arXiv:2302.02928v2 [cs.CV] 4 Sep 2023

生成连续驾驶空间中的证据式BEV地图

Yunshuang Yuan[a]、Hao Cheng[b]、Michael Ying Yang[b]、Monika Sester[a]

[a] *Institute of Cartography and Geoinformatics, Leibniz University Hannover, Germany*
[b] *Scene Understanding Group, ITC Faculty, University of Twente, The Netherlands*

摘要

安全对于自动驾驶至关重要，而提升安全性的一个方面是准确捕捉感知系统的不确定性，尤其是识别未知信息。不同于仅提供确定性或概率性结果的方法，例如，概率目标检测，这类方法只能为感知场景提供部分信息。我们提出了一种名为 GevBEV 的完整概率模型。该模型将二维驾驶空间表示为概率鸟瞰视图（BEV）地图，其中包含基于点的空间高斯分布；利用这些分布，可以为连续驾驶空间中任意新的采样点提取证据，并将其作为分类狄利克雷分布的参数。实验结果表明，在仿真和真实世界驾驶场景下的协同感知 BEV地图解释基准测试集 OPV2V 和 V2V4Real 上，GevBEV 不仅能提供更可靠的不确定性量化，性能也优于以往工作。协同感知的一个关键因素是通过通信信道传输的数据量。GevBEV 根据学习到的不确定性，仅选择最重要的信息进行共享，从而减少通信开销；在性能仅略有下降的情况下，平均通信信息量减少了 87%。我们的代码已发布于https://github.com/YuanYunshuang/GevBEV。

关键词：证据深度学习、语义分割、协同感知、鸟瞰视图

1. 引言

近几十年来，人们为自动驾驶汽车（AV）的感知系统以及众多摄影测量和遥感任务开发了大量算法，例如，[1, 2, 3, 4, 5], 还开发了许多其他算法。得益于开源数据集，例如，[6, 7, 8]，以及用于解读所采集数据的不同标准化感知任务的相应基准测试集，例如，目标检测和语义分割，我们能够通过将这些算法的预测结果与人工标注的真实值进行比较，来评估其性能。然而，算法在这些基准测试集上取得更好的性能，是否就是我们应该追求的唯一目标?

显然不是；这一目标不足以确保系统能够可靠地部署于现实世界中。

自动驾驶汽车系统必须准确评估其对驾驶环境的解读是否可信，而不能只关注相较于真实标注的准确性。这是因为真实道路驾驶情境复杂，噪声不可避免，传感器和算法存在局限性，而且遮挡也使自动驾驶汽车系统不可能做到尽善尽美。例如，在自动驾驶汽车的感知范围内，遮挡不可避免。在这种情况下，将被遮挡区域预测为可行驶路面可能会提高整体准确率，因为在过去收集的数据中，路面出现得更频繁；但这并不一定反映真实世界的情况。基于从历史数据得出的先验分布，对未观测目标作出缺乏依据的猜测，忽视了观测的不确定性，甚至可能导致严重事故。因此，感知算法可能会为了在这些基准测试集上取得更高分数，而变得过度自信并预测出危险的假阳性。由此可见，我们需要对驾驶场景作出更安全、更可信的解读。

在考虑感知系统中的不确定性的同时实现安全驾驶是一项艰巨的任务。在现实世界中，不确定性来自不同的来源 [9], ，例如传感器噪声、缺陷和感知模型失效。由于视场有限、遮挡，或传感器精度和分辨率较低，测量值不足时，往往会出现较高的不确定性。一种常见情况是，视场有限的自车自动驾驶汽车无法可靠地感知驾驶区域，从而导致危险情况。在这种情况下，自车自动驾驶汽车应减速并等待该不确定区域的障碍清除。另一种选择是利用协同感知，即利用其他道路使用者从不同视角获取的信息。

本文提出一种考虑不确定性的协同感知方法，以解决视场受限和安全问题。协同感知（协同感知）的概念是通过车对车通信，在自动驾驶汽车之间共享并融合所谓的集体感知消息（CPM），使车辆能够看到自车视野之外的区域。具备通信能力的自动驾驶汽车称为网联自动驾驶汽车（CAV）（示例见图5）。协同感知技术在现实世界中应用的一大瓶颈，是实时通信产生的通信开销和时延。因此，在网联自动驾驶汽车之间共享大量数据，例如，原始传感器数据，并非最佳方案。尽管近期基于共享神经网络学习得到的深层特征的研究 [10, 11, 12, 13] 已证明协同感知能够显著提升感知系统的性能，但网联自动驾驶汽车应仅共享自车网联自动驾驶汽车所需的最重要信息，以减少通信负担。这是因为拥塞网络会丢弃消息，进而导致协同感知系统的性能大幅下降。

本文提出通过带有不确定性量化的鸟瞰图来解读协同感知场景中的驾驶空间。鸟瞰图中行驶表面上的每个二维点都会被归入一个预定义类别。与

目标检测和语义分割相比，这种鸟瞰图解读能够更全面、更密集地展现驾驶环境，帮助自动驾驶车辆制定更安全的行驶计划。本文使用证据深度学习（EDL）[14] 来量化分类的不确定性。从本质上说，我们提出的鸟瞰图还为逐点分类结果提供了量化的不确定性值。

更具体地说，我们采用连续驾驶空间中的可学习逐点空间高斯分布来表征驾驶场景，而非使用离散网格，因此观测区域内任意新采样的点都可以从这些高斯分布中获取密度。每个高斯分布描述了邻近点属于与分布中心点相同类别的可能性。我们假设这些新采样点的分类分布服从狄利克雷分布，以描述点属于某一特定类别的概率及其不确定性。随后，将从高斯分布中采样得到的密度视为狄利克雷分布的证据。因此，通过控制新样本的狄利克雷分布，可以联合学习这些高斯分布的参数。我们的方法为每项预测结果提供可靠、可追溯且可解释的不确定性。为简便起见，我们将高斯证据鸟瞰图方法命名为 **GevBEV**。

此外，学习得到的证据式BEV地图能够全面表征驾驶环境。具体而言，除了通过逐点分类结果可靠地检测可行驶路面和其他车辆外，自动驾驶系统还能够识别自身不确定的内容（不确定性较高的检测结果）以及完全未知的内容（没有测量的未观测区域）。在协同感知网络中，证据式BEV地图是关键依据，可用于确定哪些具体区域需要通过更高效的通信从其他网联自动驾驶汽车获取额外信息，以实现协同感知。基于这一依据，通过求取与各网联自动驾驶汽车在自车网联自动驾驶汽车局部坐标系中的检测结果相关联的证据地图的交集，提炼出网联自动驾驶汽车之间共享的最重要信息。如此可避免协同感知消息中的冗余数据，防止通信网络过载和数据包丢失。

总而言之，我们提出的证据 GevBEV 的关键贡献如下：

- 我们提出了一种基于高斯分布的框架，用于在任意分辨率的连续空间中学习整体鸟瞰图，这与以往受限于训练数据所提供地图分辨率的研究不同。
- 我们利用采用狄利克雷证据的证据深度学习来量化分类不确定性，并且与传统确定性模型相比，生成了校准效果更好的不确定性。
- 我们的模型 GevBEV 在协同感知基准测试集 OPV2V [13]（包含模拟驾驶场景）以及 V2V4Real [15]（真实道路驾驶）上取得了新的最先进性能，以显著优势超越了排名第二的模型 CoBEVT [16]。
- 据我们所知，我们首次将证据式BEV地图应用于协同感知任务。分类不确定性是一个关键标准

用于在网联自动驾驶汽车之间有效选择和共享协同感知消息，并显著降低通信开销。

2. 相关工作

本节将从三个方面讨论相关工作：驾驶空间解释、协同感知的最新进展，以及鸟瞰图的不确定性估计。

2.1. 驾驶空间解释

目标检测 [17, 18] 是解释自动驾驶汽车驾驶空间的一种典型方式。检测到的物体通常通过来自摄像头/激光雷达数据的二维/三维边界框进行表征。然而，这种对驾驶场景的解释可能并不完整，因为未检测到或被遮挡的空间没有得到解释。这些空间可能是某些可行驶区域、不可行驶区域，或被未检测到或未观测到的物体占据。因此，自动驾驶汽车可能无法仅依据目标检测的输出做出可靠的驾驶决策。

语义分割是解释驾驶空间的另一种常见方法。它将每个测量点——图像中的像素或激光雷达反射中的点——分类为特定的语义类别。为了进一步区分属于同一语义类别但属于不同物体实例的点，该方法扩展为全景分割 [19]；除了语义标签外，还会为测量点分配实例标识。尽管从自车自动驾驶汽车的视角来看，即在距离视图中，语义分割和全景分割具有整体性且密集，但在鸟瞰图中它们是局部且稀疏的，而自动驾驶汽车通常依据鸟瞰图制定驾驶规划 [20]。

通常，为缓解上述局限性，对驾驶空间的解读会进一步扩展到鸟瞰图地图分割。驾驶环境被表示为一幅鸟瞰图二维图像 [21, 16] ，鸟瞰图中的每个像素都标注有语义标签，从而为车辆建图和规划提供行驶表面的整体视图 [22]。在以往基于图像的研究中，鸟瞰图解读也通过占据栅格建图 [23]、跨视角语义分割 [24], 或地图视角语义分割 [21]来实现。这些方法通过显式几何变换或隐式学习变换，将图像特征从图像坐标转换到鸟瞰图的正交坐标 [21, 16, 25]。与图像相比，包含三维信息的点云数据更直接地生成此类鸟瞰图，例如 PIXOR [26]、PointPillars [27], 和 VoxelNet [28]等方法会压缩正交方向上的信息。我们同样借助鸟瞰图来全面呈现二维驾驶空间。然而，由于远处测量值稀疏且存在遮挡，在不考虑可观测性的情况下，从图像或点云生成稠密鸟瞰图并不可靠。例如，一些被遮挡车辆占据的区域可能会被分类为可行驶区域，仅仅是因为从历史数据中学到的分类分布表明

鸟瞰图中不可见的点更可能属于可行驶区域，而非车辆。因此，为避免在未观测区域作出缺乏依据的预测，我们提出仅根据测量点的几何位置，在观测区域内绘制结果。

2.2. 协同感知的最新进展

随着车对车通信的发展，以及仿真工具能够生成高保真协同检测数据[29, 30, 13], 协同感知将感知系统从单个自车的视角扩展到纳入邻近车辆的感知信息。现代深度学习架构，如图神经网络和 Transformer [31],被用于融合网联自动驾驶汽车的感知信息，主要是检测网络骨干网络提取的深层特征。例如，V2Vnet [32]和 DiscoNet [33] 利用图模型汇总邻近车辆的检测信息。AttFuse [13], V2XViT [34], 和 CoBEVT [16] 提出在鸟瞰图场景下，利用带有自注意力机制的 Transformer 网络，促进网联自动驾驶汽车之间的协作与信息融合。FCooper [35]保留从点云中学习到的特征图的稀疏性，以降低协同感知的通信开销，并使用 Maxout 融合这些特征。然而，这些方法均未探索不确定性估计，以过滤网联自动驾驶汽车之间共享的无信息数据，进一步提高协同感知系统的通信效率。

2.3. 鸟瞰图的不确定性估计

深度神经网络中的不确定性不可避免，因此有必要对其进行估计，尤其是在安全关键应用中。深度神经网络输出的不确定性称为预测不确定性 [9]。这种不确定性主要通过两种方式进行量化：一是对模型中的系统性不确定性进行建模，即认知不确定性；二是对观测中的随机噪声进行建模，即偶然不确定性 [36]。此外还有其他方法，例如先验网络[37]，它通过对分布不确定性进行建模来量化预测不确定性；这种分布不确定性由训练数据与新的推理数据之间的分布不匹配造成。

