

综述

深度学习在调频连续波雷达里程计和自定位中的应用方法综述

Marvin Brune * , 托比亚斯·梅森  和 安德烈·庞普 

数字转型技术与管理部, 沃普塔尔大学, 高斯街20号,
德国沃普塔尔42119; meisen@uni-wuppertal.de (T.M.); pomp@uni-wuppertal.de (A.P.)

* 通讯作者: marvin.brune@uni-wuppertal.de

摘要: 本文深入综述了深度学习技术在解决基于调频连续波 (FMCW) 雷达传感器的里程计与全局自定位挑战方面的应用。特别地, 我们关注里程计的预测, 即通过外部传感器确定系统的自运动, 以及回环检测, 即专注于在现有地图上确定自身位置。我们首先强调了这些任务在雷达传感器背景下的重要性, 并阐述了其背后的动机。后续章节深入探讨了深度学习方法的实际应用, 这些方法策略性地设计用于有效应对上述挑战。我们主要关注自动驾驶领域内的旋转和汽车雷达配置。此外, 我们介绍了在解决这些挑战中发挥关键作用的部分公开数据集, 并分析了当前用于基于雷达的里程计与定位方法的必要性和挑战。最后, 本文突出了所讨论任务与其他雷达感知应用的差异, 同时讨论了它们与替代传感器模式所提出挑战的不同之处。这些发现为通过深度学习方法提升雷达传感器能力的持续讨论做出了贡献, 特别是在增强自动驾驶应用中的里程计和自定位方面。

关键词: 雷达; 同步定位与地图构建 (SLAM); 自动驾驶; 自定位; 里程计



引用: Brune, M.; Meisen, T.;
Pomp, A. 深度学习在调频连续波雷达
里程计和自定位中的应用方法综述。
应用科学 2024, 14, 2267。
<https://doi.org/10.3390/app14062267>

学术编辑: Atsushi Mase

收到: 2024年1月4日 修改:
2024年2月28日 接受:
2024年2月29日 发布:
2024年3月8日



版权: © 2024 由作者所有。许可方:
MDPI, 瑞士巴塞尔。本文是开放获取
文章, 根据知识共享署名 (CC BY)
许可协议 (<https://creativecommons.org/licenses/by/4.0/>) 的
条款和条件分发。

1. 引言

同步定位与地图构建 (SLAM) 是机器人和自主系统领域的一项基础性技术。其主要功能在于使自主系统能够不仅构建周围环境的完整地图, 还能准确确定自身在该地图中的位置。任何 SLAM 算法的核心都包含两个关键概念: 里程计, 它预测系统的自运动; 以及自定位, 即利用外部传感器数据确定其在地图中的位置。

被动和主动感知传感器, 例如相机、激光雷达、声纳和雷达, 通常应用于SLAM方法中 [1]。这些传感器具有从观测中获取精度的优势。可以利用这些传感器构建地图, 以获得对环境的空间理解。在理想情况下, 来自主动传感器 (如激光雷达、声纳和雷达) 的信息可以表示测量的绝对精度, 而系统的里程计和自定位可以仅从此信息中确定。然而, 由于各种因素, 检测结果并不总是始终一致或准确 [2]。被动位置传感器, 如全球导航卫星系统 (GNSS) 或惯性导航系统 (INS), 可以提供有关系统位置和动态的额外信息, 但无法单独捕获空间环境。此外, GNSS传感器由于信号接收不良, 在常见应用区域 (如室内场景、隧道以及高楼之间) 性能较差 [3]。在这些场景中, 基于感知的自定位更受青睐。

GNSS和INS等位置传感器，除了感知传感器（稍后将讨论）之外，通常也包含在自动驾驶车辆应用中。

基于感知的SLAM方法可以根据其依赖的传感器类型进行分类，主要区别在于主动距离测量和被动视觉方法。视觉SLAM（VSLAM）方法特别使用车载相机来描绘其周围环境。尽管对VSLAM技术进行了大量研究，但它们在恶劣天气条件（如大雨或雾）以及变化的光照情况（如强烈阳光或夜间）中会遇到困难。这些挑战扩展到关于精确深度感知 [4,5] 的问题。

基于距离的技术利用激光雷达、声纳和雷达等传感器进行操作。声纳主要在水下场景 [6,7] 中使用，而激光雷达和雷达主要在室内和室外应用中使用，特别是在自动驾驶车辆领域。虽然激光雷达提供更密集的检测能力，但雷达也带来其独特的优势。与雷达相比，激光雷达通常更昂贵，并且在恶劣天气条件下通常不太耐用。尽管存在利用雷达和激光雷达进行SLAM任务的混合方法，但文献和行业的重点仍然在于使用单一的距离传感器 [5,8]。在实践中，基于距离的传感器通常与视觉传感器集成在一起，以提供深度信息并协助视觉障碍场景。将多个传感器集成到完整的SLAM中的另一个动机是避免使用单一感知传感器时出现的遮挡问题。当物体被另一个物体遮挡，或者物体移出某个传感器的帧时，另一个传感器可能能够发现该物体。雷达对于此类任务特别相关，因为雷达可以检测到在激光雷达或视觉数据中看似被遮挡的物体 [9]。然而，雷达数据非常容易受到“幽灵检测” [9] 的影响，即雷达检测到场景中不存在的物体。这些检测主要是由多径传播 [10] 效应引起的。与计算机视觉中的遮挡问题类似，深度神经网络可以学习一个掩码来过滤掉无关的物体和区域。额外的传感器，如激光雷达或相机，也可以向系统提供额外的空间信息。

SLAM方法的分类框架于1995年通过Durrant-Whyte等人的开创性工作引入， [11,12]。在基于感知的SLAM方法领域，一个核心概念涉及手动设计特征以促进图像或点云表示中关键点的对齐。然而，与自然语言处理 [13] 和计算机视觉 [14]，等各个其他领域所见证的进展类似，深度学习方法已被开发用于解决SLAM方法所面临的挑战。

在计算机视觉领域，深度学习已被全面应用于解决SLAM问题的各种问题，包括关键点检测、自运动估计、运动结构恢复和自定位 [15–18]。针对激光雷达数据已开发出类似方法，这些方法基于点云表示而非网格表示 [19–21]。深度学习架构类型包括从标准空间和中间层（如卷积神经网络（CNNs） [22] 和循环神经网络（RNN），如长短期记忆层（LSTMs） [23] 和门控循环单元（GRUs） [24]，）到复杂的图神经网络（GNNs） [25] 和Transformer架构 [26] 等一系列模块。其他优化方法，如制定适当的损失函数、多任务学习和优化学习技术，已被全面研究用于相机和激光雷达。确切的架构和训练流程取决于可用数据以及系统中存在的传感器 [27–29]。

鉴于雷达作为一种稳健、经济高效且可靠的感知传感器日益凸显的重要性， [5,9,30]，有必要对专门针对基于雷达的SLAM方法论的深度学习方法进行全面综述。

我们的主要贡献如下：

1. 我们介绍了SLAM问题和雷达的公式化方法。
2. 我们深入研究了雷达里程计和定位中最新采用的深度学习技术。
3. 我们探讨了将雷达数据集成到深度学习模块中的当前方法，以增强视觉/激光雷达里程计和自定位精度。

图1概述了本研究中提出的综述论文。第2节介绍了基于传感器的里程计和自定位的总体问题。接下来，第3节参考了现有的综述论文，为我们的工作设定了背景。在第4节中，我们的注意力转向了基础雷达处理，探索了雷达数据的多样化设置和数据表示方法。第5节通过考察深度学习方法在基于雷达的里程计和自定位管道中的应用，整合了最新趋势。值得注意的是，我们扩展了讨论范围，不仅限于雷达数据，还包括融合和级联网络结构的集成，以应对完整可靠的SLAM方法的需求。我们还比较了模型的评估方法。最后，第6节提供了该领域的全面总结和未来潜在趋势的讨论。

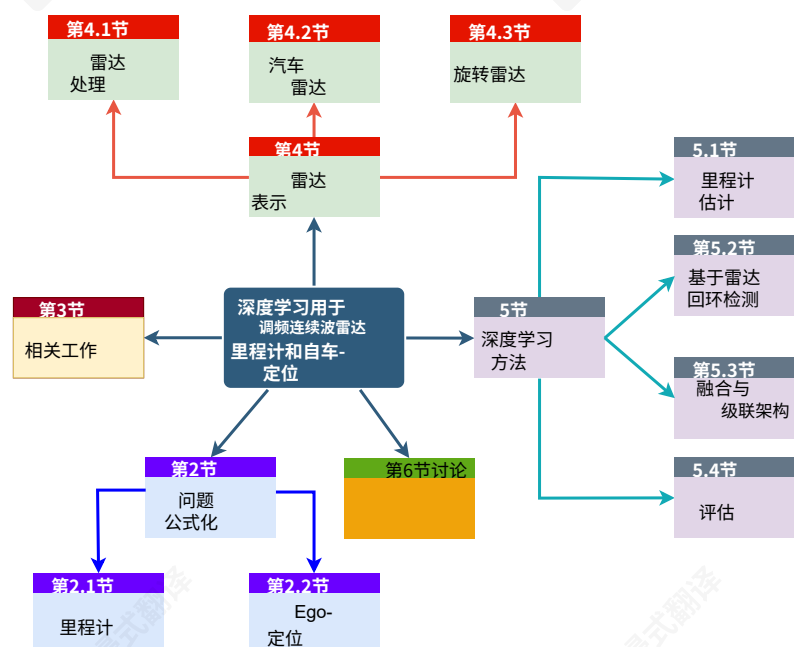


图1. 提出综述论文的概述。

2. 问题公式化

在本节中，我们介绍了里程计估计和基于雷达的SLAM中的自定位所涉及的基本概念和顺序阶段。此外，我们还重点介绍了经典雷达SLAM中采用的关键方法。

SLAM构成一个估计挑战，其目标是优化在时间1到t和全局地图M中记录的所有自定位姿态集合 $X_{1:t} = \{X_1, \dots, X_t\}$ 以及n个地标位置 $M = m_{1:n} = \{m_1, \dots, m_n\}$ 。这种优化通过利用包含感知 $Z_{1:t} = \{Z_1, \dots, Z_t\}$ 、收集到的里程计读数 $U_{1:t} = \{u_1, \dots, u_t\}$ 和初始起点 X_0 的观测数据来完成。从数学上讲，概率公式化由Grisetti等人 [31] 定义为：

$$p(X_{1:t}, M | Z_{1:t}, U_{1:t}, X_0) \quad (1)$$

为方便起见，从方程 (1) 得出的起始位置 X_0 从现在起将不再提及。自身位置和里程计测量通常使用 2D

或 3D 特殊欧几里得群变换矩阵 $T \in SE(2)$ 或 $T \in SE(3)$ 。此类变换矩阵由一个平移向量 t 和一个旋转矩阵 R 组成，该旋转矩阵可以用齐次坐标在方程 (2) 中表示为

$$T = \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \quad (2)$$

任何 SLAM 方法的根本目标都是确定轨迹 $X_{1:t}$ 和地标地图 M 的最大后验概率 (MAP) 状态。这一目标可以正式表示为:

$$\{X^*, M^*\} = \operatorname{argmax}_{\{X, M\}} p(X_{1:t}, M | Z_{1:t}, U_{1:t}) \quad (3)$$

求解公式 (3) 的常用方法包括假设静态全局地图并利用马尔可夫性质。这使得 SLAM 问题可以被概念化为一个动态贝叶斯网络 (DBN) [31]。在此框架内，初始位姿 X_0 、观测 $Z_{1:t}$ 和里程计估计 $U_{1:t}$ 被视为可观测变量。同时，模型的隐变量包括 $X_{1:t}$ (序列位姿) 和地图 M 。DBN 的结构由两个主要组件特征：一个表示新位姿取决于先前位姿和当前里程计测量的状态转移模型，记为 $p(X_t | X_{t-1}, U_t)$ ；第二个组件是观测模型，记为 $p(Z_t | X_t, M)$ ，表示观测依赖于当前位姿和整体地图配置。

一个贝叶斯网络动态 (DBN) 也可以使用图形表示来描绘，例如在 [32] 中，节点对应于位姿 X_t 。在这个图中，边作为空间约束连接位姿。这些约束可以基于位姿之间的里程计测量 U_t 建立，或者通过在不同位姿处获取的两个不同观测 Z_t 的对齐来建立。在实际实现中，观测边在图形表示中起着至关重要的作用，它们作为调节器。其目的是纠正仅使用里程计边可能产生的任何潜在误差累积或漂移。

图优化是一种普遍采用的技术，用于确定考虑扫描匹配多重结果的优化全局解。根据 SLAM 算法的具体特性，通过里程计和观测累积的边可能无法无缝对齐。这种差异可能由于各种因素引起，包括硬件和软件的复杂性。尽管存在这些潜在的不一致性，最终目标仍然是建立考虑所有观测和信息可用性的全局一致轨迹。典型的 SLAM 流程如图 2 所示。

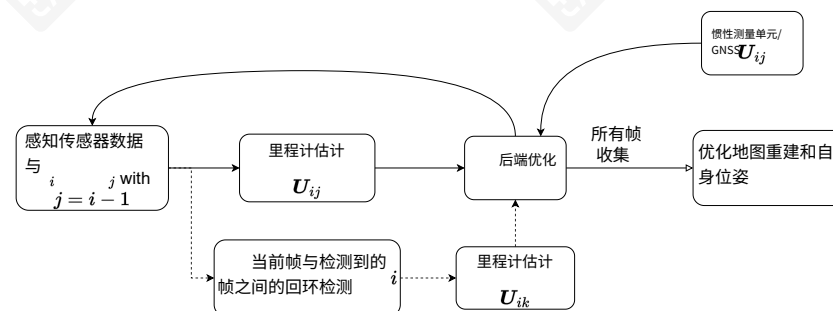


图2。SLAM方法的典型设置。传感器数据用于估计帧之间的变换矩阵。如果检测到回环，则会在后端优化方案中添加检测帧和当前帧之间的变换矩阵。最后，在后端优化过程后，会生成优化的地图和自身位姿。惯性测量单元/全球导航卫星系统数据可以作为额外的边插入后端优化中。

2.1. 里程计

各种传感器可用于在图表示中建立里程计边。惯性测量单元 (IMU) 通过利用内部传感器如陀螺仪、加速度计和磁力计来实现这一点。类似地, GNSS传感器如aGPS、差分GPS (DGPS) 或集成GPS的INS也可以通过在两个不同位置计算出的姿态差异提供里程计测量。然而, 被动IMU和GNSS传感器缺乏环境上下文, 因此容易受到外部因素的影响, 例如信号干扰或错误读数, 例如由于车辆卡住而检测到的车轮运动, 尽管IMU检测到了车轮运动。

另一方面, 基于扫描的里程计方法通过分析雷达、激光雷达和/或视觉等感知传感器获取的环境数据来推算车辆的自运动。与惯性测量单元预测的里程计不同, 基于传感器的里程计通过对齐连续扫描来生成里程计估计, 从而得到一个基于外部环境的、一致的里程计估计。然而, 依赖外部传感器会带来其自身的挑战, 具体估计的准确性可能因所用传感器的不同而有所差异。在实际应用中, 惯性测量单元/全球导航卫星系统数据与扫描匹配结果通常采用卡尔曼滤波方法进行融合 [9]。在此需要指出的是, 由于机械老化、环境变化或温度变化等多种因素, 任何惯性测量单元都需要定期重新校准 [33]。因此, 在SLAM方法中, 其他传感器的里程计估计至关重要, 能够实现自运动的精确预测。

在基于雷达的扫描匹配领域, 一种常见的技巧涉及基于距离的特征匹配 [9]。雷达数据通常以点云或二维鸟瞰图 (BEV) 图像的形式表示。目标是使连续的扫描对齐, 以识别最可能的变换。虽然深度学习方法可用于确定连续雷达扫描之间的自车变换, 但传统技术仍广泛使用。这些传统方法涉及使用手动设计的特征。例如, 用于视觉数据的定向FAST和旋转BRIEF (ORB) -特征 [34] 以及尺度不变特征变换 (SIFT) -特征 [35], 以及为雷达数据量身定制的保守滤波高效精确雷达里程计 (CFEAR) [36] 特征。这些特征表示为点检测。扫描匹配的过程涉及在单个扫描中识别这些特征, 并在连续扫描之间进行匹配。这通过迭代最近点 (ICP) [37,38]方法实现, 该方法通常与随机样本一致性 (RANSAC) [39] 技术结合使用, 以丢弃异常值。

对于激光扫描观测, 通常采用正态分布变换 (NDT) [40] 等替代方法来计算连续扫描之间的自运动变换矩阵。为了精确对齐传感器测量数据, 必须仅包含相关的静止特征。将非静止特征引入ICP过程会扭曲车辆的自身运动估计, 导致结果不准确 [41]。因此, 过滤掉来自无关特征的检测至关重要。虽然这个问题已针对相机和激光雷达传感器得到了广泛解决, 但由于捕获了多普勒特征, 雷达传感器在很大程度上避免了这一问题。尽管多普勒分量仅提供特征的径向速度, 但可以用来推算检测的真实速度。这允许在应用基于ICP的方法之前移除非静止特征, 从而确保更精确的结果 [41,42]。

