

基于 Resinv-Unet 的图像特征点检测方法^{*}

高梓皓^{1,2} 张巧芬¹ 王桂棠^{1,2} 温腾腾^{1,3} 庞亮雨¹ 贾林锋¹ 吴铭扬¹ 李云飞¹

(1. 广东工业大学机电工程学院 广州 510006; 2. 佛山沧科智能科技有限公司 佛山 528225;

3. 广东药科大学医药信息工程学院 广州 510006)

摘要:针对传统特征点检测算法需人为制定检测机制和基于深度学习的特征点检测网络泛化能力不强的问题,引入灰度不变量和残差结构,设计并实现具备像素级特征点检测能力的残差不变量神经网络(residual-invariant neural network, Resinv-Unet)。采用自标注的方式,在真实场景图像数据集的基础上构建用于训练神经网络的数据集。实验结果表明,Resinv-Unet 相较于现有的特征点检测算法和特征点检测网络,在真实场景图像上具有更强的泛化能力和鲁棒性,在平均精确度、精确度和召回率上均取得更好的性能指标,其中,平均精确度达到 0.715 5、精确度达到 0.776 2、召回率达到 0.713 7。

关键词:图像处理;特征点提取;三维重建;机器视觉;神经网络

中图分类号: TP391.9 **文献标识码:** A **国家标准学科分类代码:** 520.6040

Image feature point detection based on Resinv-Unet

Gao Zihao^{1,2} Zhang Qiaofen¹ Wang Guitang^{1,2} Wen Tengting^{1,3} Pang Liangyu¹ Jia Linfeng¹

Wu Mingyang¹ Li Yunfei¹

(1. School of Electromechanical Engineering, Guangdong University of Technology, Guangzhou 510006, China;

2. Cangke Intelligent Technology Co., Ltd., Foshan 528225, China; 3. School of Medical Information Engineering,

Guangdong Pharmaceutical University, Guangzhou 510006, China)

Abstract: Aiming at the problem that the traditional feature point detection algorithm needs to develop the detection mechanism manually and the weak generalization ability of feature point detection network based on deep learning. The Resinv-Unet neural network with pixel-level feature point detection capability is designed and implemented by introducing gray-value invariant and residual. The data source for training neural networks is constructed by self-labeling on the basis of the real scene image data set. The experimental results show that compared with the existing feature point detection algorithms and feature point detection networks, Resinv-Unet has stronger generalization ability and robustness in real scene images, and has better performance in terms of average accuracy, accuracy and recall, with average accuracy of 0.715 5, accuracy of 0.776 2, and recall 0.713 7.

Keywords: image processing; feature point detection; 3D reconstruction; machine vision; neural network

0 引言

三维重建通过运用图像处理等相关技术,在二维图像数据的基础上还原出物体或场景的三维模型,是获取研究对象三维信息的重要手段,被广泛应用于机器人视觉导航、工业制造、医学等领域。

三维重建算法流程的首要步骤是检测二维图像中稳

定的特征点的位置,要求所检测得到的特征点在光照变化、尺度缩放以及物体位移变化等情况下保持不变性。

现有的特征点检测方法包括传统特征点检测器和基于深度学习的特征点检测网络,两者均具有一定的局限性。其中,传统的特征点检测器,如尺度不变特征变换(scale-invariant feature transform, SIFT)^[1-2]、加速稳健特征(speeded up robust features, SURF)^[3]、Fast 角点检