为估计认知不确定性，贝叶斯神经网络（BNNs）[38, 39] 通过直接推断网络参数的分布，自然地解释了不确定性。尽管如此，由于计算数百万个参数的后验分布不可行，因此很难将其用于深度神经网络。因此，人们开发了蒙特卡洛（MC）Dropout [40] 和深度集成 [41] 等近似方法。MC Dropout 表明，训练基于丢弃的神经网络等价于优化网络输出的后验分布。然而，为了推断不确定性，必须在启用丢弃的情况下进行多次前向运行，这种方法效率低下且耗时。因此，本研究不采用该方法。深度集成通过训练多个模型来近似网络参数的分布。它也需要对每个训练好的模型进行多次前向运行，因此出于同样的原因，本研究也不采用该方法。

为捕捉偶然不确定性，直接建模得到了广泛应用 [42, 43,44, 45, 46]。与 MC Dropout 和深度集成相比，直接建模假设网络输出服从某种概率分布，并直接预测该分布的参数。因此，只需进行一次前向运行即可获得不确定性，效率更高。对于分类问题，传统的确定性深度神经网络对输出 logits 应用 Softmax 函数，将分类分布建模为多项分布。然而，Softmax 输出通常过度自信，校准效果不佳 [14, 47]。

相反，通过 例如，ReLU 激活将输出 logits 转换为正数，以参数化狄利克雷分布，能够更好地量化类别概率和不确定性。例如，先验网络 [37] 通过显式建模分布不确定性，并最小化确定性（分布内）数据上的预测与尖锐狄利克雷分布之间的期望库尔贝克-莱布勒（KL）散度，以及不确定性（分布外）数据上的预测与平坦狄利克雷分布之间的期望库尔贝克-莱布勒（KL）散度，从而捕获预测不确定性。然而，训练此类网络以区分分布内和分布外样本，需要额外的分布外样本。在目标检测和语义分割等复杂视觉问题中，获取足够多的样本以覆盖无限的分布外空间是难以实现的。

不同的是，证据神经网络 [14] 将网络输出视为遵循证据理论与 Dempster-Shafer 理论 [48]的信念，然后推导狄利克雷分布的参数，以建模认知不确定性。与贝叶斯神经网络相比，该方法通过导致预测结果的证据集合来量化分类的不确定性，这意味着可以根据证据量轻松量化分类的认知不确定性。证据神经网络不是最小化预测分布与预定义真实标注分布之间的差异，而是将损失定义为狄利克雷分布基本损失（例如，交叉熵）的期望值。因此，不需要额外数据或真实标注分布。基于此，我们提出应用该方法对二维驾驶空间中各点的分类分布进行建模。

此外，我们没有直接使用狄利克雷分布对不确定性进行建模，而是为测量点引入空间高斯分布，并根据预测的高斯参数为鸟瞰图绘制狄利克雷分布。这种配置模仿了条件随机场算法 [49], ，该算法可通过考虑相邻结果来平滑分割结果。需要指出的是，本文更关注认知不确定性（即神经网络模型所缺乏的知识），以帮助我们理解网联自动驾驶汽车中感知系统的输出，并在协同感知步骤中利用这种不确定性量化，从网联自动驾驶汽车之间共享的信息中提炼出最重要的内容。

3. 方法

3.1. 问题表述

我们将自车协同自动驾驶车辆与协作协同自动驾驶车辆之间的协同感知设定下，结合目标检测和语义分割生成高斯证据鸟瞰图（GevBEV）的任务表述如下。我们遵循 OPV2V 基准 [13]中的协同感知设定。给定自车，$N_{\text{coop}} < 6$ 自车的通信范围内有协同CAV以及其他一些车辆参与者。所有CAV都可以通过请求-响应的方式相互发送协同感知消息：自车协同自动驾驶车辆首先发送一条协同感知消息请求，指定所需区域的信息；收到该请求的协作协同自动驾驶车辆仅在具备请求信息时，才会以相应消息进行响应。

本文选择使用点云数据来展示我们提出的驾驶空间解释。需要指出的是，只要测量点能够投影到鸟瞰图上，这种解释就可以应用于任意模态或多种模态，不仅适用于自动驾驶，也适用于具有多个测量来源的其他制图任务。我们将这部分工作留待未来开展。点云中每个点的输入特征向量表示为 $f^{\text{in}} = [x, y, z, d,$ cos θ, sin θ, i], 其中 x, y, z 是自车激光雷达坐标系中的局部坐标， d 是该点到激光雷达原点的距离， θ是该点相对于激光雷达坐标系 x 轴的角度， i 是激光雷达反射强度。利用该输入特征向量训练基于 U-Net 的 [50]端到端多任务网络。具体而言，所提出的 GevBEV 的主要输出包括目标检测结果、驾驶表面（即道路）和动态道路目标（即车辆）的鸟瞰图，以及通过分类狄利克雷分布量化的预测鸟瞰图不确定性。

在以下小节中，我们首先介绍生成 GevBEV 地图的详细框架：网联自动驾驶汽车之间共享所有感知信息，不采用任何过滤策略来提取最重要的信息。随后，我们介绍一种基于不确定性的 CPM 选择方法，通过处理生成的证据鸟瞰图来提高网联自动驾驶汽车之间的通信效率。

3.2. 框架

GevBEV 地图框架的概览如图1所示。该框架以自车和协作车辆的点云数据为输入，并训练多个头以进行多任务学习。具体而言，基于U-Net的骨干网络首先从输入点云中学习并聚合不同分辨率的深层特征。不同尺寸体素的聚合特征记为 P_s，其中 s 是相对于预定义输入体素尺寸的下采样率。 P_0 表示点云中每个点学习到的特征。不同体素尺寸对应的这些特征分别输入道路头和目标头。道路头指定为分布头，用于生成静态路面的GevBEV地图层；目标头指定为

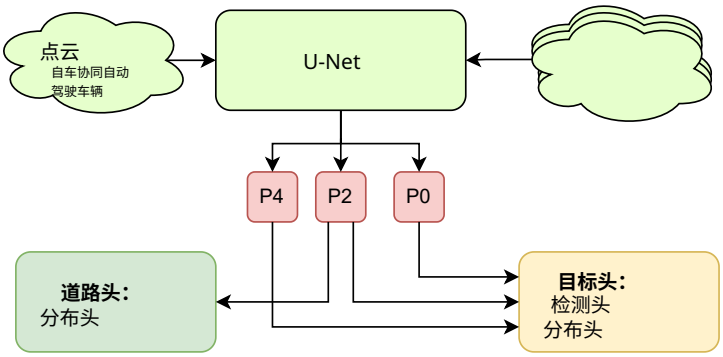

图 1：GevBEV 地图框架概览。模型以点云数据作为输入。道路头学习路面（绿色）的分布，目标头（橙色）则检测物体并学习相应边界框的分布。P_s：以步幅 $s = 0, 2, 4$ 下采样的体素的中间学习特征。

分为检测头和分布头，用于检测动态物体并生成物体层的GevBEV地图。我们将GevBEV地图分为两层，因为与路面相比，物体通常尺寸较小。否则，由于表面上的点数量占主导，小物体可能会被这些点平滑掉。下面将详细介绍框架的各个模块，即基于*U-Net*的骨干网络、检测头和分布头。

U-Net。 鉴于 U-Net 在逐点特征提取和表征方面表现出色，[50] 我们将其用作模型的骨干网络。图1 展示了采用定制编码的 U-Net 总体结构。首先，使用多层感知机（多层感知机）对输入特征进行体素化，如公式（1）所述，

$$\bar{f}_v = \frac{\sum_{i \in \mathbf{v}} f_i^{\text{in}}}{|\mathbf{v}|}, \quad \tilde{f}_{vi} = \text{mlp}([f_{vi}, (f_{vi} - \bar{f}_v)]), \tag{1}$$

其中$\mathbf{v}$ 是属于某个体素的点的集合，f_i^{in}表示点 i的输入特征，$[\cdot, \cdot]$ 表示拼接操作。随后，通过对属于该体素的编码点特征$\tilde{f}$求平均值，$f_{\mathbf{v}}$ 计算体素特征 f_{vi} ，如公式（2）所示。

$$\tilde{f}_{\mathbf{v}} = \frac{\sum_{i \in \mathbf{v}} \tilde{f}_{vi}}{|\mathbf{v}|}. \tag{2}$$

随后，将编码后的体素输入四个卷积块。具体而言，卷积层1仅包含一个用于处理输入的层。卷积层2至卷积层4各由三个卷积层组成，其中前一层会对稀疏张量进行下采样。每个稀疏卷积层后都接批归一化和 Leaky ReLU激活。在上采样层中，转置卷积层的结构与下采样层中的对应层相似。

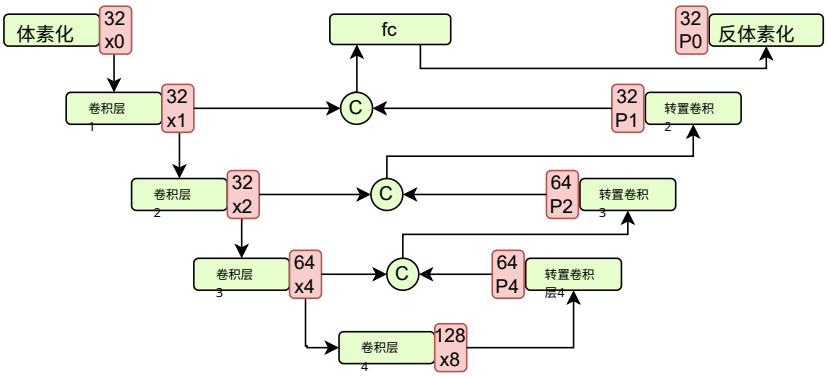

图2：基于U-Net的骨干网络。蓝色方框表示网络组件，橙色方框表示组件的输出通道和体素步幅。绿色椭圆表示拼接操作。

卷积层快捷连接中的特征均与转置卷积层的特征拼接。在最后，我们将体素特征与编码点特征~拼接起来，f_{vi} 以对步幅为一的体素进行反体素化，并获得 P_0 每个点的特征。值得一提的是，该网络中的所有卷积层均通过 Minkowski Engine [51]在稀疏体素上实现，以降低计算负载。

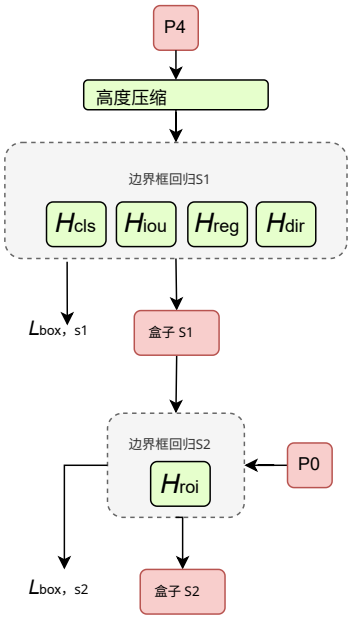

图3：检测头。

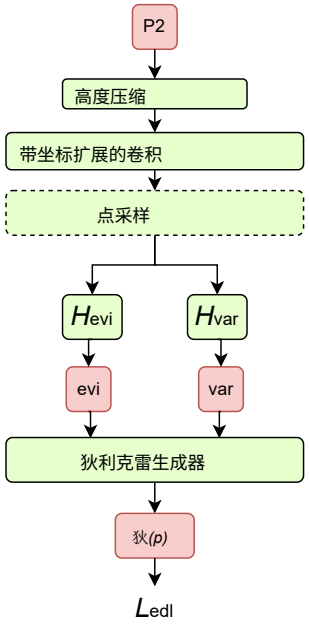

图4：分布头。

检测头。我们使用该检测头来展示目标检测所预测的边界框与 GevBEV 地图之间的对齐情况。如图3所示，在该检测头中，我们使用步幅为四的体素 P_4 来生成预定义的参考边界框，即所谓的锚框，并在此基础上进行进一步回归。首先，我们沿垂直维度使用步长卷积层来压缩三维稀疏张量的高度，从而得到BEV特征图。然后，使用该特征图进行边界框分类和回归。我们遵循 [10] 所采用的协同感知两阶段策略：第一阶段（BoxReg S1）根据自车CAV的局部信息生成候选框，第二阶段（BoxReg S2）通过融合邻近CAV的信息来优化这些边界框。

