生活在材料世界：从全波形闪光激光雷达数据中学习材料属性以进行语义分割

Andrej Janda[1], Pierre Merriaux[2],Pierre Olivier[2],和 Jonathan Kelly[1]

[1]多伦多大学空间与地面自主机器人系统实验室，加拿大多伦多{andrej.janda,jonathan.kelly}@robotics.utias.utoronto.ca[2]LeddarTech 公司，加拿大魁˙北克{pierre.merriaux,pierre.olivier}@leddartech.com

arXiv:2305.04334v2 [cs.CV] 4 Sep 2023

摘要——激光雷达技术的进步让三维点云的采集变得快捷而简单。尽管大多数激光雷达传感器会在提供距离测量值的同时返回每个点的强度（或反射率）值，闪光激光雷达传感器还能够提供有关回波脉冲形状的信息。回波波形的形状受多种因素影响，包括光脉冲传播的距离以及光线与表面的入射角。重要的是，回波波形的形状还取决于反射表面的材料属性。在本文中，我们研究能否根据全波形响应确定材料类型或类别。首先，作为概念验证，我们证明，如果能够准确获知额外的材料类别信息，就可以提升场景理解任务（如语义分割）的性能。接下来，我们训练了两种不同的全波形材料分类器：随机森林分类器和时间卷积神经网络（TCN）分类器。我们发现，在某些情况下，不同材料类型可以被区分开来；而在更广泛的材料范围内，TCN 通常表现更佳。然而，入射角、材料颜色以及材料相似性等因素可能会降低整体性能。

I. 引言

点云是一种关键的数据模态，因为它们可用于精确重建三维场景的几何结构。可靠的几何重建使机器人能够与环境进行有意义的交互，执行路径规划、避障和物体操作等任务。点云的价值，加上三维传感器（例如立体相机和激光雷达传感器）日益普及，使三维数据在许多机器人任务中的重要性不断提升。虽然采集点云的方法多种多样，但激光雷达已被证明是许多户外应用中的重要工具；在这些应用中，其他传感器（如相机）的探测距离不足，且在复杂光照条件下功能有限。

场景重建完成后，常见的下一步是使用具有语义意义的标签对场景中的各个部分进行分割。这些语义标签使机器人能够在人类的指引下与环境交互。将原始数据分割为不同的语义类别，仍然是一个充满挑战且尚未解决的问题。近期研究表明，

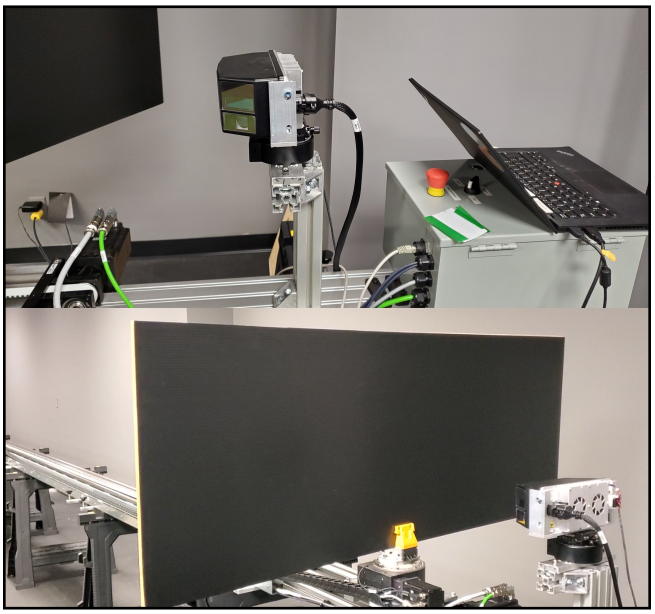


图 1：数据采集的实验装置。全波形 LeddarTech Pixell 激光雷达传感器固定在校准台上，线性平台带动一块安装有待测材料的矩形板移动。该板也可以绕垂直轴旋转。

