

基于不一致性的主动学习用于激光雷达目标检测

Esteban Rivera¹, Loic Stratil¹ and Markus Lienkamp¹

Abstract—深度学习模型在自动驾驶中的目标检测方面最近取得了显著的性能提升，并且已经在世界各地的车辆中得到应用。然而，当前的模型需要越来越大的数据集进行训练。获取和标注这些数据成本高昂，因此需要开发新的策略来优化这一过程。主动学习是一种广泛研究的图像领域内的有前景的方法。在我们的工作中，我们将这个概念扩展到激光雷达领域，并通过开发几种基于不一致性的样本选择策略来评估它们在不同设置下的有效性。我们的结果显示，使用一种基于检测框数量的简单不一致性方法，我们能够以 50% 的标注数据达到与随机采样策略相同的 mAP。

I. 介绍

近年来，几家公司在全球部署了自主车队 [1]。尽管取得了这些进展，自动驾驶的广泛应用仍然是一个挑战。从感知的角度来看，主要的挑战之一是检测模型需要处理来自自主车辆遇到的各种环境和场景中的大量传感器数据。捕捉、处理、存储和标注这些数据的成本可能变得越来越高昂，即使是行业领导者也是如此。因此，在业界和学术研究中，一个重要焦点要么是减少对标注数据的依赖，要么是优化标注预算的使用，以获取每花费美元的最大性能提升。

主动学习 (AL) 是一种旨在解决数据稀缺和标注预算限制的战略方法，不仅适用于计算机视觉模型，也适用于一般领域。它利用应用于实际任务的深度学习模型，重点在于训练特定任务的模型也可以识别最有价值的未标记样本以进行标注 [2]。这种方法不随机标注数据，而是优先考虑那些能提供最大信息增益或改进潜力的样本。其基本假设是这种有针对性的标注策略可以减少所需的数据量，并因此降低达到给定任务可接受性能水平所需的预算。主动学习在计算机视觉领域得到了广泛的应用并取得了显著的成功，这一点从各种研究中可以看出 [4], [2]。同样地，用于 AL 的一致性方法，通过评估模型预测之间的分歧来选择标

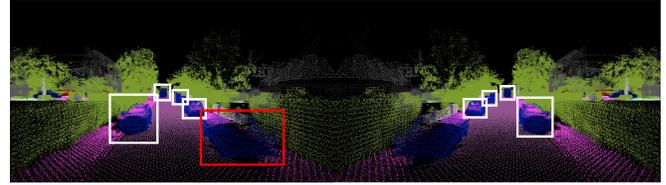


Fig. 1. 点云及其增强版本并排显示，白色方框代表在两个点云中都被 3d 检测器一致识别的对象，红色方框代表只在一个点云中识别的一辆车。这是一个有趣的标签示例，因为模型在其表示上缺乏鲁棒性。来自 KITTI 数据集的图像 [3]。

注的数据点，也展示了它们的有效性 [5], [6], [7], [8]。本文旨在探索这些方法在激光雷达领域进行三维物体检测的潜力。据我们所知，这一应用之前并未得到广泛的研究。

几种方法利用了训练好的检测模型来识别最适合作为标注的样本。专注于不确定性和熵的方法在文献 [9] 中已有详细记录。此外，基于多样性的方法也显示出有希望的结果 [10]。另一个值得注意的策略是集成，或称为委员会查询方法 [11]，它作为样本不确定性的代理。该方法认为如果多个模型对某一样本产生矛盾的预测，则该样本被认为是不确定的。这样的不一致可能表明该样本偏离了训练中使用的分布，或者靠近决策边界，使得分类变得困难。鉴于当前可用的各种模型和数据集，这一策略近年来越来越受欢迎。

不一致性可以通过两种策略进行评估：一方面，可以利用多个模型来比较它们对相同输入的输出。另一方面，也可以仅使用一种模型并通过应用某些类型的增强使得输出结果保持不变，或者变化是可以复制的。基于后一种策略，我们建议采用简单的镜像增强处理，具体针对激光雷达领域，遵循 [5]。在此基础上，我们预测原始点云和镜像点云的边界框，然后通过几种评分策略比较它们，以量化两个预测之间的不一致性水平，请参见 Figure 1。为了量化不一致性，我们提出了两种复杂度逐渐增加的方法，这些方法能够对选择算法中样本级别的信息进行量化，用于主动学习。之后，

This project has been developed in the scope of the BFS-DAD (Bayerische Forschungsförderung - Data-based Autonomous Driving) project

¹Institute for Automotive Engineering, Munich Institute of Robotics and Machine Intelligence Technical University of Munich, name.lastname@tum