收稿日期:2023-02-26

^{*} 基金项目:国家自然科学基金(61705045)、佛山市 2021 年高校教师特色创新研究项目(2021DZXX15)资助

测^[4-5]、以及定向 FAST 和旋转 BRIEF (oriented FAST and rotated BRIEF, ORB) 局部特征描述子^[6], 均采用人为制定特定检测机制的方式, 实现有监督的特征点自动检测, 因而存在泛化能力及鲁棒性不强, 在光照和视角剧烈变化等恶劣场景下表现不佳的局限。而基于深度学习技术设计的特征点检测网络^[7-9]中, 学习不变特征变换 (learned invariant feature transform, LIFT) 网络^[10]使用 SIFT 特征点检测器检测得到的特征点位置作为真值进行有监督的训练。Magic Point 网络特征点网络^[7]和 Super Point 特征点提取与匹配网络^[7]针对特征点检测任务而设计, 在现有特征点检测网络中具有优秀的检测效果, 但采用合成数据集进行训练, 相对于真实场景图像而言缺乏复杂性和多样性, 且 Magic Point 网络采用非独立像素分类的方式来解决特征点检测的位置定位问题, 使其先验假设了图像中的特征点数目, 在特征点分布密集的情况下存在丢失特征点的风险。D2-Net 网络^[11]针对昼夜转换和大视角变动等光照环境和拍摄视角变化剧烈的场景而设计, 但由于对特征图进行下采样, 导致特征点的定位精度低。

综上所述, 现有的特征点检测方法缺乏通过自主学习的方式, 完成像素级别的特征点检测任务的能力, 无法实现真实场景数据上的强泛化和高鲁棒性。

针对上述问题, 本文对 Unet 网络^[12-15]结构进行改进, 结合 Unet 在独立像素级分类任务的优秀性能和不变量^[16]在特征点检测领域的优势, 设计并实现 Resinv-Unet 网络, 用于进行图像特征点检测。通过对比实验验证本文所设计网络的优秀性能, 通过消融实验验证本文所做改进的必要性。

1 灰度不变量构建方式

灰度不变量在高斯滤波和高斯偏导滤波的作用下, 整合灰度图像中局部区域内灰度值的信息。在高斯滤波的帮助下, 获取了图像局部区域结构, 从而实现了尺度不变性。高斯偏导滤波的加入模拟了图像的差分结构, 从而可以有效地捕获图像的纹理信息。通过强度值归一化的方式, 抑制了局部区域强度的变化。因此, 灰度不变量在视点变化、物体几何形状变化、光照变化的情况下仍具有不变性。文献^[17-20]表明, 在特征点检测任务中, 基于不变量所提供的额外信息构建的特征描述符相比于仅基于像素值信息构建的特征描述符而言, 具有显著的改进效果。

灰度图像经高斯卷积或高斯偏导卷积的数学表达式

如下:

$$\hat{E}_j(x, y) = E(x, y) * G_j \quad (1)$$

式中: $\hat{E}_j(x, y)$ 为灰度图像经高斯卷积或高斯偏导卷积得到的强度值图像中坐标 (x, y) 处的强度值; $E(x, y)$ 为灰度图像中坐标 (x, y) 处像素点所具有的灰度值; G_j , $j \in \{0, x, y\}$ 表示基于二维高斯函数或二维高斯函数偏 j 方向的导数构建的卷积核; $*$ 表示卷积操作。

梯度图像的计算如下:

$$\hat{E}_w(x, y) = \sqrt{\hat{E}_x(x, y)^2 + \hat{E}_y(x, y)^2} \quad (2)$$

采用经高斯卷积后的强度值归一化每个像素处的梯度值, 可以抑制区域强度的变化, 得到灰度不变量:

$$\hat{W}(x, y) = \frac{\hat{E}_w(x, y)}{\hat{E}(x, y)} \quad (3)$$

2 Resinv-Unet 网络结构设计

2.1 灰度不变量融合网络

灰度不变量的效果如图 1 所示, 灰度不变量图像通过强化原灰度图像中的边缘和角点特征, 使得图像信息更利于边缘和角点处特征点的检测, 但弱化了原灰度图像中灰度变化缓慢的区域, 存在丢失该部分区域内特征点的风险。

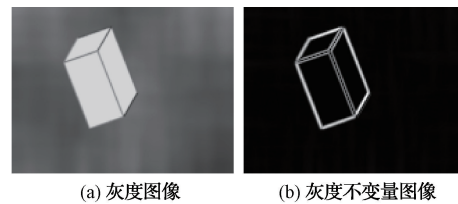


图 1 灰度不变量效果

本文采用固定权重的卷积核代替灰度不变量的计算过程, 将灰度不变量的计算纳入神经网络结构的一部分, 通过特征融合的方式, 使网络同时结合原灰度图像与所生成的灰度不变量图像, 在获得灰度不变量的强化特征信息的基础上, 弥补灰度不变量在灰度变化缓慢的区域所缺失的特征信息。用于生成灰度不变量图像并进行特征融合的网络结构如图 2 所示。