我们使用以下公式（3）-（5）对边界框进行编码：

$$\mathcal{B}_{loc} = [\frac{x_g - x_a}{d_{xy}}, \frac{y_g - y_a}{d_{xy}}, \frac{z_g - z_a}{h_a}], \tag{3}$$

$$\mathcal{B}_{dim} = [\log(\frac{l_g}{l_a}), \log(\frac{w_g}{w_a}), \log(\frac{h_g}{h_a})], \tag{4}$$

$$\begin{aligned} \mathcal{B}_{dir} = [&\cos\theta_g - \cos\theta_a, \sin\theta_g - \sin\theta_a, \\ &\cos\theta_g - \cos(\theta_a + \pi), \sin\theta_g - \sin(\theta_a + \pi)]. \end{aligned} \tag{5}$$

$\mathcal{B}_{loc}$ 和 $\mathcal{B}_{dim}$ 是边界框位置和尺寸的常用编码。下标 g 和 a 分别表示真实标注和锚框。 x、y、z是边界框中心的坐标，而d_{xy}是边界框在 XY平面上的对角线长度。 l、w、h分别表示边界框的长、宽和高。我们使用式（5）所示的正弦和余弦编码来表示边界框方向 θ，以避免方向翻转问题。 $\mathcal{B}_{dir}$ 的前两个元素是原始方向角的编码，后两个元素则对应反向方向。加入反向方向的编码后，当回归源锚框与目标真实标注框方向相反时，正弦和余弦编码之间的距离会减小，从而有助于方向回归。

BoxReg S1 中所有检测头的总损失记作 $\mathcal{L}_{box,\ s1}$，如图 3所示。具体而言，BoxReg S1 遵循 CIASSD [52]，采用四个检测头 $\mathcal{H}_{cls}$、$\mathcal{H}_{iou}$、$\mathcal{H}_{reg}$、$\mathcal{H}_{dir}$来生成边界框。 $\mathcal{H}_{cls}$使用二元交叉熵，其他检测头在此阶段使用 Smooth-L1 损失。 $\mathcal{H}_{iou}$ 回归检测到的边界框与真实边界框之间的交并比（IoU）。它用于校正来自 $\mathcal{H}_{cls}$ 的分类分数，从而在非极大值抑制（NMS）过程中保留质量更高的边界框。方向检测头 $\mathcal{H}_{dir}$回归真实边界框与对应锚框之间经过编码的角度偏移量。在边界框解码阶段，选择来自 $\mathcal{H}_{dir}$ 且数值较小的预测角度偏移量，作为边界框的正确方向。

在 BoxReg S2 中，不同 CAV 检测到的候选边界框会先共享给自车协同自动驾驶车辆，再通过非极大值抑制进行融合，随后将其用作

进一步优化的参考锚框。具体而言，遵循 FPV-RCNN[10], 的方法，CAV 仅共享属于候选框的关键点，然后利用这些关键点优化融合边界框。此外，在本文中，我们使用 Minkowski Engine [51] 简化感兴趣区域（RoI）头的结构 $\mathcal{H}_{roi}$。也就是说，首先将融合边界框变换到规范局部框坐标系中，并将关键点坐标记为 $\mathbf{kpt}_i^{in}$。然后，我们对关键点进行体素化，生成一个 $6 \times 6 \times 6$ 网格，其体素化操作与 U-Net 中使用的类似，其中 f_i^{in}变为关键点的规范坐标 $\mathbf{kpt}_i^{in}$。 $[x, y, z]$ 拼接中的第一个元素 f_{vi} 成为关键点的学习到的特征，式（1）拼接中的第二个元素则成为关键点坐标的学习到的位置编码。为便于理解，我们将式（1）改写为式（6）。

$$\begin{aligned} \bar{\text{kpt}}_v &= \frac{\sum_{i \in \mathbf{v}} \text{kpt}_i^{in}}{|\mathbf{v}|}, \\ \tilde{f}_{vi} &= \text{mlp}([f_{vi}, \text{mlp}(\text{kpt}_{vi} - \bar{\text{kpt}}_v)]). \end{aligned} \tag{6}$$

根据公式（2）汇总每个体素的特征后，网格˜中的所有体素会通过体素特征 $f_{\mathbf{v}}$ 的加权平均进行聚合。˜权重由多层感知机根据 $f_{\mathbf{v}}$ 学习得到。随后，将每个盒子网格的聚合特征输入一个 IoU 头和一个边界框回归头，以生成用于细化融合候选框的参数。解码后的边界框记为盒子 S2，如图 3所示。两个头均使用 Smooth-L1 损失。此阶段各项损失之和记为 $L_{box,\ \ s2}$。

分布头。 如图 1所示，我们设计了一个分布头，用于生成路面和物体的基于点的分布。为了平衡计算负载与点分辨率之间的取舍，我们使用步长为二的体素 P_2 来生成分布。分布头的整体结构如图 4所示。首先沿垂直方向压缩输入的三维体素，以获得二维逐点深层特征。该高度压缩由两个卷积核尺寸和步长尺寸相同的稀疏卷积层组成，因此垂直方向上的所有体素都可以合并为一个。在此模块中，道路和物体分布头共享权重。

此外，我们在稀疏卷积过程中扩张体素坐标，以填补离散测量点之间的空隙。我们将此称为坐标扩展。尽管我们提出只对可观测区域预测分布，但所有传感器都以离散方式测量连续空间，因此在测量距离较远时，不可避免地会出现空隙。如上所述，一方面，U-Net 中使用的基础卷积层会保持体素的稀疏性，以降低计算负载。另一方面，对于远处可观测区域中由激光雷达的距离视图造成的相邻两次激光测量之间的空隙，这些卷积层无法推断其中的信息。然而，只要测量密度足够，模型也应该能够推断两个离散测量值之间的信息，

并以可控的方式进行。因此，我们使用若干坐标可扩展稀疏卷积，通过预定义核大小控制扩展范围，从而谨慎地填补空隙。详细设置见第 4.5 节。

为进一步加快训练过程，如图4中虚线框所示，我们使用点采样模块对点的总数量进行下采样。为便于说明，我们将所有这些选定的点称为中心点。该模块是可选的，后续实验结果表明，它不会对训练结果产生明显的负面影响。

我们假设每个中心点 c_i 服从狄利克雷分布，以建模点分类分布，并服从空间各向同性高斯分布，以建模该点的邻域一致性。随后，分别通过头部 $\mathcal{H}_{\text{cls}}$ 和 $\mathcal{H}_{\text{var}}$回归这两个分布的参数。两个头部均由两个经 ReLU 激活的全连接层组成，以约束两个分布的参数为正值。其输出表示为

$$\mathbf{o}_{\text{cls}} = [o_{\text{cls}}^{\text{fg}}, o_{\text{cls}}^{\text{bg}}], \tag{7}$$

$$\mathbf{o}_{\text{var}} = [o_{\sigma x}^{\text{fg}}, o_{\sigma y}^{\text{fg}}, o_{\sigma x}^{\text{bg}}, o_{\sigma y}^{\text{bg}}], \tag{8}$$

其中，fg 表示前景，bg 表示背景。$\mathbf{o}_{\text{cls}}$ 被视为该点属于前景或背景的证据。$\mathbf{o}_{\text{var}}$ 是该点在 x-轴和 y-轴上的回归方差。为确保每个点都能产生贡献，我们在预测结果中加入一个较小的初始方差。因此，得到的方差为$\sigma_{x,\ y}^2 = \mathbf{o}_{\text{var}} + \sigma_0^2$。对于邻域中给定的任意新目标点$\mathbf{x}_j$ ，若其位于 BEV空间中的中心点$\mathbf{c}_i$ 附近，则可以通过公式（$\phi(\mathbf{x}_j)_i$ 计算该新点属于特定类别的概率密度，见公式（11）。

$$\Sigma_i = \begin{bmatrix} \sigma_x^2, 0 \\ 0, \sigma_y^2 \end{bmatrix}, \tag{9}$$

$$m_{ji} = (\mathbf{x}_j - \mathbf{c}_i)^T \Sigma_i^{-1} (\mathbf{x}_j - \mathbf{c}_i), \tag{10}$$

$$\phi(\mathbf{x}_j)_i = \frac{\exp\left(-0.5 \cdot m\right)}{\sqrt{2\pi^d |\Sigma_i|}}, \tag{11}$$

其中， Σ_i 是前景或背景分布的中心点$\mathbf{c}_i$ 的协方差， m 是点$\mathbf{x}_j$ 到中心点$\mathbf{c}_i$的马氏距离平方， d 是分布的维数。在我们的案例中， $d = 2$。

为了获得点$\mathbf{x}_j$处的整体狄利克雷证据 $e(\mathbf{x}_j)$，我们汇总最大分布范围 ν 内所有相邻中心点$\text{nbr}(j)$所产生的归一化加权概率质量，得到

$$\begin{aligned} e_k(\mathbf{x}_j) &= \sum_{i \in \text{nbr}(j)} \frac{\phi(\mathbf{x}_j)_i}{\phi(\mathbf{c}_i)} \cdot o_{cls}^k, \\ &= -\frac{1}{2} \sum_{i \in \text{nbr}(j)} m_{ji} \cdot o_{\text{cls}}^k, \end{aligned} \tag{12}$$

其中 $\phi(\mathbf{c}_i)$ 是中心点处的概率密度，而 $k \in \{\text{fg,bg}\}$。下文将式（12）推导如下

$$\begin{aligned}
\log(\phi(\mathbf{x}_j)_i) &= -\frac{1}{2}m_{ji} - \frac{1}{2}\log(2\pi^d|\Sigma|), \\
&= -\frac{1}{2}(d\log(2\pi) - \log|\Sigma|) - \frac{1}{2}m_{ji}, \\
\log(\phi(\mathbf{c}_i)) &= -\frac{1}{2}m_i - \frac{1}{2}\log(2\pi^d|\Sigma|), \\
&= -\frac{1}{2}(d\log(2\pi) - \log|\Sigma|), \\
\frac{\phi(\mathbf{x}_j)_i}{\phi(\mathbf{c}_i)} &= \exp(\log\frac{\phi(\mathbf{x}_j)_i}{\phi(\mathbf{c}_i)}), \\
&= \exp(\log\phi(\mathbf{x}_j)_i - \log\phi(\mathbf{c}_i)), \\
&= -\frac{1}{2}m_{ji},
\end{aligned} \tag{13}$$

其中，中心点$\mathbf{c}_i$到其自身的马氏距离平方为 $m_i = 0$。根据 [14], 期望概率$p_{j,\ k}$——点$\mathbf{x}j$ 属于类别 $\mathbf{k}$——以及该分类结果的不确定性 u_j，可得

$$\hat{p}_{j,k} = \frac{\alpha_{j,k}}{S_j} = \frac{e_k(\mathbf{x}_j) + 1}{\sum_{k\in\{\text{fg,bg}\}}(e_k(\mathbf{x}_j) + 1)}, \tag{14}$$

$$u_j = \frac{K}{S_j}, \tag{15}$$

其中，$\alpha_{j,\ k}$是类别 k对应于 $k = 1, ..., K$的浓度参数，S_j也是点$\mathbf{x}j$的狄利克雷分布强度。对于分布头的损失 $\mathcal{L}_{edl}$，我们采用 [14], 中推荐的损失函数，该函数被表述为平方损失之和与库尔贝克-莱布勒（KL）散度正则化的期望，从而防止网络生成过高的证据。该损失的最终表达式如下

$$\begin{aligned}
\mathcal{L}_{edl} = &\sum_{j=1}^{N}\sum_{k\in\{\text{fg,bg}\}}[(y_{j,k} - \hat{p}_{j,k})^2 + \frac{\hat{p}_{j,k}(1-\hat{p}_{j,k})}{S_j + 1}] \\
&+ \lambda_t\sum_{j=1}^{N}\text{KL}[\text{Dir}(\mathbf{p}_j|\tilde{\alpha_j}) \parallel \text{Dir}(\mathbf{p}_j|\mathbf{1})],
\end{aligned} \tag{16}$$

