DOI:10.16136/j.joel.2023.04.0407

YOLO-NKLT 视觉 SLAM 回环检测方法

刘 玮,温显斌*

(天津理工大学 计算机视觉与系统教育部重点实验室和天津市智能计算及软件新技术重点实验室,天津 300384)

摘要:针对存在明显光照变化或遮挡物等室外复杂场景下,现有基于深度学习的视觉即时定位与 地图构建(visual simultaneous localization and mapping,视觉 SLAM)回环检测方法没有很好地利用 图像的语义信息、场景细节且实时性差等问题,本文提出了一种 YOLO-NKLT 视觉 SLAM 回环 检测方法。采用改进损失函数的 YOLOv5 网络模型获取具有语义信息的图像特征,构建训练集, 对网络重训练,使提取的特征更加适用于复杂场景下的回环检测。为了进一步提高闭环检测的 实时性,提出了一种基于非支配排序的 KLT 降维方法。通过在 New College 数据集和光照等变化 更复杂的 Nordland 数据集上进行实验,结果表明:室外复杂场景下,相较于其他传统和基于深度 学习的方法,所提方法具有更高的鲁棒性,可以取得更佳的准确率和实时性表现。 关键词:室外复杂场景;深度学习;损失函数;回环检测;降维

中图分类号:TP391 文献标识码:A 文章编号:1005-0086(2023)04-0405-08

YOLO-NKLT visual SLAM loopback detection method

LIU Wei, WEN Xianbin*

(Key Laboratory on Computer Vision and System, Ministry of Education of China, Key Laboratory on Intelligence Computing and Novel Software Technology of the City of Tianjin, Tianjin University of Technology, Tianjin 300384, China)

Abstract: Aiming at the problems that the existing deep learning based visual simultaneous localization and mapping (visual SLAM) loopback detection methods do not make good use of the semantic information of images, scene details and poor real-time performance in complex outdoor scenes with obvious illumination changes or occlusion objects, this paper proposes an YOLO-NKLT visual SLAM loopback detection method. The YOLOv5 network model with improved loss function is used to obtain image features with semantic information, construct the training set, and retrain the network to make the extracted features more suitable for loopback detection in complex scenes. In order to further improve the real-time performance of closed-loop detection, a KLT dimensionality reduction method based on non-dominated sorting is proposed. Through experiments on the New College dataset and Nordland dataset with more complex changes such as illumination, the results show that compared with other traditional and deep learn-based methods in outdoor complex scenes, it has higher robustness and can achieve better accuracy and real-time rate performance.

Key words: complex outdoor scene; deep learning; loss function; loopback detection; dimensionality reduction

0 引 言

视觉即时定位与地图构建(simultaneous localization and mapping, SLAM)是通过移动载体上的 视觉传感器获取图像信息来实时构建地图并确定 自身位姿,由于硬件问题和算法误差,不可避免地 会出现位姿漂移问题^[1]。

减少视觉 SLAM 累积误差的重要环节之一就 是回环检测。在受到复杂场景影响时,不同场景 中的图像存在很多相似图像,导致了误匹配的增

 ^{*} E-mail:xbwen@tjut.edu.cn
 收稿日期:2022-05-30 修订日期:2022-09-18
 基金项目:天津市高等学校自然科学研究项目(2017ZD13)资助项目

回环检测方法起初是采用人工提取特征来检测是否回环。RUSSELL等^[3]最先提出先提取图像的特征信息,然后计算并描述这些信息出现的频次,此方法就是后来广泛应用的"词袋模型BOW",应用的典型代表有SIFT^[4]、SURF^[5]和ORB^[6]等。2021年,YU^[7]等利用四叉树设计了一种均匀ORB特征点的提取算法,利用更均匀分布的特征点来完成匹配,在牺牲了一定时间性能的情况下,提高了回环检测的平均准确率。2022年,ZHAO等^[8]提出一种融合ORB和VLAD特征的回环检测方法,此算法设计一种基于二值特征的VLAD量化算法(Binary-VLAD)提取全局特征,并在全局粗搜索阶段,改进倒排索引结构,有效地减少了计算和存储量,在提高回环检测的准确率和时效性上取得了显著成效。

