使用三帧 RGB 扫描编码的实时二维激光雷达目标检测

Soheil Behnam Roudsari[1],Alexandre S. Brand˜ao[2],和 Felipe N. Martins[1]1传感器与智能系统组，汉泽应用科学大学工程学院，格罗宁根，荷兰 s.behnam.roudsari@st.hanze.nl, fe.nascimento.martins@pl.hanze.nl2机器人专业化研究中心，电气工程系，维索萨联邦大学，维索萨，米纳斯吉拉斯州 36570-900，巴西 alexandre.brandao@ufv.br

arXiv:2602.02167v1 [eess.SP] 2 Feb 2026

摘要—**室内服务机器人需要具备稳健的感知能力，同时比 RGB 视频更能保护隐私，并且能够在嵌入式硬件上运行。我们提出一种无需摄像头的二维激光雷达目标检测处理流程，将连续三次扫描堆叠为 RGB 通道，以编码短期时序上下文，从而生成紧凑的 YOLOv8n 输入；该方法无需构建占据栅格，同时保留角度结构和运动线索。在 Webots 中针对 160 个随机室内场景进行评估，并严格采用场景级留出验证，结果显示，该方法在四类目标上的 mAP@0.5 达到 98.4%（mAP@0.5:0.95 为 0.778），精确率为 94.9%，召回率为 94.7%。在树莓派 5 上，该方法可实时运行；预热后，包含扫描编码和后处理在内的端到端延迟平均为每帧 47.8 毫秒。与报告中在同一平台上运行的紧密相关占据栅格激光雷达–YOLO 流程相比，所提出的表示方式对应的报告端到端延迟显著更低。尽管结果基于仿真，但表明轻量级时序编码可实现准确、实时的纯激光雷达检测，适用于嵌入式室内机器人技术，且无需采集 RGB 外观。**

索引词—**二维激光雷达、目标检测、YOLOv8、嵌入式机器人学、时间编码、注重隐私的感知**

I. 引言

室内服务机器人必须能够在小型机载计算机上运行，同时可靠地检测物体并作出响应。现代基于摄像头的检测器具有很高的准确率 [1], ，但视觉容易受到光照影响，而且在有人居住的室内空间中会引发隐私顾虑 [2]。在许多实际部署场景（家庭、医院、办公室）中，感知还必须比持续采集 RGB 视频更加注重隐私。这些限制促使人们开发减少对 RGB 外观依赖、同时又足够高效以便在嵌入式 CPU 上运行的感知处理流程。

激光雷达无需捕捉视觉外观即可提供几何结构，因此是适用于室内机器人的实用且注重隐私的选择。高端多线束 3D 激光雷达能够支持丰富的感知能力，但通常会增加硬件成本、体积、功耗和集成复杂度，因此往往不适用于低成本服务平台。在相比之下，2D 激光雷达广泛可用且价格低廉，但其稀疏的测量数据使多类别目标检测比基于摄像头的检测更具挑战性。

近期基于学习的方法表明，二维激光雷达能够支持语义感知 [3]。然而，仍存在一个关键难题：如何在嵌入式硬件上实现准确的多类别检测，同时保持较低的端到端延迟。特别是，将扫描转换为稠密占据栅格的处理流程会增加预处理开销和输入尺寸，从而可能限制实时性能。

本文提出了一种无需摄像头的二维激光雷达目标检测处理流程，通过紧凑的多帧表示对短期时序上下文进行编码。我们将连续三帧二维扫描栅格化为二值图像，并将其堆叠为 RGB 通道，形成与 YOLOv8n 兼容的张量。该方法保留了扫描的角度结构，并提供短期运动线索，同时避免了占据栅格构建。

主要贡献如下：

1）一种紧凑的多帧二维激光雷达编码方法，将连续三帧扫描堆叠为 RGB 通道，无需位姿对齐即可注入短期时间线索；2）一种基于 YOLOv8n 的纯激光雷达检测器，并在仿真中通过严格的场景级留出验证进行评估；3）一种端到端嵌入式实现，展示了其在低成本硬件上的实时性能；4）与占据栅格激光雷达–YOLO基线进行比较，揭示了所提编码在延迟和表示成本方面的影响。