2.2 Resinv-Unet

Resinv-Unet 网络在 Unet 网络的基础上, 创新性地添

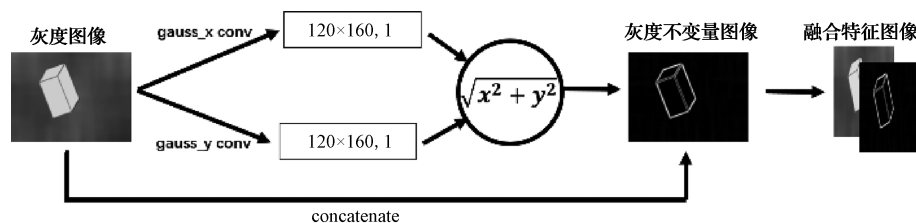


图 2 灰度不变量融合网络

加灰度不变量融合结构,使网络具有特征点预测能力。并通过加入残差结构^[21]的方式,避免了随着卷积层次的加深所带来的网络性能的退化问题,防止在训练过程中出现梯度消失和梯度爆炸问题。

Resinv-Unet 由头部的灰度不变量融合结构、左侧的编码器路径、右侧的解码器路径、中间的通道连接组成和最终的预测网络组成。

其中,编码器路径主要由 4 组尺寸大小不同的双层残差卷积模块和一系列下采样模块组成,通过卷积和下采样操作将输入数据尺寸从 120×160 压缩至 15×20 ,输入通道数从 2 扩展到 512,用于提取输入数据的高维特征。编码器路径的基本单元为双层残差卷积结构如图 3 所示,每个单元对输入数据进行两次卷积,在尺寸不变的同时扩充通道数,将图像特征抽象化,形成高维特征。

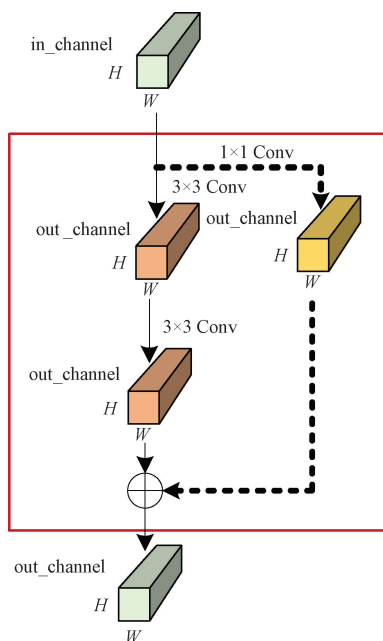


图 3 编码器单元

解码器路径则由 3 组同样的双层残差卷积模块和一系列转置卷积组成,通过卷积和转置卷积将编码器路径得到的高维特征从 15×20 的低分辨率空间映射到 120×160 的高分辨率空间,将通道数从 512 降低到 64,主要作用是将低分辨率下的高维特征逐步解码,并最终映射到高分辨率的像素空间。解码器路径的基本单元如图 4 所示,首先对输入数据进行转置卷积,使数据的通道数降为原来的 $1/2$,尺寸扩大成原来的两倍。然后对输入数据进行两次卷积,在尺寸不变的同时扩充通道数,单元中加入残差结构。

通道连接通过将编码器的输出特征层与解码器的输出特征层进行特征融合,使解码器能共享编码器路径中每一层次提取到的浅层特征。预测网络将编码路径和解码路径相互连接作用得到的最终具有原图尺寸大小的 64 通道特征层通过 1×1 卷积进行通道数的调整,把输出通道