学习算法可作为一种有效手段，用于分割图像和点云数据。语义分割的难度促使我们充分利用现有传感器技术所能带来的任何优势。

近期文献中大多被忽视的一种潜在的额外有用信息来源，是全波形闪光激光雷达数据的可用性。场景中物体的反射会引起激光雷达脉冲波形的变化，而这种变化可能包含有关物体本身的信息。由于波形会受到信号反射表面与激光之间相互作用的影响，因此有可能利用完整波形来确定表面材料属性。这一应用尚未得到充分探索，并且有望提高性能。

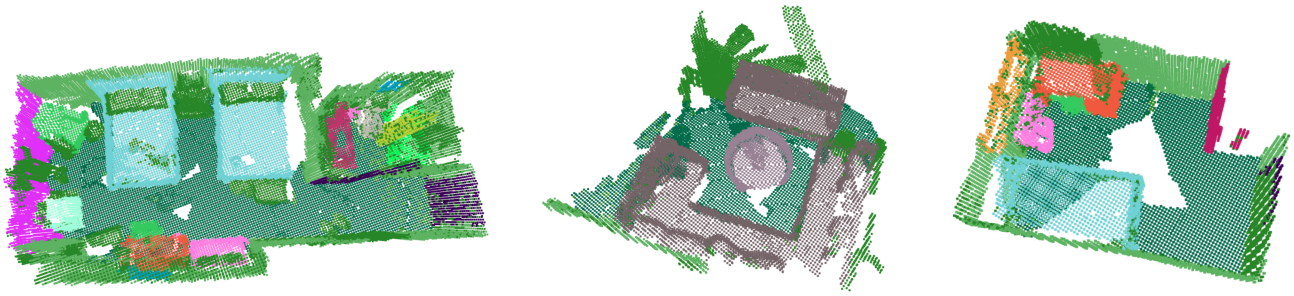


图 2：将材料类别标签添加到 ScanNet 数据集中三个不同场景的重建点云中的可视化结果。每种颜色代表一个独立的材料类别。

现有语义分割算法（通过添加额外的特征通道）。因此，我们研究了利用全波形闪光激光雷达数据确定材料属性的可行性。简而言之，本研究做出了以下贡献：

- 对从各种材料反射回来的全波形激光雷达脉冲数据进行详细分析；
- 展示了结合材料类别信息后分割性能的提升；
- 对两种基于学习的模型进行评估，用于根据激光雷达脉冲数据进行材料分类；并且
- 对波形中哪些部分对材料分类最具信息量的分析。

II. 相关工作

从原始传感器测量值中提取材料类别或属性的问题，已针对许多常见的机器人传感模态开展研究。以往利用激光雷达数据进行材料分类的工作，主要关注如何利用激光脉冲的反射强度，该强度由单个标量值测得。Song 等人 [2] 证明，强度测量值可用于区分各种地面表面类别，例如沥青和草地。Yuan 等人 [3]发现，利用激光数据中的材料反射率和粗糙度以及颜色信息训练的集成方法，对各种建筑材料的分类准确率可超过 97%。类似地，Zahiri 等人 [4]使用多光谱图像和激光束强度值训练支持向量机分类器，以区分不同类型的建筑地基。Tatoglu 和 Pochiraju [5] 应用第一性原理，利用已知的反射率模型计算各类材料的预期强度值。随后，将预期强度与观测到的强度进行比较，以得出最终的材料类别预测。与此不同，我们的工作着重分析回波激光雷达信号的完整波形；与单次强度测量相比，它提供了丰富得多的信息。

除激光雷达外，其他传感器模态也被用于材料分类。[6]、[7]中介绍的方法利用毫米波雷达传感器的全波形测量值，以高精度对材料类型进行分类。

Lu 等人 [8] 使用探地雷达对地下材料进行分类。Saponario 等人 [9] 则利用红外相机捕获的热属性进行材料分类。此外，也有一些研究尝试利用触觉传感进行材料分类 [10], [11]。Zheng 等人 [12] 结合触觉与视觉信息来提取材料类别。