其中

$$\begin{aligned}
\tilde{\alpha_j} &= \alpha_j \odot (1 - \mathbf{y}_j) + \mathbf{y}_j, \\
\lambda_t &= \min(1, A_{\text{epoch}}/A_{\max}).
\end{aligned}$$

式（16）中的第一项是目标标签$y_{j,\ k}$ 与预测值 $p_{j,\ k}$之间平方损失的期望和。N 表示样本总数，且$j \in \{1, ..., N\}$。第二项是按退火

系数 λ_t 加权的 KL 散度，该系数随训练轮次 A_{epoch}与最大退火步数 $A_{\max}$之比而变化。$\mathbf{y}_j = \{y_{j,k}\}_{k\in\{\text{fg,bg}\}}$ 是点 $\mathbf{x}j$的类别真实标签。 $\tilde{\alpha_j}$ 是 α_j 的滤波版本，以确保 KL 散度只惩罚对 α_j的误导性预测。

我们提出了一种基于高斯的方法，在连续空间中训练证据 BEV 地图的分布头。为实现这一目标，监督学习的目标点不局限于点云的原始观测点，也不局限于由离散点构成的特定分辨率下的中心点；它们可以是连续 BEV 平面中的任意点。更具体地说，我们将原始观测目标点视为种子点，并以此为基础，在受控范围内通过从其原始观测位置进行随机偏移来生成连续目标点，例如，偏移距离服从正态分布 $\mathcal{N}(0,\ 3)$ m。通过这种方式，可以从中心点生成预定义数量 n_{tgt} 的目标点，以控制密度，并且仅利用观测区域中生成的目标点来训练模型。为避免训练期间内存溢出和计算时间过长，还会对随机采样的目标点进行进一步下采样。通过体素下采样，道路头仅使用数量有限的目标点进行监督——N_{tgt}，前景样本和背景样本均如此。然而，目标头中目标点的重要性和数量在前景与背景之间更加不均衡。由于背景目标点数量占多数但重要性较低，训练时仅纳入少量背景目标点。相比之下，所有扩展至真实边界框边缘特定范围内的样本，都会被用于训练前景类别，从而更好地描述边界框周围的细节。

3.3. 协同感知

本节介绍一种通过处理我们生成的证据鸟瞰图，实现基于不确定性的 CPM 选择的协同感知方法。我们不再简单地在 CAV 之间共享所有信息（这可能导致通信网络拥塞），而是利用生成的鸟瞰图量化 CAV 的感知不确定性，从而提炼出最重要的信息。如图 5所示，红色自车协同自动驾驶车辆首先对根据公式（12）生成的不确定性地图进行阈值处理，生成请求二值掩码，阈值为 u_{ego} （底部的不确定性掩码），并且仅请求其不确定性较高区域内的感知信息（左侧红色请求掩码）。随后，绿色协作协同自动驾驶车辆使用自身经过掩码处理的证据图（绿色响应掩码）作出响应。具体而言，将协作协同自动驾驶车辆的不确定性地图按阈值 u_{coop} 进行阈值处理所生成的二值掩码，与从自车协同自动驾驶车辆接收到的请求掩码求交集，从而得到一个响应掩码，其中包含协作协同自动驾驶车辆的不确定性较低的区域。之后，仅选择响应掩码所覆盖区域内的证据，并利用所得响应掩码提炼协作协同自动驾驶车辆发送给自车协同自动驾驶车辆的 CPM。我们将这种共享策略记为 CPM_{all}，因为它将自车协同自动驾驶车辆预定义视场内的所有区域都纳入信息共享范围。然而，在驾驶空间中，CAV 更关注道路上的状况

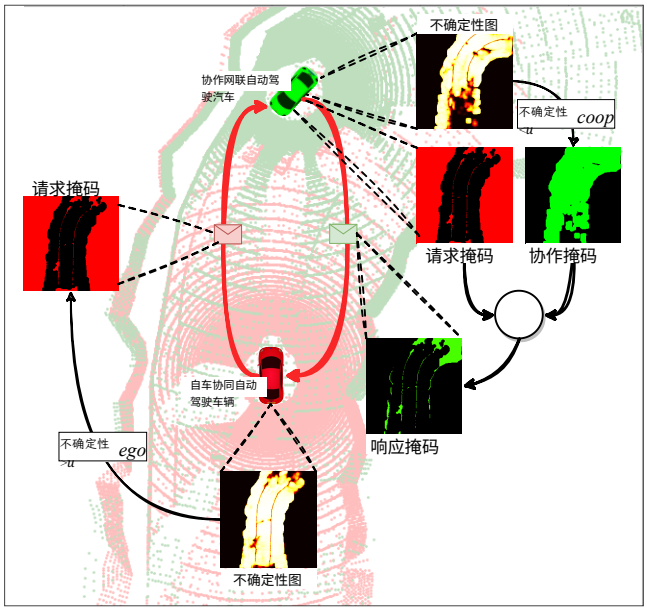

图5：网联自动驾驶汽车之间共享信息的示例。自车协同自动驾驶车辆及其请求掩码中，不确定性较高的区域以红色标示；协作协同自动驾驶车辆及其响应掩码中，不确定性较低的区域以绿色标示。在不确定性地图中，浅色表示不确定性较低，黑色表示没有测量值。

路面。因此，可以根据当前场景的道路表面几何进一步限制待共享的信息。在实际应用中，可以从地图等先验信息中获取这种几何信息。我们注意到，协同感知基准也提供了预先采集的高精地图。作为概念验证研究，我们将当前场景与高精地图进行配准，以进一步排除掩码区域中的非路面区域。当额外的高精地图已提供给 CAV 时，我们将这种共享策略记为 CPM_{road}。为简便起见，本文在实验中固定协作协同自动驾驶车辆的不确定性阈值 $u_{\text{coop}} = 1.0$，仅改变自车协同自动驾驶车辆的阈值 u_{ego}，以评估其有效性。

4. 实验

下文将介绍用于评估我们所提出模型的数据集、数据增强、评估指标以及详细实验。

4.1. 数据集

在本研究中，我们在两个多智能体协同感知基准上进行了实验：OPV2V [13], ，这是一个由 CARLA 生成的仿真数据集 [29]，以及 OpenCDA [30, 53], 和 V2V4Real [15], ，这是一个真实数据集，采集于两辆车在真实场景中行驶期间。

OPV2V数据集包含 73 个场景，涵盖九座城市的六种道路类型。该数据集包含 12K 帧激光雷达点云，以及每帧对应的标注三维边界框。该数据集的检测范围设为 $[-50, 50]\,m$，适用于 x- 和 y-坐标； $[-3, 3]\,m$ 适用于 z 坐标，与基线工作 CoBEVT [16]相同。V2V4Real数据集覆盖 410 km 的驾驶区域，包含约 10K 帧标注的激光雷达帧。我们将该数据集的检测范围设为 $[-102.4, 102.4]\,m$，适用于 x-坐标， $[-38.4, 38.4]\,m$ 适用于 y-坐标，以及 $[-5, 3]\,m$ 适用于 z-坐标，因为大多数标注车辆都位于此范围内。训练集和测试集均遵循两个数据集的官方划分。具体而言，OPV2V包含 44 个训练场景和 16 个测试场景，V2V4Real包含 33 个训练场景和 9 个测试场景。

4.2. 数据增强

点云是表示激光雷达传感器所感知几何信息的最常见方式之一。它由激光雷达射线击中物体表面时产生的反射点构成。考虑到在点云中，射线击中远处物体时得到的点非常稀疏，我们提出以下激光雷达数据增强方法，以帮助完成感知任务。

自由空间增强。在点云中，自由空间——即射线经过且未被任何障碍物占据的空间——往往会被忽略。然而，这些信息同样是测量结果的一部分，对于识别驾驶空间的占用情况和可见性至关重要。因此，我们通过从激光雷达射线路径中采样点来增强点云数据，并将这些点称为自由空间点 f_{si}，其中 $i \in \mathbb{N}$。如图6所示，激光雷达（橙色圆柱体）发射一条射线（红线），该射线击中地面表面上的点 fs_0，并且只将这个反射点记录到点云中。沿着射线路径，我们在距命中点 f_{si} 有限的距离 d_{fs} 范围内采样自由空间点 f_{s0}，采样步长较大 s_{fs}。为限制计算开销，我们只在地面上高度有限的区域内采样点，即 $z \leq h_{fs}$蓝色区域，该区域对驾驶至关重要。最后，使用给定的体素尺寸 v_{fs}进行体素下采样，再次对这些点进行下采样，以获得均匀且稀疏分布的自由空间点。随后，将这些增强后的自由空间点添加到原始点云中，并将其强度值设为 $\mathrm{i} = -1$ 指示标记。

几何增强。 我们通过随机旋转、沿 x 轴和 y 轴翻转，然后缩放点云的几何尺寸，来增强点云数据。此外，我们还为所有点添加少量高斯噪声 $[\delta x, \delta y, \delta z]$。这种增强有助于提高模型在测试过程中应对域偏移的鲁棒性。

4.3. 评估指标

我们使用交并比（IoU）作为评估整体预测性能的指标，并使用校准图分析预测不确定性的质量。尽管我们提出的模型 GevBEV 能够生成

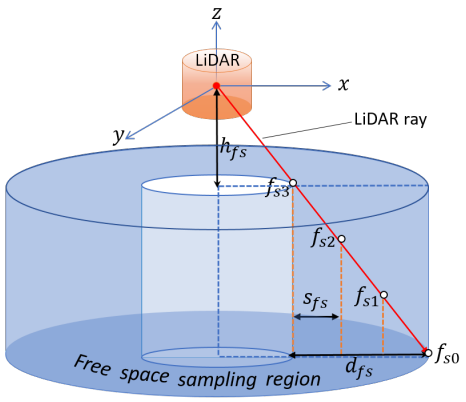

图6：自由空间点采样。红点是激光雷达坐标系的原点， z-轴表示驾驶空间中的垂直方向， x-轴表示水平方向。 f_{s0} 是激光雷达射线（红色虚线）与地面的交点。 f_{s1}、 f_{s2}和 f_{s3} 是沿射线路径采样得到的自由空间点。

任意分辨率的鸟瞰图，但我们仅评估分辨率为 $0.4\,m$ 的预测鸟瞰图，以便与其他只能生成这种预定义预测分辨率鸟瞰图的方法进行比较。

交并比。我们分别报告预定义预测范围（IoU_{all}）内所有感知区域的总体结果，以及可观测区域（IoU_{obs}）内的结果，这与仅在观测区域内进行预测才可靠这一理念相符。由于我们的模型设计上不会对不可观测区域进行预测，因此计算 IoU_{all}时，不可观测区域中的所有点均视为假阳性。从数学上说，当$\mathbf{x}j$ 满足 $\exists i \in \{i| \parallel \mathbf{x}_i - \mathbf{x}_j \parallel_2 < \nu\}$时，该点为可观测点，这意味着该点位于任意中心点的 $\nu = 2\,m$ 范围内。这些可观测区域上的交并比根据公式（18）-（21）计算。

$$\mathbf{X} = \{\mathbf{x}_j | u_j < u_{\text{thr}}\}, \tag{17}$$

$$\mathbf{X}^{\text{fg}} = \{\mathbf{x}_j | \hat{p}_j^{\text{fg}} > \hat{p}_j^{\text{bg}}, \mathbf{x}_j \in \mathbf{X}\}, \tag{18}$$

$$\mathbf{Y} = \{\mathbf{x}_j | \text{lbl}(\mathbf{x}_j) = \text{fg}\}, \tag{19}$$

$$\mathbf{Y}^{\text{fg}} = \{\mathbf{x}_j | \mathbf{x}_j \in \mathbf{Y}, \mathbf{x}_j \in \mathbf{X}\}, \tag{20}$$

$$IoU = \frac{|\mathbf{X}^{\text{fg}} \cap \mathbf{Y}^{\text{fg}}|}{|\mathbf{X}^{\text{fg}} \cup \mathbf{Y}^{\text{fg}}|}. \tag{21}$$