2.2. 自定位

里程计估计在精确地映射环境连续扫描中起着至关重要的作用。然而, 无论里程计测量的精度和使用的传感器如何, 随着时间的推移, 总会不可避免地出现全局漂移。这种漂移会导致生成的地图出现扭曲, 从而导致全局姿态估计不准确 [9]。为了减轻这种漂移问题, 采用自定位的形式进行回环检测, 以识别之前访问过的位置。当检测到已经访问过的位置时, 会在当前扫描和过去扫描之间进行扫描匹配程序, 或

地图，从而导致根据匹配结果调整当前姿态。在图优化方法中，回环关闭变换被作为附加边包含在图结构中。

在检测潜在回环检测中，会采用多种技术。一些方法使用位置传感器（如GPS），而另一些方法将新扫描与存储扫描集中的扫描进行匹配，以确定是否将边添加到图优化中。一种常见的方法是使用词袋模型（BoW）方法 [43]。这种技术在高效比较感知扫描方面非常有价值，而感知扫描是SLAM和类似系统中回环检测的关键方面。BoW技术将每个扫描视为视觉词的集合，简洁地表示为单个向量。这些视觉词包含通过SIFT、SURF、ORB或训练网络提取的特征，如边缘、角点和纹理区域。

最初，通过从不同扫描中聚合特征并通过聚类分析（如k均值聚类）进行组织，来构建字典。每个聚类中心随后作为不同的“视觉词”。当记录新扫描时，其特征被提取并与字典中的视觉词或聚类中心进行匹配。这导致扫描被表示为一个向量，展示了扫描中每个视觉词的频率。在自定位过程中，使用BoW来识别当前扫描与已知场景之间的相似性，从而估计相对于先前访问位置的位置。如果当前向量与字典中向量的相似性超过先前定义的阈值，则在这两个扫描之间执行扫描匹配，以确定变换矩阵，并将边添加到图中。然而，这些边必须仔细选择，因为错误的边会导致生成的地图严重退化 [11,12]。

3. 相关工作

大量综述论文深入探讨了各种扫描匹配和里程计估计方法，涵盖了不同的传感器和方法类型，如表1所示。这些综述论文是通过结合“Survey”或“Review”等任务或传感器特定术语（如“SLAM”、“VSLAM”、“激光雷达”、“Survey”、“自定位”、“定位”、“深度学习”、“雷达”、“里程计”）使用谷歌学术和Web of Science找到的。此外，通过在找到的综述上进行逆向搜索，发现了更多论文。

虽然有许多专门针对SLAM中的里程计和定位的综述论文，但没有一篇专门深入探讨深度学习在雷达模态中的应用。Louback da Silva Lubanco等人最近的一项贡献 [44] 为经典雷达里程计的基础理解奠定了基础。

Geng等人 [45] 讨论了深度学习在雷达感知中的应用进展，尽管他们的重点并不特别集中于里程计或位置识别。相反，姚等人 [4] 进行的综述专门围绕雷达-视觉融合用于目标检测和分割。

周等人 [9] 和Harlow等人 [5] 撰写的综述论文深入探讨了雷达预处理方案，并详细概述了自动驾驶领域雷达感知任务中的前沿方法。值得注意的是，这些综述主要集中在对目标检测和分割的雷达感知。它们仅简要提及了少数基于深度学习的里程计估计和定位方法。

表1. 相关综述论文概述。主要传感器：R、L和C分别代表雷达、激光雷达和相机传感器。✓表示综述包含SLAM组件，(✓)表示其部分包含，×表示其不包含。

论文	Year	主要传感器	深度学习重点	里程计	定位
[44]	2022	R	×	✓	×
[46]	2018	R,L,C,GPS	×	×	✓
[47]	2021	L,C	✓	×	✓
[48]	2021	L,C	✓	×	✓
[49]	2022	R,L,C	✓	×	✓
[29]	2023	L,C	✓	✓	✓
[45]	2021	R	✓	×	×
[4]	2023	R,C	✓	×	×
[50]	2022	L,C	×	✓	✓
[51]	2022	L,C	✓	✓	✓
[52]	2021	L	✓	✓	✓
[53]	2021	L	×	✓	✓
[54]	2022	C	✓	✓	✓
[55]	2015	C	×	×	✓
[56]	2020	C	✓	✓	×
[57]	2019	C	✓	✓	✓
[58]	2023	L,C	(✓)	✓	✓
[8]	2020	L,C	✓	✓	✓
[27]	2023	C	✓	✓	✓
[59]	2023	C	✓	✓	✓
[28]	2023	C	✓	✓	✓
[9]	2022	R	(✓)	✓	✓
[5]	2023	R	(✓)	✓	✓
Ours	2023	R	✓	✓	✓

Saleem等人 [29] 的研究与我们的综述论文高度契合，因为这些作者深入探讨了SLAM问题的基本原理，并提出了大量应用于激光雷达和视觉传感器的深度学习技术。然而值得注意的是，他们对雷达的覆盖范围仅限于与激光雷达或视觉传感器的潜在融合场景。

除了这些关注雷达的综述论文，许多其他综述论文则关注基于视觉或激光雷达的SLAM方法。Kuutti等人 [46] 专注于非学习型定位方法，适用于各种感知传感器。在此基础上，Arshad等人 [47]研究了深度学习技术如何促进激光雷达和VSLAM中的回环检测。转换焦点，Roy等人 [48] 专门研究针对室内定位和嵌入式系统的深度学习技术。此外，Yin等人 [49] 全面覆盖了跨传感器模态和应用领域的地点识别与回环检测。

在传感器融合领域，徐等人 [50] 强调了基于激光雷达的传感器与其他模态的集成，主要利用视觉特征。类似地，Chghaf等人 [51] 专注于激光雷达和VSLAM，简要提及了雷达与其他模态的融合。

黄等人 [52] 和Khan等人 [53] 的综述主要集中在对激光雷达-SLAM的研究。在VSLAM的背景下, 包括 [54–57,60], 在内的一些综合性综述论文提供了详细的概述, 但并未提及雷达传感器。

Placed 等人的工作中探讨了深度学习在不同传感器模态上的应用, 而 [58] 和 Chen 等人的研究则聚焦于激光雷达和 VSLAM, 特别强调信念空间规划和深度强化学习。在相关背景下, Zhang 等人 [27] 和 Favorskaya 等人 [59] 深入研究了深度学习在 VSLAM 方法中的作用, 并调查了其在基础 SLAM 任务 (如位姿优化和地图构建) 中的应用。

Mokssit 等人 [28] 专注于 VSLAM 方法的特定阶段, 包括深度估计、光流、里程计、回环检测和端到端学习。此外, Zeng 等人 [61] 的综述探讨了 Transformer 架构在点云表示中的 VSLAM 上下文应用。

4. 雷达表示

在本节中, 我们关注雷达处理以及雷达数据可以表示的多种方式。在深入雷达标准信号处理时, 我们重点关注自动驾驶车辆最常用的两种雷达配置: 汽车雷达和旋转雷达。

4.1. 雷达处理

雷达传感器自20世纪30年代以来已在各种应用中发挥作用, 并因其鲁棒性和相对较低的成本而特别适用于感知任务。这种偏好源于其鲁棒性和相对较低的成本 [5,9]。

雷达传感器通过发射和接收调频连续波 (FMCW) 工作, 通常在以24 GHz或77 GHz为中心的频率带内, 频率呈线性增加。为此, 采用了一个频率混频器, 将发射和接收信号组合起来以管理高速采样要求。随后, 应用低通滤波器消除求和分量, 从而得到一个中频信号 (IF)。这个IF信号可以被模数转换器 (ADC) 捕获, 产生一个离散时间复数 ADC信号。确定物体距离的过程涉及分析拍频, 这是由于发射和接收之间的时间差而产生的 [9]。

同时, 通过检查IF信号中的相位偏移来确定多普勒速度。这种相位偏移是由于连续生成的FMCW波 (也称为线性调频脉冲) 而产生的。单个线性调频脉冲内的采样点维度称为快时间, 单个线性调频脉冲索引在单个帧内多个线性调频脉冲中的维度称为慢时间。在以下假设下, 即由于帧时间短, 目标运动在慢时间中引起的距离变化可以忽略不计, 并且通过利用大频率调制来忽略快时间中的多普勒频率, 距离和多普勒计算可以解耦。该过程的一个关键成果是通过在快时间维度和慢时间维度上执行两个离散傅里叶变换获得的一个复数值数据矩阵。称为距离-多普勒 (RD) 地图, 该矩阵描绘了对应于特定距离和多普勒值的功率读数。该地图提供了有关雷达范围内检测到的物体存在和运动的宝贵信息 [5]。

每部调频连续波雷达都采用所概述的方案来计算距离和多普勒分量。然而, 在确定目标到达方向 (DOA) 时, 技术会分道扬镳。这种区别取决于雷达的定位和配置, 导致出现两种主要的方法: 汽车雷达 (AR) 和旋转雷达 (SR)。图3说明了每种雷达传感器配置如何在车辆上布置。

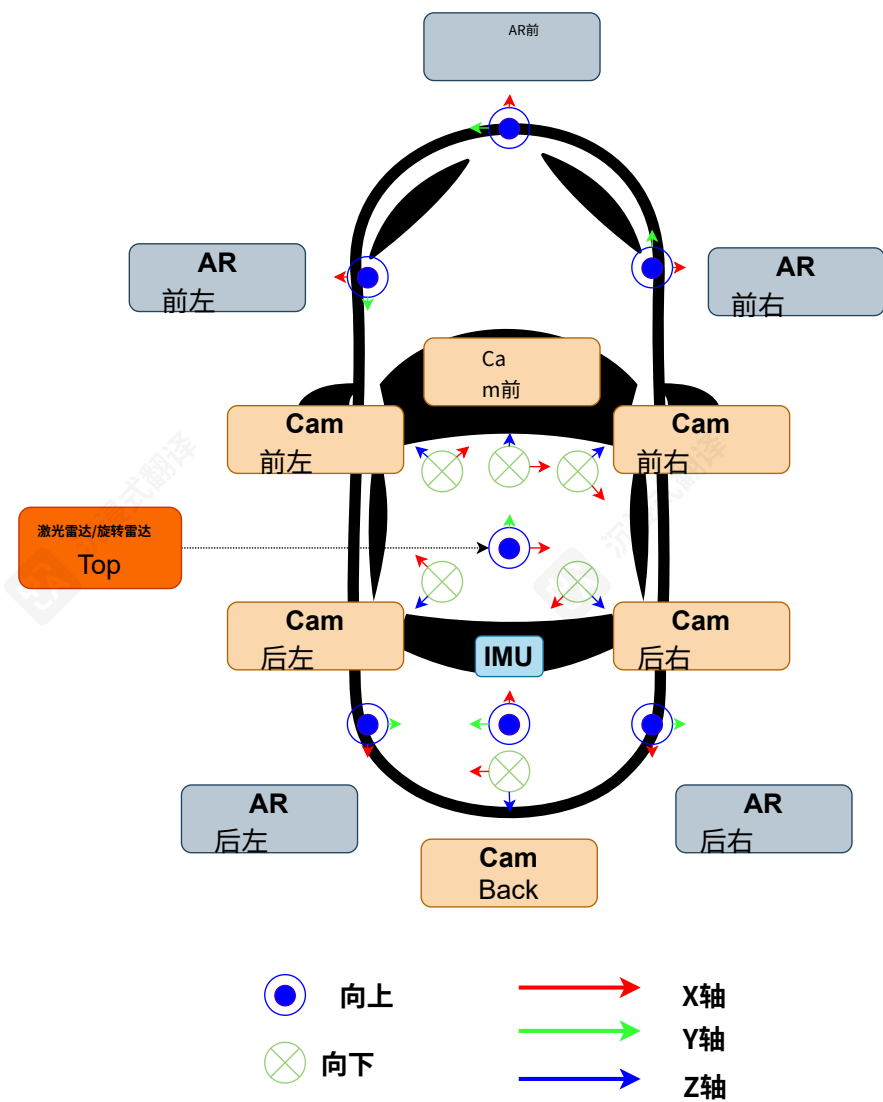


图3。汽车雷达传感器简称AR，旋转雷达传感器简称SR。单个旋转雷达传感器安装在车辆顶部，而在汽车雷达配置中，多个雷达传感器则安装在车辆前方、角点或侧面。

旋转雷达配置的优势在于，它们仅由一个传感器组成，安装在车辆顶部。使用一个外部校准矩阵将目标投影到车辆参考系中。对于汽车雷达配置，雷达传感器（通常为4-8个小芯片）被放置在车辆周围的不同位置。在此，传感器记录的信息需要转换到同一坐标系中，通常转换到车辆坐标系，以匹配不同传感器记录的相同检测结果。因此，在处理AR配置时，点云表示形式更受欢迎 [5,9]。

表2概述了目前用于雷达传感器进行自运动或自定位目的而明确采用的开源雷达数据集。这些数据集通过GitHub仓库 [62]，发现，并补充了第5节中识别的其他数据集。

表 2. 应用于深度学习 SLAM 方法的雷达数据集。类型：SR 和 AR 分别代表旋转雷达和汽车雷达。场景：U、S、H、P、I、T、W 分别代表城市、郊区、高速公路、停车场、室内、隧道和水域。其他传感器：C、L、O、S 分别代表相机、激光雷达、内部里程计传感器和合成孔径雷达（SAR）。✓ 数据集包含多普勒测量，而 × 则不包含。

Name	Type	数据表示	多普勒	场景	其他传感器
牛津 [63]	SR	RA	×	U	CLO
Boreas [64]	SR	RA	×	S	CLO
MulRan [65]	SR	RA	×	US	LO
RADIATE [66]	SR	RA	×	USHP	LO
Nuscenes [67]	AR	PC	✓	USH	CLO
View-of-Delft [68]	AR	4D PC	✓	USH	CLO
ColoRadar [69]	AR	模数转换器, 个人计算机	✓	SIT	CLO
USVInland [70]	AR	PC	✓	W	CLO
Zendar [71]	AR	模数转换器, RAD, 个人计算机	✓	U	CLOS
欧盟长期 [72]	AR	PC	✓	U	CLO
Astryx [73]	AR	3D PC	✓	SH	CLO
K-Radar [74]	AR	RAD, 4D PC	✓	USHP	CLO
径向 [75]	AR	ADC	✓	USH	CLO

虽然存在关于IF信号、距离、多普勒速度和DOA等变量的数学表达式和推导，但这些细节在此不涵盖。此遗漏是因为基于雷达的里程计和自定位方法通常在恒虚警率（CFAR）处理之后应用，并且已在其他作品中广泛涵盖 [5,9]。此外，原始数据表示在当前公开研究的数据集中很少包含，导致模型依赖于每个受尊重数据集的预处理。

在此，还需要注意的是雷达数据的预处理是一个活跃的研究领域。雷达传感器的一个进一步问题是互相关，尤其是在考虑AR配置时。随着雷达传感器数量的增加，雷达之间可能会相互干扰，从而可能产生幽灵目标检测并降低信噪比 [76]。这可能是由于同一车辆的雷达，甚至可能是来自不同车辆的雷达传感器造成的。存在减少传感器之间干扰的方法 [77]，但这些方法很少在公共数据集中讨论。

重要的是要强调自动驾驶车辆中提到的雷达配置设置之间的差异，因为第5节中发现的许多方法只适用于其中一种或另一种雷达配置。这些差异将在下文中讨论。

4.2. 汽车雷达

汽车雷达采用使用多个天线的策略，无论是单发多收（SIMO）配置还是多发多收（MIMO）配置，来确定到达方向（DOA）。这种确定依赖于评估接收天线之间的相位变化，并结合它们之间的物理空间分离。通过在接收器维度上执行角度快速傅里叶变换（FFT），可以推导出到达方向，从而得到一个名为范围-方位-多普勒（RAD）图的3D张量。该张量可以与之前讨论的范围-多普勒（RD）图连接，从而得到每个反射的方向和多普勒分量的全面表示。

传统上，仅解析方位角。然而，由于工程技术的进步，现代雷达传感器能够解析方位角和

仰角 [78,79]。这种增强的能力催生了 4D 雷达地图，涵盖距离、方位角、仰角和多普勒特征。

通过在RAD张量上使用CFAR检测器，雷达检测可以被转换为点云。然而，这些点云经常会受到道路杂波、干扰和多径效应的影响，这些效应源于噪声中频信号。为了改进检测，可以采用额外的时空滤波方法，例如基于密度的空间聚类应用与噪声（DBSCAN）[80] 或卡尔曼滤波器。

除了多普勒速度，雷达检测的另一个重要特征是雷达散射截面（RCS），它取决于反射目标的特性。这是被物体反射并由接收天线记录的强度。强度取决于物体的材料、尺寸和方向。RCS用于区分不同的物体。例如，汽车和厚墙具有高RCS，而行人和小物体通常具有低RCS [9]。

目前，大多数基于雷达的SLAM方法缺乏高度分辨率。因此，点云通常表示为2D BEV图像。为了提高物体可见性和检测能力，通常会在车辆上安装多个SIMO/MIMO传感器，以实现360度视野，这些传感器通常位于车辆的角点。然而，由于空间和时间上的差异，结合这些多个传感器的特征需要进行进一步处理，最终形成统一的表示。每个点云检测通常包含空间信息、多普勒信息和RCS分量。