随着深度学习在特征提取中不断取得优异表 现,视觉 SLAM 的研究者们开始将卷积神经网络 融合进回环检测方法。结合深度学习的回环检测 方法在 2015 年被提出^[9]。此算法是利用了 Over Feat^[10]提取图像的特征,相比传统方法而言,不仅 节省了大量的人工计算且在复杂场景下具有更佳 的性能表现。2021年,GAO^[11]等系统介绍了结合 深度学习的动态 SLAM 方法,为后期研究者们优 化 SLAM 方法打开了思路。2022年, CHEN^[12]等 利用 Resnet 网络提取图像特征,通过改进关键帧 选择策略和 PCA 降维,进一步提高了大规模复杂 场景下回环检测的实时性。YANG^[13]利用改进损 失函数的 Darknet 网络提取图像特征,采用 NS-GA-MPCA 进一步降低提取的特征维度,能够在 复杂场景下取得极佳准确率的同时,具有不错的 实时性。

随着深度学习技术的不断成熟,现有网络框架对图像的特征描述更加充分,且在时间性能上

更为突出。同时,也有越来越多的人考虑将语义 信息融入到回环检测中。YOLO系列网络以 YOLOv5为代表,YOLOv5是一种单阶段目标检 测算法,该算法在YOLOv5是一种单阶段目标检 测算法,该算法在YOLOv4的基础上添加了新的 改进思路,使其速度与精度都得到极大的提升。 由于平衡了速度快的特点,丧失了部分精度,因为 没有进行区域采样,所以在小范围的信息上表现 较差,存在识别物体位置精准性差,召回率低等 问题^[14]。

由此,本文提出改进 YOLOv5 的损失函数,对 其重训练,更加充分利用了图像的语义信息和场 景细节,提取出区分度更佳、维度更低的特征描述 子。针对耗时高的问题,提出了基于非支配排序 的改进 KLT 降维方法,YOLO-NKLT 方法能够大 幅提升大规模复杂场景下视觉回环检测的性能。

1 本文方法

虽然现今复杂场景下,基于深度学习的回环检 测方法,通过卷积神经网络提取图像特征,能够很好 地应对光照、气候等复杂环境变化,但当存在明显光 照变化或遮挡物等室外复杂场景,其实际效果依然 有些不尽人意。YOLOv5 在 YOLOv4 的基础上添 加了新的改进思路,使其速度与精度都得到极大的 提升。针对现行回环检测问题中存在的场景细节和 语义信息利用不足问题,本文改进 YOLOv5 的损失 函数,使其更加适应回环检测特性。又针对回环检 测中提取的特征向量维度较高,耗时较长,设计了一 种基于非支配排序的 NKLT 降维方法。综上,本文 提出的 YOLO-NKLT 视觉 SLAM 回环检测方法,总 体流程如图 1 所示。





1.1 YOLOv5 网络结构框架

本文所提出的回环检测方法以 YOLOv5 作为主 干网络提取图像特征。YOLOv5 网络模型整体框架 和 YOLOv3、YOLOv4 相同,主要均由输入、Backbone、Neck 和输出4部分组成。在输入模块,采用数 据增强、自适应图片缩放和锚框计算等手段对图像 进行预处理,将图像调整到网络的输入大小。Backbone 是用来提取一些通用的特征表示, YOLOv5同时使用了CSPDarknet53结构与一个 6×6大小的卷积层作为基准网络。其中,本文所用 的 v6.1版本的 YOLOv5 和之前版本相比,将先前的 Focus 模块换成了一个 6×6大小的卷积层。对于当 前的一些 GPU 设备及其优化算法而言,使用 6×6 大小的卷积层比使用 Focus 模块更加高效。Neck 使 用 SPP 模块,将输入并行通过多个不同大小的 Max-Pool,然后做进一步融合,能在一定程度上解决目标 多尺度问题。将 FPN+PAN 模块放在网络的中间 位置,利用此两个模块进一步提升了特征的鲁棒性 并增加了特征类别^[15]。最后通过 Head 输出部分来 完成检测结果的输出,包括分类和回归两个分支,利 用距离度量损失函数(*GIOU_Loss*)来计算定位损 失,增加了相交尺度的衡量,从而进一步提升方法的 准确率^[4]。YOLOv5 网络结构框架如图 2 所示。