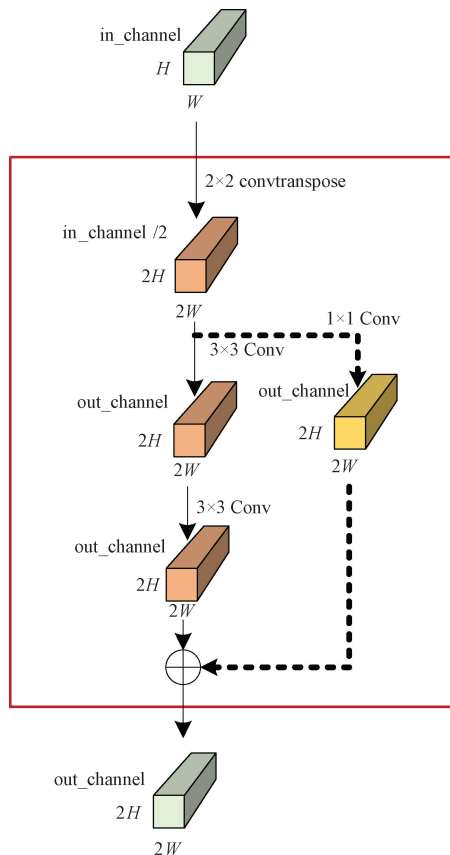


图 4 解码单元

数调整成本文设定的分类的数量 2,将特征点检测任务转化成像素级别的二分类任务,对每一像素进行独立的预测。

Resinv-Unet 的完整结构示意图如图 5 所示。

3 实验与结果分析

3.1 实验数据集

本文通过自标注的方式,在 PASCAL VOC^[22] 的基础上构建训练集。采用合成数据集 Synthetic Shapes Dataset^[7] 和自标注的真实场景图像数据集 PASCAL VOC 共同组成验证集的方式,分别计算实验对象的性能指标。

Synthetic Shapes Dataset 是由具有无争议特征点位置的几何图形组成的数据集,数据集共 4 000 张图像,每张图像的尺寸均为 160×120 。图像由程序渲染生成,包含棋盘格、立方体、椭圆、线段、多边形等共 8 种简单几何图形,如图 6 所示。所对应的真值特征点位置包括线段的端点和交点、棋盘格的交点、多边形的顶点等在简单几何图形中无争议的特征点位置,在图像生成时由程序记录并保存成特征点标注信息。采用此合成数据集的方式避免了复杂繁琐且容易产生位置争议的图像特征点标注工作。

PASCAL VOC 数据集是计算机视觉中常用的数据集,由真实场景图像组成,主要用于分类、检测和分割任务,本文采用 PASCAL VOC 数据集中的 17 115 张图像组