4.3. 旋转雷达

与汽车雷达不同，旋转雷达使用单个发射和接收天线。类似于激光雷达传感器，它们被安装在车辆顶部并围绕其垂直轴旋转。像任何调频连续波雷达一样，这些雷达利用发射和接收波之间的频率偏移来计算物体的距离。返回信号的力量提供了有关物体反射率、大小、形状以及相对于接收器的方向的宝贵信息。然而，到达方向由旋转雷达的当前角度决定，从而产生极化功率谱 [81,82]。

这种二维极谱能够生成比汽车雷达生成的经过重度处理的点云更稠密的雷达表示。然而，由于这种密集表示，多普勒和雷达散射截面（RCS）分量通常不包括在旋转雷达公开数据集中。另一个问题是，由于缺乏天线，旋转雷达设置无法确定任何高度信息。因此，旋转雷达通常用于二维场景 [81,82]。

需要注意的是，雷达数据集不仅在其类型（即旋转雷达和汽车雷达）上不同，而且在其精确的传感器配置上也不同。例如，尽管两个系统使用汽车雷达配置，但它们在工作范围和分辨率上可能差异很大，通常取决于传感器的价格。旋转雷达也是如此。具有更高旋转频率的雷达将提供更高分辨率的距离-角度图像，从而产生更高质量的数据。

5. 深度学习方法

在本节中，我们讨论并比较已应用于雷达里程计与定位的各种深度学习架构。随后，我们深入探讨雷达数据在深度学习方法背景下如何与其它传感器输入无缝集成的方式。

对各种方法的调查涉及在Google Scholar和Web of Science上进行彻底的搜索。最初，通过在Google Scholar上使用关键词搜索，确定了与基于雷达的里程计估计或位置识别相关的方法。然后，通过回溯最初确定的参考文献，找到了其他相关方法。为了补充Google Scholar的搜索，在Web of Science上进行了类似的搜索，使用了与Google Scholar搜索中相同的查询。该搜索使用了不同的关键词排列和选择，例如“深度学习”、“基于学习”、“（无）监督”、“雷达”、“自定位”、“定位”、“扫描匹配”、“场景流”、“里程计”、“位置识别”、“基于点”、“传感器融合”和“SLAM”，这些关键词和短语使用连词“和”和“用于”连接。鉴于深度学习在基于雷达的SLAM中的相对较新引入，搜索特别考虑了从2016年起发表的方法。这个截止标准是由Wang等人[1]在2017年发表的关于使用深度学习的视觉里程计的著名论文所驱动的。图4显示了使用深度学习和雷达传感器进行里程计估计和定位调查的论文数量。在这里，方法也根据雷达设置进行了区分，如果该方法还使用了其他模式。

由于深度学习在基于雷达的SLAM中的引入相对较新，因此搜索特别考虑了从2016年起发表的方法。这个截止标准是由Wang等人 [83]在2017年发表的关于使用深度学习的视觉里程计的著名论文所驱动的。图4显示了使用深度学习和雷达传感器进行里程计估计和定位调查的论文数量。在这里，方法也根据雷达设置进行了区分，如果该方法还使用了其他模态。

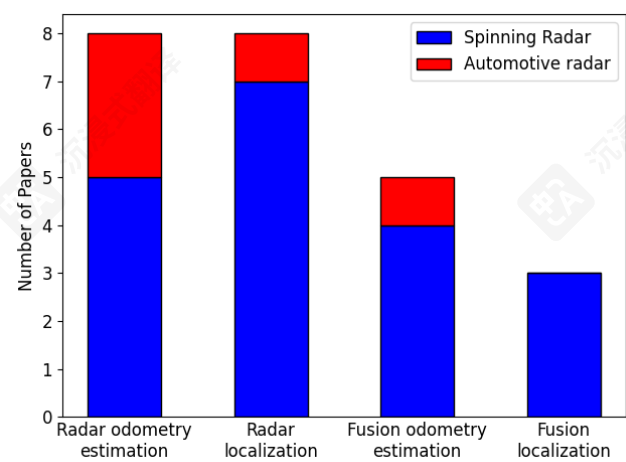


图 4。使用深度学习进行雷达里程计估计和定位的论文数量，以及使用雷达传感器的融合方法。

由于雷达非常经济实惠，它们可以作为感知传感器集成到各种系统中。这些系统在可用的计算硬件和使用场景方面可以有很大差异。例如，自动驾驶汽车有更多空间携带用于处理的图形处理单元（GPU），而清洁机器人内存非常有限。另一方面，汽车运行在更具挑战性的环境中，其运行规模和自车速度可能远大于清洁机器人，导致更大的运动畸变。因此，介绍的方法并不完全可相互比较，需要根据其对应的数据集和使用场景进行解释。

所研究方法、雷达表示、数据集、其任务以及深度学习的使用情况概述如表3所示。

表 3。应用于雷达里程计和定位的深度学习方法列表。SR 代表旋转雷达，AR 代表汽车雷达设置。

方法	数据表示	数据集	Task	模型架构和使用
[84]	SR笛卡尔扫描	牛津	里程计	用于屏蔽无关检测的卷积神经网络
[85]	SR笛卡尔扫描	牛津	里程计	用于屏蔽无关检测的卷积神经网络
Fast-MbyM [86]	SR笛卡尔扫描	牛津	里程计	卷积神经网络用于屏蔽无关检测
UnderTheRadar [87]	SR笛卡尔扫描	牛津	里程计与定位	使用带多个解码器的U-Net提取关键点位置、权重和描述符。
HERO [88]	SR笛卡尔扫描	牛津，博雷亚斯	里程计	使用具有多个解码器的U-Net提取关键点位置和权重。
RaFlow [89]	AR 4D点云	内部	里程计，场景流估计	包含多尺度编码器、成本体积层和流解码器等多个模块，用于预测场景流。

表 3. 继续

方法	数据表示	数据集	Task	模型架构与使用
4DRO [90]	AR 4D点云	View-of-Delf ,In-房子	里程计	由特征编码、初始位姿估计和位姿扭曲三个模块组成，用于预测自运动。
Milli-RIO [91]	AR 3D点云	内部	里程计	使用LSTM层作为运动模型。
自动放置 [92]	AR 3D 点云	Nuscenes	定位	卷积神经网络作为空间编码器，LSTM作为时间编码器。NetVLAD用于局部描述符。
绑架雷达 [93]	SR极角扫描	牛津	定位	VGG-16编码器和NetVLAD描述符。
看看你周围的 [94]	SR极角扫描	牛津	定位	VGG-16编码器和NetVLAD描述符用于序列数据。
[95]	SR笛卡尔扫描	牛津	定位	VGG-19或ResNet-152作为特征编码器，MLP作为描述符生成器。
雷达LCD [96]	SR笛卡尔扫描	牛津, MulRan	定位	U-Net特征图和关键点生成。PointNetVLAD用于描述符。
Off the Radar [97]	SR笛卡尔扫描	牛津, MulRan	定位	变分自编码器用于将特征表示解耦为不确定性来源和不变描述符。
RadarLoc [98]	SR笛卡尔扫描	牛津	全局姿态预测	自注意力与DenseNet模块特征提取。用于姿态估计的多层感知机。
RSL-Net [99]	SR笛卡尔扫描和卫星影像	牛津	里程计	用于卫星到雷达的cGAN。多个卷积神经网络用于旋转和平移分量的解耦。
GRAMME [100]	SR笛卡尔扫描和相机序列	牛津, RADIATE	里程计	四个模块，包括用于融合的注意力模块，为每个传感器和自运动预测自运动。
L2R [101]	SR笛卡尔扫描	牛津	激光雷达地图定位	激光雷达到雷达GAN用于执行经典ICP和粒子滤波进行定位。
RaLL [102]	SR笛卡尔扫描和激光雷达map	牛津, MulRan	里程计	U-Net 模块用于计算激光雷达嵌入与雷达扫描之间的差异张量。差分卡尔曼滤波器。
指引方向 [103]	SR 笛卡尔扫描, 雷达检测, (激光雷达)地图	Boreas	里程计	U-Net 生成检测地图的掩码。与地图数据的差分 ICP。
[104]	SR 极坐标扫描和激光雷达地图	牛津, MulRan	定位	雷达、激光雷达地图和正负样本之间共享 U-Net 编码器。FFT 作为最终描述符生成。
RaLF [105]	SR笛卡尔扫描和激光雷达子图	牛津 , MulRan , Boreas	定位	RAFT特征编码器和流场预测。CNN用于描述符生成。
Hidden-gems [106]	AR 4D点云、激光雷达扫描、相机输入	View-of-Delf	里程计和场景感知	设置卷积点编码器、用于场景流预测的成本体层、自运动头、用于时间更新的GRU。
Seraloc [107]	SR笛卡尔地图	内部	里程计	点云语义分割。

5.1. 里程计估计

在用于里程计估计的深度学习领域，特别是从点云和图像表示中，一个核心挑战在于识别相关特征并在连续帧中有效跟踪它们。如果这些相关特征能够被成功识别并持续跟踪，则可以使用可微分的点匹配机制（如Kabsch算法 [108]）来确定帧之间的变换矩阵 $T \in SE(3)$ 。然而，问题的核心仍然在于识别这些特征并建立准确的点关联，这构成了从传感器数据中推导出的里程计估计的核心挑战。

模型从点云和图像中提取关键点的能力已得到深入探究，特别是在VSLAM和基于激光雷达的SLAM方法中 [8,28,29]。值得注意的是，与其他传感器相比，雷达由于数据稀疏且不可靠而面临独特挑战。

一种常见的确定变换参数的方法涉及扫描匹配方法，通常假设 $T \in SE(2)$ 。通过学习一个最大化无噪声表示重叠的掩码来估计最优自运动参数，并使用形式为 [8]的损失进行训练：

$$\mathcal{L} = ||\hat{t} - t||_2 - \alpha ||\hat{R}R^T - \mathbf{I}||_2 \tag{4}$$

其中 \hat{t} 和 \hat{R} 是方程(4)中预测的平移和旋转矩阵。 α 是一个固定或可学习的参数，用于调整平移和旋转权重的误差。 I 是单位矩阵。

Barnes 等人 [109] 在 U-Net 编码器-解码器架构中采用了一种掩码网络，以有效过滤深度图像中的无关区域，具体用于视觉里程计。受此启发，Aldera 等人 [84] 采用类似的 U-Net 架构来掩码笛卡尔雷达表示图像中的特定区域。学习到的掩码能有效排除无关元素，如噪声、虚假检测和移动物体。然而，需要注意的是，该过程忽略了检测的多普勒速度。在此方法中，U-Net 架构是唯一负责区分相关地标的组件。

Barnes 等人 [85] 还采用了一种卷积神经网络（CNN）来预测旋转雷达扫描的掩码。该模型学习从雷达扫描中过滤掉所有非必要检测。为了确定两个扫描之间的相对位姿，会为参考扫描（标记为 S1）和另一个后续扫描 S2 学习掩码。每个掩码随后与其对应的扫描相乘。扫描 S2 然后使用可能的位姿偏移 $\Delta t = [\Delta x, \Delta y, \Delta z]$ 进行调整。最后，Barnes 等人采用了一种暴力关联匹配技术来识别最合适的变换。该过程涉及计算 S1 的 FFT 和 S2 的 FFT 之间的离散互相关，该 FFT 使用上述掩码和位姿偏移进行了调整。

在 [85], Weston 等人进行的研究所奠定的基础上，[86]该方法通过解耦旋转和变换分量的搜索过程向前迈进了一步。这种策略性的分离有助于减轻相关扫描匹配技术中固有的计算复杂度。最初，通过评估极坐标中掩码扫描的互相关来识别最佳角度。该角度随后成为后续相关扫描匹配方法中的一个关键因素。通过将旋转角的搜索与变换分量的搜索分开，计算负载显著降低。与 [85]提出的方法相比，这种计算效率的提高仅在性能略有下降的情况下实现。

生成自运动预测的另一种方法涉及直接从数据中提取关键点。Barnes 等人 [87] 实现了一种具有两个解码器分支的 U-Net 编码器-解码器架构，以洞察关键点、权重和描述符。这些组件共同有助于确定两个帧之间的变换矩阵。该架构设计擅长为这些关键点生成关键点、相应的权重和描述符。随后，应用匹配过程来识别表现出最显著相似性的关键点。从这些匹配的关键点中，使用奇异值分解（SVD）技术推导出变换。他们方法的一个值得注意的方面是关键点匹配和姿态估计过程中的嵌入式可微分性。这种设计选择确保了模型参数的可微分性，从而便于与优化过程无缝集成。

与 [87], Burnett 等人 [88] 所采用的方法类似，也利用了 U-Net 架构的能力。该架构被用于从旋转雷达扫描中提取关键点、权重和描述符，最终通过无监督方法促进对连续扫描的里程计预测。作者将通用期望最大化（GEM）方案集成到他们的方法中，其主要目的是最大化证据下界（ELBO）。这种综合策略使他们能够提高其无监督里程计预测过程的效率和有效性。

$$\mathcal{L} = \underbrace{\int_{-\infty}^{\infty} q(\mathbf{x}) \ln \left(\frac{p(\mathbf{x}|\mathbf{z}, \theta)}{q(\mathbf{x})} \right) d\mathbf{x}}_{\leq 0} - \underbrace{\int_{-\infty}^{\infty} q(\mathbf{x}) \ln \left(\frac{p(\mathbf{x}, \mathbf{z}|\theta)}{q(\mathbf{x})} \right) d\mathbf{x}}_{\text{upper bound}} \quad (5)$$

在寻找参数更新规则时，仅优化方程 (5) 的上界。当假设后验近似 $q(\mathbf{x}) = \mathcal{N}(\mu, \Sigma)$ 为多元高斯概率分布时，它会导致一个新的损失。

$$V(q|\Theta) = \mathbb{E}_q[\phi(\mathbf{x}, \mathbf{z}|\theta)] + \frac{1}{2} \ln(|\Sigma^{-1}|) \quad (6)$$

其中 \mathbf{x} 表示潜在轨迹， \mathbf{z} 代表数据测量值，而 θ 表示网络参数。在 EM 训练过程中，联合因子 $\phi(\mathbf{x}, \mathbf{z}|\theta)$ 的参数被拆分为运动先验因子和测量因子。对于运动先验因子，Burnett 等人应用了 Anderson 等人 [110] 提出的加速度白噪声先验，而测量因子则使用从 U-Net 架构中学习到的关键点、权重和描述符进行计算。在 GEM 的 E 步中，模型参数 θ 保持固定，同时优化方程 (6) 以获得后验分布 $q(\mathbf{x})$ ，而在 M 步中，后验分布保持固定，同时优化方程 (6) 以获得网络参数 θ 。与 [87] 中提出的方法的一个显著区别是采用了移动窗口策略。这种方法涉及将不仅仅是两个连续的雷达扫描纳入分析中。

另一种使用外部传感器进行里程计估计的方法是通过场景流预测，该方法涉及预测动态场景的三维运动场。这种方法捕捉了感知扫描环境中静态和动态对象的空间和时间方面。这通常使用成本体积层 [111] 来聚合视差或视差和光流的跨信息，从而促进预测的细化。该层计算不同视差下点的匹配成本，使任何网络都能捕捉复杂的细节，并通过考虑不同维度之间的成本关系来提高场景流预测的准确性。

Ding 等人 [89] 介绍了 RaFlow，如图 5 所示，这是一个专门为通过 4D 自动驾驶雷达检测估计车辆场景流而设计的深度学习网络。该方法涉及在雷达导向流估计模块中处理两个连续的点云。随后，静态流细化模块采用 Kabsch 算法 [108]，来估计车辆的自我运动。利用确定的自我运动，模型随后确定动态检测的场景流。该方法将自我运动作为一个损失，目标是确定完整的场景流。

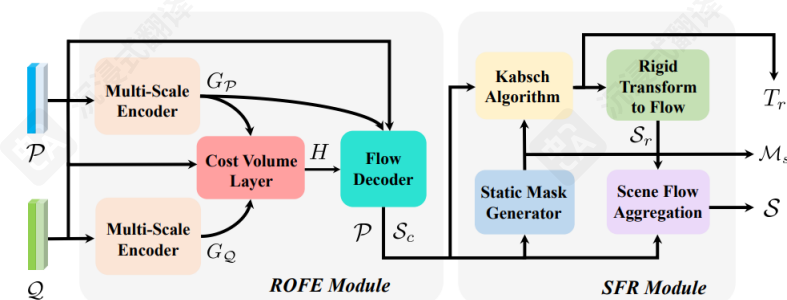


图 5. RaFlow [89] 概述。两个点云 P 和 Q 被输入并首先使用多尺度编码器进行编码。在流解码器模块中，使用点云及其特征的成本体积层来近似整个场景的场景流。此外，使用一个静态掩码生成器来过滤可能影响自我运动估计的动态对象，通过生成掩码 M_s 。最后，应用 Kabsch 算法生成自我运动 T_r 。此变换用于通过预测最终场景流 S 进一步细化剩余场景。经 [89] 授权转载。

Lu 等人 [90] 提出了 4DRO-Net，这是一种从 4D 雷达数据中预测雷达里程计的新方法。这项技术采用由粗到细的层次优化

提取使用 PointNet 风格的特征编码器 [112] 从 4D 雷达数据中提取分离特征，能够融合特征以生成每个扫描的点和全局特征。全局特征对于生成初始位姿估计至关重要，而点特征则有助于利用连续缩放点特征之间的速度感知注意力成本体积层实现的位姿回归模块。校正后的位姿应用于初始变换假设，从而在两个点云之间产生最终的里程计，并在每个研究的尺度上产生自车变换。