我们使用不同的训练范式（微调、重新训练）所选的样本进一步描述哪些场景、训练策略和不一致性能带来更大的改进。

我们的贡献可以概括如下：

- 我们开发了一种基于扩增不一致性仅使用激光雷达的样本选择策略。
- 我们将随机基线 mAP 提高了 2.5%，适用于完整数据范围。
- 我们设计、分析和评估了几种基于不一致性分数采样点云的策略。

II. 相关工作

A. 主动学习

主动学习是优化标注数据的时间和资金预算的最流行策略之一。尽管其概念早已确立 [11]，最近的进步可追溯到深度学习的爆炸性发展。已经开发出多种变体和技术 [2] 以适应所有可能的数据领域。特别是对于计算机视觉，主动学习方法已被应用于提高分类任务的效率 [12], [13]。具体来说，在 2D 检测任务中，这些方法结合了定位和分类以提高所选样本的信息得分。例如，Choi[14] 提出了使用混合密度网络来估计每个定位和分类头输出的概率分布。Aithal[15] 利用一种不确定性度量，该度量考虑了检测、分类和分布统计信息以采样具有更高对象预测得分的图像，旨在实现样本中的平衡分布。为了利用样本的所有可能的预测边界框，Wu[9] 提出了基于熵的非极大值抑制方法，为每个边界框分配一个不确定性分数而不是整个图像，丢弃具有最多冗余信息的样本。最后，Haussman[4] 强调了一个用于实际应用场景（即自动驾驶）的主动学习管道实现的细节和挑战。

B. 不一致性分析

不一致性策略旨在通过多个模型的矛盾输出，或同一场景的不同视角来识别样本。或者，当使用应保持不变的增强输入进行查询时，一个模型预测出不一致的结果。Seung[11] 提出了委员会查询方法，在该方法中，训练简单感知机的样本是从预训练模型之间的最大分歧中确定的。他们还研究了与随机抽样相比，这种方法获得的理论信息增益。具体来说，在计算机视觉领域，几种方法已经使用一致性策略以半监督的方式训练模型 [6], [7]，其中在训练过程中强制执行增强样本之间的一致性。后来，通过同时将不一致性策略

应用于主动学习和半监督学习策略 [5], [16], [17], [18], [19], [20] 来实现了对训练预算的进一步优化。因此，不一致性被主动学习方法用来找到最好的标注样本，并由半监督学习方法来确保对抗增强的鲁棒性。

C. 三维主动学习

主动学习已经在自动驾驶领域中的三维物体检测这一计算机视觉的关键领域中得到应用。这里，同时利用了相机图像和激光雷达点云——它们更能够捕捉环境的三维特性。Schmidt[21] 比较了几种二维和三维主动学习方法在相机目标检测方面的性能。Kao[22] 探讨了物体的定位信息如何影响所选样本的质量。

以下工作利用了自动驾驶应用中更广泛的传感器可用性，而不是仅使用相机数据。例如，Liang[10] 使用 GPS 数据中的信息来选择记录在不同地图位置的样本，确保它们代表各种场景。此外，一些研究试图通过相机和激光雷达之间的信息蒸馏来选择最佳标签样本。Hekimoglu[23] 利用激光雷达点云的三维信息来找到最能提供信息的样本以供相机训练使用。另一方面，Gunnard[24] 提出同时改进激光雷达和相机模型的样本选择。最后，Rivera[25] 提出了比较相机和激光雷达在二维自车平面中的不一致性，以弥补相机检测器中缺乏深度估计，并找到有信息量的样本。

III. 方法

我们方法的核心是开发一种新的策略，在主动学习设置中对 LiDAR 点云进行采样。因此，我们在下面定义了一个主动学习周期和不一致性的概念来量化给定样本的有用性。

A. 主动学习循环

在本文中，我们提出了以下主动学习周期：使用 KITTI 数据集，我们首先用 10% 的可用标注数据训练一个 3D 检测模型，即 371 个点云。这作为主动学习周期的起点。首先，该模型用于对剩余 90% 的主动学习数据集或工作集进行边界框预测。其次，根据开发的策略选择样本，每个周期占原始数据集的 10%，即从工作集中选出 371 个样本。在此过程中，使用真实标签来模拟人类标注过程。第三，这 371 个样本随后被添加到标注训练集中，并用它来训练物体检测器。这就是一个完整的主动学习周期。该周期重复进行直到所有可用的标注数据都被完全使用，在每次迭代中都会增加原始数据集的 10%。随机基线用于对不同的样