我们只评估不确定性低于不确定性阈值 u_{thr} 的点$\mathbf{X}$ 的结果。$\mathbf{X}^{\text{fg}}$ 是正类预测的集合，即根据预测分类概率被分类为前景的点。$\text{lbl}(\mathbf{x}j)$ 是点$\mathbf{x}j$的类别标签，因此$\mathbf{Y}$ 是真正例点的集合。$\mathbf{Y}^{\text{fg}}$ 是真实标注中所有正例点的集合。然后，计算真正例点与真实标注正例点之间的交并比。

校准图。 我们绘制分类准确率与不确定性的关系图，以展示二者预期的相关性。我们的模型会为每个点的分类生成不确定性，从而能够对所有类别进行整体评估。然而，不同类别的样本数量不均衡会导致评估结果产生偏差。例如，由于缺乏来自邻居的证据，某个点的分类不确定性较高，但由于其邻居属于主导类别，最终分类准确率却可能较高。为避免这种有偏评估，我们根据特定类别的样本总数比例为每个样本赋权。然后，我们将不确定性 $u \in [0, 1]$ 划分为十个区间，每个区间的长度为 0。1。随后，我们计算不确定性处于该区间内所有样本的加权平均分类准确率。理想的校准图是一条对角线，表示分类准确率与不确定性之间存在最强的负相关，即 高准确率对应低不确定性。

4.4. 基线模型与对比模型

我们在协同感知基准测试集 OPV2V 和 V2V4Real 上，针对仿真和真实驾驶场景，将所提出的 GevBEV 模型与当前最先进的摄像头和激光雷达协同感知模型进行比较，以评估其有效性，具体包括 $FCooper$ [35], $AttFuse$ [13], $V2VNet$ [32], $DiscoNet$ [33],$V2XViT$ [34] 以及 即 ***CoBEVT*** [16]。与 GevBEV 不同，上述模型均未提供感知不确定性估计，也无法提炼传递给自车协同自动驾驶车辆的关键信息。因此，我们仅在感知性能方面将 GevBEV 与这些模型进行比较。在协同感知应用中，我们不会进一步从 V2V 通信的 CPM 大小方面将它们与 GevBEV 进行比较。

此外，我们开展了一系列消融研究，以分析 GevBEV 所提出模块的有效性。

- BEV 是去除了基于点的空间高斯和证据损失 $\mathcal{L}_{edl}$ 的所提出模型，这使我们的模型从概率模型变为确定性模型。它使用交叉熵训练相应的分布头，以对鸟瞰图中的点进行分类。我们将此模型作为基线模型。

- $EviBEV$ 仅去除了基于点的空间高斯。它仍使用$\mathcal{L}_{edl}$ 训练分布头。 • $GevBEV^{-}$ 是我们提出的模型，但训练时未使用自由空间增强。

4.5. 实现细节

在所有实验中，我们将输入体素尺寸设为 0.2 m，以平衡计算开销与性能。OPV2V 上的自由空间点按照配置 $h_{fs} = -1.5\,m, d_{fs} = 1.5m, s_{fs} = 7.5\,m$ 进行采样。不过，由于 V2V4Real 数据集在 x 方向上的检测范围更大，因此我们将其 s_{fs} 增至 $9m$ 。在几何增强过程中，点云坐标会在 $[0.95, 1.05]$ 范围内随机缩放，然后添加

服从正态分布的 $\mathcal{N}(0, 0.2)$ m 噪声。对于检测头，我们在步长为四的鸟瞰图上的每个观测点处生成两个参考锚框。这两个锚框的尺寸相同：OPV2V 为 $[l, w, h] = [4.41, 1.98, 1.64]\, m$，V2V4Real 为 $[l, w, h] = [3.90, 1.60, 1.56]\, m$。不过，它们的角度不同，分别为 $0°$ 和 $90°$。与真实边界框的交并比大于 0.4 的锚框被视为正样本，小于 0.2 的则视为负样本；其他锚框在训练过程中计算损失时不予考虑。负样本远多于正样本，导致训练过程面临困难。因此，我们仅随机采样 512 个负样本用于训练。

我们从头开始训练整个网络，共训练 50 个轮次，权重衰减设为 0。01，并使用 Adam 优化器 [54] 进行参数化，参数为 $\beta = [0.95, 0.999]$和 $\gamma = 0.1$。我们使用多步学习率调度器，初始学习率为 $lr = 10^{-3}$，并在第 20 和第 45 个训练轮次设置两个里程碑。每到一个里程碑，学习率都会降低至原来的 0.1 倍。

对于采样生成器，道路头使用 $n_{\text{tgt}} = 10$，目标头使用 $n_{\text{tgt}} = 1$。道路头的体素下采样分辨率设为 0。4 米，并将 N_{tgt} 设为 3000。对于目标头，保留所有与任一真实标注框边缘的最小距离小于 4 米的生成目标点。从背景中采样 $N_{\text{tgt}} = 50 \cdot n_{\text{gt}}$ 个点作为负样本，其中 n_{gt} 为真实边界框的数量。对于这些采样得到的目标点，根据公式（12）在最大分布范围 $\nu = 2\, m$内绘制证据。上述参数均通过经验设定，代码和训练好的模型已发布于 https://github.com/YuanYunshuang/GevBEV，以便复现。

4.6. 结果

4.6.1. 定量分析

与最先进模型相比。 我们提出的 GevBEV 模型在仿真 OPV2V 数据集 [13] 和真实数据集 V2V4Real [15]上，与用于协同感知的最先进模型进行了基准测试。表 1列出了道路和动态物体分割的结果。可以看出，基于相机数据的模型表现不如基于激光雷达数据的模型。这是因为激光雷达数据提供了更准确的三维信息，而这对于将信息投影到鸟瞰图以执行感知任务至关重要。GevBEV 优于所有其他模型，包括为确保公平比较而使用与 GevBEV 相同激光雷达数据的模型。与 OPV2V 基准测试集上排名第二的模型 CoBEVT 相比，我们的模型借助分布头，在动态物体分割上的交并比提高了 23.7%，在道路表面分割上的交并比提高了 4.7%。在 V2V4Real 基准测试集上，令人惊讶的是，与 V2XViT 相比，我们的模型将交并比提高了 54.8%。这些提升表明，基于点的空间高斯能够有效地提供关于每个表面点邻域的更平滑信息，从而在两个基准测试集上都取得更准确的结果。此外，我们提出的训练采样方法能够更好地抵御真实标注中的误差。这使得模型在真实道路驾驶中标签不准确的真实数据集 V2V4Real 上取得了显著提升。

模型	模态 相机	激光雷达	OPV2V Road	物体	V2V4Real 物体
AttFuse [13]	✓		60.5	51.9	-
V2VNet[32]	✓		60.2	53.5	-
DiscoNet [33]	✓		60.7	52.9	-
CoBEVT [16]	✓		63.0	60.4	-
Fcooper [35]		✓	70.3	52.1	25.9
AttFuse [13]		✓	75.3	52.0	25.5
V2XViT [34]		✓	75.0	50.4	<u>29.9</u>
CoBEVT [16]		✓	<u>75.9</u>	<u>52.3</u>	29.6
GevBEV（我们的方法）		✓	**79.5**	**74.7**	**46.3**

表 1：在 OPV2V 和 V2V4real 数据集上与最先进模型的比较。最佳值以粗体标出，次佳值以下划线标出。

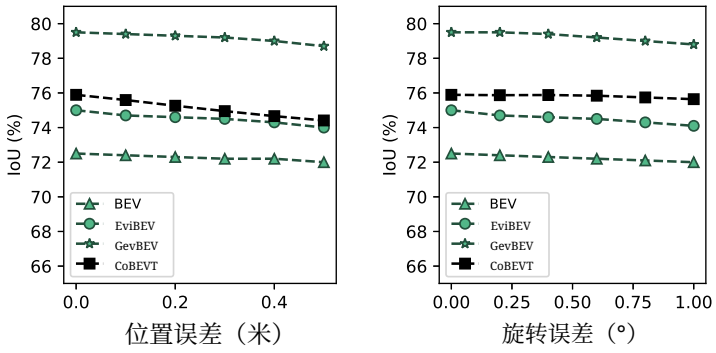

图 7：存在定位噪声时的 OPV2V 道路分割结果。

然而，真实世界驾驶场景给协同感知任务带来了更多挑战，因为不可避免地会引入定位误差；这一点可从 OPV2V（74.7%）与 V2V4Real（46.3%）在目标分割性能上的巨大差距看出。此外，V2V4Real 数据集仅提供两辆联网车辆之间的通信，少于仿真 OPV2V 数据集。因此，V2V4Real 上的感知性能远不如在 OPV2V 上测试的结果。

对定位噪声的敏感性。在此前针对模拟数据的实验中，我们假设定位信息完美无误。为了评估模拟结果对定位噪声的敏感性，我们引入由正态分布生成的定位误差，其中位置的标准差范围为 0 至 0.5 米，方向的标准差范围为 0 至 1 度。图7 表明，随着位置误差和旋转误差增大，所有模型在路面分割任务中的性能均略有下降。尽管如此，我们提出的

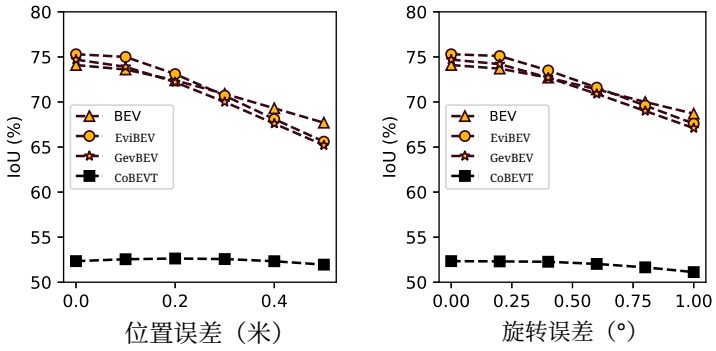

图8：定位噪声下的 OPV2V 目标分割结果。

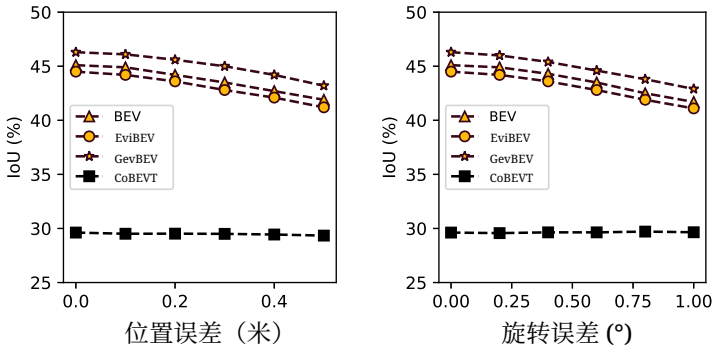

图9：V　　2V4Real 目标分割结果及定位　　n 噪声。

GevBEV 模型优于 CoBEVT，在所有误差配置下均保持最佳性能，领先幅度约为 4%。与分割性能相比，物体由于尺寸较小，对定位误差更加敏感。在这种设置下，正如图8 和9所示，我们的模型在 OPV2V 和 V2V4Real 数据集上仍然优于排名第二的模型 CoBEVT。这表明，与 CoBEVT 相比，我们提出的方法能够更好地应对定位噪声，具有更强的鲁棒性。