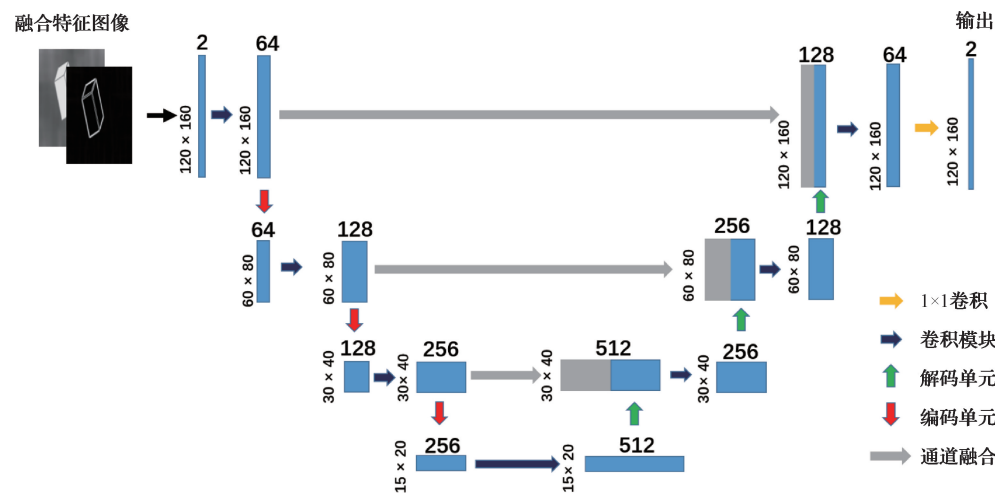


图5 Resinv-Unet 结构示意图

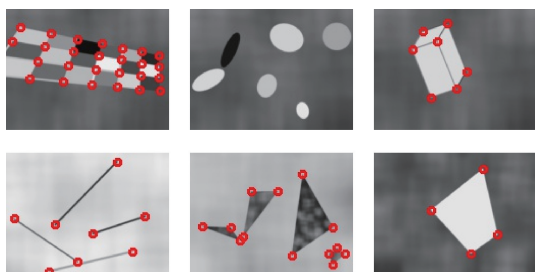


图6 Synthetic Shapes Dataset 合成数据集

成数据集,并按 8:2 的比例将数据集分成训练集和验证集。由于 PASCAL VOC 数据集中图像的尺寸普遍大于 160×120 ,为了使图像尺寸与 Synthetic Shapes Dataset 数据集相对应,同时不损失图像像素信息,对 PASCAL VOC 数据集中的图像采用图 7 所示的图像尺寸预处理操作。

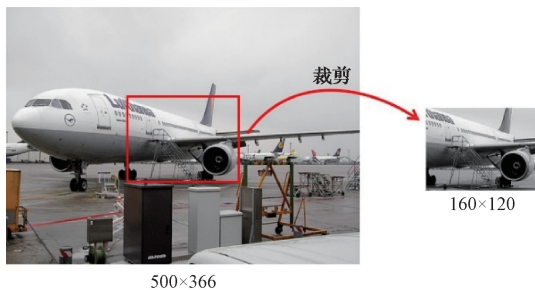


图7 图像尺寸预处理

PASCAL VOC 数据集的标注中并不包含用于特征点检测任务的标注,而人工标注的方式工作繁琐且价格昂贵,因此本文采用 Harris 算法^[23]检测图像中特征点,并通过非极大值抑制的方式,过滤真值特征点附近的大量伪特征点,生成简介精确的特征点标注信息,将标注信息作为特征点真值,用于网络的特征点检测训练。

数据集自标注如图 8 所示,图 8(a)是 Harris 算法得

到的特征点检测结果,图 8(c)是对应的特征点标注信息,在真值特征点附近存在大量伪特征点。图 8(b)是经非极大值抑制算法过滤后得到的特征点检测结果,非极大值抑制过滤掉大量伪特征点,图 8(d)是对应的特征点标注信息,只保留检测响应最高的真值特征点,使神经网络能准确学习到特征点检测机制。

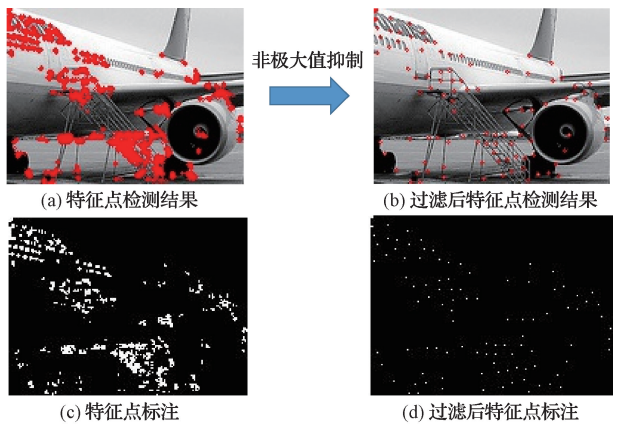


图8 数据集自标注

3.2 评价指标

本文采用精确度 (precision)、召回率 (recall)、平均精确度 (average precision, AP) 作为评价模型性能的指标。这 3 个评价指标均是基于混淆矩阵进行计算,混淆矩阵如表 1 所示。

表 1 混淆矩阵

样本	正样本	负样本
正样本	真阳性 (true positive, TP)	假阴性 (false negative, FN)
负样本	假阳性 (false positive, FP)	真阴性 (true negative, TN)