本选择方法进行定量比较，每个周期都采用随机样本选择，并在每次迭代时从头开始训练。

B. 不一致定义

为了选择主动学习循环中的有趣样本，提出了一种基于不一致性的方法。该方法基于 Query-by-committee 的想法，与使用多个模型对单一样本进行预测不同，我们建议增强样本，在我们的案例中是一个 LiDAR 点云，并获取原始样本和增强样本的预测结果。这一想法背后的原理类似于 Elezis 的工作 [5]，即模型应该普遍对弱增强具有鲁棒性，如镜像或输入的小位移。因此，对于那些模型不稳健且检测结果不一致的样本应考虑进行真实标签标注。与纯粹的 2D 图像情况相比，点云增强的多样性有限，因为亮度、对比度或着色增强不再适用。因此，我们只关注点云的水平反射，因为这种增强代表了实际生活中的场景。

下一个需要考虑的方面是如何比较原始点云和增强后的点云以找到不一致之处。交集比并集 (IoU) 已经在文献中被提出用于相机 2D 和 3D 检测 [23]；并且类别不符已在 2D 检测中使用 [5]。为此，我们比较单模态激光雷达方法中的框数 [25]。从此比较中，不一致得分 S_{NoB} 为：

$$S_{NoB} = \frac{|N_o - N_a|}{\max(N_o, N_a)}, \quad (1)$$

其中 N_o 是原始点云检测到的框数，而 N_a 是增强后的点云检测到的框数。该分数被归一化至检测到的最大框数，以获得相对不一致性的概念。我们丢弃那些 N_o 和 N_a 都为 0 的样本。相对不一致性指标可以通过以下示例来说明：一个样本中有两个框被检测到，在增强版本中四个框被检测到，相对来说比另一个样本更不一致，其中 20 个框被检测到，而在增强版本中 22 个框被检测到。尽管在两种情况下绝对边界框不一致性都为两，第一种情况的相对不一致性是 0.5，第二种情况则是 0.09。我们称这种方法为箱数不一致分数 (NoB)。

为了完整性，我们还评估了一个基于 IoU 的不一致性得分的表现，利用 3D IoU 进行跨模态比较。只有当该值超过阈值时，我们才接受该框为匹配项。匹配的数量被累加为 N_m ，然后得分 S_{IoU} 被计算为：

$$S_{IoU} = \frac{\max(N_o, N_a) - N_m}{\max(N_o, N_a)}. \quad (2)$$

IV. 实验

除了测试不同的不一致性计算方法外，我们还测试了两种使用选定的不一致性进行模型训练的策略。一方面，在**刮擦**设置中，我们在当前周期内选取的样本以及之前周期中选取的样本的基础上从头开始训练模型。另一方面，在**重新训练**设置中，我们将前一次迭代中的已训练检查点作为起点，并对每次主动学习周期进行额外的 80 个 epoch 的训练。我们测试了**重新训练**设置对于 S_{NoB} 和 S_{IoU} 两种不一致性类型。

A. 模型

所有实验均使用 PointPillars 架构进行，因为其训练速度较快。鉴于我们不追求最佳性能，而是寻找最大的相对改进，快速模型使我们在有限的时间窗口内能够执行更多的实验。实验中的训练设置保持不变，并定义为：周期：80，学习率：0.001，优化器：ADAM。实验中使用了 KITTI 数据集。仅使用带有标签的原始训练集，包含 7481 个点云。将其分为两部分：3712 用于主动学习循环，3769 用于测试。

V. 结果

A. 不一致排序

使用定量不一致度量进行样本选择允许将它们按降序或升序排列。因此，作为一种概念验证，我们设计了一个实验来初步比较两种可能的排序方式，即降序或升序。在这两种排序中，首先使用不一致度大于 0 的样本。当没有更多这样的样本时，使用剩余的样本完成训练。

我们评估了升序与降序采样策略的有效性，采用了一个伪主动学习周期。其功能类似于主动学习周期，主要区别在于样本仅在第一个周期中排序一次。然后保持这一样本顺序直到所有样本被选完。因此，在随后的每个周期中，最合适的 10%，即下一批次的样本会被添加到训练集中，直至所有样本都被使用。性能结果见 Figure 2。对于随机基线，标准差也被绘制成一个区域。

与我们的预期相反，在低数据量情况下 (20%-40%)，按降序利用样本的表现比随机基线更差。当使用超过 50% 的可用数据时，其表现与随机基线相似。对于升序方法，与随机采样相比没有明显改进，但在整个数据范围内是稳定的。根据这些发现，我们在后续的主动学习周期中采用升序选择样本。