尽管材料分类已在文献中得到广泛研究，但利用闪光激光雷达传感器获取的全波形数据执行这一任务，尚未得到充分探索。值得注意的是，波形信息已被证明在其他应用中具有实用价值。例如，完整的激光雷达波形已被用于：从地形扫描中测量树冠时，从单个脉冲中提取多个检测结果[13]；通过超分辨率提高激光雷达扫描的点密度 [14]；以及通过融合激光雷达扫描与视觉数据来改进土地覆盖分类 [15]。此外，多项研究也发现，基于学习的方法在机载激光扫描（ALS）点云分类方面具有很高的有效性[16]–[19]。全波形测量已在其他应用中展现出的实用价值，正是我们在本文中探究如何利用全波形数据，通过材料识别实现对场景的理解的原因。

III. 材料类别对语义分割有用吗?

在考察能否从全波形激光雷达数据中区分材料之前，我们首先研究这类信息是否有助于语义分割任务。据作者所知，目前不存在同时包含语义标签和材料标签的数据集，因此我们假设每个语义类别对应一种材料类型，以模拟这类信息。我们使用 ScanNet [1]作为基础数据集。ScanNet 是一个室内数据集，包含约 1,600 个重建场景，以及 20 个不同类别的语义标签。我们为每个语义标签指定的材料类别见表 I。我们的映射力求尽可能贴近各类别在现实世界中的材料类型。因此，多个物体可能被映射为同一种材料。图 2 展示了材料类别的可视化结果

材料	类别
石膏板	Wall
乙烯基地板	地板
花岗岩	台面
玻璃	窗户
纸	图片
搪瓷	浴缸、马桶、水槽、冰箱
织物	床、沙发、窗帘、浴帘
Wood	橱柜、椅子、桌子、门、书架、书桌、其他家具

表 I：ScanNet 数据集中每个语义类别的模拟材料关联 [1]。

参数	值
卷积核大小	1
丢弃率	0.05
通道数	32, 32, 32, 64, 64, 64, 128, 128
输出层	线性 + SoftMax
批次大小	32
训练迭代次数	4000
优化器	Adam
学习率	2e-3
动量	0.9

表III：我们的 TCN 分类器模型的网络参数。

其中包括三个不同场景的重建点云。

在此实验中，我们使用 [20], [21] 中的 3D 稀疏 UNet 架构进行语义分割，并采用与 [22] 相同的数据增强和训练参数。在训练和推理阶段，都会将材料类别作为一个额外通道，附加到每个点输入特征图中已有的颜色信息上。我们在表 II 中比较了材料标签对语义分割性能的影响。结果表明，与仅使用颜色信息相比，加入材料类别后，平均交并比提升了 +5.7。这一结果促使我们开展工作，从全波形激光雷达数据中提取材料类别。

IV. 实验设置

在本节中，我们将介绍与本研究相关的细节。首先讨论数据采集流程、所用材料以及标注方法。随后介绍如何将接收到的激光雷达脉冲映射到正确的材料类别。最后说明用于评估结果的具体评估指标。

A. 数据采集与标注

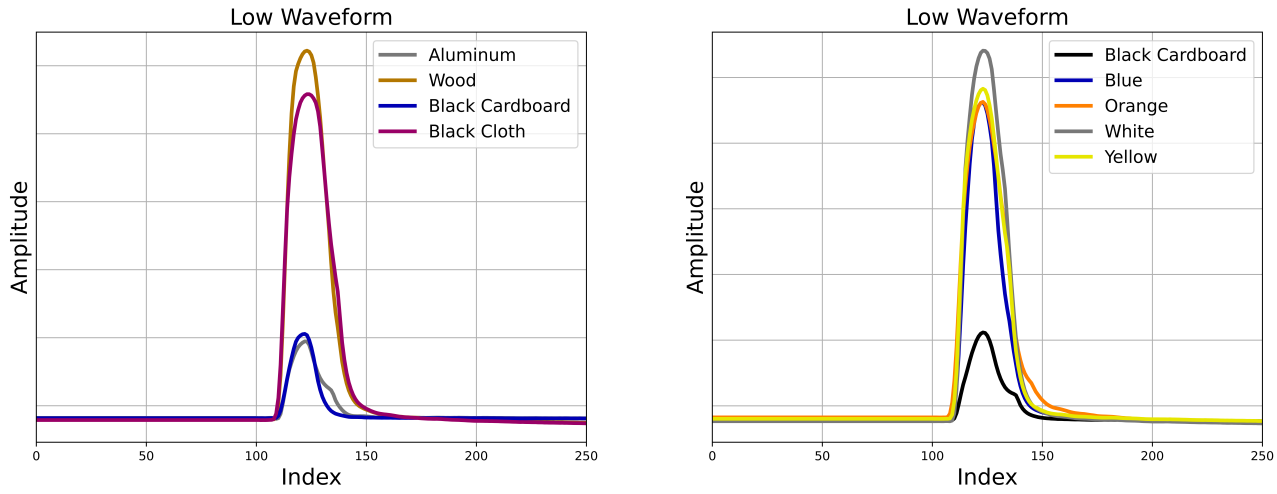
测量是在测试台上进行的，闪光式激光雷达单元固定不动。测试板（板材）安装在线性运动平台上，可向前移动（朝向激光雷达）；板材也可以绕垂直轴旋转（见图1）。为简单起见，板材固定在1米距离处，并在 15°范围内以 $-60°$ 为增量旋转至 $+60°$。每个旋转角度下分别采集五次测量值。激光雷达传感器同时使用高功率和低功率信号。我们选择