本文其余部分安排如下：第 II 节回顾相关工作；第 III 节介绍所提出的方法；第 IV 节展示实验结果；第 V 节总结全文。

II. 相关工作

早期面向移动机器人的二维激光雷达感知通常依赖聚类等几何预处理 [4] 和手工设计的规则 [5]。这些方法高效且易于解释，但需要仔细调参，并且难以扩展到多样化的环境和更丰富的目标类别。

基于学习的方法表明，当保留扫描的有序角度结构时，二维激光雷达可支持语义感知——

补充材料（代码、训练权重和演示视频）可在以下地址获取：https://github.com/soheilbh/2d-lidar-identification 和 https://www.youtube.com/watch?v=sjO1z04g8Jg 本文已提交至 IEEE，供其考虑发表。版权可能在不另行通知的情况下转让，届时此版本可能不再可访问。

时间上下文也已被证明能够提升稀疏距离数据的鲁棒性。例如，DROW 和 DR-SPAAM 会融合多个连续扫描，从而在遮挡和视野不完整的情况下稳定检测，而无需进行昂贵的扫描匹配 [6], [7]。不过，这些系统通常只关注一小类目标（往往是行人），并未直接解决采用边界框的 YOLO 式多类别目标检测问题。

将 CNN 检测器应用于二维激光雷达的一种策略，是将扫描转换为密集的占据栅格图像。Najem 等人（2025）采用了这种方法，通过栅格化扫描并应用 YOLOv8n 进行检测，但他们报告称，该方法在 Raspberry Pi 5 上的端到端延迟较高，凸显了栅格构建中的预处理和输入尺寸所带来的开销 [3]。相比之下，其他研究使用轻量级模型（如 2DLaserNet [8]和 TinyLidarNet [9], ）以更接近原生表示的方式处理扫描，表明紧凑型架构能够在受限硬件上运行。不过，这些方法并非为图像式（YOLO）框架下的多类别二维目标检测而设计。

与此同时，三维激光雷达检测通常采用鸟瞰图式编码，并能取得出色的结果，但往往要求使用昂贵的多线束传感器和更高的计算资源，这通常与低成本室内服务机器人的需求不相容。最后，室内环境中的隐私限制促使人们采用避免采集 RGB 外观的感知方式；在此背景下，二维激光雷达提供纯几何测量，通常被视为比视频更有利于隐私保护。

我们的工作沿着这些方向展开，避免构建占据栅格，并引入最精简的时间编码（将三帧扫描堆叠为 RGB 通道），既能兼容快速 CNN 检测器，又能保持输入紧凑。

III. 方法

本节介绍所提出的纯二维激光雷达目标检测流程，包括输入表示、时间编码、数据集生成、训练配置和部署。

A. 系统概述

我们提出了一种完全基于二维激光雷达的高效目标检测流程，旨在嵌入式 CPU 上实时运行。与先构建占据栅格 [3], 的既有激光雷达–YOLO 流程相比，我们的方法直接将扫描编码为紧凑的 RGB风格张量，可由标准 CNN检测器处理。

核心思路是将连续三帧二维扫描堆叠为小型栅格图像的三个通道。这样既保留了扫描有序的角度结构，又能融入短期时间线索（运动和视角变化），而无需进行里程计对齐。由此得到的输入紧凑（$64 \times 384 \times 3$），并避免了占据栅格流程通常需要的预处理成本和较大的输入尺寸（通常会调整至 640×640）。

算法 1 总结了端到端运行时流程。

算法 1 在线二维激光雷达推理流程

输入： 时间 t 的原始二维激光雷达扫描
状态： FIFO缓冲区 $\mathcal{B}$，保存最近 3 个栅格图
输出： 时间 t 的边界框和类别分数
将扫描栅格化并填充至 64×384
使用最新栅格图更新 $\mathcal{B}$ （保留 3 个）
将 $\mathcal{B}$ 堆叠为RGB张量（$64 \times 384 \times 3$）
在RGB张量上运行YOLOv8n推理
返回检测结果（箱子和类别分数）