图 2 网络框架 YOLOv5 Fig. 2 Network frame of YOLOv5

1.2 ED-CIOU 损失函数

YOLOv5采用 GIOU_Loss 来计算定位损失,使 用二进制-交叉熵损失函数(BCE_Loss)计算所有样 本的网络预测的目标边界框与 GT_Box 的重叠率参 数(CIOU),并计算正样本的分类损失,计算式为:

$$GIOU_Loss = \sum_{i=0}^{S^2} \sum_{j=0}^{B} 1_{ij}^{obj} (1 - GIOU_{ij}), \quad (1)$$

式中:

$$GIOU_{ij} = \frac{X}{A} - \frac{S-A}{S} , \qquad (2)$$

$$A = \hat{\omega}_i \cdot \hat{h}_i + \omega_i \cdot h_i - X , \qquad (3)$$

式中, GIOU_Loss 为目标位置损失函数;1^{bi}_j 为单元 格 *i* 产生的先验框*j* 包含的目标, X 为边框交集面积, A 为边框并集面积, S 为边框最小外界矩形面积, $\hat{\omega}_i$ 和 \hat{h}_i 分别为预测框的长与宽。

当预测框和真实框相距较远时,GIOU趋近于 -1,即损失值越大。当两者重合时,GIOU 取最大值 1。但对于存在遮挡物体的检测,预测框与真实框会 出现宽和高相同或出现包含关系的情况,该条件下, 差值为0,损失函数不可导,无法收敛。基于此,本文 增加了对 CIOU 的计算和对预测框中心点欧拉距离 的计算。将其作为偏差指标,如图3所示。其中,(h1, w1)、(h,w)分别表示了预测框与真实框的宽与高, b、b1分别表示预测框和真实框的中心点,p表示两个 中心点间的欧拉距离,c表示最小外界矩形对角线的 距离。则相应的表达式为:

$$\begin{cases} CIOU_{ij} = \frac{J}{U} - \frac{\rho^2(b, b^{gt})}{c^2} - \alpha v \\ v = \frac{4}{\pi^2} \left(\arctan \frac{\omega^{gt}}{h^{gt}} - \arctan \frac{\omega}{h} \right)^2 \end{cases}, \tag{4}$$

式中, α 为权值函数。

则损失函数改进为:

$$ED-CIOU = \sum_{i=0}^{S^2} \sum_{j=0}^{B} 1_{ij}^{aj} (1 - CIOU_{ij}) .$$
(5)
ED-CIOU 损失函数增加了对中心点距离的度

量,可以计算出两个目标框的最小距离,大大提高了 收敛速度。同时避免了出现真实框与预测框间包含 关系等时的不收敛的情况,可优化边框间的相互关 系,有效提高对存在遮挡物体检测时的准确率。