最后，Almalioglu 等人 [91] 实现了 Milli-RIO，这是一种基于雷达的里程计方法，基于点云对齐，A NDT [40] 在无迹卡尔曼滤波 (UKF) 方法中用于点云对齐，并采用复杂的点匹配程序。一个 RNN 以 LSTM 层的形式实现为运动模型。给定系统状态为

$$\mathbf{x}_t = [\mathbf{p}_t, \mathbf{q}_t, \mathbf{v}_t, \mathbf{b}_t^a]^T \quad (7)$$

其中 \mathbf{x}_t 是状态， \mathbf{p}_t 是位置， \mathbf{q}_t 是旋转， \mathbf{v}_t 是速度， \mathbf{b}_t^a 是时间 t 上的角速度偏差在公式 (7) 中。假设传感器运动模型 $f(\cdot)$ 的转换函数和角速度偏差的恒定偏差，连续扫描之间的系统状态由

$$\mathbf{x}_t = [\mathbf{p}_{t-1} + f(\mathbf{x}_{t-1}), \mathbf{q}_t \cdot \Delta \mathbf{q}, \mathbf{v}_{t-1}, \mathbf{b}_{t-1}^a]^T \quad (8)$$

使用一个包含 256 个节点的双向-LSTM 来近似公式 (8) 中的转换函数 $f(\mathbf{x}_{t-1})$ 。选择双向 LSTM，因为它应该能够捕捉运动模型的正向和反向依赖关系。该网络使用真实轨迹进行训练，预测状态被输入到 UKF 中。

如所见，用于雷达里程计的深度学习主要关注旋转雷达数据或汽车雷达数据，具体取决于可用的表示形式。处理步骤和实现方式因表示形式而异。

对于汽车雷达，数据通常表示为稀疏检测点云。RCS 和多普勒值被用作每个检测的进一步特征。汽车雷达的趋势是预测场景中所有检测的流动，并通过静止物体预测自运动。各种不同层被用于最佳近似这种转换，包括 PointNet 风格的特征增强器、线性层、变形层和成本体积层。

对于旋转雷达数据，重点在于从类似图像的雷达扫描中识别相关特征和关键点。这些模型能够在极坐标或笛卡尔空间中运行，利用旋转雷达的更密集表示。CNN 和 Transformer 架构被集成用于空间特征提取，而 LSTM 和 GRU 层被实现以在空间特征之间建立时间连接。

对于这两种雷达设置配置，重点仍然是如何找到一个 2D 变换矩阵 $T \in SE(2)$ 。大多数论文还会利用某种掩码模块，在其中过滤掉无关的匹配或特征。只有最近的模型尝试利用 4D 汽车雷达来找到一个 3D 自车变换矩阵 $T \in SE(3)$ 。

5.2. 基于雷达的回环检测

深度学习在回环检测领域获得了显著 prominence (突出地位)。它们通常集成到现有的定位框架中，经常与诸如局部聚合描述符向量 (VLAD) 层等组件一起使用。

VLAD 层最初由 Arandjelovic 等人提出，[113]，将输入的局部特征总结为紧凑的表示。局部特征最初从给定的输入数据集中提取。然后使用聚类算法 (如 k 均值聚类) 将这些提取的特征组织成簇。这个聚类步骤生成一个视觉上有意义的词的隐藏字典，也称为“视觉词”。对于每个局部特征，会确定它在字典中与哪个视觉词最接近。这个分配是基于选择的距离度量 (如余弦相似度或欧几里得距离) 进行的。分配完成后，将分配的视觉词与原始特征之间的差异累积并归一化。这个聚合过程产生一个紧凑的描述符，该描述符编码了差异在视觉词上的分布。描述符的最终维度计算为字典中视觉词数量 (C) 与局部特征初始维度 (D) 的乘积，得到一个大小为 [113] 的描述符。

对于每个局部特征，会确定它在字典中与哪个视觉词最接近。这个分配是基于选择的距离度量（如余弦相似度或欧几里得距离）进行的。分配完成后，将分配的视觉词与原始特征之间的差异累积并归一化。这个聚合过程产生一个紧凑的描述符，该描述符编码了差异在视觉词上的分布。描述符的最终维度计算为字典中视觉词数量（C）与局部特征初始维度（D）的乘积，得到一个大小为 $VLAD = C \times D$ 的描述符。

基于这一概念，NetVLAD层 [114] 通过在VLAD聚合之前引入一个可学习的分配阶段来扩展该方法。这涉及结合一个深度神经网络，该网络学习视觉词和聚类的编码。NetVLAD层的训练通常采用边缘损失，该损失可以是三元组损失 [115] 或对比损失 [116]。这些损失函数通过鼓励学习到的描述符空间中的有意义表示来指导训练过程。

$$\mathcal{L}_{\text{triplet}}(A, P, N) = \max(0, d(A, P) - d(A, N) + \eta) \quad (9)$$

公式(9)表示一个标准三元组损失，其中变量的含义如下：A指的是锚点或查询描述符，P对应于物理上靠近查询位置的正面描述符，而N表示物理上远离查询位置的负面描述符。 η 表示边缘。正面通常选择为在查询内距离小于5米的地点，而负面则是在查询位置15米以外。

Cai 等人 [92] 将深度时空编码器与动态点移除方案融合用于预处理，并采用基于雷达散射截面（RCS）的重新排序分数进行后处理。该架构如图6所示。过滤后的表示通过NetVLAD层，并使用三元组损失进行优化以生成唯一的场景描述符。最后，使用RCS重新排序来匹配具有相似RCS值的场景，过滤掉非相关的静止特征，如果扫描与在指定关键姿态下的扫描的RCS值直方图差异过大，则将地点作为潜在回环检测点拒绝。作者在NuScenes数据集上测试了该方法。

Suaftescu 等人 [93] 采用VGG-16骨干网络与NetVLAD层，从旋转雷达捕获的RA地图中提取全局不变描述符。该方法利用扫描的极坐标表示，结合圆形填充、抗混叠模糊和方位角方向最大池化技术，以实现局部描述符的旋转不变性。网络使用三元组损失进行优化，增强数据中不同类别之间的分离，旨在提取用于基于雷达的地点识别的鲁棒特征。基于这项工作，Gadd 等人 [94] 创建了参考轨迹和实时轨迹上地点嵌入之间的对比增强距离矩阵。地点识别随后通过基于不同地点距离矩阵的最邻近搜索执行。

在后续工作中，Gadd 等人 [95,117] 专注于无监督雷达位置识别，采用嵌入学习方法，确保不同实例的特征被分离，而增强实例的特征保持不变。他们使用 VGG-19 作为前端骨干网络，提取局部特征，这些特征应与总增强量无关。类似地，基于邻域搜索，为定位创建了 [94] 差异矩阵。

袁等 [97] 实现了一种生成模型，即变分自编码器（VAE） [118]，用于创建一种结合不确定性的位置识别方法，以突出正样本对之间的差异。VAE用于将特征表示分解为预测无关的不确定性来源的方差和语义不变描述符。对于负样本对，语义不变描述符之间的距离被最大化。对于正样本对，使用VAE找到的均值和方差采样高斯分布，并将其添加到描述符中以考虑不确定性。这在高度噪声的雷达扫描中尤其相关。

不确定性还允许保持一致的映射，拒绝与现有地图数据相比具有高不确定性的样本。

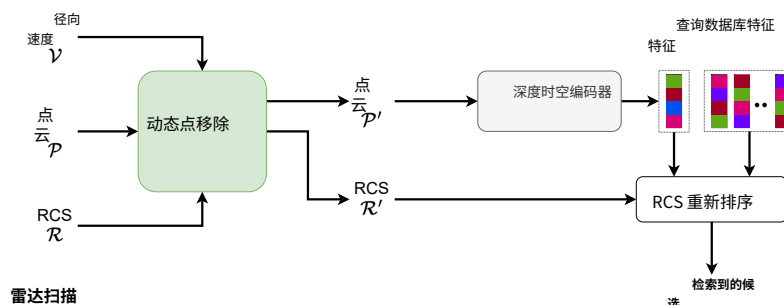


图6. Autoplace模型 [92]的相似位置检索。初始时，一个2D点云，附带其径向速度和RCS评分，作为输入。第一步涉及过滤点云以消除移动目标。得到的过滤点云通过一个时空编码器进行处理，该编码器使用CNN作为空间编码器，使用LSTMs作为时间编码器。最终表示随后通过一个NetVLAD层，促进提取关键特征和空间关系。随后，NetVLAD层确定的相似位置通过根据每个RCS直方图的Kullback–Leibler散度进行重新排序进一步细化。这个过程增强了模型根据雷达特征理解和区分场景的能力。

Usuelli 等人 [96] 通过使用与相关关键点采样局部描述符来执行位置识别，扩展了Burnett 等人 [88] 的工作，如图7所示。这些描述符随后通过一个针对点式全局定位进行适配的NetVLAD层进行处理，该层最初是为激光雷达方法 [119] 实现的。他们还使用了余弦距离度量来评估三元组损失，增强了提取的特征的判别能力。

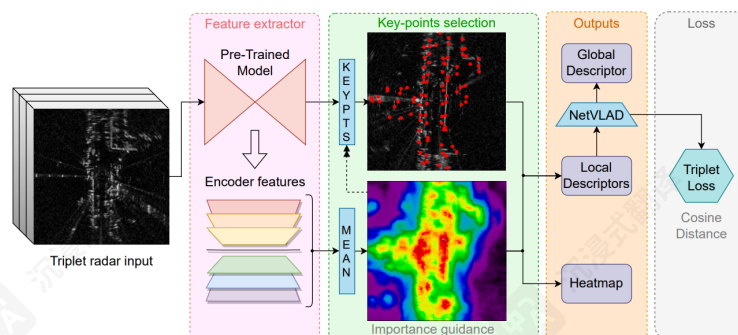


图7. RadarLCD的回环检测训练流程 [96]。该模型采用了来自 [87] 的关键点提取和编码结构，并在其中进行了修改，即计算描述符图的均值以生成重要性引导图。该重要性引导图随后在关键点位置进行采样，并通过NetVLAD层处理以生成全局描述符。此流程确保了关键点信息与重要性引导的全局表示相结合，从而有效利用雷达回环检测

经 [96] 许可转载。

与先前的工作不同，Wang 等人 [98] 实现了 RadarLoc，该系统从旋转雷达扫描中获取单个笛卡尔地图作为输入，并基于 GT 信息输出单个姿态 T。卷积网络的架构包括一个自注意力掩码模块，随后是一个 DenseNet [120] 特征提取器，最后是两个用于解耦确定平移和旋转分量的独立多层感知机 (MLP) 层。这种配置促进了特征图的提取

从笛卡尔地图输入。值得注意的是，一个通用的最大池化层 [121] 被用作池化机制，有助于在 DenseNet 模块中提取全局描述符。

Adolfsson 等人 [122] 提出了一种针对旋转雷达数据的复杂回环验证方案，该方案采用了两个逻辑回归分类器。尽管这种方法不是通过深度学习进行训练的，但它仍然与最先进的研究相关。该过程涉及从笛卡尔鸟瞰图（BEV）雷达扫描中提取手工制作的 CFEAR [36] 特征作为全局描述符。然后利用这些全局描述符来识别回环候选。对两个扫描进行对齐测量，从合并的点云中提取各种特征，包括联合和单独熵、雷达峰值检测以及重叠度量。这些特征作为训练逻辑分类器的输入，反映了两个点云基于其重叠度的相似性。最后，这种相似性结合了点云在里程计和全局描述符距离上的相似性，被整合到逻辑回归分类器方案中，以确定一个经过验证的回环候选。这种综合方法增强了旋转雷达数据回环验证的鲁棒性。

基于雷达的自定位的焦点可分为两类：一类是依赖训练模型生成独特局部位置描述符的方法，另一类是希望从单次扫描中预测全局位姿 $T \in SE(3)$ 的方法。生成局部描述符的模型通常在端到端方法中使用 NetVLAD 层进行训练。在某种程度上，里程计方法中的相同模块被集成到位置识别方法中，包括用于空间和时间特征提取的 CNN 和 LSTM。从训练数据中提取相似和不同的位置，以迫使相似位置的特征描述符具有较小的距离度量，而物理上相距较远的位置之间存在较大的差异。这导致对所选的相似和不同距离标准以及所研究的 datasets 产生很大依赖。

执行定位的另一种主要方法是直接从局部扫描中预测自位姿 $T \in SE(3)$ 。与利用 BOW 方法相比，这些方法的研究要少得多，因为它们不依赖于任何地图的使用，因此非常容易过拟合。然而，当向系统中添加额外的传感器时，这些方法更为普遍。

5.3. 融合与级联架构

5.3.1. 视觉融合

在许多方面，视觉融合与雷达数据结合具有最大的潜力，因为雷达能为相机提供非常需要的深度感知，同时与激光雷达相比，成本相对较低，适合经济用途。同时，视觉也能为雷达提供非常需要的空间上下文，在“正常”光照条件下 [5,9]。

Yu 等人 [123] 实现了一种多模态深度定位方案。该架构包含两个主要分支：一个图像分支，使用 CNN 模块对相机图像进行编码；一个点云分支，使用 PointNet [112] 主干网络。激光雷达与 CNN 分支之间的融合通过空间对应、逐点图像特征提取，以及每个图像编码器块与生成的点云表示块的特征融合来实现，然后编码到点云特征中。最后，在最后一个融合层上应用 NetVLAD 层来训练网络。

Almalioglu 等人 [100] 将相机数据与距离传感器数据（即雷达或激光雷达）融合，用于在所有天气条件下进行名为 GRAMME 的自监督几何感知多模态自运动估计网络。该网络由多个模块组成，用于在各种天气和光照条件下估计车辆的自运动。一个 DepthNet U-Net 从相机输入中生成深度图像，而一个 VisionNet 则根据连续的相机帧预测自运动。一个 MaskNet 从距离输入中生成输入掩码，同时一个 RangeNet 基于距离传感器输入预测自运动。RangeNet 和 VisionNet 的输出在 FusionNet 模块中融合，以预测

自运动。最后, DepthNet、MaskNet 和 FusionNet 的结果在一个空间变换器中组合, 以完全监督的方式训练模型。

Zhuo 等人 [124] 实现了一种相机-雷达融合模型, 用于基于 4D 雷达数据 (称为 4DRVO, 如图 8 所示) 进行视觉-雷达里程计估计。该模型采用多尺度方法融合特征, 其中相机特征在金字塔网络中计算, 雷达特征则采用 PointNet++ [125] 的扩展版本 Radar-PointNet++ 以金字塔风格计算, 该版本除了检测的坐标位置外, 还集成了额外的雷达散射截面 (RCS) 和多普勒信息。视觉和雷达特征在多个尺度上使用外部相机-雷达校准矩阵将雷达特征投影到 RGB 图像的特征图上。所得投影和雷达特征图进一步使用基于可变形注意力的空间交叉注意力机制进行融合, 其中雷达特征识别 RGB 特征图中的相关物体。引入成本体积层以在连续点云之间建立关联, 其中使用速度引导的点置信度估计为每个雷达检测分配置信度校准, 本质上学习对非平稳检测进行掩码处理。多层感知机 (MLPs) 在每个尺度上实现以生成里程计估计。最后, 实现了一个姿态扭曲细化模块, 以融合当前尺度和下一尺度的嵌入、置信度校准和姿态估计, 以进一步细化姿态估计。4DRVO 模型能够比单独的雷达或相机数据更好地预测车辆的里程计, 甚至能够生成比单一激光雷达传感器更好的结果, 证明了 4D 雷达数据的重要性。

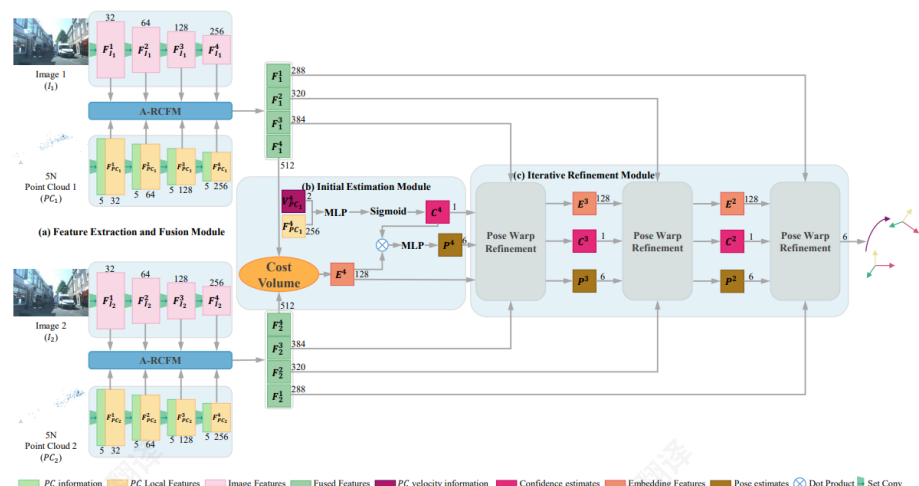


图8. 4DRVO-Net [124]包含三个主要组件。第一个是用于雷达和视觉数据的多尺度提取块, 每个尺度都有一个自适应4D雷达-相机融合模块 (A-RCFM) 融合块。从视觉和4D雷达数据中分别提取图像特征和4D点特征。第二个是一个初始位姿估计模块, 它使用最粗糙尺度的两个4D点云及其特征来生成置信度分数、嵌入特征和位姿估计。最后一个组件是一个位姿细化模块, 它使用先前尺度的置信度、嵌入特征和预测位姿来逐步改进每个尺度的估计