B. 时序数据编码

二维激光雷达扫描是由角度–距离测量值 $\{(\theta_i,\ \rho_i)\}_{i=1}^{N}$组成的序列，其中 θ_i 是方位角，ρ_i 是测量距离。

栅格化。 每次扫描被转换为稀疏二值栅格$\mathbf{I}_t \in \{0,255\}^{64\times360}$，其中列表示角度分箱，行表示离散化的距离分箱。我们使用360个角度分箱（每个分箱1度）。角度分箱索引的计算方式为

$$a_b \;=\; \mathrm{mod}(\mathrm{round}(\theta\,[^\circ]),\,360)\,, \tag{1}$$

即，将角度四舍五入到最接近的度数，并按360取模进行回绕。

对于距离轴，我们将其离散为 64 个分箱，并采用固定的满量程值 $R_{\max} = 4.0$ m。模拟器传感器量程为 0.12–3.5 米；对于无穷大或超出量程的测量值，将其替换为哨兵值 $R_{\max}$ （此哨兵值仅用于统一处理无效回波，并不表示量程范围外的测量值）。距离分箱索引的计算方式如下:

$$r_b \;=\; \left\lfloor \mathrm{clip}\!\left(\rho \cdot \frac{63}{R_{\max}},\,0,\,63\right)\right\rfloor, \tag{2}$$

其中 ρ 以米为单位。对于每个测量值，将坐标为 $(r_b,\ a_b)$ 的像素设为 255，其余所有像素均保持为 0。

坐标轴约定。 在栅格中，水平轴（列）对应角度分箱 $a_b \in [0,359]$，垂直轴（行）对应距离分箱 $r_b \in [0,63]$。标签投影和边界框均采用这一约定。

填充以匹配 CNN 步长。 YOLO骨干网络通常要求空间维度可被 32 整除。为避免缩放扫描栅格（这会扭曲角度结构），我们在$\mathbf{I}_t$ 中填充 24 个零列，得到 $\mathbf{I}_t \in \{0,255\}^{64\times384}$。

三帧时序堆叠。 为编码短期时序上下文，我们维护一个先进先出（FIFO）缓冲区，其中存储了时间为 $t-2$、$t-1$ 和 t 的连续三帧栅格图。随后将它们堆叠为一个 RGB张量 $\mathbf{I}_t^{\mathrm{RGB}} \in \{0,255\}^{64\times384\times3}$:

$$\mathbf{I}_t^{\mathrm{RGB}}(r,a,:) \;=\; \big[\mathbf{I}_{t-2}(r,a),\,\mathbf{I}_{t-1}(r,a),\,\mathbf{I}_t(r,a)\big]. \tag{3}$$