1.3 基于改进 *ED-CIOU* 损失函数的 YOLO-ED 网 络训练

通过在 Nordland 数据集上对 YOLO-ED 网络进行训练,以获取最优的参数,达到更好的回环检测效果。

构造好训练集后,可对本文提出的 YOLO-ED 进行训练,其训练流程如图 4 所示。



图 3 损失函数 ED-CIOU 边框图 Fig. 3 Border diagram of loss function ED-CIOU

为了验证本文改进的损失函数的优越表现,也 为了及时发现 YOLO-ED 网络因到达误差阈值而过 早停止训练,将本文提出的方法和 GIOU、BCE_Loss 两种损失函数进行对比。从表 1 中也可以看到,本文 提出的损失函数的损失值的下降速度最快,在网络 迭代 250 次之内就有一个较大幅度的下降,使用 ED-CIOU 损失函数的损失值变化更加稳定且更趋 近于零。



图 4 YOLO-ED 网络训练流程

Fig. 4 YOLO-ED network training process

表1 4种迭代次数下各损失函数的损失值对比表

Tab. 1 Loss value comparison table of each loss

function under four iterations

Iterations	GIOU	BCE	ED-CIOU
250	0.3196	0.8731	0.2534
300	0.1275	1.5969	0.0931
350	0.1914	2.7306	0.0743
400	0.1563	2.3765	0.0711
Average	0.1987	1.8943	0.1230

改进后的YOLO-ED网络的训练参数如表2 所示。

表 2 YOLO-ED 网络参数配置情况

Tab. 2 YOLO-ED network parameter settings

Parameter	Default		
Batch	64		
Subdivisions	16		
Width	608		
Height	608		
Channels	3		
Hue	0.1		
Learning_rate	0.001		
Burn_in	100		
MomenNordland	0.9		
Decay	0.0005		
Angle	0		
Saturation	1.5		
Exposure	1.5		
Max_batches	500 200		
Steps	400 000,450 000		
Scales	0.1,0.1		

1.4 基于非支配排序的 NKLT 降维方法

大型室外场景中场景图像复杂度高,为了获取 更加准确的数据,需要尽可能多地考虑其空间和波 段等信息。本文通过改进 NKLT 方法进行降维处 理,通过对获取的图像信息进行张量化,来避免向量 化过程中对空间结构和数据之间结构关系的破坏。 通过本文改进后的方法进行降维处理,在降低数据 维度的同时还可以最大程度保持张量不同方向上的 数据信息和结构关系,最大程度保留数据信息和潜 在信息。

传统 KLT 算法是将 K 阶高维张量通过 K 个投 影矩阵映射为一个 K 阶低维张量,计算的关键是求 解和确定每一阶的投影矩阵。

考虑到大型室外复杂场景中获取的图像,其每 一模式下张量所包含的数据信息和意义存在很大差 异,对应的投影空间选择也应该不同。本文采用非 支配排序的思想,设计并优化了一个权值矩阵,在原 算法的投影矩阵基础上,对信息量较大的特征向量 保留更多的信息,又尽可能减少信息量较少的特征 向量维度。保证信息量大的特征向量在投影过程中 占主导地位,又能充分保留特征小的特征向量的作 用。实现既可最大程度保留空间信息又减少特征维 度的目的。本文提出的 NKLT 降维方法的计算过程 如下:

1) 初始张量样本输入: $\{K_m \in R^{I_1 \times I_2 \times \cdots I_n}, m = 1, 2, \cdots, M\}$ 。

2) 样本中心值的计算。如式(6)和式(7)所示: $\overline{K}_m = K_m - \overline{K}$, (6)

$$\overline{K} = \frac{1}{M} \sum_{m=1}^{M} K_m \quad . \tag{7}$$

3) 投影矩阵初始化。对矩阵 $\Phi^{(n)^*} = \sum_{m=1}^{M} \overline{K}_{m(n)}$.