经 [124] 授权转载。

也可以将雷达与其他基于视觉的传感器耦合。Tang 等人 [99,126] 使用卫星影像在人工雷达扫描和真实雷达扫描之间进行扫描匹配。作者实现了一个生成对抗网络 (GAN) [127] 将卫星图像转换为人工雷达图像。他们使用了一种条件生成对抗网络, 或 cGAN, 以便在先验卫星图像上对人工雷达图像进行条件化。训练分为两部分。第一部分涉及训练 cGAN 以生成人工雷达表示。网络的第二部分包括找到变换矩阵 $T \in SE(2)$ 的平移和旋转分量。由于旋转雷达扫描或顶视图图像中均未包含高度信息, 因此无法推断出高度变换信息。该网络通过确定原始雷达数据与人工扫描旋转版本之间在傅里叶域中最大的相关性来分割平移和旋转分量的搜索角度。最后, 使用线性层和卷积层来确定平移偏移量。这种方法需要近似正确的自位姿才能获取卫星图像。

关于高度变换的信息可以推断出来。该网络通过在傅里叶域中确定原始雷达数据与人工扫描旋转版本之间的最大相关性来确定旋转角度，从而将平移和旋转分量的搜索分开。最后，使用线性层和卷积层来确定平移偏移量。这种方法需要近似正确的自位姿才能获得卫星图像。

5.3.2. 激光雷达融合

近年来，许多作者已实现深度生成模型，从定制输入数据生成特定合成数据，其中雷达到激光雷达的生成经常被执行，因为昂贵的激光雷达数据可以通过生成模型与廉价的雷达进行交换。生成模型不仅执行传感器之间的转换，如激光雷达到雷达 [128] 或雷达到激光雷达 [101]，还执行不同天气条件下激光雷达表示之间的转换 [129]。

基于生成建模，Yin 等人 [101] 引入了一种 GAN，用于将旋转雷达扫描转换为合成激光雷达地图。该网络不是共享极坐标特征，而是从雷达扫描生成合成激光雷达 BEV 图像。合成激光雷达输出旨在消除雷达扫描中的杂波，同时生成类似于激光雷达扫描的特征。然后，作者使用该合成激光雷达扫描以及先验地图，在粒子滤波定位方案中进行里程计和自定位。里程计使用连续激光雷达扫描之间的点对平面 ICP 进行估计。定位使用基于匹配到地图的匹配点数量的度量进行执行。

Yin 等人 [102] 使用图 9 中所示的可微分卡尔曼滤波方法，在已知位置的先验激光雷达地图上执行里程计。一个笛卡尔 2D 雷达扫描和激光雷达子图通过多个 U-Net 架构输入，从每个传感器中提取两个独立的特征嵌入。激光雷达嵌入在一系列角度 $\delta\theta$ 内旋转，并通过一系列可能的解因子 Δx 和 Δy 进行平移。雷达嵌入与每个转换后的激光雷达嵌入之间的差异张量被分割成块，每个块通过另一个网络输入以获得基于块的差异。这个差异被平均处理，并取其 Softmin，最终为每个解候选 δx 、 δy 和 $\delta\theta$ 生成边缘概率分布。该模型使用三种不同的损失函数进行训练。第一个是使用地面真实参数的 one-hot 编码来计算所找到的边缘概率分布的交叉熵损失。第二个损失是地面真实参数与从边缘分布和相应偏移量计算出的估计参数之间的均方误差。最后，最后一个损失基于可微分的卡尔曼滤波器，以在似然性和后验中较低的不确定性之间进行平衡。

与 [101,102]，Lisus 等人提出的方法类似，[103] 利用深度学习将雷达扫描定位在激光雷达地图上。具体而言，采用 U-Net 网络从雷达扫描中提取掩码，类似于 [85] 的方法。同时，使用 BFAR 检测器 [130]，从雷达扫描中提取点云，并从掩码中采样特征作为提取点的权重值。该方法的一个重要部分依赖于差分 ICP 后端，该后端将提取的点与参考激光雷达地图进行匹配。

Sun 等人 [131] 在激光雷达数据中实现了一个 GAN 网络用于移动物体检测。虽然雷达数据不直接被深度学习模块使用，但雷达扫描的多普勒速度可作为验证工具，以确保动态物体已被准确识别。此外，激光雷达数据用于过滤雷达数据中的幽灵物体。

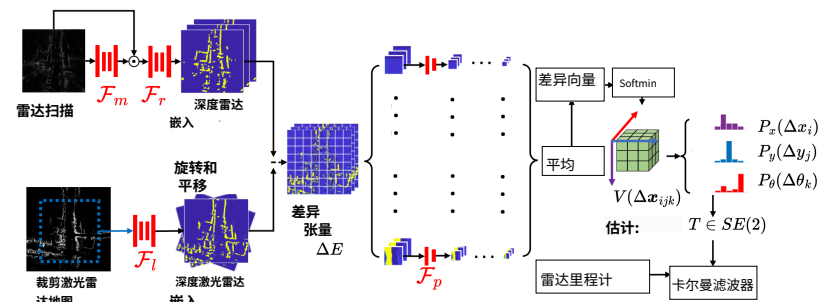


图 9. RaLL 架构概述 [102]。三个网络，标记为 \mathcal{F}_m 、 \mathcal{F}_l 和 \mathcal{F}_r ，共同在一个预定义的搜索空间中生成差异张量 ΔE 。随后，另一个网络 \mathcal{F}_p ，称为补丁网络，被用于为每个补丁和潜在偏移组合构建一个单值数字。这可以用单个张量 $V(\Delta x_{ijk})$ 表示。然后，基于此张量在各个维度上形成概率分布。最后，从这些分布中生成姿态偏移。获得姿态偏移可用于卡尔曼滤波方法，并结合惯性测量单元数据。该综合架构集成了多个网络，以有效生成和利用姿态信息，用于增强定位。经许可转载自 [102]。

Yin 等人 [104] 采用共享编码器-解码器架构，对雷达扫描的极坐标表示和激光雷达子图进行联合位置识别。输出特征被转换为鸟瞰图（BEV）表示，并应用二维快速傅里叶变换（FFT）生成最终向量描述符。该网络采用联合训练方案进行训练，针对每个查询样本（锚点、正样本、负样本）使用不同传感器的损失进行优化。

与Yin等人等人的工作类似，Nayak等人等人的工作也利用激光雷达子图在激光雷达地图上进行基于雷达的定位。然而，与 [104] 不同，单个雷达和激光雷达主干网络使用循环全对场变换器（RAFT）特征提取器 [132] 来生成 BEV 特征表示和每个传感器的唯一描述符。该方法利用位置识别头进行回环检测，并使用度量定位头在激光雷达地图上找到 3D 位姿 $T \in SE(3)$ 。在训练过程中，位置识别头使用当前雷达描述符与激光雷达子图的正面和负面示例之间的三元组损失进行训练，以强制执行不同传感器记录的相似位置之间的描述符相似性。有趣的是，应用最终的 NetVLAD 层并没有产生比最终 CNN 层更好的定位结果。编码的正面激光雷达子图进一步用于度量定位头，以预测雷达场景中的完整光流。类似于 [89]，该方法也利用两个 BEV 编码之间的相关体积作为 GRU 的输入来预测场景光流。利用此光流和定位头，系统在激光雷达地图中的自位姿可以从单个雷达扫描中推断出来。

Ding 等人 [106] 使用 4D 汽车雷达传感器、激光雷达传感器、相机图像和惯性测量单元里程计传感器实现了一个跨模态模型，用于监督场景流方法。主要目标是为扫描预测完整场景流，包括刚性的自运动和其他动态物体的运动。从通过包含集合卷积层 [133]、初始流估计头和运动分割头的骨干网络处理的 4D 雷达点云中估计连续扫描的相关静止点和相应权重。类似于 [89]，对找到的坐标应用奇异值分解来导出刚性变换矩阵，该矩阵使用 IMU 传感器找到的变换进行训练。

5.3.3. 级联架构

接下来，我们讨论两种采用级联技术的典型方法，这些技术涉及利用其他模型的结果，例如在传感器融合方法中进行目标检测和标记时对雷达点的标注。这是为了推动深度学习网络的运用，例如将其作为预处理网络，通过非学习型匹配方法协助过滤掉无关的检测结果。

深度学习网络，例如作为预处理网络，能够通过非学习型匹配方法协助过滤掉无关的检测结果。

Isele 等人进行的工作 [107] 整合了来自其他传感器的语义分割结果，将雷达检测分类到不同的类别中。随后，对分割后的雷达检测进行过滤，仅保留属于相关类别的检测。标注后的雷达点云随后被用于 NDT 扫描匹配以进行对齐。值得注意的是，深度学习并未直接应用于雷达数据；相反，深度学习通过整合其他传感器的信息来增强雷达特征。这种方法展示了来自不同传感器信息的协同集成，利用深度学习将雷达特征集丰富化，从而获得语义分割结果中的洞察。级联方法增强了雷达数据的区分能力，最终有助于在存在协同传感器信息的情况下提高匹配精度。

Cheng 等人 [134] 使用 GAN 架构从 RD 地图创建合成点云。基于 CNN 的生成器为 RD 矩阵创建掩码，而基于 CNN 的判别器则被训练以区分生成器掩码和从激光雷达与同步雷达数据中创建的真实值掩码。生成的合成点云可用于目标检测任务或跟踪，以及在 SLAM 方法中进行里程计估计或自定位。

在融合架构中，雷达数据通常被用于通过 GAN 或 VAE 等技术生成更密集的合成激光雷达表示，用于里程计与自定位。这类方法使得可以使用针对更密集激光雷达表示的合成数据进行里程计与定位方法的设计。然而，生成网络作用于预处理后的雷达表示，如点云或扫描表示。这种预处理可能会导致原始表示中的相关信息丢失，被合成数据替代，从而可能导致结果不准确。

5.4. 评估

在本小节中，我们根据其雷达设置比较所介绍的方法，即汽车雷达和旋转雷达设置，用于里程计和自身定位任务。我们试图从它们的表现好坏方面比较这些方法，同时尝试通过分析推理速度和用于测试模型的设备来突出所使用硬件的重要性。我们还介绍了相关的经典方法以供比较，以查看深度学习模型与成熟方法相比的表现如何。这些数据直接来自所引用的论文。

在所提及的方法之间进行公平比较，找到一个共同的基础，面临着相当大的挑战。为了确保公平的评估，理想情况下，方法应该在相同的数据集和相同条件下进行测试。然而，由于不同任务中数据集和条件的固有差异，实现这一点很复杂，尤其是在比较为旋转雷达传感器设计的方法与为汽车雷达数据集量身定制的方法，反之亦然时。

此外，评估推理速度对于里程计和定位任务至关重要，特别是在自动驾驶等实时应用中。模型的计算时间理想情况下不应是主要瓶颈；相反，重点应放在感知传感器的数据收集速度上，例如雷达数据生成。这意味着公平的比较需要对里程计或定位模型在相同的硬件上进行测试。

然而，实际挑战阻碍了实现这一理想场景。数据生成速度不仅可能因不同类型的传感器（例如，汽车传感器与旋转传感器）而异，甚至同类型传感器之间也可能存在差异。对于汽车传感器而言，额外的预处理步骤可能会在不同配置的传感器或型号之间引入显著的数据生成速度差异。统一采用单一传感器也不切实际，因为不同的应用需要不同的传感器规格，导致分辨率、准确性和价格等方面的差异。

此外，系统的计算硬件也可能限制模型的选择。如果系统因空间或价格限制而缺乏支持现代GPU的能力，那么实现大型模型可能并不可行。选择与系统环境相匹配的适当模型又增加了另一层复杂性，使得在不同环境中比较类似方法变得具有挑战性。尽管存在这些挑战，但建立标准化的评估指标和实践的努力将有助于该领域进行更有意义的比较。

5.4.1. 里程计

在点匹配或位姿估计评估中，确定预测位姿与真实位姿之间位姿差异的平移和旋转分量是一个基本方面。绝对轨迹误差（ATE）是一种常用指标，表示平移和旋转分量的平均误差。通常以米（m）为单位测量平移，以度（deg）为单位测量旋转，ATE的计算公式如下，用于一对位姿：

$$E_{tra} = \|\hat{t} - t\|_2 \quad E_{rot} = \|\hat{R}R^T - \mathbf{I}\|_2 \quad (10)$$

在公式(10)中， \hat{t} 和 t 是预测和真实的平移分量， \hat{R} 和 R 是预测和真实的旋转分量。ATE将这些单独分量组合成一个平移和旋转误差的单一度量。

然而，在里程计估计的背景下，其中在特定空间范围内的全局漂移通常比连续帧之间的微小位姿差异更关键，ATE可能不是最合适的指标。如果网络学习预测两次扫描之间的变换，那么了解网络在经过一定空间或时间范围后的错误程度比了解网络在连续扫描之间的表现更有意义。这在慢速系统且数据生成频率很高的情况下尤其相关。例如牛津雷达数据集，其工作频率为4 Hz。在这种情况下，微小位姿差异可能无法准确捕捉在显著范围内的全局漂移。

为解决此问题，KITTI评估基准 [135] 引入了一个指标，该指标对各种长度的子序列（例如，100米、200米、...、800米）上的相对位置和旋转误差进行平均。得到的平移和旋转ATE被报告为平移分量的百分比偏移和每1000米的角度。这提供了一个更全面的评估指标，考虑了系统在较长距离上的性能，并为里程计任务提供了一个相关的指标。

旋转雷达

表4中提出的方法展示了四种使用相同数据和相同指标评估的不同方法；牛津雷达数据集使用旋转雷达上的KITTI指标。Fast-MbyM [86] 展示了最快的推理速度，与MbyM方法相比，仅以微小的精度牺牲实现了显著的加速。在深度学习模型中，RaLL [102] 脱颖而出，成为最佳表现者，尽管它需要在训练期间使用额外的激光雷达数据，以便使用GAN有效地在雷达和激光雷达表示之间进行转换。

观察表4可知，相关模型通常优于关键点提取和匹配方法。值得注意的是，像 [85–88] 这样的方法还有使用学习到的关键点和权重进行里程计估计和定位的优势，使其在端到端方法中更具通用性。然而，值得一提的是，[87,88] 的方法推理时间更高，需要更强大的GPU支持。尽管如此，考虑到牛津机器人数据集中旋转雷达的4Hz旋转频率，所有方法在实时场景中仍然适用。

表4. 已用于雷达里程计并测试在平均牛津机器人测试数据集上的深度学习方法列表。性能评估来自各自的引用。带(*)标记的模型表示涉及额外激光雷达数据的训练，而带(**)标记的需要额外相机数据。带(***)标记的方法代表最先进的经典雷达里程计方法。表现最佳的模型用粗体突出显示。

方法	KITII (%/deg km)	推理速度 (s)	设备
MbyM [85]	1.16/3.0	0.0747	Nvidia Titan Xp 显卡 (圣克拉拉, 加州, 美国)
Fast-MbyM [86]	2.00/6.3	0.0063	Jetson 显卡 (Nvidia, 圣克拉拉, 加州, 美国)
RaLL * [102]	1.09/3.7	0.243	Nvidia Titan X 显卡 (圣克拉拉, 加州, 美国)
UnderTheRadar [87]	2.06/6.7	0.034	-
HERO [88]	1.95/6.5	0.180	Nvidia Tesla V100 显卡 (圣克拉拉, 加州, 美国)
GRAMME ** [88]	1.98/5.1	-	NVIDIA GTX 1080Ti (圣克拉拉, 加州, 美国)
CFEAR-3 *** [136]	1.09/3.6	0.167	i7-6700K CPU (英特尔, 圣克拉拉, 加州, 美国)

与CFEAR [90], 深度学习方法相比, 在准确性和计算时间方面表现出可比的性能。然而, 需要注意的是, CFEAR评估是在Intel i7-8700k CPU上进行的, 而其他模型需要GPU支持进行评估。

汽车雷达

表5展示了为汽车雷达量身定制的里程计方法集合, 并在相同的View-of-Delft数据集上进行了评估。汽车雷达数据以其稀疏性而闻名, 通常用于目标检测和自定位, 而非里程计。鉴于稀疏雷达点云的固有挑战, 在此背景下, 重点在于里程计估计, 特别是利用View-of-Delft数据集中雷达点云的4D特性。值得注意的是, 表5中的所有模型均在单个NVIDIA 2080Ti GPU上进行了评估, 这确保了比先前比较更可靠的推理比较。

表5. 在View-of-Delft (VoD) 数据集上评估的用于雷达里程计的深度学习方法列表。这些方法在VoD Seq. 4上进行了评估, 数据来源于 [90] 值得注意的是, (*) 在训练过程中整合了相机、激光雷达和惯性测量单元数据, 提供了一种多模态学习方法。此外, (**) 代表了一种基于正态分布变换的经典方法。评估是在单个 NVIDIA 2080Ti GPU 上进行的。表现最佳的模型以粗体标出。