这种编码无需扫描匹配或位姿对齐，即可注入运动线索和短期视点变化，同时保持预处理轻量。

图 1 展示了逐通道堆叠和一种简单的融合视图。

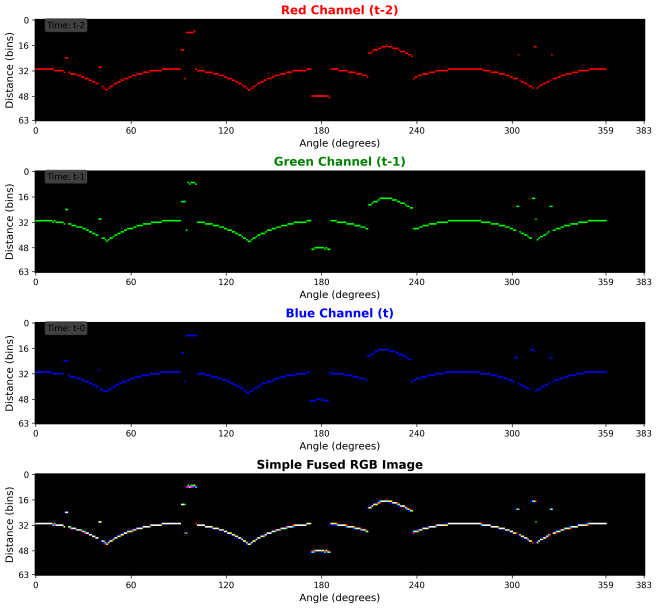

图 1. RGB 编码的激光雷达输入：将三个连续扫描栅格化并堆叠（红色：$t-2$，绿色：$t-1$，蓝色：t），形成供 YOLOv8n 使用的紧凑张量。

C. 替代编码方案探索

在设计迭代过程中，我们评估了一种更复杂的编码方案，尝试补偿扫描之间的机器人旋转。该方案使用 5 帧缓冲区，应用基于偏航角的列偏移，并将 (i) 命中计数、(ii) Sobel 边缘和 (iii) 点密度编码到三个通道中（如图 2 所示）。

虽然这使时序融合在视觉上更加清晰，但也增加了预处理成本和内存占用。在一项消融实验中（训练集/验证集划分及 YOLOv8n 设置相同），对齐的 5 帧编码并未提升平均精度均值，反而增加了预处理时间；因此，为了最大限度地提高端到端效率，我们在最终系统中采用未对齐的三帧堆叠。

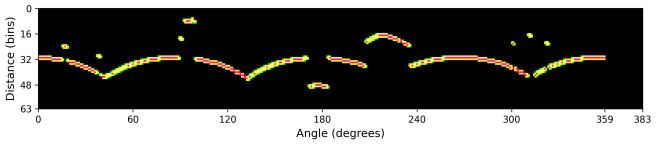

图 2. 使用 5 帧缓冲区探索的“对齐融合”编码可视化（命中计数、边缘和密度）。由于增加了成本却未带来明确的精度提升，因此最终处理流程未采用此方案。

D. 自动化数据集生成

我们使用 Webots R2023b 在仿真环境中生成训练数据。每次数据采集运行由每个场景包含的 N 个随机场景和 M 个预定义机器人位置组成，共产生 $N \times M$ 个采样回合。在每个位置，机器人执行一段短距离运动，我们记录多个连续扫描，从而为每个回合生成多个标注帧。图 3

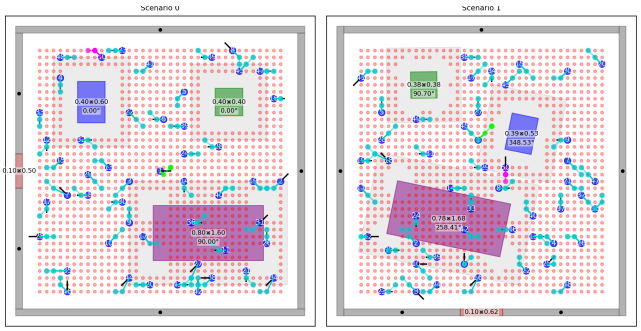

图 3. 包含 90 个预定义机器人位置的随机场景示例。物体类型以不同颜色的矩形表示（箱子：蓝色，椅子：绿色，书桌：紫色，门框：红色），机器人航点以蓝点表示。

展示了仿真场景示例，每个场景都包含预定义的机器人位置和随机化的物体配置（椅子、书桌、箱子、门框）。图 4 展示了仿真环境的三维视图。

为提升多样性，可移动 物体（椅子、书桌、箱子）的平面平移和偏航角均随机化，其平面占地尺寸按 $\pm 20\%$ 缩放（缩放系数从 $[0.8, 1.2]$中均匀采样）。在 Webots 中，椅子和书桌类别使用内置的 Table 模型（由四条腿支撑的矩形桌面）实例化；我们采用不同的标称占地尺寸（椅子：0.4×0.4 m，书桌：0.8×1.6 m）和腿部厚度设置（feetSize 0.03 m 与 0.05 m），椅子采用各向同性缩放，书桌则采用宽高独立缩放。箱子类别使用 WoodenBox 长方体模型，标称占地尺寸为 0.4×0.6 m，并采用宽高独立缩放。门框对应围墙中的门洞；在不同场景中，房间布局按 $90°$的整数倍旋转，因此门洞可能出现在任意一侧。我们使用以门洞中心为中心的薄矩形代理来标注门框，其厚度固定为 0.1 m，高度从 $[0.5, 1.0]$ m中均匀采样。每个训练输入均按上述方式叠加三个连续扫描。