消融研究。表2 展示了各消融模型的性能。总体而言，不含基于点的空间高斯（G.s.）和证据损失的鸟瞰图基线模型 $\mathcal{L}_{edl}$ 逊于其他模型。这表明，通过优化交叉熵训练的传统确定性模型不如概率模型。相反，通过用空间连续高斯分布对每个点进行建模，我们能够弥补点云稀疏性造成的差距，并生成更平滑的鸟瞰图。对于所有区域和观测到的区域，依据交并比衡量表面时，采用基于点的空间高斯的 EviBEV 表现优于基线。然而，它的

物体的性能略有下降。

模型	模块 G.s.	$\mathcal{L}_{edl}$	Ag.	OPV2V（全部） 道路	OPV2V（全部） Obj	V2V4Real（全部） Obj	OPV2V（观测） Rd	OPV2V（观测） Obj	V2V4Real（观测） Obj
BEV			✓	72.5	74.1	45.1	76.1	75.8	46.1
EviBEV		✓	✓	75.0	**75.3**	44.5	78.3	**76.3**	45.3
GevBEV$^-$	✓	✓		59.7	73.1	46.0	62.5	73.2	46.7
GevBEV	✓	✓	✓	**79.5**	74.7	**46.3**	**83.1**	76.1	**46.9**

表2：所提出模块的消融研究。所有交并比结果均以百分比表示。最佳值以粗体突出显示。G.s.：基于点的空间高斯，$\mathcal{L}_{edl}$：证据损失，Ag.：自由空间增强；Rd：道路，Obj：物体。

此外，如表 [16]1 IoU_{all} 所示，在道路头的方面，鸟瞰图和 EviBEV 的表现不如 CoBEVT。这是因为我们的框架基于全稀疏卷积网络，不会在未观测区域进行运算。为了便于与稠密卷积模型比较，在计算 IoU_{all}时，我们的框架会将某些路面视为观测不足，因此将其作为错误预测处理。与 GevBEV 不同，由于缺少高斯尾部，我们的鸟瞰图和 EviBEV 模型都将额外区域标记为未观测，因此导致交并比数值较低。不过，值得一提的是，消融模型的物体分类性能显著优于排名第二的模型 CoBEVT。这可归因于两个因素。首先，我们谨慎地控制模型中的网络，使其仅在特定范围内扩展坐标，并只对可观测区域进行预测。坐标扩展模块能够覆盖大部分物体区域，使这些区域获得预测结果，而不是被视为未观测区域。其次，得益于训练期间从连续驾驶空间进行动态采样，我们的模型在对车辆点作出正类预测时表现得较为谨慎。这种谨慎的方式使我们能够更有效地捕捉车辆的边缘细节，从而提升整体物体分类性能。

对比 GevBEV$^-$ 与 GevBEV，交并比的提升（尤其是道路方面）表明，自由空间增强（Ag.）为射线路径上的未占用空间提供了明确线索，并提升了检测性能。此外，该模块在缓解点云稀疏性问题方面发挥着重要作用，大幅提升了预测性能。总体而言，GevBEV 在所有测量指标上均优于消融模型，这验证了所提出各模块的有效性。

4.6.2 视觉分析

用于自动驾驶的整体鸟瞰图。 借助我们提出的概率模型，我们生成了GevBEV地图，并在图10中将复杂驾驶场景下的结果进行了可视化。从左到右，第一行的三个子图分别展示了不确定性、分类置信度以及路面真实标注的结果。在不确定性地图中，颜色越浅

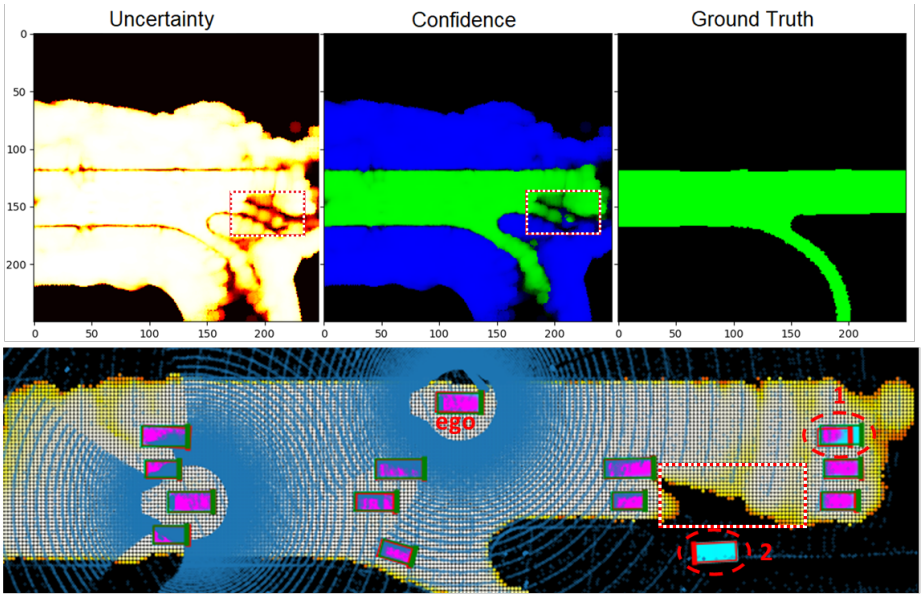

图10：GevBEV地图的结果。在 *Uncertainty map*中，较低的不确定性以较浅的颜色表示。在 *Confidence map*中，道路置信度以不同深浅的绿色表示，背景置信度则以不同深浅的蓝色表示。底部子图以更大的范围展示了更详细的检测结果。具体而言，原始输入点云以蓝色点表示；不确定性低于0.7阈值的道路点以浅色突出显示；物体以边界框表示，其中红色表示检测结果，绿色表示对应的真实标注。边界框前方的粗线表示行驶方向。此外，边界框内填充了从目标头输出的点置信度：洋红色点表示置信度较高，青色则表示置信度较低。

表示不确定性越低，而黑色区域则被视为不可观测区域。相应地，置信度地图给出了前景（路面）和背景的置信度分数。底部子图展示了叠加在同一图中的路面和物体的详细检测结果。在底层中，只有被分类为道路且不确定性低于0.7 阈值的点，才会以浅色突出显示，以展示可行驶区域的情况。预测边界框及其对应的真实边界框分别以红色和绿色绘制。此外，边界框中填充了从物体头绘制出的点置信度，其中洋红色点表示高置信度，而青色点则表示相反情况。

如图10所示，GevBEV 地图被视为整体鸟瞰图，可提供可靠的信息来源，支持自动驾驶车辆进行决策。首先，这些信息可以追溯到原始测量点。例如，由于遮挡和检测范围较远，红色虚线矩形标出的区域是自车协同自动驾驶车辆无法确定的区域。如果自车协同自动驾驶车辆需要驶入该不确定区域，就应向其他网联自动驾驶汽车请求缺失的信息，或减速等待障碍清除

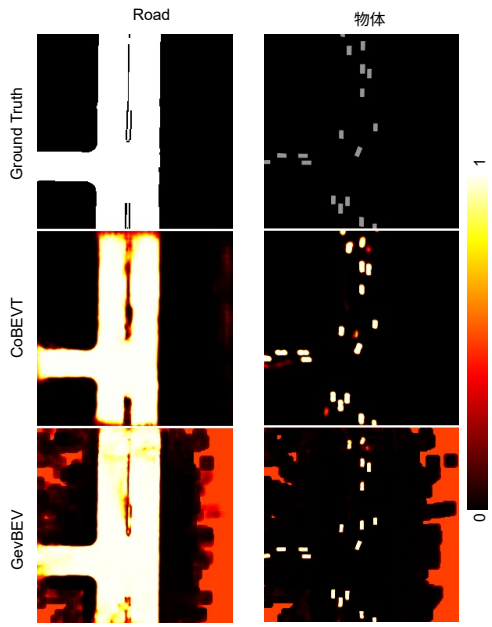

图11：GevBEV 与 CoBEVT 在 OPV2V数据集上的分类置信度对比。

该不确定区域。其次，物体的头部会生成关于可能被车辆占据区域的证据。因此，这些证据是验证边界框检测的可靠且可解释的线索。例如，如底部子图所示，除红色椭圆标出的两个框外，大多数预测框都与真实标注对齐良好。椭圆 1 中的车辆仅被部分观测到，没有足够的证据支持这项检测。同样，椭圆 2 中几乎没有证据支持这项假阳性检测。因此，可以直接移除这项检测。

与基线的比较。图11 和12 展示了分类置信度，颜色由深到浅变化，表示置信度值从 0 到 1。在图11 的右列和图12的底行中，红色表示我们的模型保留了有关未观测区域的信息。在缺少观测的情况下，模型不会偏向于预测前景类别或背景类别。此外，我们的模型对道路和车辆边缘都能生成更精细的细节。值得注意的是，在图12中，CoBEVT 往往会产生更多假阳性预测，在某些情况下，车辆甚至会连在一起。相比之下，得益于精准的边缘描述，我们的模型能够准确地区分所有车辆。

物体点分布的证据。此外，我们使用平均证据分数，更好地展示检测质量与

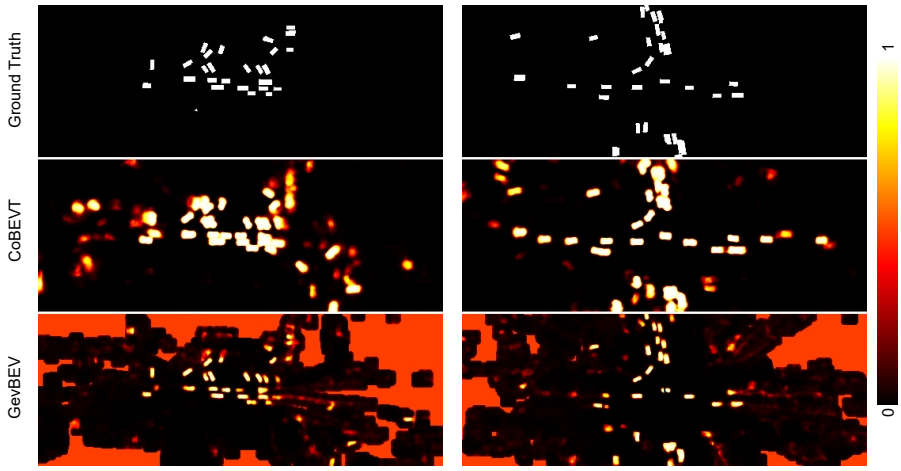

图12：GevBEV 与 CoBEVT 在 V2VReal 数据集上的分类置信度对比。

相应物体点分布之间的关系。受 Wang 的研究启发等人 [55], 除了检测区域上的常规交并比外，我们还利用生成的不确定性来计算 JIoU。JIoU 是交并比的概率版本，能够更好地评估目标检测的概率特征。与 [55] 将 JIoU 定义为检测结果与真实边界框所覆盖概率质量之间的交并比略有不同，我们将其定义为检测边界框内证据之和与描述该物体的所有证据质量之和之间的交并比。然后，我们对单个检测边界框的平均证据分数进行校准，该分数是边界框内各点所绘制置信度的均值。这样定义的 JIoU 可确保：当预测框被强证据完全填满，却未覆盖描述该物体的所有证据时，其证据分数较低。

图 13中的结果表明，平均证据分数与检测质量高度相关。如最后一行所示，检测效果越差，预测边界框内的证据往往越少。此外，JIoU 反映了预测边界框与真实边界框相对于证据质量的对齐程度。对于测量提供的线索不足的物体，即使通过人工标注，也很难定义出完美的确定性真实值，因此也难以据此公平地评估模型的预测。在这种情况下，检测边界框和真实边界框在 GevBEV 地图上的证据覆盖率都很低。这样一来，这两个边界框之间的概率并集会更小，因此得出的 JIoU 更高。与交并比相比，这种指标更合理，因为 JIoU 将模型的不完美性与测量的不完美性区分开来。例如，左下子图中的检测边界框交并比较低，但 JIoU 相对较高，这是因为观测点稀疏导致证据过少，即测量不完美性所致。相比之下，在右下子图中，与真实标注相比，检测结果