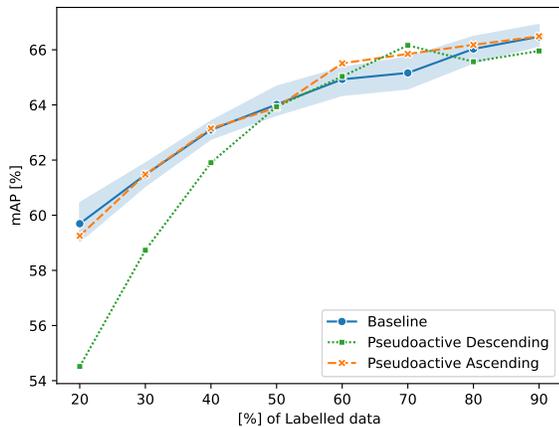


Fig. 2. 伪主动学习循环结果。基线的偏差显示为蓝色区域

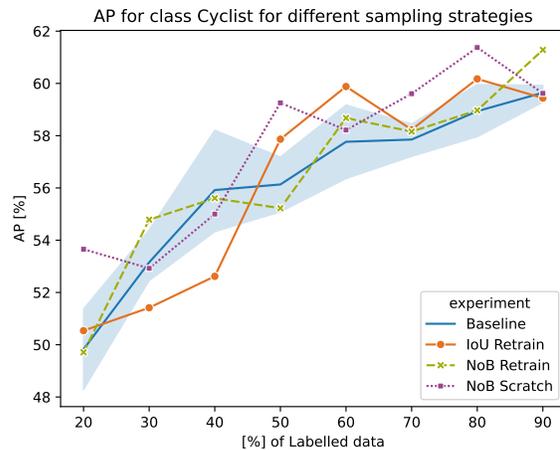


Fig. 4. 自行车运动员的结果

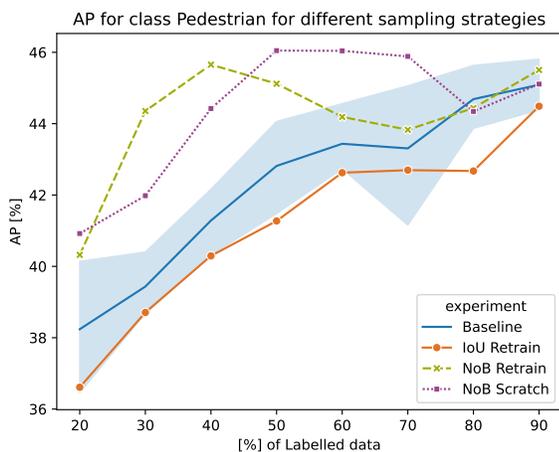


Fig. 3. 行人结果

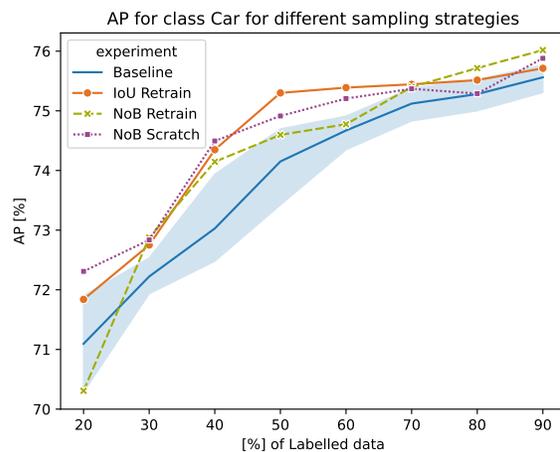


Fig. 5. 汽车的结果

B. 主动学习循环

结果按类别展示，即行人、骑车人和汽车，因为 AP 曲线在标注数据范围内的表现不同。在每个图中，展示了 Section IV 中解释的三种设置以及随机基线。每次实验用不同的种子运行三次，并绘制最佳值。

对于 Figure 3 中的行人案例，在 NoB Scratch 和 NoB Retrain 设置中都观察到了一个有趣的趋势。在低至中等数据范围 (20%-70%)，它们的表现都优于基线，显示出在可用数据的 30% 到 40% 之间有急剧的增长。对于 NoB Retrain 设置，平均精度 (AP) 在 40% 的标记数据处达到峰值，随后逐渐下降至基线值。另一方面，NoB Scratch 设置则一直持续到 70% 的数据才达到平稳期。最后，在整个标记数据范围内，IoU 设置的表现均低于基线。

骑行者的结果，在 Figure 4 中显示，从标签效率角度来看是有趣的，因为该类在完整的 KITTI 数据集中代表性不足，仅有约 4.67% 的标签。因此，观察所提出的策略在期望类别实例更少时的表现是很有趣的。可以看出，对于 40% 或更少的数据，三种设置的 AP 值都低于或等于随机基线。对于两种 NoB Retrain 设置，这一趋势在整个范围内持续存在。另一方面，NoB Scratch 设置略好于随机基线，平均高出 1-2%。

