16] ZHOU X, WANG D, KRÄHENBÜHL P. Objects as points[J]. ArXiv Preprint,2019,ArXiv:1904.07850,.
- [17] LAW H, DENG J. Cornernet: Detecting objects as paired keypoints [C]. Proceedings of the European Conference on Computer Vision (ECCV), 2018:734-750.
- [18] DING X, GUO Y, DING G, et al. Acnet: Strengthening the kernel skeletons for powerful cnn via asymmetric convolution blocks[C]. Proceeding of the IEEE/CVF International Conference on Computer Vision,2019:1911-1920.

作者简介

- 顾东泽,硕士研究生,主要研究方向为数字图像处理。
E-mail:dongze1997@foxmail.com
- 王敬东,副教授,主要研究方向为计算机视觉。
E-mail:wjdl291@nuaa.edu.cn
- 姜宜君,硕士研究生,主要研究方向为数字图像处理。
E-mail:616529744@qq.com
- 廖元晖,硕士研究生,主要研究方向为数字图像处理。
E-mail:1060661417@qq.com