可用特征	mIOU
颜色	62.5
颜色 + 材料标签	68.2

表II：在 ScanNet 数据集 [1] 上进行语义分割时，将合成材料类别标签添加到输入特征中的效果。

重点关注低功率信号，因为板材距离过近，在高功率模式下会导致传感器饱和。低功率波形由256个独立测量值组成，我们将其串接为模型的输入向量。图3展示了不同材料的低功耗波形示例。测量时，板上固定一种材料，且材料覆盖板材的整个表面。测试的材料包括：铝、木材、黑色硬纸板和黑布。除不同材料外，我们还采集了不同颜色硬纸板的反射波形。测试的颜色包括：黑色、白色、蓝色、橙色和黄色。将属于板材周围紧邻区域的所有传感器测量值标记为“代表”所测材料类型。其他点则视为未知背景类别。

学习模型： 为了将原始波形数据映射到材料类别，我们采用基于学习的模型。我们使用学习技术，因为不存在简单的判定标准，能够根据各自的材料类别区分单个波形，这一点从图3中显而易见。不同类别的许多波形特征彼此重叠，并且会随反射角度的变化而显著变化。基于学习的方法已证明擅长对复杂的非线性决策边界进行建模。具体而言，我们选择规模相对较小、训练速度较快的模型，因为最终算法需要实时运行，才能使其预测结果可供下游分割模块使用。我们研究的第一种算法是随机森林（RF），这是一种通过多个小型决策树达成共识的集成方法。关于随机森林的详细背景介绍，请参阅 [23]。我们的模型使用 100 棵独立的决策树，每棵树的最大深度为 50。我们研究的第二种算法是时间卷积神经网络（TCN），它通过学习连续的一维卷积层的参数来进行建模。关于 TCN 的详细背景介绍，请参阅 [24]。模型所用的参数列于表III。模型最后一层的输出会经过 softmax 分类函数。模型使用交叉熵损失进行训练。



(a) (b)

图3：使用（左）不同材料的样本和（右）不同颜色的纸板样本，对多次测量得到的低功耗波形取平均值。请注意，振幅是 Pixell传感器输出的原始专有值，为便于阅读，故未予显示。两张图的索引轴均表示传感器捕获的连续测量的位置，从首次发射激光雷达波时开始。每张图的起始处都有一段平坦区域，在此期间，发射的波尚未从表面反射并返回传感器。

评估指标： 为了评估各方法在材料分类方面的性能，我们使用交并比（*IOU*）指标，该指标通常用于评估语义分割任务的性能。IOU 将精确率（即预测正确的比例）和召回率（即正确预测的数据点比例）综合为一个数值。对于涉及多个类别的问题（本研究即属此类），该指标取各类别的平均值，称为平均交并比（mIOU）。