与自行车手相反，汽车是数据集中最占优势的类别，约占所有标签的 82.5%。如 Figure 5 所示，所有设置的行为都比之前更稳定。在这种情况下，三种设置都表现出优于基线的 AP。NoB Scratch 和 IoU Retrain 在中低数据量下表现更强，其 AP 至少比基线高出 1%，而 NoB Retrain 在高数据量下的表现略好一些。

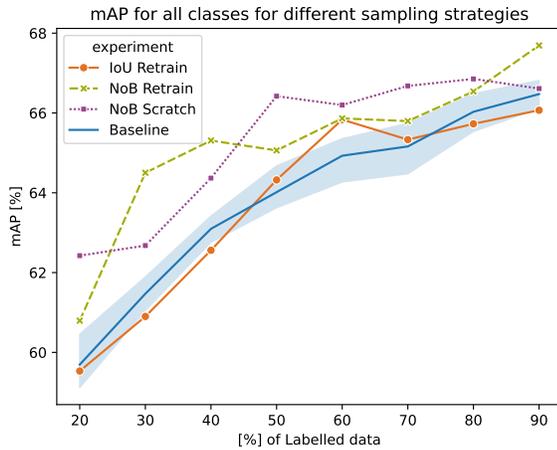


Fig. 6. 所有类别的 mAP 结果

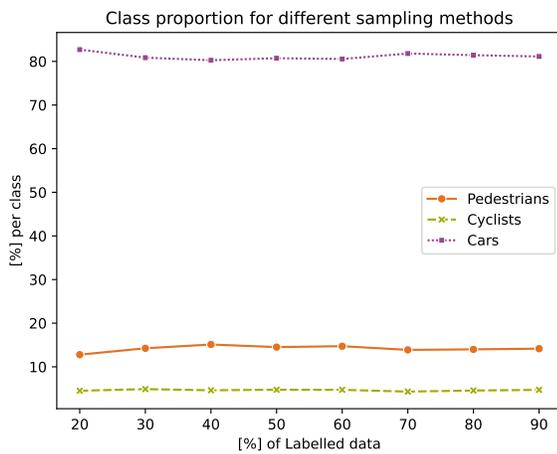


Fig. 7. 循环中的类别分布

在 Figure 6 中，可以观察到所有类别的 mAP 结果。NoB Scratch 设置与基线相比效果最佳，改进幅度在整个范围内介于 1% 到 2% 之间。从经济角度来看，它实现了与随机抽样相同的 mAP，但只需要一半的标注点云。NoB Retrain 设置也优于随机基线，在数据量较低的情况下改进更加明显，仅使用 30% 的数据即可达到随机基线 60% 或 NoB Scratch 40% 数据的性能。另一方面，基于 IoU 的设置大多数数据范围内表现低于随机基线，除了在 50% 到 70% 的数据之间。

C. 类别比例

在 Figures 3 to 5 中，可以观察到所有类别的行为在整个可用数据范围内并不恒定。对于行人情况 Figure 3，主动学习方法在低至中等数据量区间的表现优于基线，在区间的末尾达到最大值；对于骑行者

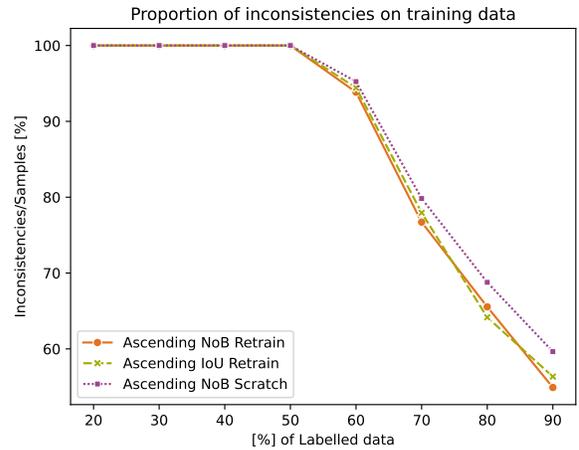


Fig. 8. 不一致性在整个周期中的分布

Figure 4，尽管曲线不够平滑，但在高数据量区间可以见到小幅提升。最后，对于汽车类别 Figure 5，在整个数据范围内存在一个恒定但微小的改进。为了说明这种行为，我们在 Figure 7 的 NoB Scratch 设置中绘制了类分布随数据范围的变化图。可以看出，类分布并未随着数据范围的变化而变化。因此，曲线沿区间的变化并非由数据分布的变化引起。