方法	ATE (m/deg)	推理速度 (s)
RaFlow [89]	0.07/0.45	0.0363
CMFlow * [106]	0.05/0.10	0.0304
4DRO-Net [90]	0.08/0.07	0.0108
NDT **	0.56 / 1.52	0.0010

在仅使用雷达的设置中, RaFlow [89] 和 4DRO-Net [90] 展现出具有竞争力的性能。有趣的是, CMFlow [106], 在其训练设置中集成了额外的信息 (如相机、激光雷达和惯性测量单元数据), 其结果也相当可观。值得注意的是, 仅使用雷达的方法, 尤其是4DRO-Net [90], 在 ATE 和推理速度方面表现突出, 即使与 CMFlow 设置相比也是如此。此外, 深度学习方法始终优于经典的 NDT 变换, 这表明尽管经典方法在里程计预测过程中无需 GPU 而具有计算效率, 但性能方面已取得显著进步。

5.4.2. 自定位

自定位方法与里程计估计相比, 需要不同的评估标准。虽然公式(10)可用于预测检测到的回环检测中的扫描匹配, 但它不足以评估模型识别已访问位置的能力。

通常，会生成一个场景描述符，类似于(Net)VLAD描述符，用于封装位置并促进位置识别。

一种评估方法涉及最近邻搜索（NN），使用定义的距离度量（例如，方程（9）中指定的NetVLAD损失）在描述符数据库中确定观测描述符的最近邻。常用召回率@1、F1分数和平均精度均值（mAP）等指标通过精度-召回曲线评估一致的回环检测。然而，挑战在于定义字典中预定义关键姿态的数量，因为关键姿态之间的距离较小会导致指标性能随着字典规模增大而变差，而未必表示模型整体较差。

类似于雷达里程计，位置识别模型主要使用旋转雷达数据进行评估。然而，与里程计方法不同，这些方法对GPU要求不那么关键，因为它们不需要连续运行。一旦在地图上确定了位置，就可以使用里程计网络进行姿态细化。此外，如果在先验地图上执行自定位，将两者部署在同一设备上具有计算效率，无需数据传输。在本地集群执行定位或使用机载硬件之间进行选择会带来不同的挑战，每种选择都有其自身的考量。在任一情况下，模型应具备足够速度以在全局漂移或检测到已知场景时实现重新定位。

表6展示了为旋转雷达设计的各种深度学习模型的定位识别性能，该性能使用牛津数据集进行评估，并采用召回率@1、F1和mAP等指标进行衡量。为了便于比较，还包含了非学习型描述符方法RaPlace [137]。然而，由于正边界定义存在差异，在评估标准上出现了挑战，其中 [94] 使用更严格的10米边界，而其他模型则采用较宽松的25米边界。

表6. 列出了应用于牛津雷达数据集旋转雷达定位的深度学习方法。值得注意的是，(*)的正边界为10米，而其他方法采用25米边界。此外，(***)在训练过程中集成了额外的激光雷达子图。表现最佳模型以粗体标出。

方法	召回率@1	mAP	F1	推理（秒）	设备
绑架雷达 [93]	90.49	0.75	0.7	-	-
环顾四周 * [94]	-	0.53	0.53	-	-
对比学习 [95]	85.39	0.95	0.89	-	-
Off the Radar [97]	93.67	0.99	0.95	-	-
RadarLCD [96]	-	0.94	-	-	-
雷达到激光雷达 ** [104]	93.18	-	-	-	-
RaLF ** [105]	97.0	-	-	-	2 英特尔 TITAN-X GPU (圣克拉拉, 加州, 美国)
RaPlace * [137]	78.0	-	-	0.24	16核英特尔 i7-11700 CPU (圣克拉拉, 加州, 美国)

不同的模型在不同的评估指标上表现优异。RaLF [105]在 Recall@1 指标上达到了最高分，超越了纯雷达模型，也超过了在训练过程中使用激光雷达子图的类似模型 [104]。Off the Radar [97] 在 mAP 和 F1 分数上表现更优，甚至超过了 [104]。值得注意的是，Recall@1 分数最高的模型未能提供其他指标，这凸显了跨模型统一定位指标的需求。

表6还揭示了模型推理速率信息的缺失，RaPlace 是唯一例外。此外，关于运行模型的硬件信息也有限。考虑到经典方法与旋转雷达频率（4 Hz）紧密匹配，评估模型能否在线环境中有效运行至关重要。

一些模型包含一个精细的姿态匹配层，通常利用激光雷达子图，在位置识别后估计姿态与关键姿态之间的变换矩阵。这些模型的评估指标与里程计评估一致。然而，在

这个背景下，姿态差异比接近单位矩阵的差值更显著，反映了姿态之间更大的平均位移，使得平移和旋转的ATE（绝对追踪误差）更具意义。

表7表明，结合扫描匹配技术的模型，如RaLL [102]和RaLF [105]，表现优于RadarLoc，强调了除扫描数据外包含额外信息的重要性。这突显了实现位置识别头的必要性，因为仅依赖单个6自由度头可能不足以有效定位，特别是在大型或可能重复的环境中。

表7. 牛津雷达数据集上用于基于雷达的度量定位的深度学习方法列表。（*）仅预测单个6自由度变换矩阵，位移误差。

方法	δx (m)	δy (m)	$\delta \theta$ (度)
RadarLoc * [98]	2.53	2.53	2.81
RaLL [102]	1.04	0.69	1.26
RaLF [105]	0.97	0.69	1.15

总之，表8突出了仅使用汽车雷达的定位模型与采用相机和激光雷达传感器的最先进方法之间的差异。后者在Recall@1和F1分数上分别比仅使用雷达的流程高出约7个百分点和4个百分点。虽然最先进的方法展现了卓越性能，但结果仍然强调了仅使用雷达模型的可靠性。进一步的研究可以探索传感器融合模型的潜在优势，这些模型已在里程计估计中证明了其效用。

表8. 在NuScenes数据集上，Autoplace [92] 与标记为(*)的视觉-激光雷达定位模型 [138]在汽车定位方面的比较。

方法	Recall@1	F1
Autoplace [92]	90.45	0.95
LCPR * [138]	97.34	0.99

6. 讨论与未来研究

深度学习技术逐渐应用于基于雷达的里程计与定位方法，主要作为空间和时间特征编码器。这些方法通常使用卷积神经网络（CNN）进行空间编码，并采用RNN网络（如LSTM和GRU）作为时间特征编码器，以获取同一物体在不同场景中的时间信息。

无论任务如何，所实现的模块和层在雷达配置上有所不同。汽车雷达配置侧重于PointNet 和 PointNet++ 特征编码器，以处理点云表示。这些特征编码器因其简单架构而被选用，PointNet 的共享 MLP 结构使其能够实现快速推理时间。PointNet 及其改进版本在其他领域也已被广泛用于各种最先进方法进行点云处理 [139–144]。旋转雷达则侧重于使用 CNN 编码表示。类似于计算机视觉中的遮挡问题，大多数模型尝试生成掩码或权重函数来处理噪声和潜在的雷达数据错误 [145]。

与其他雷达感知任务相比，深度学习在此背景下的应用 somewhat 有限。尽管使用了最先进的深度学习技术，但用于雷达里程计与定位的论文数量远低于用于雷达目标检测和分割的论文数量 [5,9]。类似于[107]，分割论文 [145–147] 的结果可用于从对系统里程计或定位无贡献的扫描中屏蔽无关的检测结果和区域。在一个完整的流程中，分割、分类、里程计和自定位应在端到端方法中执行，其中某一模块的结果

集成, 例如作为掩码, 用于另一个模块。类似地, 用于无监督分类的语义网络已在计算机视觉 [148] 和自然语言处理 [149] 领域成功应用。

尽管存在如 BEV 图像或点云等经过处理的雷达表示, 这些表示对于里程计和定位任务比预-CFAR 表示更直观, 但深度学习方法直接从较少处理的表示中解决雷达里程计或定位的应用仍然有限。虽然 ADC 或 RAD 表示可能包含无关信息, 如杂波或动态物体, 但深度神经网络有能力学习屏蔽这些区域, 并专注于基于这些表示的自运动和独特的全局描述符。由于屏蔽是旋转雷达里程计的重要组成部分, 因此令人惊讶的是, 没有类似的方法应用于汽车雷达的预-CFAR 表示。这种方法有可能克服在预处理过程中丢弃相关物体的问题, 在雷达数据可用为 RAD 张量或 ADC 信号的场景中 (如在 K-Radar 或 RadIAL 数据集中) 具有优势。

相比之下, 其他雷达感知任务中的一个显著趋势是转向较少处理的 2D/3D 数据 (例如 BEV 或点云表示), 而转向 ADC [150] 或 RAD [74] 等较少处理的表示。这一趋势背后的原理是使深度学习模型能够在输入空间中识别相关特征并屏蔽无关元素。虽然这一趋势在雷达里程计中可见, 特别是在处理旋转雷达的 RA 地图时, 但汽车雷达领域尚未遵循类似的轨迹, 因为它利用 CFAR 后的检测进行 PointNet 特征编码。此外, 虽然 4D 雷达数据在目标检测和分割等任务中的使用正在增加, 但它们在里程计和定位方法中的应用仍然有限, 仅限于少数方法。

另一个观察到的趋势是, 在所研究的传感器设置中, 旋转雷达配置占据主导地位。对于里程计与定位任务而言, 使用旋转雷达数据的论文数量超过了使用汽车雷达的论文数量。然而, 鉴于汽车雷达成本较低且易于集成到现有车辆中, 因此存在对这类雷达传感器进行更大研究关注的潜力。

此外, 我们观察到, 许多在相同数据集上评估的模型遵循截然不同的评估标准, 无论是在评估方法还是模型推理速度方面。与目标检测和分割任务不同——在这些任务中, NMS和IOU分数被广泛接受——里程计和位置识别任务缺乏标准化的评估标准, 导致基于外部参数 (如确定真实识别位置的真实位姿参数) 的评估结果存在差异。

在自动驾驶定位系统领域, 大多数论文提到该系统的性能速度高于录制频率。然而, 在实时应用中, 需要考虑其他因素, 例如位置识别过程中的数据传输和位姿细化。对于里程计应用而言, 各论文使用的硬件组件差异显著, 导致难以直接比较结果。对于深度学习方法, 推理时间在优化和微调 [151,152] 方面尤为重要。如果模型在平台上无法实时运行, 可以采取多种步骤, 例如降低网络中的参数数量或更换系统硬件。此外, 系统中的某些硬件可能会升级, 或者雷达的某些特性会发生变化。然后, 模型需要快速适应这些变化, 并且需要控制以确保网络仍然与新数据的特性兼容, 这突显了迁移学习的重要性。

此外, 与雷达里程计和定位相比, 更广泛的传感器感知领域正关注先进的深度学习技术。这些技术包括变体转换器 [153,154], 各种点特征编码器 [112,155] 架构的使用, 以及通过模型架构而不是训练期间的增强将旋转不变性纳入深度学习模块 [156,157]。

这些进步不仅影响着通用雷达感知，也影响着其他模态的里程计和定位模型。

作者贡献：概念化：M.B. 和 A.P.；初稿撰写，M.B.；监督，A.P. 和 T.M.；审阅，A.P. 和 T.M.；项目管理，A.P. 和 T.M. 所有作者均已阅读并同意手稿的发表版本。

资助：本研究未获得外部资助。

机构审查委员会声明：不适用。

知情同意声明：不适用。

数据可用性声明：不适用数据共享。

利益冲突：作者声明无利益冲突。

参考文献

1. Bavle, H.; Sanchez-Lopez, J.L.; Cimorelli, C.; Tourani, A.; Voos, H. From slam to situational awareness: Challenges and survey. *Sensors* **2023**, *23*, 4849. [\[CrossRef\]](#)
2. Ryde, J.; Hillier, N. Performance of laser and radar ranging devices in adverse environmental conditions. *J. Field Robot.* **2009**, *26*, 712–727. [\[CrossRef\]](#)
3. He, G.; Yuan, X.; Zhuang, Y.; Hu, H. An integrated GNSS/LiDAR-SLAM pose estimation framework for large-scale map building in partially GNSS-denied environments. *IEEE Trans. Instrum. Meas.* **2020**, *70*, 1–9. [\[CrossRef\]](#)
4. Yao, S.; Guan, R.; Huang, X.; Li, Z.; Sha, X.; Yue, Y.; Lim, E.G.; Seo, H.; Man, K.L.; Zhu, X.; et al. Radar-camera fusion for object detection and semantic segmentation in autonomous driving: A comprehensive review. *arXiv* **2023**, arXiv:2304.10410.
5. Harlow, K.; Jang, H.; Barfoot, T.D.; Kim, A.; Heckman, C. A New Wave in Robotics: Survey on Recent mmWave Radar Applications in Robotics. *arXiv* **2023**, arXiv:2305.01135.
6. Zhang, J.; Xie, Y.; Ling, L.; Folkesson, J. A Fully-automatic Side-scan Sonar SLAM Framework. *arXiv* **2023**, arXiv:2304.01854.
7. Steiniger, Y.; Kraus, D.; Meisen, T. Survey on deep learning based computer vision for sonar imagery. *Eng. Appl. Artif. Intell.* **2022**, *114*, 105–157. [\[CrossRef\]](#)
8. Chen, C.; Wang, B.; Lu, C.X.; Trigoni, N.; Markham, A. A survey on deep learning for localization and mapping: Towards the age of spatial machine intelligence. *arXiv* **2020**, arXiv:2006.12567.
9. Zhou, Y.; Liu, L.; Zhao, H.; López-Benítez, M.; Yu, L.; Yue, Y. Towards deep radar perception for autonomous driving: Datasets, methods, and challenges. *Sensors* **2022**, *22*, 4208. [\[CrossRef\]](#)
10. Visentin, T.; Hasch, J.; Zwick, T. Analysis of multipath and DOA detection using a fully polarimetric automotive radar. *Int. J. Microw. Wirel. Technol.* **2018**, *10*, 570–577. [\[CrossRef\]](#)
11. Durrant-Whyte, H.; Bailey, T. Simultaneous localization and mapping: Part I. *IEEE Robot. Autom. Mag.* **2006**, *13*, 99–110. [\[CrossRef\]](#)
12. Bailey, T.; Durrant-Whyte, H. Simultaneous localization and mapping (SLAM): Part II. *IEEE Robot. Autom. Mag.* **2006**, *13*, 108–117. [\[CrossRef\]](#)
13. Lewis, M.; Liu, Y.; Goyal, N.; Ghazvininejad, M.; Mohamed, A.; Levy, O.; Stoyanov, V.; Zettlemoyer, L. Bart: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension. *arXiv* **2019**, arXiv:1910.13461.
14. Voulodimos, A.; Doulamis, N.; Doulamis, A.; Protopapadakis, E. Deep learning for computer vision: A brief review. *Comput. Intell. Neurosci.* **2018**, *2018*, 7068349. [\[CrossRef\]](#)
15. Sarlin, P.E.; DeTone, D.; Malisiewicz, T.; Rabinovich, A. Superglue: Learning feature matching with graph neural networks. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 20–25 June 2020; pp. 4938–4947.
16. Zhu, A.Z.; Liu, W.; Wang, Z.; Kumar, V.; Daniilidis, K. Robustness meets deep learning: An end-to-end hybrid pipeline for unsupervised learning of egomotion. *arXiv* **2018**, arXiv:1812.08351.
17. Liu, X.; Chen, S.W.; Aditya, S.; Sivakumar, N.; Dcunha, S.; Qu, C.; Taylor, C.J.; Das, J.; Kumar, V. Robust fruit counting: Combining deep learning, tracking, and structure from motion. In Proceedings of the 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Madrid, Spain, 1–5 October 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 1045–1052.
18. Sarlin, P.E.; Cadena, C.; Siegwart, R.; Dymczyk, M. From coarse to fine: Robust hierarchical localization at large scale. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 12716–12725.
19. Filatov, A.; Rykov, A.; Murashkin, V. Any motion detector: Learning class-agnostic scene dynamics from a sequence of lidar point clouds. In Proceedings of the 2020 IEEE International Conference on Robotics and Automation (ICRA), Paris, France, 31 May–31 August 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 9498–9504.
20. Mohammadi, M.E.; Watson, D.P.; Wood, R.L. Deep learning-based damage detection from aerial SfM point clouds. *Drones* **2019**, *3*, 68. [\[CrossRef\]](#)