V. 材料分类

本节分析随机森林和时序卷积网络根据全波形数据对材料类型进行分类的性能。我们将材料测试分为三组。第一组是仅使用铝和黑布的简单测试。选择这两种材料是因为它们的反射率差异很大，应该很容易区分（见表IVa）。第二组测试使用了所有可用材料（见表IVb）。最后，我们考察了能否利用波形区分同一种材料的不同颜色（见表IVc）。每项实验均采用（1）不旋转（偏航角）和（2）以 15° 为增量从 -60° 旋转至 $+60^\circ$ 两种设置（标记为 全部）。分类后的点云可视化结果见图4和图6。

我们的结果表明，在第一种也是最简单的情形下（表IVa），分类性能可以相当准确，尤其是在不考虑旋转时。当将木材和黑色硬纸板加入可能的类别列表后，性能会下降。当考虑所有可能的类别和旋转时，两种模型的平均交并比均低于 50%，这对现有的语义分割方法几乎没有提升价值。我们的结果还表明，两种模型都难以区分同一种材料的不同颜色，如表IVc所示。纳入不同纸板样本的所有旋转后，三组实验中的性能最差。因此，我们得出结论：同一种材料类型在不同颜色下产生的波形本质上无法区分。这在一定程度上是意料之中的，因为激光雷达回波会受到反射率和颜色的影响。

除了上述结果外，我们还强调几个重要观察结果。首先，随机森林和 TCN 的表现相似。由于 TCN 是一个表达能力强得多的模型，我们认为，局限在于数据，而非特定模型的选择。其次，我们发现，非零偏航角会导致模型分类性能大幅下降。反射角会改变回波波形，进而可能使其与其他材料在不同角度下测得的波形有很大重叠。

VI. 相对特征重要性

随机森林的一项特殊属性是，它们能够识别特征向量中哪些部分对决策的影响最大。每个元素的影响力根据其在各个决策树的每个节点被选作“分裂”依据的频率来衡量。我们在图中直观地呈现了这种重要性

模型	角度	mIOU
RF	0	86.0
	All	65.0
TCN	0	87.4
	All	64.1

（a）铝和黑布。

模型	角度	mIOU
RF	0	76.4
	All	44.3
TCN	0	76.3
	All	49.7

（b）铝、黑布、木材和黑色硬纸板。

模型	角度	mIOU
RF	0	57.1
	All	17.0
TCN	0	49.8
	All	30.3

(c) 不同颜色的硬纸板。

表 IV：材料分类结果。

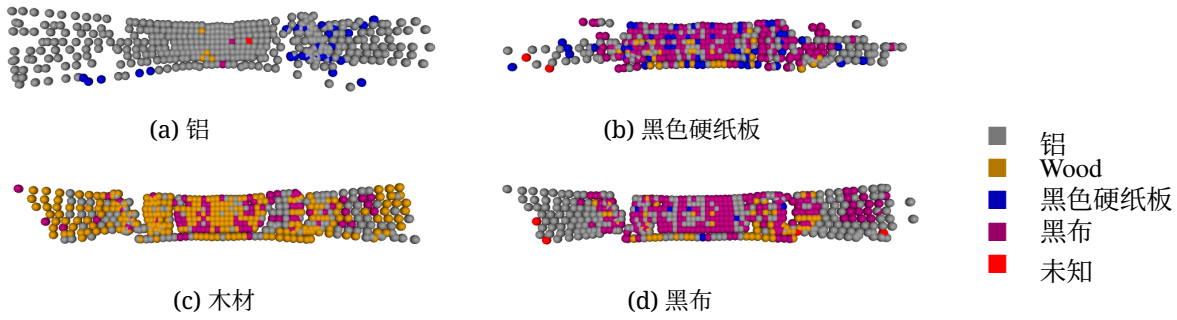


图 4：随机森林模型在所有角度和材料上的分割结果可视化。每个图展示了由单一材料类型构成的板材的分割结果。

图5。与图3中的波形相比，我们可以看到，波形起始处的平台对最终性能没有影响。这仅仅是因为传感器尚未接收到反射信号，因此没有有用的信息。相比之下，峰值之后的区域对最终结果的影响更大。在振幅降至零之前，波形的“尾部”呈现出变化的信号，而这一区域包含一些关于材料的有限信息。然而，最重要的区域是波形达到其