D. 不一致比例

可以从结果看出，相对于随机基线采样，在低、中和高数据范围内的相对性能是不同的。可能的原因之一可能是每次迭代中发现的不一致样本比例不同。与总可用数据相比，训练集越大，剩余需要找到的不一致样本就越少。因此，随着主动学习周期的增加，即使包含了一致样本以满足每个周期的选择要求，最终使用的仍然是随机采样策略。在 Figure 8 中显示了三个提议实验的不一致性比例。值得注意的是，所有实验变体的行为相似，主动递增方法在范围结束时有稍微多一些的不一致情况。主要观察结果是，训练集仅由不一致样本组成，直到可用数据的 50%。超过这个分布后，就会选择一致样本用于主动学习周期。然后，即使模型在每次迭代中重新训练，也不会再找到更多的不一致样本，并开始直接使用一致样本。

E. 进一步的实验

我们还评估了训练类型、排序或不一致性的类型如何影响采样策略的性能，如 Table I 所示。关于不一致性类型，可以看到，在相同的重新训练设置下，朴素的 NoB 不一致性比 IoU 不一致性表现好近 3%，其

TABLE I
相对于随机基线的几种设置改进

Ordering	Inconsistency	Training	Normalised	Improvement (%)
Ascending	NoB	Retrain	Yes	2.43
Ascending	NoB	Scratch	Yes	2.58
Ascending	NoB	Fine-tuning	Yes	-5.76
Ascending	IoU	Retrain	Yes	-0.13
Ascending	NoB	Retrain	No	-0.11
Descending	NoB	Retrain	Yes	-0.05

中 IoU 不一致性实际上比随机基线差。使用 NoB 不一致性，但在每个周期中仅对新出现的不一致性进行微调的模型比使用新样本加上上一次迭代中的样本进行训练的表现差 8% 左右。重新训练的情况下，即使用上一次迭代的检查点作为起点，其表现比最佳情况低 0.15%，因此仍然优于随机基线。最后，跳过对不一致分数的归一化会使表现低于随机基线 0.11%。

VI. 讨论

在审查了呈现的结果后，出现了一些见解。首先，检查 Equation (1)，可以看出检测到的对象数量较高的样本表现出较低的不一致性评分，而对对象较少的样本则评分较高，这归因于得分的标准化。其次，对 Figure 2 和 Table I 的分析突显了一个明确的趋势：最有效的采样策略是按升序排列，优先考虑分数较低的样本而不是分数较高的样本。因此，这些发现表明，最优的采样策略是在前几个周期中优先考虑包含更多对象的样本，并将不一致性评分作为对象数量的代理指标。这种做法在需要为每帧而非每个边界框支付标注费用的情况下尤其具有经济效益。

一类性能比较导致了进一步有趣的观察。对于“汽车”类别，改进在整个标注数据范围内，包括随机基线，仍然相对微小，徘徊在 1% 左右。这表明该类别的性能接近饱和，即使是超过基线的小幅提升也显得有希望。此外，所有采样策略的平均精度 (AP) 保持恒定并稳定在随机基线之上，表明观察到的改进不仅仅是数据集或模型内部随机波动的结果。

相比之下，“骑车人”和“行人”类别在标记数据范围内经历了更显著的性能提升，初始和最终数据点之间的 AP 差异约为 10%。值得注意的是，“骑车人”的 AP 改进波动较大，在高数据范围 (70%-90%) 内观察到主要增益。相比之下，“行人”类——采用两种数量框方法——在整个数据范围内始终优于基线。重要的是，从类别分布的角度来看，“骑车人”类别严重不

足——与“汽车”甚至“行人”类别相比。这表明存在一个较低的样本量限制，在此之下可用的数据可能无法提供足够的信息以允许使用我们的采样策略进行改进。然而，“行人”类尽管相对于“汽车”类别而言代表性不足，但仍取得了显著改善。

这一发现支持以下观点：为了使我们的主动学习采样策略高效工作，一个类别必须拥有足够的未使用信息或熵水平，以允许对检测器进行有影响力的训练周期。在我们的情况下，“汽车”类在初始检测器训练后已经被很好地识别，因为“汽车”的样本集非常大，导致信息增益较低。“行人”类最初由检测器无法稳健地检测到，但可用的样本集足够大，可以进行有影响力的学习，从而导致信息增益较高。如上所述，“骑车人”类代表性不足，导致信息增益较低。

有趣的是，研究在相同训练方法论下两种不一致类型的差异是有洞察力的，特别是关于边界框的数量和 (IoU)。正如几个先前的研究 [23], [21], [24] 所指出的，3D 检测方法的一致性策略通常依赖于样本间边界框之间的 IoU 匹配进行比较。类似于 [25]，我们的结果表明，基于边界框数量的一致性策略在所有数据范围内都超过了基于 IoU 的那些策略，平均提升了 2.56%。我们假设这种优势源于 IoU 不一致性的内在局限性，这与边界框的精确定位密切相关。对于距离 LiDAR 较远的对象，IoU 匹配极易受到微小变化的影响，使其产生的不一致性过于嘈杂，掩盖了任何可能有助于在定位误差噪声中进行训练的有用信息。另一方面，使用边界框数量作为不一致性的直接度量标准平滑了用于选择样本的得分。因此，该策略对局部化的小变化更加稳健，并专注于寻找使样本对于训练而言具有信息价值的特征。