标签生成。真实标签对应时间$t-1$（三帧缓冲区中的中间扫描帧）。利用模拟器元数据和 $t-1$时的机器人位姿，将目标角点变换到机器人坐标系，并投影到栅格坐标（距离 bin、角度 bin）。随后在归一化栅格坐标中生成轴对齐的 YOLO 格式边界框（x_c，y_c，w，h）。边界框会裁剪至角度边界$[0, 359]$；因此，跨越 $0°/359°$ 不连续点的目标会以裁剪后的边界框表示，而不是拆分成两个边界框。图5展示了一个 RGB 融合输入示例，以及为全部四种目标类别自动生成的边界框。

在所有回合中，该流程从 160 个场景生成了 768,897 个编码样本，每个场景在 90 个机器人位置进行评估。

E. 模型训练

我们使用 PyTorch 中的 Ultralytics 实现（v8.3.152）训练了 YOLOv8n 检测器。YOLOv8n 是 nano

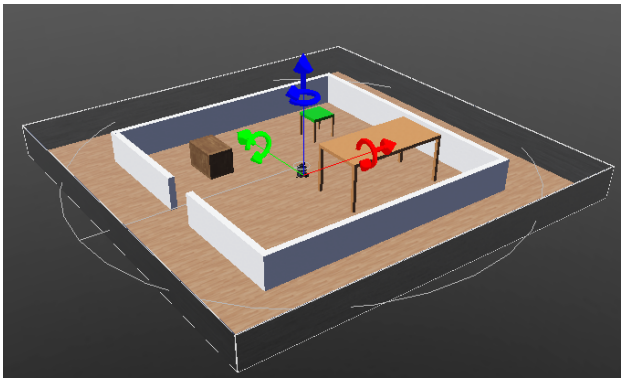

图4. Webots仿真环境的三维视图，展示了一个典型的室内场景，其中包含物体（箱子、桌子、椅子、门框）和房间结构。

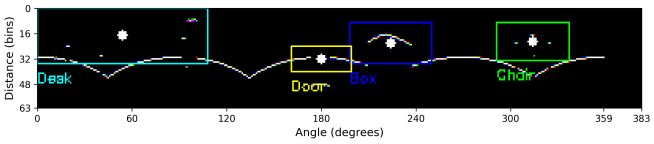

图5. RGB融合激光雷达输入示例，以及自动生成的四个目标类别的边界框。

版 YOLOv8 变体，因其嵌入式可行性而被选用（约 300 万参数）。输入分辨率固定为64×384，并启用矩形训练（rect=True），以保留原生激光雷达纵横比。

由于激光雷达栅格的几何结构受到限制，我们禁用强图像增强（马赛克增强、MixUp、CutMix、复制粘贴增强）。我们仅采用轻量级翻转：沿角度轴左右翻转（概率为50%），以及沿距离轴上下翻转（概率为20%）。对于固定的二维传感器，距离轴翻转在物理上无法实现；我们仅将其视为图像空间正则化项，并在第IV-D节中将其作为仿真到现实风险加以讨论。

训练在一台配备NVIDIA H100 GPU（80 GB显存）的服务器上进行。模型训练了120轮，批量大小为1,024（得益于GPU显存和数据集缓存）。优化采用AdamW，初始学习率为0.001，并使用余弦学习率调度，根据验证集平均精度均值进行提前停止。