本文的特征点检测任务中,像素被分类成特征点属于正样本,被分类成非特征点属于负样本。鉴于特征点检测任务在特征点的位置定位上具有一定的宽容性,因此,本文设定被识别成特征点的像素点是否的确属于正样本 TP (被正确识别为正样本)由式(4)判定。

$$TP(y) = \begin{cases} 1, & (\min_{j \in \{1, \dots, K\}} \|y - x_j\|) \leq d \\ 0, & (\min_{j \in \{1, \dots, K\}} \|y - x_j\|) > d \end{cases} \quad (4)$$

式中: y 为图像中被预测成特征点的像素点; $\{x_1, \dots, x_K\}$ 为该副图像中的特征点标签中所标注的属于特征点的像素点; d 为距离判断阈值,本文为 2。当被预测成特征点的像素点距离真实特征点集的最小距离小于等于距离判断阈值 d 时,此预测被统计成 TP。

在混淆矩阵的基础上,精确度和召回率的计算方法如下:

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

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

平均精确度通过对精确度-召回率曲线的积分求得,计算公式如下:

$$AP = \int_0^1 p(r) dr \quad (7)$$

式中: p 表示精确度; r 表示召回率。

3.3 实验设备与网络训练

本文使用 NVIDIA GeForce GTX3090 GPU 进行模型的训练与测试,利用 CUDA 10.2 加速训练过程,使用 Python 3.6 作为编程语言,以 PyTorch 1.10.0 搭建深度学习运行环境。

本文实验中使用适应性矩估计(adaptive moment estimation, Adam)算法作为优化算法,学习率设置为 0.001,一阶矩估计的指数衰减率(beta1)设置为 0.9,二阶矩估计的指数衰减率(beta2)设置为 0.999。网络训练时的 batch size 设置为 64,epoch 设置为 500。

图 9 所示为 Resinv-Unet 训练过程中损失函数(loss)的变化曲线,纵坐标为损失函数值,横坐标表示每个 batch

计算一次损失函数值。

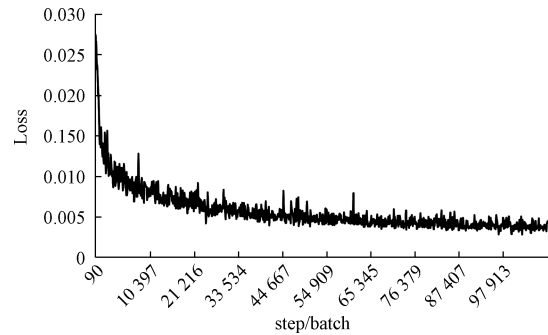


图 9 Resinv-Unet 损失函数变化曲线

图 10 所示为 Resinv-Unet 训练过程中,预测结果的平均精确度的变化曲线,纵坐标为平均精确度的数值,横坐标表示每个 epoch 计算一次平均精确度。

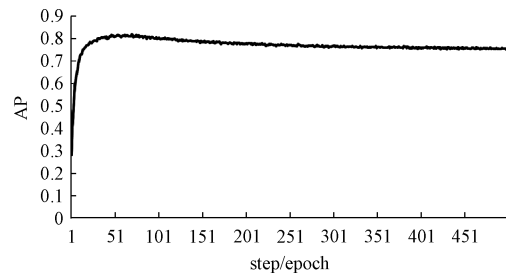


图 10 Resinv-Unet 平均精确度变化曲线

3.4 实验结果与分析

本文在 Synthetic Shapes Dataset 验证集中挑选 1 600 张图像,在 PASCAL VOC 验证集中挑选 1 600 张图像,共同构成混合验证集,仅有当模型确切学习到特征点检测机制且能在真实场景图像上具有优秀的泛化能力和鲁棒性时,才能在此验证集上得到较高的指标数据。

分别检验 Magic Point 网络、无改进的 Unet 网络、传统特征点检测算法中最成熟应用最广泛的算法以及经本文改进后的 Resinv-Unet 网络在此混合验证集上的表现,计算性能评价指标,结果如表 2 所示。