最后，我们概述了几个未来工作的方向以进一步提高对该领域的理解。首先，我们计划整合更多的数据集。尽管 KITTI 数据集在社区内仍然备受推崇，但

其相对较小的规模限制了分析的深度，特别是在数据预算约束方面。其次，结合半监督学习方法和主动学习已经被证明可以在有限的数据预算下显著提升性能[5]。此外，考虑到框数不一致性指标的相对成功，直接将该指标整合到损失函数中可能会带来益处。这种整合将在训练过程中强制保持检测框数量的一致性，可能带来更稳健的模型。最后，为了评估我们提出方法的可迁移性，适应现有的最先进的三维检测主动学习策略以使用基于框数的一致性度量而非目前依赖的IoU基础不一致性测度将非常有趣。

VII. 结论

本文中，我们介绍了一种在主动学习框架下选择激光雷达样本的新策略，该策略基于点云与其水平镜像增强之间的一致性差异。我们探讨了两种一致性得分：一种基于框数的方法，另一种基于匹配框之间的IoU，将两者与随机采样基线进行比较。我们的分析将结果分为KITTI数据集的三个类别，提供了对观察到的行为的见解。此外，我们还研究了训练方法和分数归一化对我们方法性能的影响。我们的发现表明，在我们的测试场景中，仅基于框数的一致性得分超过了基于IoU和随机基线的得分。这一见解有可能改进当前依赖于IoU得分和定位损失的最先进的方法。

致谢

埃斯特万·里韦拉作为第一作者，开发了该方法，实现了代码并撰写了论文；洛伊克·斯特拉蒂尔对分析做出了贡献并撰写了论文，马库斯·莱因坎普对该研究项目的设计作出了重要贡献。他对论文进行了重要的学术内容的批判性修订。马库斯·莱因坎普批准了即将发表的版本，并对工作的各个方面表示同意。作者们想感谢慕尼黑机器人与机器智能研究所（MIRMI）的支持。

REFERENCES

- [1] Waymo. (2023) Waymo significantly outperforms comparable human benchmarks over 7+ million miles of rider-only driving. Accessed: 01, 11, 2024. [Online]. Available: <https://waymo.com/blog/>
- [2] P. Ren, Y. Xiao, X. Chang, P.-Y. Huang, Z. Li, B. B. Gupta, X. Chen, and X. Wang, "A survey of deep active learning," *ACM Comput. Surv.*, vol. 54, no. 9, oct 2021. [Online]. Available: <https://doi.org/10.1145/3472291>
- [3] A. Geiger, P. Lenz, and R. Urtasun, "Are we ready for autonomous driving? the kitti vision benchmark suite," in *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2012.
- [4] E. Haussmann, M. Fenzi, K. Chitta, J. Ivanecy, H. Xu, D. Roy, A. Mittel, N. Koumchatzky, C. Farabet, and J. M. Alvarez, "Scalable active learning for object detection," in *2020 IEEE Intelligent Vehicles Symposium (IV)*. IEEE Press, 2020, p. 1430 – 1435. [Online]. Available: <https://doi.org/10.1109/IV47402.2020.9304793>
- [5] I. Elezi, Z. Yu, A. Anandkumar, L. Leal-Taixé, and J. M. Alvarez, "Not all labels are equal: Rationalizing the labeling costs for training object detection," in *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022, pp. 14472–14481.
- [6] K. Sohn, D. Berthelot, C.-L. Li, Z. Zhang, N. Carlini, E. D. Cubuk, A. Kurakin, H. Zhang, and C. Raffel, "Fixmatch: Simplifying semi-supervised learning with consistency and confidence," in *Proceedings of the 34th International Conference on Neural Information Processing Systems*, ser. NIPS'20. Red Hook, NY, USA: Curran Associates Inc., 2020.
- [7] M. Gao, Z. Zhang, G. Yu, S. Ö. Arık, L. S. Davis, and T. Pfister, "Consistency-based semi-supervised active learning: Towards minimizing labeling cost," in *Computer Vision – ECCV 2020*, A. Vedaldi, H. Bischof, T. Brox, and J.-M. Frahm, Eds. Cham: Springer International Publishing, 2020, pp. 510–526.
- [8] A. Ghita, B. Antoniussen, W. Zimmer, R. Greer, C. Creß, A. Møgelmoose, M. M. Trivedi, and A. C. Knoll, "Activeanno3d – an active learning framework for multi-modal 3d object detection," 2024.
- [9] J. Wu, J. Chen, and D. Huang, "Entropy-based active learning for object detection with progressive diversity constraint," in *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Los Alamitos, CA, USA: IEEE Computer Society, jun 2022, pp. 9387–9396. [Online]. Available: <https://doi.ieeecomputersociety.org/10.1109/CVPR52688.2022.00918>
- [10] Z. Liang, X. Xu, S. Deng, L. Cai, T. Jiang, and K. Jia, "Exploring diversity-based active learning for 3d object detection in autonomous driving," *ArXiv*, vol. abs/2205.07708, 2022. [Online]. Available: <https://api.semanticscholar.org/CorpusID:248811445>
- [11] H. S. Seung, M. Opper, and H. Sompolinsky, "Query by committee," in *Proceedings of the Fifth Annual Workshop on Computational Learning Theory*, ser. COLT '92. New York, NY, USA: Association for Computing Machinery, 1992, p. 287 – 294. [Online]. Available: <https://doi.org/10.1145/130385.130417>
- [12] R. Caramalau, B. Bhattarai, D. Stoyanov, and T.-K. Kim, "Mobyv2al: Self-supervised active learning for image classification," in *33rd British Machine Vision Conference 2022, BMVC 2022, London, UK, November 21-24, 2022*. BMVA Press, 2022. [Online]. Available: <https://bmvc2022.mpi-inf.mpg.de/0674.pdf>
- [13] H. Ranganathan, H. Venkateswara, S. Chakraborty, and S. Panchanathan, "Deep active learning for image classification," in *2017 IEEE International Conference on Image Processing (ICIP)*, 2017, pp. 3934–3938.
- [14] J. Choi, I. Elezi, H.-J. Lee, C. Farabet, and J. M. Alvarez, "Active learning for deep object detection via probabilistic modeling," 2021.