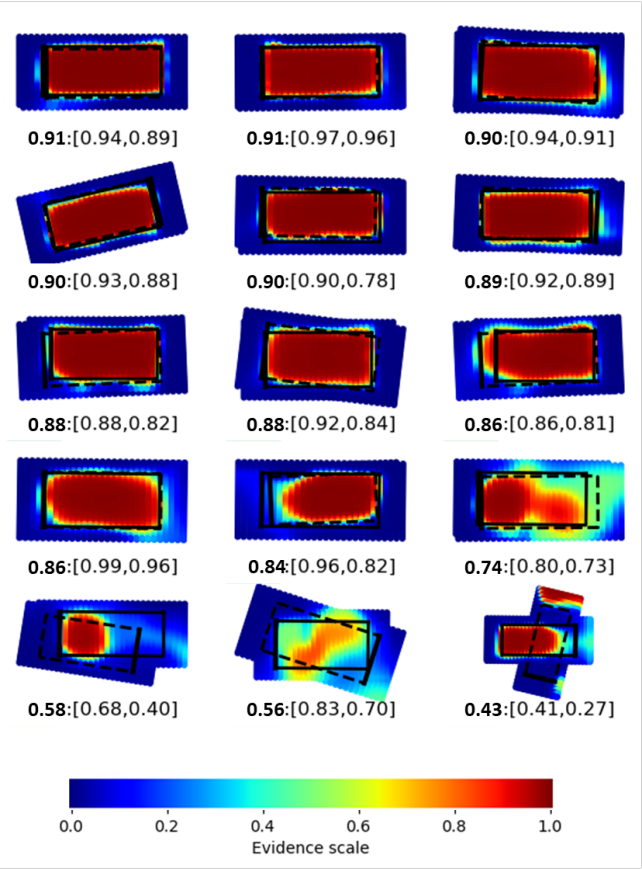

图13：物体点分布。虚线边界框表示检测结果，实线边界框表示对应的真实标注。每个边界框上的粗线段表示行驶方向。每个子图下方的统计数据表示平均证据和 [JIoU、IoU]。

证据充足，但 JIoU 和交并比都较低，这表明检测效果不佳并非测量所致，而是模型性能有限。

通过校准图呈现期望置信水平。我们使用校准图（图14）以及该图与完美不确定性-准确率线（黑色虚线）之间的平均偏移量（表3），作为分析基线模型 CoBEVT 和 BEV、采用 $\mathcal{L}_{edl}$ 损失的 EviBEV 模型，以及我们的完整概率模型 GevBEV 所生成的不确定性质量的汇总指标。由于基线模型仅为每个类别生成 Softmax分数，我们将这些分数转换为熵以量化不确定性，并将其与另外两个基于狄利克雷分布计算证据不确定性的模型进行比较。如图14所示，在道路和物体分类任务中，相较于基线模型，GevBEV 和 EviBEV 的置信度图都更接近完美校准线（即对角虚线）。

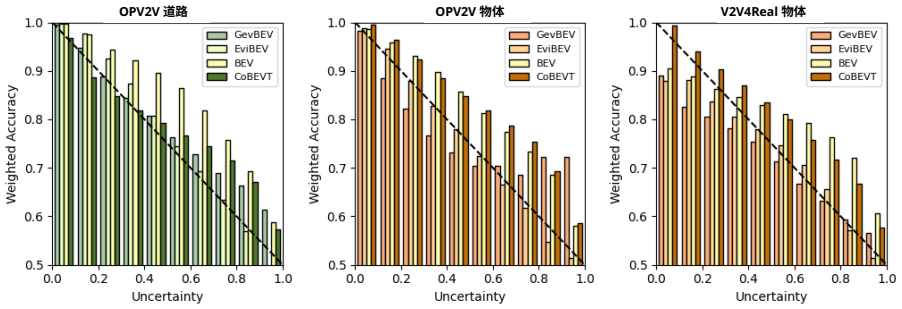

图 14：不同模型的校准图。对角虚线表示完美校准线。

模型	OPV2V Road	OPV2V 物体	V2V4Real 物体	平均值
BEV	0.095	0.072	0.076	0.081
EviBEV	**0.031**	**0.010**	**0.030**	**0.023**
CoBEVT	0.044	0.075	0.056	0.058
GevBEV	0.044	0.066	0.040	0.050

表3：校准图与完美校准线之间的平均偏差。最佳值以粗体突出显示，次佳值以下划线标出。

基线模型 BEV 和 CoBEVT 在道路和物体分类任务中的表现亦是如此。这两个基线模型似乎比另外两个模型更高估了不确定性，这印证了我们的担忧：未专门考虑不确定性的确定性模型，最终可能会生成不那么可信的分数，进而影响驾驶决策。另一方面，结果表明，假设鸟瞰图中点类别服从狄利克雷分布，可以为地图提供更可靠的概率特征，因此在自动驾驶汽车感知系统中使用更安全。

有趣的是，GevBEV 的不确定性质量比 EviBEV 差，尤其是在物体分类方面。这可能是由于相邻中心点贡献的证据求和出现饱和所导致的。此外，GevBEV 的物体头输出中高度不确定的点往往表现得不够自信。我们推测，部分车辆由于遮挡只能被观测到一部分。我们有限的坐标扩展（$1.2\,m$）只能覆盖这些车辆的一部分。因此，只有这些扩展后的中心点对应的分布瓦片能够覆盖车辆车身的其余部分。这可能会导致较高的不确定性，但同时具有较高的正例率。尽管 GevBEV 的不确定性比 EviBEV 略微更保守，但如表 2中更高的交并比所示，学习到的空间高斯分布能够生成更平滑的鸟瞰图，并绘制连续二维鸟瞰空间中任意点的分类分布。

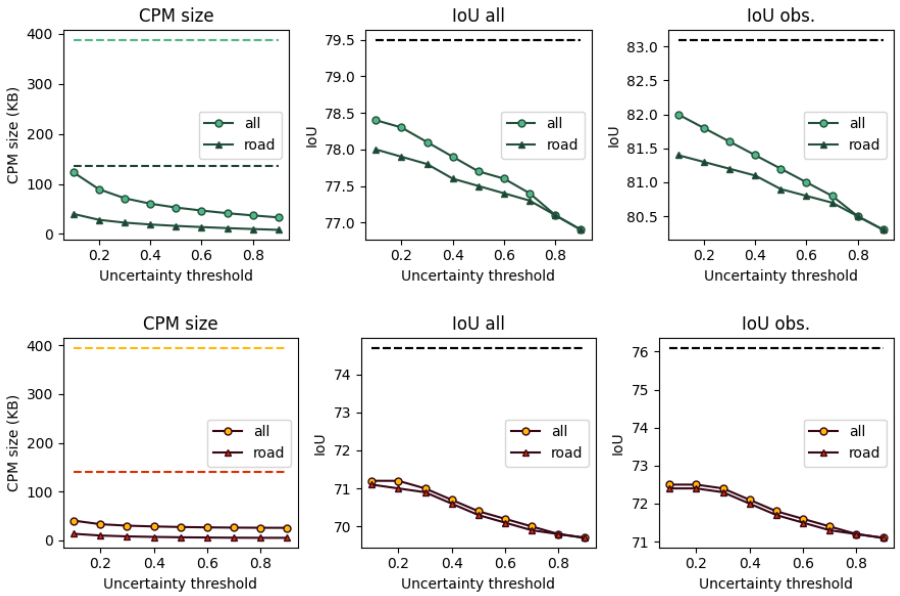

图15：协同感知中不同 CPM 共享策略的比较。第一行显示道路头的结果（偏绿色），第二行显示目标头的结果（偏橙色）。

4.7. GevBEV 在协同感知中的应用

利用不同不确定性阈值生成的证据型 GevBEV 地图，被用于选择网联自动驾驶汽车之间传输的信息。在不同不确定性阈值下，基于不确定性的信息选择前后协同感知消息的大小，以及分类结果在感知范围（IoU_{all}）内所有区域和可观测区域（IoU_{obs}）内相应的交并比，均绘制于图15中。第一行展示道路头的结果，第二行展示物体头的结果。如第 3.3节所述，我们针对两种 CPM共享策略进行了实验：一种是共享所有感知区域的掩码证据图（浅绿色和橙色），另一种是仅共享由现有高精地图约束的道路区域（深绿色和橙色）。虚线表示未进行信息选择时共享协同感知消息的基线结果，在不同不确定性阈值下显示为水平线。

第一列的图表显示，在所有不确定性阈值下，经过信息选择后，协同感知消息的大小相较于相应基线均显著减少。例如，当共享所有感知区域（浅绿色）时，在不确定性阈值为 0.5 的情况下，道路头的协同感知消息大小从 388 KB 降至 52 KB，减少约 87%。相应地，IoU_{all} 和 IoU_{obs} 仅减少约 2%。在相同配置下，目标检测的协同感知消息大小从 395 KB 降至 27 KB，减少约 93%，而交并比仅下降约 4%。仅考虑共享道路区域时，协同感知消息大小还可进一步降至道路头约 16 KB、

在不确定性阈值为 0.5 时，目标头的协同感知消息大小可降至约 6 KB，如深绿色和橙色实线所示，交并比仅下降 0.5% 以内。根据 V2V通信协议 [56]，在不考虑其他通信开销的情况下，数据吞吐率可达到 27 Mbps。因此，基于我们从道路区域生成的 GevBEV地图进行数据选择后，发送两个头的协同感知消息的时延从约 28 ms 降至 0.8 ms。如此显著缩短的时延对于实时 V2V通信至关重要。

5. 结论

本文提出了一种新颖的方法，利用可观测的高斯证据鸟瞰图来解释驾驶环境。这些地图以可追溯的方式解释激光雷达感知数据，使每项预测都由原始观测点提供的证据所支持。此外，我们设计了一种基于 U-Net 的概率分类模型，为这种解释生成统计信息。该模型假设预定义分辨率下的每个体素均服从空间高斯分布，因此连续驾驶空间中的任意点都可以根据从相邻体素的空间高斯分布中获取的证据，为自身生成一个狄利克雷分类分布。

我们在 OPV2V 和 V2V4Real 这两个基准测试集上，分别针对仿真和真实世界驾驶场景中的协同感知鸟瞰图解释任务，对所提出的 GevBEV 进行了测试。实验表明，GevBEV 大幅优于基于图像的鸟瞰图基线。通过分析预测不确定性，我们还证明，与同一模型的确定性对应模型相比，证据分类能够以过度自信程度更低、校准效果更好的方式评估分类结果。此外，事实也证明，为每个可观测点分配空间高斯分布，有助于在可控范围内弥补稀疏点云之间的空缺，并平滑鸟瞰图。借助这一空间分布，可以为连续驾驶空间中的任意点生成狄利克雷分类结果。这一概率结果能够利用测量证据量化不确定性，从而为自动驾驶制定更安全的决策。

尽管本文的模型仅应用于协同感知场景中的点云数据，但将其用于其他模态的数据源或场景也十分直接，例如用于协同感知的图像，或遥感领域中的卫星图像和机载点云。我们将这些应用留待未来工作。

6. 致谢

本研究得到德国科学基金会（DFG，德国研究基金会）项目 227198829/GRK1931 以及 MSCA 欧洲博士后奖学金项目 101062870——VeVuSafety 的资助。