表 2 混合验证集实验结果

	Resinv-Unet	Magic Point	Unet	SIFT	FAST	ORB
AP	0.715 5	0.612 7	0.283 6	—	—	—
precision	0.776 2	0.761 9	0.507 3	0.108 7	0.667 0	0.311 4
recall	0.713 7	0.614 4	0.306 1	0.152 8	0.044 8	0.110 9

本文所设计并实现的 Resinv-Unet 在平均精确度、精确度、召回率 3 个性能指标上均获得最优表现。实验结果表明,Resinv-Unet 模型性能相比于 Magic Point、Unet 以及各广泛应用的传统特征点检测算法均具有显著的提升,在真实场景图像上具有更强的泛化能力和鲁棒性。

如图 11 所示,随机选用一张尺寸为 160×120 的真实场景图像用于对比展示特征点检测效果。图 12 所示是各模型及特征点检测算法的特征点检测效果对比,其中,本文提出的 Resinv-Unet 模型对特征点的检测效果最准确全面,漏点错点最少。



图 11 灰度图

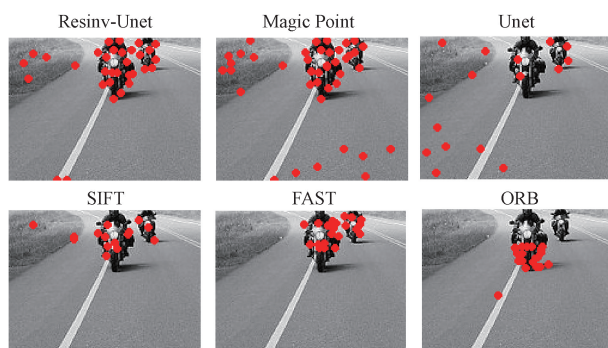


图 12 特征点检测效果

4 结 论

本文提出一种融合灰度不变量和双层残差卷积的 Resinv-Unet 模型,对输入图像进行像素级别的独立分类,使网络具有更精确的特征点检测能力。使用自标注的真实场景数据集对网络进行训练,使网络具有在真实场景图像上的泛化能力和鲁棒性。

消融实验结果验证了本文对 Unet 网络所做改进的必要性,原 Unet 网络无法用于完成特征点检测任务。对比实验结果表明,Resinv-Unet 模型在与各传统特征点检测算法和特征点检测网络的对比中均取得显著优势,且相较于比较对象中性能最好的 Magic Point 模型,平均精确度从 0.612 7 提升到 0.715 5,提升高达 16.78%;精确度从 0.761 9 提升到 0.776 2,提升 1.87%;召回率从 0.614 4 提升到 0.713 7,得到 16.16% 的提升。Resinv-Unet 模型在精度和性能上均有所提升,对真实场景图像的检测结果良好,在特征点密集且分布无规律的真实图像数据集上仍能有效检测出图像中的特征点。