 $\overline{K}_{m(n)}^{\mathrm{T}}$ 进行特征分解,分别得到 I_n 个特征值和特征向量,取最大的 D_n 个特征值对应的特征向量,得到投影 矩 $\overline{H}^{(n)}(\overline{H}^{(n)} \in R^{I_n \times D_n})$ 。

4)局部最优化投影矩阵的计算。对矩阵 $\Phi^{(n)} = \sum_{m=1}^{M} (K_{m(n)} - \overline{K}_{(n)}) \cdot \overline{H}_{\Phi^{(n)}} \cdot \overline{H}_{\Phi^{(n)}}^{\mathsf{T}} (K_{m(n)} - \overline{K}_{(n)})^{\mathsf{T}}$ 进行 特征分解。通过循环计算 1 - N 阶张量 *t* 次,分别得到 *I*_n 个特征值和特征向量,取最大的 *P*_n 个特征值对应 的特征向量,得到投影矩阵 $\overline{H}^{(n)} (\overline{H}^{(n)} \in \mathbb{R}^{I_n \times P_n})$ 。

计算 { \overline{L}_m , $m = 1, 2, \dots, M$ } 和 Ψ_{Y_k} 。迭代 t 次后, 跳出循环,执行步骤 5)。整体计算过程如式(8) 和式 (9) 所示:

$$\{ \overline{L}_{m} = \overline{K}_{m} \times_{1} \overline{H}^{(1)T} \times_{2} \overline{H}^{(2)T} \times \cdots \times_{N} \overline{H}^{(N)T}, m = 1, 2, \cdots, M \},$$
(8)

$$\Psi_{L_0} = \sum_{m=1}^M \| \overline{L}_m \|_F^2 \, . \tag{9}$$

5)采用非支配排序思想计算权值矩阵。设Y个 参考点 Zs,父代种群 Pt,种群大小为 N。

① 设置归档集 $S_t = 0, i = 1;$

② 通过交叉变异获得子代种群 Q_i ;

③ 合并父代 P_t 和子代 Q_t 构成种群 R;

④ 对 Rt 进行非支配排序, 划分若干非支配层 F1, F2, F3,...;

⑤判断如果 | S_t | < N,将优先级高的非支配层存入 S_t 中, i = i+1,并将临界层设置为 F1 = Fi;判断如果 | S_t | = N,那么 P_{t+1} = S_t,否则 P_{t+1} = S_t;

⑥ 从 Ft 中选择 X 个个体,其中 X = N − | P_t +
 1 |,不断迭代,直到满足终止条件,得到的最优解集
 即为最优权值;

⑦ 将最优权值对角化表示,如式(10) 所示:

$$W_{(n)} = diag(w_{n1}, w_{n2}, \dots w_{nP(n)}), n = 1, 2, \dots N, \quad (10)$$

6) 投影矩阵的计算 。如式(11)所示:

 $\{L_{m} = K_{m} \times (W_{(1)} \times_{1} \overline{H}^{(1)T}) \times (W_{(2)} \times_{2} \overline{H}^{(2)T}) \times \cdots \times (W_{(3)} \times_{N} \overline{H}^{(N)T}), m = 1, 2, \cdots, M\}_{\circ}$ (11)

7) 输出降维后的样本: $\{L_m \in R^{P_1 \times P_2 \times \cdots P_n}, m =$

 $1,2,\cdots,M$.

1.5 相似度计算

对于回环检测算法而言,获取更为准确特征向 量是重要基础,而如何高效计算图像之间的相似性 才是实现回环检测算法的核心。本文采用了余弦相 似度进行计算。设当前图像和其之前某一帧图像分 别为 p_n 、 p_m ,通过 YOLO-NKLT 模型得到特征向量 v_n 和 v_m 。两个向量间相似度的计算表达式为:

$$\cos\theta = \frac{\sum_{i=1}^{K*D} v_n v_p}{\sqrt{\sum_{i=1}^{K*D} (v_n)^2} \times \sqrt{\sum_{i=1}^{K*D} (v_m)^2}} \quad (12)$$

2 仿真实验与实验分析

2.1 实验环境和实验步骤

本实验所用硬件配置为:处理器 Intel(R)Core (TM)i7-10500CPU@3.10 GHz,操作系统 Windows10,显卡 NVDIAGTX1080Ti,运行内存12 GB。