参考文献

[1] Y. Lin、G. Vosselman、Y. Cao、M. Y. Yang，《用于机载激光扫描点云语义分割的局部与全局编码器网络》，《ISPRS摄影测量与遥感》176（2021年）151–168。1[2] D. Feng、A. Harakeh、S. L. Waslander、K. Dietmayer，《自动驾驶中概率目标检测的综述与比较研究》，IEEE智能交通系统汇刊 23（8）（2022）9961–9980。1[3] Y. Zhang、A. Carballo、H. Yang、K. Takeda，《恶劣天气条件下自动驾驶车辆的感知与传感：综述》，《ISPRS摄影测量与遥感》196（2023年）146–177。1[4] L. Fang、Z. You、G. Shen、Y. Chen、J. Li，《结合点云与多视图的联合深度学习网络：基于激光雷达点云的道路侧目标分类》，《ISPRS摄影测量与遥感》193（2022）115–136。1[5] A. Zang、R. Xu、Z. Li、D. Doria，《从卫星影像中提取车道边界》，载于：第1届 ACM SIGSPATIAL 自动驾驶车辆高精度地图与智能应用研讨会论文集，2017年，第1–8页。1[6] A. Geiger、P. Lenz、R. Urtasun，《我们准备好自动驾驶了吗？KITTI视觉基准套件》，载于：计算机视觉与模式识别会议（CVPR），2012年。1[7] H. Caesar、V. Bankiti、A. H. Lang、S. Vora、V. E. Liong、Q. Xu、A. Krishnan、Y. Pan、G. Baldan、O. Beijbom，《nuScenes：面向自动驾驶的多模态数据集》，载于：计算机视觉与模式识别会议（CVPR），2020年。1[8] P. Sun、H. Kretzschmar、X. Dotiwalla、A. Chouard、V. Patnaik、P. Tsui、J. Guo、Y. Zhou、Y. Chai、B. Caine、V. Vasudevan、W. Han、J. Ngiam、H. Zhao、A. Timofeev、S. Ettinger、M. Krivokon、A. Gao、A. Joshi、Y. Zhang、J. Shlens、Z. Chen、D. Anguelov，《自动驾驶感知的可扩展性：Waymo开放数据集》，载于：计算机视觉与模式识别会议（CVPR），2020年，第2443–2451页。1[9] J. Gawlikowski、C. R. N. Tassi、M. Ali、J. Lee、M. Humt、J. Feng、A. M. Kruspe、R. Triebel、P. Jung、R. Roscher、M. Shahzad、W. Yang、R. Bamler、X. Zhu，《深度神经网络不确定性综述》，arXiv abs/2107.03342（2021年）。2，5

[10] Y. Yuan, H. Cheng, M. Sester，基于关键点的深度特征融合，用于自动驾驶中的协同车辆检测，IEEE Robotics and AutomationLetters PP（2022）1–1。 2，10，11

[11] R. Xu, J. Li, X. Dong, H. Yu, J. Ma，弥合多智能体感知的域差距，arXiv预印本 arXiv:2210.08451（2022）。 2[12] J. Cui, H. Qiu, D. Chen, P. Stone, Y. Zhu，Coopernaut：面向联网车辆的协同感知端到端驾驶，载于：计算机视觉与模式识别会议（CVPR），2022，第17231–17241页。 2[13] R. Xu, H. Xiang, X. Xia, X. Han, J. Li, J. Ma，Opv2v：用于车辆间通信感知的开放基准数据集与融合流水线，载于：IEEE国际机器人与自动化会议（ICRA），2022。 2、3、5、7、15、18、19、20[14] M. Sensoy, L. Kaplan, M. Kandemir，证据深度学习：量化分类不确定性，神经信息处理系统进展（NeurIPS）31（2018）。 3、6、13[15] R. Xu, X. Xia, J. Li, H. Li, S. Zhang, Z. Tu, Z. Meng, H. Xiang, X. Dong, R. Song, H. Yu, B. Zhou, J. Ma，V2v4real：用于车辆间协同感知的真实世界大规模数据集，载于：IEEE/CVF计算机视觉与模式识别会议（CVPR），2023。 3、15、19[16] R. Xu, Z. Tu, H. Xiang, W. Shao, B. Zhou, J. Ma，Cobevt：采用稀疏Transformer的协同鸟瞰图语义分割，载于：机器人学习会议（CoRL），2022。 3、4、5、16、18、20、22[17] L. Jiao, F. Zhang, F. Liu, S. Yang, L. Li, Z. Feng, R. Qu，基于深度学习的目标检测综述，IEEE Access 7（2019）128837–128868。 4[18] J. Li, R. Xu, J. Ma, Q. Zou, J. Ma, H. Yu，雾天天气下自动驾驶的域自适应目标检测，载于：IEEE冬季计算机视觉应用会议（WACV），2022，第612–622页。 4[19] A. Kirillov, K. He, R. B. Girshick, C. Rother, P. Doll´ar，全景分割，计算机视觉与模式识别会议（CVPR）（2018）9396–9405。 4[20]H. Qiu, B. Yu, D. Tao，GFNet：用于三维点云语义分割的几何流网络，机器学习研究汇刊（2022）。 4[21] B. Zhou, P. Kr¨ahenb¨uhl，跨视角Transformer：用于实时地图视角语义分割，载于：计算机视觉与模式识别会议（CVPR），2022。 4[22] A. Loukkal, Y. Grandvalet, T. Drummond, Y. Li，在扁平移动物体之间驾驶：利用单目相机生成鸟瞰图占用栅格，以实现整体轨迹规划，IEEE冬季计算机视觉应用会议（WACV）（2021）51–60。 4

[23] C. Lu, M. J. G. van de Molengraft, G. Dubbelman，采用卷积变分编码器-解码器网络的单目语义占据栅格建图，IEEE机器人与自动化快报 4（2019年）445–452。4[24] B. Pan, J. Sun, H. Y. T. Leung, A. Andonian, B. Zhou，用于周围环境传感的跨视角语义分割，IEEE机器人与自动化快报 5（2020年）4867–4873。4[25] Z. Li, W. Wang, H. Li, E. Xie, C. Sima, T. Lu, Q. Yu, J. Dai，BEVFormer：通过时空Transformer从多摄像头图像学习鸟瞰视图表示，载于：欧洲计算机视觉会议，2022。4[26] B. Yang, W. Luo, R. Urtasun，PIXOR：基于点云的实时三维目标检测，计算机视觉与模式识别会议（CVPR）（2018）7652–7660。4[27] A. H. Lang, S. Vora, H. Caesar, L. Zhou, J. Yang, O. Beijbom，PointPillars：用于点云目标检测的快速编码器，载于：计算机视觉与模式识别会议（CVPR），2019，第12689–12697页。4[28] Y. Zhou, O. Tuzel，VoxelNet：面向基于点云的三维目标检测的端到端学习，计算机视觉与模式识别会议（CVPR）（2018）4490–4499。4[29] A. Dosovitskiy, G. Ros, F. Codevilla, A. Lopez, V. Koltun，CARLA：一个开放的城市驾驶模拟器，载于：机器人学习会议，PMLR，2017，第1–16页。5, 15[30] R. Xu, Y. Guo, X. Han, X. Xia, H. Xiang, J. Ma，OpenCDA：一个集成协同仿真的开放协同驾驶自动化框架，载于：国际智能交通系统会议（ITSC），2021，第1155–1162页。5, 15[31] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, I. Polosukhin，Attention Is All You Need，神经信息处理系统进展（NeurIPS）30（2017）。5[32] T.-H. Wang, S. Manivasagam, M. Liang, B. Yang, W. Zeng, R. Urtasun，V2VNet：用于联合感知与预测的车对车通信，载于：欧洲计算机视觉会议，Springer，2020，第605–621页。5, 18, 20[33] Y. Li, S. Ren, P. Wu, S. Chen, C. Feng, W. Zhang，学习用于多智能体感知的蒸馏协作图，神经信息处理系统进展（NeurIPS）34（2021）29541–29552。5, 18, 20

[34] R. Xu, H. Xiang, Z. Tu, X. Xia, M.-H. Yang, J. Ma，V2X-ViT：采用视觉变换器的车路协同感知，arXiv预印本 arXiv:2203.10638（2022）。5, 18, 20 [35] Q. Chen，F-Cooper：一种基于特征的协同感知方法，用于采用三维点云的自动驾驶车辆边缘计算系统，第四届ACM/IEEE边缘计算研讨会论文集（2019年）。5, 18, 20[36] A. Kendall, Y. Gal，贝叶斯深度学习中计算机视觉需要哪些不确定性？，载于：神经信息处理系统进展（NeurIPS），2017。5 [37] A. Malinin, M. J. F. Gales，通过先验网络进行预测不确定性估计，载于：神经信息处理系统进展（NeurIPS），2018。5, 6[38] D. J. C. Mackay，反向传播网络的实用贝叶斯框架，神经计算（1991）。5[39] R. M. Neal，神经网络的贝叶斯学习，1995年。5[40] Y. Gal，深度学习中的不确定性，载于：博士论文，2016。5[41] B. Lakshminarayanan, A. Pritzel, C. Blundell，利用深度集成进行简单且可扩展的预测不确定性估计，载于：神经信息处理系统进展（NeurIPS），2017。5[42] D. Feng, L. Rosenbaum, F. Timm, K. C. J. Dietmayer，利用异方差偶然不确定性实现稳健的实时激光雷达三维目标检测，IEEE智能车辆研讨会（IV）（2019年）1280–1287。6[43] G. P. Meyer, A. G. Laddha, E. Kee, C. Vallespi-Gonzalez, C. K. Wellington，LaserNet：一种用于自动驾驶的高效概率三维目标检测器，载于：计算机视觉与模式识别会议（CVPR），2019，第12669–12678页。6[44]D. Miller, F. Dayoub, M. Milford, N. S¨underhauf，评估目标检测中基于采样的不确定性技术的融合策略，国际机器人与自动化会议（ICRA）（2019）2348–2354。6[45] H. Pan, Z. Wang, W. Zhan, M. Tomizuka，迈向自动驾驶汽车三维目标检测中更优性能和更具可解释性的不确定性，国际智能交通系统会议（ITSC）（2020年）1–7。6[46] D. Feng, Y. Cao, L. Rosenbaum, F. Timm, K. C. J. Dietmayer，在自动驾驶中利用不确定性进行深度多模态目标检测，IEEE智能车辆研讨会（IV）（2020年）877–884。6

[47] V. T. Vasudevan, A. Sethy, A. R. Ghias，关于提升神经模型置信度估计的研究，2019年，第7335–7339页。6[48] A. P. Dempster，贝叶斯推断的一种推广，载于：Dempster-Shafer信念函数理论经典著作，2008年。6[49] J. D. Lafferty, A. McCallum, F. Pereira，条件随机场：用于分割和标记序列数据的概率模型，载于：国际机器学习会议，2001年。6[50] O. Ronneberger, P. Fischer, T. Brox，U-Net：用于生物医学图像分割的卷积网络，载于：国际医学图像计算与计算机辅助干预会议，Springer，2015年，第234–241页。7，8[51] C. Choy, J. Gwak, S. Savarese，四维时空卷积网络：Minkowski卷积神经网络，载于：计算机视觉与模式识别会议（CVPR），2019年，第3075–3084页。9，11[52] W. Zheng, W. Tang, S. Chen, L. Jiang, C.-W. Fu, CIA-SSD：基于点云的高置信度感知交并比单阶段目标检测器，载于：AAAI，2021年。10[53] R. Xu, H. Xiang, X. Han, X. Xia, Z. Meng, C.-J. Chen, J. Ma, 面向协同驾驶自动化研究的OpenCDA开源生态系统，arXiv预印本 arXiv:2301.07325（2023年）。15[54] D. P. Kingma, J. Ba，Adam：一种随机优化方法，载于：ICLR，2014年。19[55] Z. Wang, D. Feng, Y. Zhou, W. Zhan, L. Rosenbaum, F. Timm, K. C. J. Dietmayer, M. Tomizuka，推断目标检测中的空间不确定性，国际智能机器人与系统会议（IROS）（2020年）5792–5799。25[56] F. Arena, G. Pau，车载通信概述，《未来互联网》11（2019年）27。29