21. Im, J.U.; Ki, S.W.; Won, J.H. Omni Point: 3D LiDAR-based Feature Extraction Method for Place Recognition and Point Registration. *IEEE Trans. Intell. Veh.* **2024**, 1–18. [\[CrossRef\]](#)
22. O'Shea, K.; Nash, R. An introduction to convolutional neural networks. *arXiv* **2015**, arXiv:1511.08458.
23. Yu, Y.; Si, X.; Hu, C.; Zhang, J. A review of recurrent neural networks: LSTM cells and network architectures. *Neural Comput.* **2019**, *31*, 1235–1270. [\[CrossRef\]](#)
24. Cho, K.; Van Merriënboer, B.; Gulcehre, C.; Bahdanau, D.; Bougares, F.; Schwenk, H.; Bengio, Y. Learning phrase representations using RNN encoder-decoder for statistical machine translation. *arXiv* **2014**, arXiv:1406.1078.
25. Zhou, J.; Cui, G.; Hu, S.; Zhang, Z.; Yang, C.; Liu, Z.; Wang, L.; Li, C.; Sun, M. Graph neural networks: A review of methods and applications. *AI Open* **2020**, *1*, 57–81. [\[CrossRef\]](#)
26. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention is all you need. *Adv. Neural Inf. Process. Syst.* **2017**, *30*.
27. Zhang, Y.; Wu, Y.; Tong, K.; Chen, H.; Yuan, Y. Review of Visual Simultaneous Localization and Mapping Based on Deep Learning. *Remote Sens.* **2023**, *15*, 2740. [\[CrossRef\]](#)
28. Mokssit, S.; Licea, D.B.; Guermah, B.; Ghogho, M. Deep Learning Techniques for Visual SLAM: A Survey. *IEEE Access* **2023**, *11*, 20026–20050. [\[CrossRef\]](#)
29. Saleem, H.; Malekian, R.; Munir, H. Neural Network-Based Recent Research Developments in SLAM for Autonomous Ground Vehicles: A Review. *IEEE Sens. J.* **2023**, *23*, 13829–13858. [\[CrossRef\]](#)
30. Bilik, I. Comparative Analysis of Radar and Lidar Technologies for Automotive Applications. *IEEE Intell. Transp. Syst. Mag.* **2022**, *15*, 244–269. [\[CrossRef\]](#)
31. Grisetti, G.; Kümmerle, R.; Stachniss, C.; Burgard, W. A tutorial on graph-based SLAM. *IEEE Intell. Transp. Syst. Mag.* **2010**, *2*, 31–43. [\[CrossRef\]](#)
32. Wang, Z.; Li, W.; Shen, Y.; Cai, B. 4-D SLAM: An efficient dynamic bayes network-based approach for dynamic scene understanding. *IEEE Access* **2020**, *8*, 219996–220014. [\[CrossRef\]](#)
33. Ru, X.; Gu, N.; Shang, H.; Zhang, H. MEMS inertial sensor calibration technology: Current status and future trends. *Micromachines* **2022**, *13*, 879. [\[CrossRef\]](#) [\[PubMed\]](#)
34. Rublee, E.; Rabaud, V.; Konolige, K.; Bradski, G. ORB: An efficient alternative to SIFT or SURF. In Proceedings of the 2011 International Conference on Computer Vision, Barcelona, Spain, 6–13 November 2011; IEEE: Piscataway, NJ, USA, 2011; pp. 2564–2571.
35. Lowe, D.G. Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vis.* **2004**, *60*, 91–110. [\[CrossRef\]](#)
36. Adolfsson, D.; Magnusson, M.; Alhashimi, A.; Lilienthal, A.J.; Andreasson, H. Cfear radarodometry-conservative filtering for efficient and accurate radar odometry. In Proceedings of the 2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Las Vegas, NV, USA, 24 October 2020–24 January 2021; IEEE: Piscataway, NJ, USA, 2021; pp. 5462–5469.
37. Segal, A.; Haehnel, D.; Thrun, S. Generalized-icp. In Proceedings of the Robotics: Science and Systems, Seattle, WA, USA, 28 June–1 July 1; Volume 2, p. 435.
38. Zhang, J.; Yao, Y.; Deng, B. Fast and robust iterative closest point. *IEEE Trans. Pattern Anal. Mach. Intell.* **2021**, *44*, 3450–3466. [\[CrossRef\]](#)
39. Derpanis, K.G. Overview of the RANSAC Algorithm. *Image Rochester NY* **2010**, *4*, 2–3.
40. Biber, P.; Straßer, W. The normal distributions transform: A new approach to laser scan matching. In Proceedings of the 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2003) (Cat. No. 03CH37453), Las Vegas, NV, USA, 27 October–1 November 2003; IEEE: Piscataway, NJ, USA, 2003; Volume 3, pp. 2743–2748.
41. Kellner, D.; Barjenbruch, M.; Klappstein, J.; Dickmann, J.; Dietmayer, K. Instantaneous ego-motion estimation using doppler radar. In Proceedings of the 16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013), The Hague, The Netherlands, 6–9 October 2013; IEEE: Piscataway, NJ, USA, 2013; pp. 869–874.
42. Kellner, D.; Barjenbruch, M.; Klappstein, J.; Dickmann, J.; Dietmayer, K. Instantaneous ego-motion estimation using multiple Doppler radars. In Proceedings of the 2014 IEEE International Conference on Robotics and Automation (ICRA), Hong Kong, China, 31 May–7 June 2014; IEEE: Piscataway, NJ, USA, 2014; pp. 1592–1597.
43. Cui, Y.; Chen, X.; Zhang, Y.; Dong, J.; Wu, Q.; Zhu, F. Bow3d: Bag of words for real-time loop closing in 3d lidar slam. *IEEE Robot. Autom. Lett.* **2022**, *8*, 2828–2835. [\[CrossRef\]](#)
44. Loubach da Silva Lubanco, D.; Schlechter, T.; Pichler-Scheder, M.; Kastl, C. Survey on Radar Odometry. In Proceedings of the Computer Aided Systems Theory–EUROCAST 2022: 18th International Conference, Las Palmas de Gran Canaria, Spain, 20–25 February 2022; Revised Selected Papers; Springer: Berlin/Heidelberg, Germany, 2023; pp. 619–625.
45. Geng, Z.; Yan, H.; Zhang, J.; Zhu, D. Deep-learning for radar: A survey. *IEEE Access* **2021**, *9*, 141800–141818. [\[CrossRef\]](#)
46. Kuutti, S.; Fallah, S.; Katsaros, K.; Dianati, M.; McCullough, F.; Mouzakitis, A. A survey of the state-of-the-art localization techniques and their potentials for autonomous vehicle applications. *IEEE Internet Things J.* **2018**, *5*, 829–846. [\[CrossRef\]](#)
47. Arshad, S.; Kim, G.W. Role of deep learning in loop closure detection for visual and lidar slam: A survey. *Sensors* **2021**, *21*, 1243. [\[CrossRef\]](#) [\[PubMed\]](#)
48. Roy, P.; Chowdhury, C. A survey of machine learning techniques for indoor localization and navigation systems. *J. Intell. Robot. Syst.* **2021**, *101*, 63. [\[CrossRef\]](#)

49. Yin, P.; Zhao, S.; Cisneros, I.; Abuduweili, A.; Huang, G.; Milford, M.; Liu, C.; Choset, H.; Scherer, S. General Place Recognition Survey: Towards the Real-world Autonomy Age. *arXiv* **2022**, arXiv:2209.04497.
50. Xu, X.; Zhang, L.; Yang, J.; Cao, C.; Wang, W.; Ran, Y.; Tan, Z.; Luo, M. A review of multi-sensor fusion slam systems based on 3D LIDAR. *Remote Sens.* **2022**, *14*, 2835. [\[CrossRef\]](#)
51. Chghaf, M.; Rodriguez, S.; Ouardi, A.E. Camera, LiDAR and multi-modal SLAM systems for autonomous ground vehicles: A survey. *J. Intell. Robot. Syst.* **2022**, *105*, 2. [\[CrossRef\]](#)
52. Huang, L. Review on LiDAR-based SLAM techniques. In Proceedings of the 2021 International Conference on Signal Processing and Machine Learning (CONF-SPML), Beijing, China, 18–20 August 2021; IEEE: Piscataway, NJ, USA, 2021; pp. 163–168.
53. Khan, M.U.; Zaidi, S.A.A.; Ishtiaq, A.; Bukhari, S.U.R.; Samer, S.; Farman, A. A comparative survey of lidar-slam and lidar based sensor technologies. In Proceedings of the 2021 Mohammad Ali Jinnah University International Conference on Computing (MAJICC), Karachi, Pakistan, 15–17 July 2021; IEEE: Piscataway, NJ, USA, 2021; pp. 1–8.
54. Chen, W.; Shang, G.; Ji, A.; Zhou, C.; Wang, X.; Xu, C.; Li, Z.; Hu, K. An overview on visual slam: From tradition to semantic. *Remote Sens.* **2022**, *14*, 3010. [\[CrossRef\]](#)
55. Lowry, S.; Sünderhauf, N.; Newman, P.; Leonard, J.J.; Cox, D.; Corke, P.; Milford, M.J. Visual place recognition: A survey. *IEEE Trans. Robot.* **2015**, *32*, 1–19. [\[CrossRef\]](#)
56. Wang, K.; Ma, S.; Chen, J.; Ren, F.; Lu, J. Approaches, challenges, and applications for deep visual odometry: Toward complicated and emerging areas. *IEEE Trans. Cogn. Dev. Syst.* **2020**, *14*, 35–49. [\[CrossRef\]](#)
57. Duan, C.; Junginger, S.; Huang, J.; Jin, K.; Thuro, K. Deep learning for visual SLAM in transportation robotics: A review. *Transp. Saf. Environ.* **2019**, *1*, 177–184. [\[CrossRef\]](#)
58. Placed, J.A.; Strader, J.; Carrillo, H.; Atanasov, N.; Indelman, V.; Carlone, L.; Castellanos, J.A. A survey on active simultaneous localization and mapping: State of the art and new frontiers. *IEEE Trans. Robot.* **2023**, *39*, 1686–1705. [\[CrossRef\]](#)
59. Favorskaya, M.N. Deep Learning for Visual SLAM: The State-of-the-Art and Future Trends. *Electronics* **2023**, *12*, 2006. [\[CrossRef\]](#)
60. Zeng, F.; Wang, C.; Ge, S.S. A survey on visual navigation for artificial agents with deep reinforcement learning. *IEEE Access* **2020**, *8*, 135426–135442. [\[CrossRef\]](#)
61. Zeng, J.; Wang, D.; Chen, P. A Survey on Transformers for Point Cloud Processing: An Updated Overview. *IEEE Access* **2022**, *10*, 86510–86527. [\[CrossRef\]](#)
62. Yi, Z. A Survey of Radar Perception—Datasets, Methods and Applications. Available online: <https://github.com/ZHOUYI1023/awesome-radar-perception> (accessed on 28 February 2024).
63. Barnes, D.; Gadd, M.; Murcutt, P.; Newman, P.; Posner, I. The oxford radar robotcar dataset: A radar extension to the oxford robotcar dataset. In Proceedings of the 2020 IEEE International Conference on Robotics and Automation (ICRA), Paris, France, 31 May–31 August 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 6433–6438.
64. Burnett, K.; Yoon, D.J.; Wu, Y.; Li, A.Z.; Zhang, H.; Lu, S.; Qian, J.; Tseng, W.K.; Lambert, A.; Leung, K.Y.; et al. Boreas: A multi-season autonomous driving dataset. *Int. J. Robot. Res.* **2023**, *42*, 33–42. [\[CrossRef\]](#)
65. Kim, G.; Park, Y.S.; Cho, Y.; Jeong, J.; Kim, A. Mulran: Multimodal range dataset for urban place recognition. In Proceedings of the 2020 IEEE International Conference on Robotics and Automation (ICRA), Paris, France, 31 May–31 August 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 6246–6253.
66. Sheeny, M.; De Pellegrin, E.; Mukherjee, S.; Ahrabian, A.; Wang, S.; Wallace, A. RADIATE: A radar dataset for automotive perception in bad weather. In Proceedings of the 2021 IEEE International Conference on Robotics and Automation (ICRA), Xi'an, China, 30 May–5 June 2021; IEEE: Piscataway, NJ, USA, 2021; pp. 1–7.
67. Caesar, H.; Bankiti, V.; Lang, A.H.; Vora, S.; Liong, V.E.; Xu, Q.; Krishnan, A.; Pan, Y.; Baldan, G.; Beijbom, O. nuscenes: A multimodal dataset for autonomous driving. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 11621–11631.
68. Palffy, A.; Pool, E.; Baratam, S.; Kooij, J.F.; Gavrila, D.M. Multi-class road user detection with 3+ 1D radar in the View-of-Delft dataset. *IEEE Robot. Autom. Lett.* **2022**, *7*, 4961–4968. [\[CrossRef\]](#)
69. Kramer, A.; Harlow, K.; Williams, C.; Heckman, C. ColoRadar: The direct 3D millimeter wave radar dataset. *Int. J. Robot. Res.* **2022**, *41*, 351–360. [\[CrossRef\]](#)
70. Cheng, Y.; Jiang, M.; Zhu, J.; Liu, Y. Are we ready for unmanned surface vehicles in inland waterways? The usinland multisensor dataset and benchmark. *IEEE Robot. Autom. Lett.* **2021**, *6*, 3964–3970. [\[CrossRef\]](#)
71. Mostajabi, M.; Wang, C.M.; Ranjan, D.; Hsyu, G. High-resolution radar dataset for semi-supervised learning of dynamic objects. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, Seattle, WA, USA, 13–19 June 2020; pp. 100–101.
72. Yan, Z.; Sun, L.; Krajník, T.; Ruichek, Y. EU long-term dataset with multiple sensors for autonomous driving. In Proceedings of the 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Las Vegas, NV, USA, 25–29 October 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 10697–10704.
73. Meyer, M.; Kusch, G. Automotive radar dataset for deep learning based 3d object detection. In Proceedings of the 2019 16th European Radar Conference (EuRAD), Paris, France, 2–4 October 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 129–132.
74. Paek, D.H.; Kong, S.H.; Wijaya, K.T. K-Radar: 4D radar object detection for autonomous driving in various weather conditions. *Adv. Neural Inf. Process. Syst.* **2022**, *35*, 3819–3829.

75. Rebut, J.; Ouaknine, A.; Malik, W.; Pérez, P. Raw high-definition radar for multi-task learning. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 18–24 June 2022; pp. 17021–17030.
76. Goppelt, M.; Blöcher, H.L.; Menzel, W. Automotive radar—investigation of mutual interference mechanisms. *Adv. Radio Sci.* **2010**, *8*, 55–60. [\[CrossRef\]](#)
77. Xu, Z. Bi-Level l1 Optimization-Based Interference Reduction for Millimeter Wave Radars. *IEEE Trans. Intell. Transp. Syst.* **2022**, *24*, 728–738. [\[CrossRef\]](#)
78. Li, X.; Zhang, H.; Chen, W. 4D Radar-based Pose Graph SLAM with Ego-velocity Pre-integration Factor. *IEEE Robot. Autom. Lett.* **2023**, *8*, 5124–5131. [\[CrossRef\]](#)
79. Zhang, J.; Zhuge, H.; Wu, Z.; Peng, G.; Wen, M.; Liu, Y.; Wang, D. 4DRadarSLAM: A 4D Imaging Radar SLAM System for Large-scale Environments based on Pose Graph Optimization. In Proceedings of the 2023 IEEE International Conference on Robotics and Automation (ICRA), London, UK, 29 May–2 June 2023; IEEE: Piscataway, NJ, USA, 2023; pp. 8333–8340.
80. Khan, K.; Rehman, S.U.; Aziz, K.; Fong, S.; Sarasvady, S. DBSCAN: Past, present and future. In Proceedings of the Fifth International Conference on the Applications of Digital Information and Web Technologies (ICADIWT 2014), Chennai, India, 17–19 February 2014; IEEE: Piscataway, NJ, USA, 2014; pp. 232–238.
81. Jose, E.; Adams, M.D. Relative radar cross section based feature identification with millimeter wave radar for outdoor slam. In Proceedings of the 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (IEEE Cat. No. 04CH37566), Sendai, Japan, 28 September–2 October 2004; IEEE: Piscataway, NJ, USA, 2004; Volume 1, pp. 425–430.
82. Jose, E.; Adams, M.; Mullane, J.S.; Patrikalakis, N.M. Predicting millimeter wave radar spectra for autonomous navigation. *IEEE Sens. J.* **2010**, *10*, 960–971. [\[CrossRef\]](#)
83. Wang, S.; Clark, R.; Wen, H.; Trigoni, N. Deepvo: Towards end-to-end visual odometry with deep recurrent convolutional neural networks. In Proceedings of the 2017 IEEE International Conference on Robotics and Automation (ICRA), Singapore, 29 May–3 June 2017; IEEE: Piscataway, NJ, USA, 2017; pp. 2043–2050.
84. Aldera, R.; De Martini, D.; Gadd, M.; Newman, P. Fast radar motion estimation with a learnt focus of attention using weak supervision. In Proceedings of the 2019 International Conference on Robotics and Automation (ICRA), Montreal, QC, Canada, 20–24 May 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 1190–1196.
85. Barnes, D.; Weston, R.; Posner, I. Masking by moving: Learning distraction-free radar odometry from pose information. *arXiv* **2019**, arXiv:1909.03752.
86. Weston, R.; Gadd, M.; De Martini, D.; Newman, P.; Posner, I. Fast-MbyM: Leveraging Translational Invariance of the Fourier Transform for Efficient and Accurate Radar Odometry. In Proceedings of the 2022 International Conference on Robotics and Automation (ICRA), Philadelphia, PA, USA, 23–27 May 2022; IEEE: Piscataway, NJ, USA, 2022; pp. 2186–2192.
87. Barnes, D.; Posner, I. Under the radar: Learning to predict robust keypoints for odometry estimation and metric localisation in radar. In Proceedings of the 2020 IEEE International Conference on Robotics and Automation (ICRA), Paris, France, 31 May–31 August 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 9484–9490.
88. Burnett, K.; Yoon, D.J.; Schoellig, A.P.; Barfoot, T.D. Radar odometry combining probabilistic estimation and unsupervised feature learning. *arXiv* **2021**, arXiv:2105.14152.
89. Ding, F.; Pan, Z.; Deng, Y.; Deng, J.; Lu, C.X. Self-supervised scene flow estimation with 4-d automotive radar. *IEEE Robot. Autom. Lett.* **2022**, *7*, 8233–8240. [\[CrossRef\]](#)
90. Lu, S.; Zhuo, G.; Xiong, L.; Zhu, X.; Zheng, L.; He, Z.; Zhou, M.; Lu, X.; Bai, J. Efficient Deep-Learning 4D Automotive Radar Odometry Method. *IEEE Trans. Intell. Veh.* **2023**, *9*, 879–892. [\[CrossRef\]](#)
91. Almalioglu, Y.; Turan, M.; Lu, C.X.; Trigoni, N.; Markham, A. Milli-RIO: Ego-motion estimation with low-cost millimetre-wave radar. *IEEE Sens. J.* **2020**, *21*, 3314–3323. [\[CrossRef\]](#)
92. Cai, K.; Wang, B.; Lu, C.X. Autoplace: Robust place recognition with low-cost single-chip automotive radar. *arXiv* **2021**, arXiv:2109.08652.
93. Săftescu, Ș.; Gadd, M.; De Martini, D.; Barnes, D.; Newman, P. Kidnapped radar: Topological radar localisation using rotationally-invariant metric learning. In Proceedings of the 2020 IEEE International Conference on Robotics and Automation (ICRA), Paris, France, 31 May–31 August 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 4358–4364.
94. Gadd, M.; De Martini, D.; Newman, P. Look around you: Sequence-based radar place recognition with learned rotational invariance. In Proceedings of the 2020 IEEE/ION Position, Location and Navigation Symposium (PLANS), Portland, OR, USA, 20–23 April 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 270–276.
95. Gadd, M.; De Martini, D.; Newman, P. Contrastive learning for unsupervised radar place recognition. In Proceedings of the 2021 20th International Conference on Advanced Robotics (ICAR), Ljubljana, Slovenia, 6–10 December 2021; IEEE: Piscataway, NJ, USA, 2021; pp. 344–349.
96. Uselli, M.; Frosi, M.; Cudrano, P.; Mentasti, S.; Matteucci, M. RadarLCD: Learnable Radar-based Loop Closure Detection Pipeline. *arXiv* **2023**, arXiv:2309.07094.
97. Yuan, J.; Newman, P.; Gadd, M. Off the Radar: Uncertainty-Aware Radar Place Recognition with Introspective Querying and Map Maintenance. *arXiv* **2023**, arXiv:2306.12556.
98. Wang, W.; de Gusmao, P.P.; Yang, B.; Markham, A.; Trigoni, N. Radarloc: Learning to relocalize in fmcw radar. In Proceedings of the 2021 IEEE International Conference on Robotics and Automation (ICRA), Xian, China, 30 May–5 June 2021; IEEE: Piscataway, NJ, USA, 2021; pp. 5809–5815.