1) 首先修改 YOLOv5 的网络损失函数。

2) 对修改损失函数后的 YOLO-ED 网络进行重训练。

3)将获取到的图像信息输入 YOLO-ED 网络得 到分类好的特征向量表示。

4) 对获取的特征向量进行降维处理。

5) 计算相似度,进行合理约束,得到回环检测结 果^[16]。

2.2 实验数据集与评估指标

本实验采用 New College 和 Nordland 两个公开的数据集进行测试。Nordland 数据集同一场景明显 光照变化示例如图 5 所示。

实际应用中,不可避免会出现误差。回环检测的结果可能会出现表 3 中的几种情况。

在实验中,通过绘制不同算法在不同配置下的 准确率和召回率曲线,验证回环检测方法的效果。 准确率和召回率的计算式如下:





图 5 Nordland 数据集同一场景明显光照变化示例图 Fig. 5 An example of a significant illumination change in the same scene from the Nordland dataset

$$Precision = \frac{TP}{TP + FP},$$
(13)

$$Recall = \frac{TP}{TP + FN} \,^{\circ} \tag{14}$$

通过计算 PR 曲线与坐标轴围成的面积,得出 (平均精度)AP 值,以此衡量本方法的平均准确率, 计算式如下:

$$AP = \int_{0}^{1} P(r) \mathrm{d}r \, \, . \tag{15}$$

2.3 在 New College 和 Nordland 上的实验结果分析
 图 6、图 7 分别为 YOLO-NKLT 方法、YOLO-

ED方法及其他3种主流的回环检测方法在New

	表 3	回环检测的结果分类
Tab. 3	Classifica	ation of loopback detection results

Algorithm	Fact is loopback	Fact is not loopback	
Is a loopback	True Positive, TP	False Positive, FP	
Is not a loopback	False Negative, FN	True Negative, TN	

College 和 Nordland 数据集中的实验对比结果。



从实验结果可以看出,本文改进的方法和当前

具有代表性的基于深度学习的方法的回环检测效果 均大幅度优于传统 ORB-SLAM2 方法。

在 New College 数据集上,本文提出的 YOLO-NKLT 与 YOLO-ED 的实验结果相近,相比现有几 个典型方法均有着明显的提升。增加了基于非支配 排序的 NKLT 降维方法后的准确率会略有下降,但 在可接受范围内,其综合效果待结合时间性能对比 后评判。FLCNN 方法的准确率略低于本文选用的 YOLOv5 特征。

在场景更复杂的 Nordland 数据集上,当召回率 超过 50%之后,基于传统的 ORB 方法的准确率下降 更加明显,而基于深度学习的方法则能够保持较好 的准确性。YOLO-ED、YOLO-NKLT 及 YOLOv5 当召回率超过 60%之后,展现了其优越性,有着更加 出色的准确率。尤其是 YOLO-ED,其在复杂场景中 的准确率更高,说明本文改进损失函数后的网络能 够获取复杂场景中更多的图像信息,更加高效准确 地检测到真实回环。

表4为5种方法分别在两个数据集中计算得出的回环检测的平均准确率。可以看出,本文提出的YOLO-NKLT相比其他深度学习网络模型有更高的准确率。其中,改进前的YOLOv5和FLCNN在两个数据集上的总体表现差距不是很大,经过改进损失函数后的YOLO-ED方法,在两个数据集中的平均准确率均有了明显的提高。融合了本文提出的基于非支配排序的NKLT降维方法的YOLO-NKLT

• 411 •

在准确率表现上略逊于 YOLO-ED,但相较于 其他几种方法依然存在较大优势,后面将结合时间 性能进行综合评估。

对回环检测平均准确率进行分析,在NewCol-

液 4 回外 極洲 中均准确学 Tab. 4 Average accuracy of loopback detection					
New College	0.691	0.794	0.803	0.853	0.841
Nordland	0.627	0.807	0.842	0.897	0.868