参 考 文 献

- [1] 李学亮,王维. 基于 CNN 和 SIFT 特征的遥感图像变化检测[J]. 电子测量技术, 2019, 42(21): 87-91.
- [2] 陈宁,刘志坚,苏雪平,等. 基于改进的 SIFT 算法的集成电路图像拼接[J]. 国外电子测量技术, 2021, 40(6): 159-164.
- [3] 胡森,吴德阳,仲美玉,等. 基于 SURF 和半色调映射加密的彩色零水印算法[J]. 激光与光电子学进展, 2021, 58(20): 228-242.
- [4] 张绍荣,张闻宇,李云,等. 基于 FAST 角点和 FREAK 描述符改进的无人机景象匹配算法[J]. 电子测量与仪器学报, 2020, 34(4): 102-110.
- [5] 张进,赵相伟,栾吉山,等. 改进 FAST 和对立颜色特征的向量场一致性匹配[J]. 测绘通报, 2020 (11): 50-54.
- [6] RUBLEE E, RABAUD V, KONOLIGE K, et al. ORB: An efficient alternative to SIFT or SURF[C]. Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2011: 2564-2571.
- [7] DETONE D, MALISIEWICZ T, RABINOVICH A. SuperPoint: Self-supervised interest point detection and description[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2018: 224-236.
- [8] SARLIN P E, DETONE D, MALISIEWICZ T, et al. SuperGlue: Learning feature matching with graph neural networks[C]. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020: 4937-4946.
- [9] 余洪山,郭丰,郭林峰,等. 融合改进 SuperPoint 网络的鲁棒单目视觉惯性 SLAM[J]. 仪器仪表学报, 2021, 42(1): 116-126.
- [10] YI K M, TRULLS E, LEPETIT V, et al. LIFT: Learned invariant feature transform [C]. Lecture Notes in Computer Science, 2016: 467-483.
- [11] DUSMANU M, ROCCO I, PAJDLA T, et al. D2-Net: A trainable CNN for joint description and detection of local features[C]. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019: 8084-8093.
- [12] RONNEBERGER O, FISCHER P, BROX T. U-Net: Convolutional networks for biomedical image segmentation [C]. Medical Image Computing and Computer-Assisted Intervention (MICCAI), 2015: 234-241.
- [13] 秦俊豪,李志忠,冯家乐. 基于 Shuffle-Unet 的视网膜血管分割研究[J]. 电子测量技术, 2022, 45(20): 117-124.
- [14] 张珊,王文爽,刘雪峰. 基于嵌入注意力机制的 UNet-DB_ECA 网络检测金具研究[J]. 电子测量技术, 2022, 45(20): 125-134.
- [15] 赵红爱,王旭智,万旺根. 一种用于车辆图像分割的 MSSA-UNet 模型[J]. 电子测量技术, 2022, 45(8): 102-107.
- [16] 魏国剑,侯志强,李武,等. 融合颜色不变量的彩色

- 图像光流估计算法[J]. 电子与信息学报, 2013, 35(12): 2927-2933.
- [17] 王梦思, 王任华. 基于 SIFT 和 RGB 特征的同图复制篡改检测算法[J]. 中国人民公安大学学报(自然科学版), 2018, 24(2): 79-83.
- [18] 吴寅初, 马戎. 基于彩色信息的尺度不变特征变换图像特征点提取与匹配[J]. 计算机应用, 2011, 31(4): 1024-1026, 1124.
- [19] 韩顺, 陶跃华, 朱英南. 彩色描述子 SIFT 的研究新进展[J]. 微型机与应用, 2010, 29(22): 6-8.
- [20] 都布, 岳雅雯. 基于 C-SIFT 特征向量图像复制粘贴篡改取证算法[J]. 电子测量技术, 2019, 42(5): 29-33.
- [21] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]. 2016 IEEE Conference on Computer Vision and Pattern Recognition(CVPR), 2016: 770-778.
- [22] EVERINGHAM M, ESLAMI S M A, GOOL L V, et al. The pascal visual object classes challenge: A retrospective[J]. International Journal of Computer Vision, 2015, 111(1): 98-136.
- [23] 高帅, 贺伟, 李涛, 等. 改进的 Harris 角点检测目标识别方法[J]. 探测与控制学报, 2022, 44(6): 81-86.

作者简介

高梓皓, 硕士研究生, 主要研究方向为三维重建及可视化、数字图像处理、计算机图形学等。

E-mail: 974725965@qq.com

张巧芬(通信作者), 副教授, 主要研究方向为导波光学与光通信技术、成像光学、机器视觉等。

E-mail: zhqf@gdut.edu.cn

王桂棠, 教授, 主要研究方向为嵌入式控制、智能制造、机器视觉等。

E-mail: wanggt@gdut.edu.cn

温腾腾, 讲师, 主要研究方向为智能感知。

E-mail: wentt@gdpu.edu.cn

庞亮雨, 硕士研究生, 主要研究方向为导波光学与光通信技术。

E-mail: 1215841672@qq.com

贾林锋, 硕士研究生, 主要研究方向为音频信号处理。

E-mail: 846526132@qq.com

吴铭扬, 硕士研究生, 主要研究方向为非成像光学。

E-mail: 15915866646@163.com

李云飞, 硕士研究生, 主要研究方向为机器视觉。

E-mail: 307241343@qq.com