99. Tang, T.Y.; De Martini, D.; Barnes, D.; Newman, P. Rsl-net: Localising in satellite images from a radar on the ground. *IEEE Robot. Autom. Lett.* **2020**, *5*, 1087–1094. [\[CrossRef\]](#)
100. Almalioglu, Y.; Turan, M.; Trigoni, N.; Markham, A. Deep learning-based robust positioning for all-weather autonomous driving. *Nat. Mach. Intell.* **2022**, *4*, 749–760. [\[CrossRef\]](#) [\[PubMed\]](#)
101. Yin, H.; Wang, Y.; Wu, J.; Xiong, R. Radar style transfer for metric robot localisation on lidar maps. *CAAI Trans. Intell. Technol.* **2023**, *8*, 139–148. [\[CrossRef\]](#)
102. Yin, H.; Chen, R.; Wang, Y.; Xiong, R. RaLL: End-to-end Radar Localization on Lidar Map Using Differentiable Measurement Model. *IEEE Trans. Intell. Transp. Syst.* **2021**, *23*, 6737–6750. [\[CrossRef\]](#)
103. Lisus, D.; Laconte, J.; Burnett, K.; Barfoot, T.D. Pointing the Way: Refining Radar-Lidar Localization Using Learned ICP Weights. *arXiv* **2023**, arXiv:2309.08731.
104. Yin, H.; Xu, X.; Wang, Y.; Xiong, R. Radar-to-lidar: Heterogeneous place recognition via joint learning. *Front. Robot. AI* **2021**, *8*, 661199. [\[CrossRef\]](#)
105. Nayak, A.; Cattaneo, D.; Valada, A. RaLF: Flow-based Global and Metric Radar Localization in LiDAR Maps. *arXiv* **2023**, arXiv:2309.09875.
106. Ding, F.; Palffy, A.; Gavrilu, D.M.; Lu, C.X. Hidden gems: 4D radar scene flow learning using cross-modal supervision. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Vancouver, BC, Canada, 17–24 June 2023; pp. 9340–9349.
107. Isele, S.T.; Haas-Fickinger, F.; Zöllner, J.M. SERALOC: SLAM on semantically annotated radar point-clouds. In Proceedings of the 2021 IEEE International Intelligent Transportation Systems Conference (ITSC), Auckland, New Zealand, 1–19 September 2021; IEEE: Piscataway, NJ, USA, 2021; pp. 2917–2924.
108. Kabsch, W. A solution for the best rotation to relate two sets of vectors. *Acta Crystallogr. Sect. Cryst. Phys. Diff. Theor. Gen. Crystallogr.* **1976**, *32*, 922–923. [\[CrossRef\]](#)
109. Barnes, D.; Maddern, W.; Pascoe, G.; Posner, I. Driven to distraction: Self-supervised distractor learning for robust monocular visual odometry in urban environments. In Proceedings of the 2018 IEEE International Conference on Robotics and Automation (ICRA), Brisbane, Australia, 21–25 May 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 1894–1900.
110. Anderson, S.; Barfoot, T.D. Full STEAM ahead: Exactly sparse Gaussian process regression for batch continuous-time trajectory estimation on SE (3). In Proceedings of the 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Hamburg, Germany, 28 September–3 October 2015; IEEE: Piscataway, NJ, USA, 2015; pp. 157–164.
111. Sun, D.; Yang, X.; Liu, M.Y.; Kautz, J. Pwc-net: Cnns for optical flow using pyramid, warping, and cost volume. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 8934–8943.
112. Qi, C.R.; Su, H.; Mo, K.; Guibas, L.J. Pointnet: Deep learning on point sets for 3D classification and segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 652–660.
113. Arandjelovic, R.; Zisserman, A. All about VLAD. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Portland, OR, USA, 23–28 June 2013; pp. 1578–1585.
114. Arandjelovic, R.; Gronat, P.; Torii, A.; Pajdla, T.; Sivic, J. NetVLAD: CNN architecture for weakly supervised place recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 26 June–1 July 2016; pp. 5297–5307.
115. Hoffer, E.; Ailon, N. Deep metric learning using triplet network. In Proceedings of the Similarity-Based Pattern Recognition: Third International Workshop, SIMBAD 2015, Copenhagen, Denmark, 2–14 October 2015; Proceedings 3; Springer: Berlin/Heidelberg, Germany, 2015; pp. 84–92.
116. Khosla, P.; Teterwak, P.; Wang, C.; Sarna, A.; Tian, Y.; Isola, P.; Maschinot, A.; Liu, C.; Krishnan, D. Supervised contrastive learning. *Adv. Neural Inf. Process. Syst.* **2020**, *33*, 18661–18673.
117. Gadd, M.; De Martini, D.; Newman, P. Unsupervised place recognition with deep embedding learning over radar videos. *arXiv* **2021**, arXiv:2106.067031.
118. Kingma, D.P.; Welling, M. Auto-encoding variational bayes. *arXiv* **2013**, arXiv:1312.6114.
119. Uy, M.A.; Lee, G.H. Pointnetvlad: Deep point cloud based retrieval for large-scale place recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 4470–4479.
120. Huang, G.; Liu, Z.; Van Der Maaten, L.; Weinberger, K.Q. Densely connected convolutional networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 4700–4708.
121. Radenović, F.; Tolias, G.; Chum, O. Fine-tuning CNN image retrieval with no human annotation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2018**, *41*, 1655–1668. [\[CrossRef\]](#) [\[PubMed\]](#)
122. Adolfsson, D.; Karlsson, M.; Kubelka, V.; Magnusson, M.; Andreasson, H. TBV Radar SLAM—trust but verify loop candidates. *IEEE Robot. Autom. Lett.* **2023**, *8*, 3613–3620. [\[CrossRef\]](#)
123. Yu, X.; Zhou, B.; Chang, Z.; Qian, K.; Fang, F. MMDF: Multi-Modal Deep Feature Based Place Recognition of Mobile Robots with Applications on Cross-Scene Navigation. *IEEE Robot. Autom. Lett.* **2022**, *7*, 6742–6749. [\[CrossRef\]](#)
124. Zhuo, G.; Lu, S.; Zhou, H.; Zheng, L.; Xiong, L. 4DRVO-Net: Deep 4D Radar-Visual Odometry Using Multi-Modal and Multi-Scale Adaptive Fusion. *arXiv* **2023**, arXiv:2308.06573.
125. Qi, C.R.; Yi, L.; Su, H.; Guibas, L.J. Pointnet++: Deep hierarchical feature learning on point sets in a metric space. *Adv. Neural Inf. Process. Syst.* **2017**, *30*.

126. Tang, T.Y.; De Martini, D.; Wu, S.; Newman, P. Self-supervised localisation between range sensors and overhead imagery. *arXiv* **2020**, arXiv:2006.02108.
127. Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative adversarial networks. *Commun. ACM* **2020**, *63*, 139–144. [[CrossRef](#)]
128. Wang, L.; Goldluecke, B.; Anklam, C. L2R GAN: LiDAR-to-radar translation. In Proceedings of the Asian Conference on Computer Vision, Kyoto, Japan, 30 November 2020.
129. Lee, J.; Shiotsuka, D.; Nishimori, T.; Nakao, K.; Kamijo, S. Gan-based lidar translation between sunny and adverse weather for autonomous driving and driving simulation. *Sensors* **2022**, *22*, 5287. [[CrossRef](#)] [[PubMed](#)]
130. Alhashimi, A.; Adolfsson, D.; Magnusson, M.; Andreasson, H.; Lilienthal, A.J. Bfar-bounded false alarm rate detector for improved radar odometry estimation. *arXiv* **2021**, arXiv:2109.09669.
131. Sun, B.; Gao, S.; Zi, H.; Wu, Q. GAN based simultaneous localization and mapping framework in dynamic environment. *J. King Saud-Univ. Sci.* **2022**, *34*, 102298. [[CrossRef](#)]
132. Teed, Z.; Deng, J. Raft: Recurrent all-pairs field transforms for optical flow. In Proceedings of the Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, 23–28 August 2020; Proceedings, Part II 16; Springer: Berlin/Heidelberg, Germany, 2020; pp. 402–419.
133. Liu, X.; Qi, C.R.; Guibas, L.J. FlowNet3D: Learning scene flow in 3D point clouds. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 529–537.
134. Cheng, Y.; Su, J.; Jiang, M.; Liu, Y. A novel radar point cloud generation method for robot environment perception. *IEEE Trans. Robot.* **2022**, *38*, 3754–3773. [[CrossRef](#)]
135. Geiger, A.; Lenz, P.; Stiller, C.; Urtasun, R. Vision meets robotics: The kitti dataset. *Int. J. Robot. Res.* **2013**, *32*, 1231–1237. [[CrossRef](#)]
136. Adolfsson, D.; Magnusson, M.; Alhashimi, A.; Lilienthal, A.J.; Andreasson, H. Lidar-Level Localization with Radar? The CFEAR Approach to Accurate, Fast, and Robust Large-Scale Radar Odometry in Diverse Environments. *IEEE Trans. Robot.* **2022**, *39*, 1476–1495. [[CrossRef](#)]
137. Jang, H.; Jung, M.; Kim, A. RaPlace: Place Recognition for Imaging Radar using Radon Transform and Mutable Threshold. *arXiv* **2023**, arXiv:2307.04321.
138. Zhou, Z.; Xu, J.; Xiong, G.; Ma, J. LCPR: A Multi-Scale Attention-Based LiDAR-Camera Fusion Network for Place Recognition. *arXiv* **2023**, arXiv:2311.03198.
139. Zhang, F.; Fang, J.; Wah, B.; Torr, P. Deep fusionnet for point cloud semantic segmentation. In Proceedings of the Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, 23–28 August 2020; Proceedings, Part XXIV 16; Springer: Berlin/Heidelberg, Germany, 2020; pp. 644–663.
140. Zhang, Y.; Zhou, Z.; David, P.; Yue, X.; Xi, Z.; Gong, B.; Foroosh, H. Polarnet: An improved grid representation for online lidar point clouds semantic segmentation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 14–19 June 2020; pp. 9601–9610.
141. Chen, X.; Jiang, K.; Zhu, Y.; Wang, X.; Yun, T. Individual tree crown segmentation directly from UAV-borne LiDAR data using the PointNet of deep learning. *Forests* **2021**, *12*, 131. [[CrossRef](#)]
142. Cao, P.; Chen, H.; Zhang, Y.; Wang, G. Multi-view frustum pointnet for object detection in autonomous driving. In Proceedings of the 2019 IEEE International Conference on Image Processing (ICIP), Taipei, Taiwan, 22–25 September 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 3896–3899.
143. Wang, L.; Chen, T.; Anklam, C.; Goldluecke, B. High dimensional frustum pointnet for 3D object detection from camera, lidar, and radar. In Proceedings of the 2020 IEEE Intelligent Vehicles Symposium (IV), Las Vegas, NV, USA, 19 October–13 November 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 1621–1628.
144. Bi, J.; Wei, H.; Zhang, G.; Yang, K.; Song, Z. DyFusion: Cross-Attention 3D Object Detection with Dynamic Fusion. *IEEE Lat. Am. Trans.* **2024**, *22*, 106–112. [[CrossRef](#)]
145. Li, F.; Li, X.; Liu, Q.; Li, Z. Occlusion handling and multi-scale pedestrian detection based on deep learning: A review. *IEEE Access* **2022**, *10*, 19937–19957. [[CrossRef](#)]
146. Orr, I.; Cohen, M.; Zalevsky, Z. High-resolution radar road segmentation using weakly supervised learning. *Nat. Mach. Intell.* **2021**, *3*, 239–246. [[CrossRef](#)]
147. Ouaknine, A.; Newson, A.; Pérez, P.; Tupin, F.; Rebut, J. Multi-view radar semantic segmentation. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, BC, Canada, 11–17 October 2021; pp. 15671–15680.
148. Zhang, C.; Cheng, J.; Tian, Q. Unsupervised and semi-supervised image classification with weak semantic consistency. *IEEE Trans. Multimed.* **2019**, *21*, 2482–2491. [[CrossRef](#)]
149. Pawar, A.; Mago, V. Challenging the boundaries of unsupervised learning for semantic similarity. *IEEE Access* **2019**, *7*, 16291–16308. [[CrossRef](#)]
150. Zhang, A.; Nowruz, F.E.; Laganieri, R. Raddet: Range-azimuth-doppler based radar object detection for dynamic road users. In Proceedings of the 2021 18th Conference on Robots and Vision (CRV), Burnaby, BC, Canada, 26–28 May 2021; IEEE: Piscataway, NJ, USA, 2021; pp. 95–102.
151. Goodfellow, I.J.; Vinyals, O.; Saxe, A.M. Qualitatively characterizing neural network optimization problems. *arXiv* **2014**, arXiv:1412.6544.

152. Subramanian, M.; Shanmugavadivel, K.; Nandhini, P. On fine-tuning deep learning models using transfer learning and hyper-parameters optimization for disease identification in maize leaves. *Neural Comput. Appl.* **2022**, *34*, 13951–13968. [[CrossRef](#)]
153. Dalbah, Y.; Lahoud, J.; Cholakkal, H. RadarFormer: Lightweight and accurate real-time radar object detection model. In Proceedings of the Scandinavian Conference on Image Analysis, Sirkka, Finland, 18–21 April 2023; Springer: Berlin/Heidelberg, Germany, 2023; pp. 341–358.
154. Lee, D.; Nam, H.; Shim, D.H. ELiOT: End-to-end Lidar Odometry using Transformer Framework. *arXiv* **2023**, arXiv:2307.11998.
155. Yoon, D.J.; Zhang, H.; Gridseth, M.; Thomas, H.; Barfoot, T.D. Unsupervised learning of lidar features for use in a probabilistic trajectory estimator. *IEEE Robot. Autom. Lett.* **2021**, *6*, 2130–2138. [[CrossRef](#)]
156. Kim, G.; Kim, A. Scan context: Egocentric spatial descriptor for place recognition within 3D point cloud map. In Proceedings of the 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Madrid, Spain, 1–5 October 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 4802–4809.
157. Kim, G.; Choi, S.; Kim, A. Scan context++: Structural place recognition robust to rotation and lateral variations in urban environments. *IEEE Trans. Robot.* **2021**, *38*, 1856–1874. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.