为评估泛化能力，我们按场景级别划分数据集：85%的场景用于训练（629,591个样本），10%用于验证（72,528个样本），5%用于测试（38,321个样本）。这确保测试集中的空间布局和物体配置在训练期间未曾出现。

F. 部署

训练好的模型（.pt）已部署到多个平台，以验证其嵌入式可行性。在 Raspberry Pi 5 和 MacBook Air M2 上，推理通过 Ultralytics运行时运行。对于 Raspberry Pi 3，模型会导出为 TorchScript，以便在资源有限的设备上进行推理（功能可用，但无法实时运行）。

部署时，模型直接接收 $64 \times 384 \times 3$ RGB激光雷达张量，并在一次前向传播中输出类别分数和边界框，仅需执行上述轻量级编码步骤。该张量通过角度量化和距离量化（公式（1）–（2）），将每个原始极坐标回波（θ， ρ）映射到一个二维栅格索引（r_b， a_b）而构成，得到一个图像平面，其中列表示角度分箱，行表示距离分箱。

IV. 实验结果

本节在仿真中评估所提出的方法，报告检测准确率和嵌入式运行时间，并结合先前工作阐述其表示成本。

A. 实验设置

所有评估数据均在 MacBook Air M2 上使用 Webots R2023b 生成。我们创建了160个独特的随机室内场景，并在每个场景中的90个预定义机器人位置进行评估，共得到14,400个采样回合。如第III-D节所述，标签根据模拟器元数据（物体位姿和尺寸）以及机器人位姿自动生成，无需人工标注。按照第III-D节所述流程，这些回合生成了768,897个带标签的RGB张量输入。端到端仿真和标注约耗时18小时。

模拟二维激光雷达传感器（LDS-01）在 10Hz 下提供 360°覆盖范围。扫描数据被分配到 360 个角度区间进行栅格化。有效测量范围为 0.12–3.5 m；无穷远或超出范围的回波会被替换为 4.0 m，并映射到最远距离区间。

我们在代表不同计算资源预算的三个平台上评估运行时：

- **树莓派 5（8 GB RAM）：** 博通 BCM2712 四核 Cortex-A76 @ 2.4 GHz；
- **MacBook Air M2（8 GB 统一内存）：** Apple M2（8 核 CPU、10 核 GPU）；
- **Raspberry Pi 3 Model B+ （1 GB RAM）：** 博通 BCM2837B0 四核 Cortex-A53 @ 1.4 GHz（Torch-Script）。

B. 检测性能

在留出测试集（严格的场景级留出验证）上，模型达到 0.984 mAP@0.5 和 0.778 mAP@0.5:0.95，表明其在未见过的场景布局中能够可靠地进行检测和分类。

表 I 报告了各类别的精确率、召回率和平均精度均值。四个类别的表现始终出色。大多数错误出现在书桌 和门框之间，因为从某些视角观察，二维激光雷达可能会生成相似的局部几何轮廓。图 6 展示了召回率–置信度曲线和归一化混淆矩阵，进一步说明了这些性能特点。背景误检仍然非常少（$< 1\%$），表明类别区分能力很强。

表 I 测试集各类别检测指标

类别	精确率	召回率	mAP@0.5	mAP@0.5:0.95
椅子	0.986	0.913	0.981	0.746
Box	0.957	0.996	0.993	0.780
Desk	0.922	0.954	0.982	0.822
门框	0.930	0.926	0.980	0.763
全部（均值）	**0.949**	**0.947**	**0.984**	**0.778**