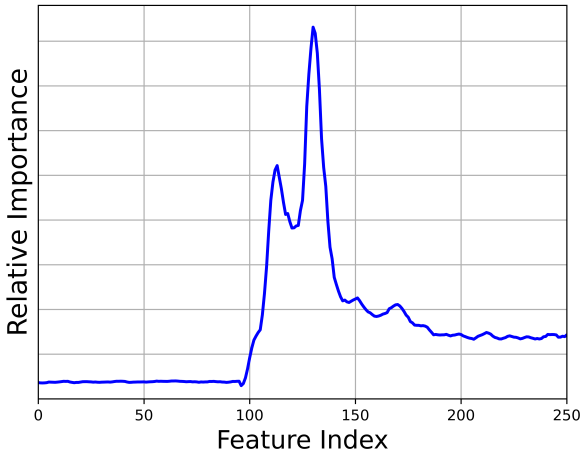


图 5：随机森林进行材料分类时各个特征元素相对重要性的分布。

最大振幅。我们观察到图中有两个峰值，一个对应于未饱和振幅，另一个反映了传感器的饱和状态。饱和现象有助于区分非反射性材料（如黑色硬纸板）和反射性更强的材料（如铝）。进一步分析特征重要性可能会改进特征选择，并能够减小特征向量的大小，同时缩短训练和推理时间。这些步骤将留待未来研究完成。

VII. 结论

在这篇短篇论文中，我们旨在确定是否可以根据全波形闪光激光雷达测量值区分物体的材料类别。我们首先展示了材料知识对于语义分割任务的附加价值。随后，我们介绍了真实世界数据集，并评估了两种不同的基于学习的模型根据波形数据区分材料类别的性能。结果表明，在某些情况下，例如关注少数几种反射率特性差异显著且入射角相近的材料时，材料分类是一种可行的策略。然而，随着不同材料类别的数量和可能的入射角增加，材料分类会变得更加困难，并且最终可能对分割任务的作用有限。

参考文献

[1] A. Dai、A. X. Chang、M. Savva、M. Halber、T. Funkhouser 和 M. Nießner，“ScanNet：丰富标注的 3D 重建-