主 1 同环协测亚协准确实

lege 数据集中,YOLO-NKLT 相较于传统 ORB 算法,准确率提升了 15%;在基于深度学习的方法中, YOLO-NKLT 相较于 YOLO-ED 的准确率略低,但 相较 YOLO-v5 方法准确率提升 3.8%;比基于 FLC-NN 的算法,准确率提升 4.7%。在 Nordland 数据 集中,本文提出的算法的准确率同样略低于 YOLO-ED;但相比于传统 ORB算法,准确率提升了 24.1%;在基于深度学习的方法中,本文方法比基于 YOLOv5 的算法的准确率提升 2.6%;比基于 FLC-NN 的算法的准确率提升 6.1%。在 Nordland 数据 集中,基于深度学习的方法取得更佳效果。因为其 场景复杂,特征数量、种类较多,能够训练提取更多 信息,更加考验算法的特征提取能力。同时,本文提 出的 YOLO-ED 方法因为能够在复杂场景中获取更 多的潜在信息,因此能够得到更佳准确的位姿信息, 取得更佳好的回环检测效果。YOLO-NKLT 网络因 为降低了数据维度,牺牲了一点准确率,后续将结合 时间性能对其综合评判。

表 5 记录了同一幅图像在两个数据集中用上述 5 种方法提取的特征检测回环所消耗的时间。

由表 5 中数据可知,本文所用网络的时间性能 优越。基于改进 YOLO-v5 的回环检测方法再结合 基于非支配排序的 NKLT 降维方法后,计算相似度 的时间得到极大的缩短。

综上所述,本文方法可以在保持较高准确率的 同时大大提高其实时性,特别是在存在明显光照变

表 5 时间性能比较(ms) Tab. 5 Time performance comparison (ms)

Dataset	ORB-SLAM2	FLCNN	YOLOv5	YOLO-ED	YOLO-NKLT
New College	58.21	45.81	39.8	38.1	33.1
Nordland	63.21	47.21	37.2	41.2	38.9

化或遮挡物等室外复杂场景时,优势更甚。

3 结 论

本文提出了一种基于改进 YOLOv5 来提取图像 特征并融合了改进的 KLT 降维方法的回环检测方 法。获取的语义信息也使得本文方法取得更高的准 确率。考虑到对 CIOU 和预测框中心点欧拉距离的 计算,改进损失函数,重新训练 YOLO-NKLT,增强 了网络对图像的局部信息和场景细节的利用率,使 本网络框架能够更加适应复杂场景下的回环检测条 件;通过非支配排序思想改进的 NKLT 降维方法,对 提取的特征向量进行降维处理。经过在两个数据集 上的实验得出,当召回率为 60%时,能够达到 91%准 确率的平均水平;本文方法提高了室外复杂场景下 回环检测的准确性,同时具有优秀的实时性。

参考文献:

- YE J H. Research on closed-loop detection method of visual SLAM based on deep learning[D]. Chengdu:University of Electronic Science and Technology of China, 2020.
 叶俊宏. 基于深度学习的视觉 SLAM 闭环检测方法研究[D].成都:电子科技大学, 2020.
- [2] GUO J Z,LIU F L,YANG X Z,et al. Visual SLAM closed-loop detection method based on deep learning [J]. Journal of Optoelectronics Laser, 2021, 32(6); 628-636.
 郭纪志,刘凤连,杨馨竹,等.基于深度学习的视觉 SLAM 闭环检测方法[J].光电子 激光, 2021, 32(06); 628-636.
- [3] Russell B C, Freeman W T, Efros A A, et al. Using multiple segmentations to discover objects and their extent in image collections[C]//IEEE Computer Society Conference

on Computer Vision and Pattern Recognition, June $17\mathchar`-22$, 2006, New York, NY, USA. New York : IEEE , 2006.