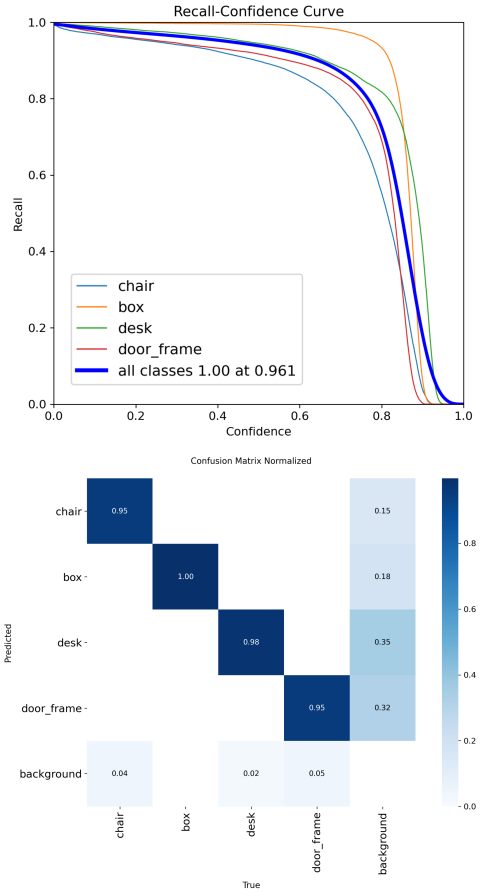

图 6. 上：召回率–置信度曲线。下：归一化混淆矩阵

C. 推理时间分析

我们在批量大小为 1、采用 FP32 推理的条件下测量端到端 CPU 延迟，其中包括扫描编码、YOLOv8n 前向传播和非极大值抑制（NMS）。对于每个平台，我们执行一次预热推理，并报告连续 100 帧的平均延迟。我们的目标是实现 < 50ms 的端到端延迟，以支持 ≈ 20 Hz 的感知更新，从而满足典型室内导航和控制回路的需求。

表 II 汇总了测得的运行时间。在树莓派 5 上，该处理流程的平均端到端延迟为每帧 47.8 ms，达到了实时目标，同时为其他机载模块（例如定位与规划）留出了计算余量。MacBook Air M2 的延迟显著更低，而 Raspberry Pi 3（TorchScript）虽仍可运行，但远未达到实时。

表 II 不同硬件平台上的推理时间

平台	平均时间 (ms)	推导帧率
MacBook Air M2	6.2	161.29
树莓派 5	47.8	20.92
Raspberry Pi 3 (TorchScript)	∼2000	∼0.50

推导帧率的计算方式为 1/平均延迟。计时包含扫描编码、推理和 NMS，并报告预热推理一次后连续 100 帧的平均值。

D. 局限性与有效性威胁

本研究在仿真中进行评估；真实的二维激光雷达传感器会引入噪声、缺失回波和校准效应，这些因素可能在域偏移下降低性能。标签来自模拟器元数据，从而消除了标注歧义，并可能使所报告的指标相对于人工标注的真实世界数据集有所偏高。该任务考虑了四个以静态为主的室内目标类别，并未评估动态智能体（例如人类）。时序堆叠在未进行位姿对齐的情况下完成，在机器人快速旋转或高度动态的场景中性能可能下降。最后，尽管二维激光雷达避免了RGB外观采集，几何信号在某些部署中仍可能支持行为推理；因此，我们将该方法描述为对隐私友好，而非隐私保护。

E. 与先前工作的比较

表 III 总结了 Najem 等人 [3]以及我们的方法所报告的结果。由于 Najem 等人使用带有人工标注的真实数据进行评估，而我们的结果基于仿真并采用自动标注，因此准确率指标无法直接比较。因此，我们着重比较各研究在各自设置下报告或测得的树莓派 5 上的表示成本和端到端延迟。这并非受控的直接基准测试；此比较旨在说明占据栅格构建与紧凑扫描编码对运行时的影响。

如表 III 所示，树莓派 5 上的占据栅格处理流程报告的端到端延迟显著高于我们的紧凑扫描编码。对此应谨慎解读，因为数据集、标注方案、软件栈和计时协议各不相同；尽管如此，这种差异方向与占据栅格构建带来的额外预处理成本以及更大的网络输入尺寸相符。