- [15] S. J. Aithal, S. P. Adhikari, M. Ghorai, and H. Misra, “Balanced sampling-based active learning for object detection,” in *Computer Vision and Machine Intelligence*, M. Tistarelli, S. R. Dubey, S. K. Singh, and X. Jiang, Eds. Singapore: Springer Nature Singapore, 2023, pp. 323–333.
- [16] J. Lim, J. Na, and N. Kwak, “Active semi-supervised learning by exploring per-sample uncertainty and consistency,” 2023.
- [17] M. Cai, B. Wei, X. Li, and L. Li, “Semi-supervised active learning based on semantic-aware crop consistency,” 2022, pp. 655–660.
- [18] M. H. Aneesh Rangnekar, Christopher Kanan, “Semantic segmentation with active semi-supervised learning.” 2023 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), 2023.
- [19] S. Chen, Y. Yang, and Y. Hua, “Semi-supervised active learning for object detection,” *Electronics*, vol. 12, no. 2, pp. 375–375, 2023.
- [20] S. Huang, T. Wang, H. Xiong, J. Huan, and D. Dou, “Semi-supervised active learning with temporal output discrepancy,” in *ICCV*, 10 2021, pp. 3427–3436.
- [21] S. Schmidt, Q. Rao, J. Tatsch, and A. Knoll, “Advanced active learning strategies for object detection,” in *2020 IEEE Intelligent Vehicles Symposium (IV)*, 2020, pp. 871–876.
- [22] C.-C. Kao, T.-Y. Lee, P. Sen, and M.-Y. Liu, “Localization-aware active learning for object detection,” in *Computer Vision – ACCV 2018*, C. Jawahar, H. Li, G. Mori, and K. Schindler, Eds. Cham: Springer International Publishing, 2019, pp. 506–522.
- [23] A. Hekimoglu, M. Schmidt, and A. Marcos-Ramiro, “Monocular 3d object detection with lidar guided semi supervised active learning,” in *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, January 2024, pp. 2346–2355.
- [24] A. Gunnargård and D. Odin, “Camera-lidar active learning for object detecting deep neural networks,” Master’s thesis, University of Chalmers, Chalmers, Sweden, June 2020, master’s thesis.
- [25] E. Rivera, A. Serra Do Nascimento, and M. Lienkamp, “Camera-lidar inconsistency analysis for active learning in object detection,” in *35th IEEE Intelligent Vehicles Symposium, IV 2024*, ser. IEEE Intelligent Vehicles Symposium, Proceedings. Institute of Electrical and Electronics Engineers Inc., 2024, pp. 97–103, publisher Copyright: © 2024 IEEE.; 35th IEEE Intelligent Vehicles Symposium, IV 2024 ; Conference date: 02-06-2024 Through 05-06-2024.