- [4] MUR-ARTAL R, TARDO J D. ORB-SLAM2: An opensource SLAM system for monocular, stereo, and RGB-D cameras[J]. IEEE Transactions on Robotics, 2017, 33 (5):1255-1262.
- [5] CHEN Z T, LAM O, JACOBSON A, et al. Convolutional neural network-based place recognition[EB/OL]. (2014-11-06)[2022-05-30]. https:arxiv.org/abs/1411.1509.
- [6] SERMANET P,EIGEN D,ZHANG X, et al. Overfeat: Integrated recognition, localization and detection using convolutional networks[EB/OL](2013-12-21)[2022-05-30]. https:arxiv.org/abs/1312.6229.
- [7] YULL.Research on loopback detection algorithm in visual SLAM [D]. Changcun: Jilin University, 2021.
 于录录.视觉 SLAM 中回环检测算法的研究[D].长春: 吉林大学, 2021.
- [8] ZHAO S,GUAN Q,DING D R,et al. COVFast-LCD: A fast loop back detection algorithm combining ORB and VLAD features[J/OL]. Journal of Chinese Computer Systems:1-8 [2022-05-30] https://kns. cnki. net/kcms/detail/21. 1106.TP.20220418.1239.008.html.

赵珊,管启,丁德锐,等. COVFast-LCD:一种组合 ORB 和 VLAD 特征的快速回环检测算法[J/OL].小型微型计算机系统:1-8[2022-05-30]. https://kns. cnki. net/kc-ms/detail/21.1106.TP.20220418.1239.008.html.

- [9] SÜNDERHAUF N, SHIRAZI S, JCOBSON A, et al. Place recognition with ConvNet landmarks: Viewpoint-robust, condition-robust, training-free[C]//Robotics: Science and Systems XI, July 13-17, 2015, Rome, Italy., 2015:1-10.
- [10] KEISUKE T, FEDERICO T. CNN-SLAM: Real-time dense monocular SLAM with learned depth prediction//2017 IEEE Computer Conference on Computer Vision and Pat-

tern Recognition, July 2126, 2017, Honolulu, HI, USA. New York: IEEE, 2017: 65656574.

 [11] GAO X B,SHI X H,GE Q F,et al. Review of visual SLAM for dynamic object scene [J]. Robot, 2021, 43(6): 733-750.
 高兴波,史旭华,葛群峰,等.面向动态物体场景的视觉

SLAM 综述[J]. 机器人,2021,43(6): 733-750.

- [12] CHEN Y, DANG S W, NIE L. Loopback detection algorithm based on ResNet model [J]. Intelligent Computer and Application, 2022, 12(8):196-199.
 陈勇,党淑雯,聂铃.基于 ResNet 模型的回环检测算法
 [J].智能计算机与应用, 2022, 12(8):196-199.
- [13] YANG X Z. Research on fast closed-loop detection method of Darknet-NVPP Visual SLAM [D]. Chongqing: Chongqing University of Technology,2022. 杨馨竹. Darknet-NVPP 视觉 SLAM 快速闭环检测方法 研究[D]. 重庆;重庆理工大学,2022.
- [14] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLOv4: Optimal speed and accuracy of object detection [EB/ OL]. (2020-04-23) [2022-05-30]. https: arxiv. org/abs/ 2004.10934.
- [15] LI G Q,YU L,FEI S M. A deep-learning real—time visual SLAM system based on multi-task feature extraction network and self-supervised feature points [J]. Measurement,2021,168:108403.
- [16] CHEN B, YUAN D, LIU C, et al. Loopclosure detection based on multi-scale deep feature fusion [J]. Applied Sciences, 2019, 9(6):1120.

作者简介:

温显斌 (1966-),男,博士,教授,博士生导师,硕士生导师,主要的 研究领域为遥感图像理解、模式识别、计算机视觉、多媒体分析、信息 隐藏等.