V. 结论与未来工作

本文提出了一种面向室内服务机器人的隐私友好型实时二维激光雷达目标检测处理流程。该方法将连续三次扫描栅格化并堆叠为 RGB 通道，以编码短期时序上下文，生成与 YOLOv8n 兼容的紧凑型 $64 \times 384 \times 3$ 输入，同时避免占据栅格构建带来的预处理开销。在 Webots 中采用严格的场景级留出验证时，该方法在四类室内目标上的 mAP@0.5 达到 98.4%（mAP@0.5:0.95 为 0.778），精确率为 94.9%，召回率为 94.7%。在树莓派 5 上，

表 III 与一种相近的二维激光雷达–YOLO 占据栅格流水线的比较

方法	表示	输入	数据条件	树莓派 5 耗时（毫秒，端到端）	推导帧率	mAP@0.5
Najem 等人 [3]	占据栅格	640×640	真实 + 人工标注	846.5	1.18	—
我们的方法	3帧RGB张量	$64 \times 384 \times 3$	模拟 + 自动标签	47.8	20.92	0.984

推导帧率值根据 1/报告时间计算。

完整处理流程可实时运行，预热后的平均端到端延迟为 47.8 毫秒/帧（包括编码和后处理），表明轻量级时间编码能够支持嵌入式硬件上的高精度纯激光雷达检测。

未来工作将使用真实二维激光雷达传感器，在实体机器人上评估仿真到现实迁移，包括对噪声、缺失回波和校准效应的鲁棒性。我们还计划扩展目标集（包括动态智能体），研究快速旋转时的失效模式（此时未对齐的时序堆叠可能导致性能下降），并探索量化和硬件加速运行时等部署优化，以进一步降低延迟和功耗。

参考文献

[1] C.-Y. Wang、H.-Y. M. Liao 和 I.-H. Yeh，“YOLOv9：利用可编程梯度信息学习所需内容”，载于欧洲计算机视觉会议（*ECCV*）论文集，2024年。[2] N. Lintvedt，“机器人学中的热成像：增强隐私或侵犯隐私的技术——系统综述”，《国际社会机器人学杂志》，第15卷，第1629–1643页，2023年。[3] A. Najem、L. Kuiper、T. P. Jansen、A. S. Brand¯ao 和 F. N. Martins，“利用二维激光雷达和YOLO进行目标分类以实现机器人导航”，载于《优化、学习算法与应用：第六届国际会议 *OL2 30，2025，2025*年*4*月，意大利塞斯特里莱万特——修订版精选论文集》，《计算机与信息科学通信》丛书，第2618卷。查姆：施普林格，2025年，第246–260页。 [在线]。获取地址：https://link.springer.com/chapter/10.1007/ 978-3-032-00140-5 17[4] A. Sec.kin，“利用二维激光雷达检测行人和移动机器人”，《欧洲科学与技术杂志》，第23卷，¯ 第583–588页，2021年。[5] L. A. Fagundes、A. G. Caldeira、M. B. Quemelli、F. N. Martins 和 A. S. Brand¯ao，“二维激光雷达数据表示与目标检测的解析形式体系：在移动机器人学中的应用”，《传感器》，第24卷，第7期，2284页，2024年。[6] L. Beyer、A. Hermans、T. Linder、K. O. Arras 和 B. Leibe，“基于二维距离数据的深度行人检测”，*IEEE RoboticsandAutomation Letters*，第3卷，第3期，第2726–2733页，2018年。[7] D. Jia、A. Hermans 和 B. Leibe，“DR-SPAAM：用于二维距离数据行人检测的空间注意力与自回归模型”，arXiv预印本 arXiv:2004.14079，2020年，arXiv:2004.14079。[8] B. Kaleci、K. Turgut 和 H. Dutagaci，“2DLaserNet：一种基于二维激光扫描、用于移动机器人位置语义分类的深度学习架构”，《工程科学与技术：国际期刊》，第28卷，101027页，2021年。[9] M. M. Zarrar、Q. Weng、B. Yerjan、A. Soyyigit 和 H. Yun，“TinyLi-darNet：用于F1TENTH自动驾驶赛车的基于二维激光雷达的端到端深度学习模型”，2024年。