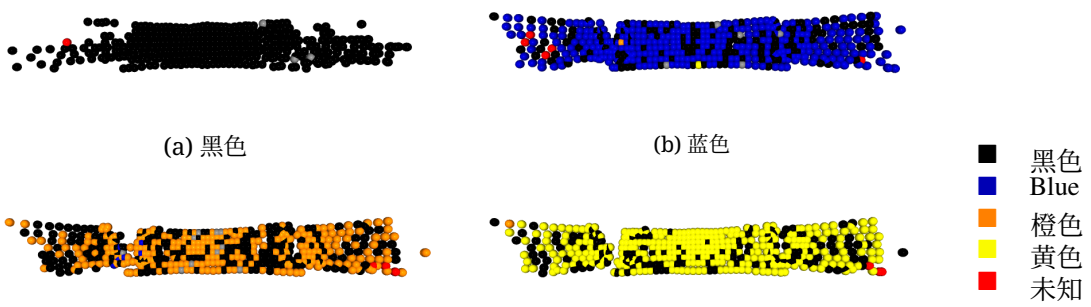


图 6：随机森林模型在所有角度和颜色下的分割结果可视化。每幅图展示了单一材料类型板材的分割结果。

室内场景”，载于*IEEE*计算机视觉与模式识别会议（*CVPR*），2017年，第2432–2443页。[2] J.-h. Song、S.-h. Han、K. Yu 和 Y.-i. Kim，《利用激光雷达强度数据评估土地覆盖分类的可能性》，国际摄影测量、遥感与空间信息科学档案，第34卷，2012年5月。[3] L. Yuan、J. Guo 和 Q. Wang，《从三维地面激光扫描数据自动分类常见建筑材料》，《自动化建筑》，第110卷，第103017页，2020年。[4] Z. Zahiri、D. F. Laefer 和 A. Gowen，《利用多光谱图像和激光雷达强度数据表征建筑材料》，《建筑工程学报》，第44卷，102603，2021年，第。[5] A. Tatoglu 和 K. Pochiraju，《基于激光雷达反射强度特征的点云分割》，载于*IEEE*国际机器人与自动化会议（*ICRA*），2012年，第786–790页。[6] J. Weiß 和 A. Santra，《利用60 GHz雷达和深度卷积神经网络进行材料分类》，载于国际雷达会议（*RADAR*），2019年，第1–6页。[7] ——，《利用毫米波雷达系统进行稳健材料分类的单样本学习》，《*IEEE*传感器快报》，第1–4页，2018年。[8] Q. Lu、J. Pu 和 Z. Liu，《从探地雷达数据中提取地下物体特征并自动分类》，《电气与计算机工程学报》，2014年11月。[9]P. Saponaro、S. Sorensen、A. Kolagunda 和 C. Kambhamettu，《利用热成像进行材料分类》，载于*IEEE*计算机视觉与模式识别会议（*CVPR*），2015年6月，第4649–4656页。[10] S. S. Baishya 和 B. B¨auml，《利用深度学习和触觉皮肤进行稳健材料分类》，载于*IEEE/RSJ*国际智能机器人与系统会议（*IROS*），2016年，第8–15页。[11] N. Jamali 和 C. Sammut，《利用表面纹理通过触觉传感进行材料分类》，载于*IEEE*国际机器人与自动化会议（*ICRA*），2010年，第2336–2341页。[12] H. Zheng、L. Fang、M. Ji、M. Strese、Y. ¨Ozer 和 E. Stein- bach，《利用触觉与视觉信息进行表面材料分类的深度学习》，*IEEE*多媒体汇刊，第18卷，第12期，第2407–2416页，2016年。[13]J. Blair、D. L. Rabine 和 M. A. Hofton，《激光植被成像传感器：一种用于植被测绘的中等高度、纯数字化机载激光高度计，以及地形，*ISPRS*摄影测量与遥感期刊，第54卷，第2期，第115–122页，1999年。

[14] G. Liu 和 J. Ke，“用于超分辨率的全波形激光雷达深度学习”，载于光电成像与多媒体技术 *VI*。SPIE，2019年11月，第38页。[15] H. Wang 和 C. Glennie，“融合波形激光雷达数据与高光谱影像进行土地覆盖分类”，*ISPRS*摄影测量与遥感期刊，第108卷，第1–11页，2015年10月。[16] Z. Yang、W. Jiang、B. Xu、Q. Zhu、S. Jiang 和 W. Huang，“一种基于卷积神经网络的三维语义标注方法，用于机载激光扫描点云”，《遥感》，第9卷，第9期，第936页，2017年9月。[17] Z. Ao、Y. Su、W. Li、Q. Guo 和 J. Zhang，“利用存在与背景学习算法对城市区域中的机载激光雷达数据进行单类分类”，《遥感》，第9卷，第10期，第1001页，2017年9月。[18] S. Zorzi、E. Maset、A. Fusiello 和 F. Crosilla，“利用卷积神经网络对全波形机载激光雷达数据进行分类”，《*IEEE*地球科学与遥感汇刊》，第57卷，第10期，第8255–8261页，2019年10月。[19] N. Chehata、L. Guo 和 C. Mallet，“利用随机森林进行城市分类的机载激光雷达特征选择”，国际摄影测量、遥感与空间信息科学档案，第38卷，2009年1月。[20] C. Choy、J. Gwak 和 S. Savarese，“四维时空卷积网络：Minkowski卷积神经网络”，载于*IEEE*计算机视觉与模式识别会议（*CVPR*），2019年，第3070–3079页。[21] C. Choy，“Minkowski引擎”，https://github.com/NVIDIA/MinkowskiEngine，2019年。[22] J. Hou、B. Graham、M. Nießner 和 S. Xie，“利用对比场景上下文探索数据高效的三维场景理解”，载于*IEEE*计算机视觉与模式识别会议（*CVPR*），2021年，第15 582–15 592页。[23] G. Biau 和 E. Scornet，“随机森林导览”，《西班牙统计与运筹学会杂志》（*TEST*），第25卷，第197–227页，2016年。[24] S. Bai、J. Z. Kolter 和 V. Koltun，“通用卷积网络与循环网络在序列建模中的实证评估”，*arXiv:1803.01271*，2018年。