

# 自动越野导航的实时语义映射

丹尼尔·马图拉娜，周波伟，青山正树和塞巴斯蒂安·谢尔

摘要本文描述了一种全地形车辆(ATV)自主越野驾驶的语义映射系统)。该系统的目标是提供比纯几何地图更丰富的环境表示,使其能够区分,例如。来自障碍物的高草。该系统构建了一个2.5D网格地图,编码几何(地形高度)和语义信息(导航相关类,如小径、草地等。))。分别从LiDAR和图像传感器数据中在线和实时估计几何信息和语义信息。使用这个语义图,运动规划者可以创建语义感知的轨迹。为了实现鲁棒和高效的语义分割,我们设计了一个定制的卷积神经网络(CNN),并为此建立了一个新的标记越野图像数据集。我们离线评估我们的语义分割,显示出与技术状态相当的性能,延迟很低。我们还展示了一个自主的ATV驾驶在具有挑战性的越野地形上的闭环场结果,使用语义图和一个简单的路径规划器。我们的模型和标签数据集将公开提供。

## 1 引言

在过去的几年里,自主车辆的三维传感能力取得了巨大的进步。成熟和健壮的LiDAR和INS技术为自动驾驶车辆提供了对其周围几何结构的准确和实时的感觉,极大地简化了与导航相关的任务。

然而,我们已经观察到,主要依赖几何信息导致令人失望的结果,自主导航在越野环境。

丹尼尔·马图拉娜-周波伟-塞巴斯蒂安·谢尔  
卡内基梅隆大学机器人研究所,电子邮件: dimatura@cmu.edu

Masahi Uenoyama  
雅马哈汽车公司美国,电子邮件: mike\_uenoyama@yamaha-motor.com



无花果。 1 我们的自动全地车)。 两个主要的传感器, 一个旋转的 3D 李 DAR 和一个立体声相机, 可以在车辆屋顶上看到。

主要原因是几何结构本身不能提供许多重要的区别轮式全地形车辆 (ATV), 如我们的, 如 1 所示。 例如, 高草可能被认为是一个障碍, 但我们的 ATV 可能穿越它, 如果需要的话。 同样, 落叶可能以岩石地形出现, 或水坑可能以孔或光滑表面出现。 所有这些都可能导致路径规划中的次优甚至危险的决策。 在越野机器人方面, 例如在 [11, 15, 22, 10] 方面, 也曾多次提出类似的意见。

在本文中, 我们描述了一个系统来解决这个问题, 通过建立一个语义图, 一个表示车辆的环境编码两个几何 (例如。 高度、粗糙度) 和语义信息 (导航相关类, 如小径、草地、障碍物等。)。 地图存储为以车辆框架为中心的 2.5D 网格, 并随着新的传感器数据的获取而不断更新。 使用这种表示, 运动规划器可以创建语义感知的轨迹。

我们的关键贡献是一个简单而有效的系统耦合一个定制的卷积神经网络体系结构, 基于全卷积网络 [14], 和一个 2.5D 以车辆为中心的语义网格地图, 融合几何和语义测量, 随着车辆的移动和获取更多的数据。 我们展示了语义分割 CNN 在离线基准中的有效性。 通过使用带有语义图的简单规划器, 我们展示了我们的系统被成功地用于导航具有挑战性的越野地形的定性示例。

作为额外的贡献, 用于培训我们的网络的越野图像标记数据集将公开提供。

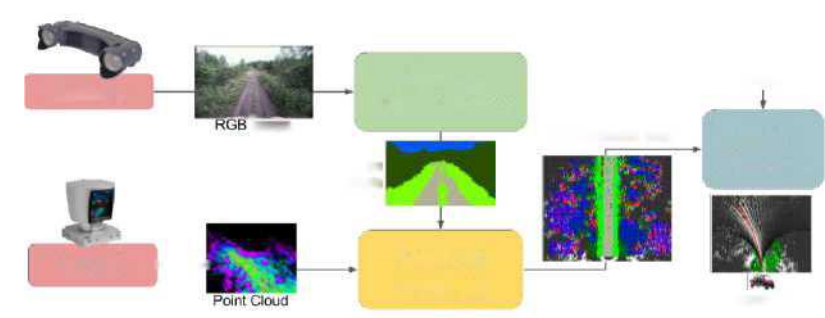
2 相关工作

我们的系统在很大程度上受到了关于越野导航任务语义方法的丰富文献的启发，可以追溯到 1990 年[7]。

十年后，各种实际系统在这种范式下显示了令人印象深刻的结果，通常是李 DAR 和图像[11, 15, 22, 25, 26]的结合。

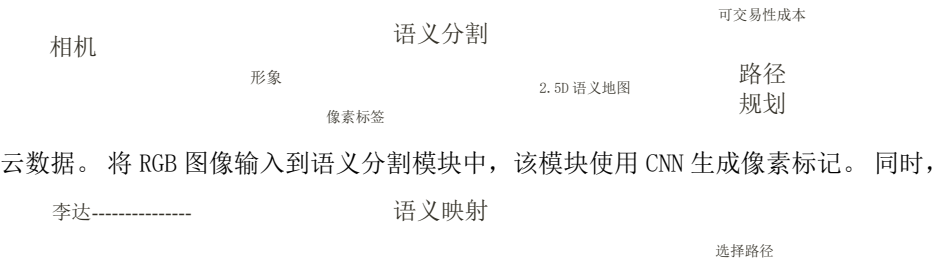
LAGR 程序的特点[10]各种高度相关的系统，如[18, 12, 24, 2]，它执行语义分类与手工设计的视觉管道。一个例外是[9]的，具有早期的深层神经网络系统进行语义分割。

在最近的工作中，[21]演示了具有轻量级语义分割系统的自主导航。与我们的系统不同，它们使用传统的视觉特征工程，导致噪声像素级预测，它们用一种新的正则化方法平滑。相反，我们的体系结构，基于完全卷积网络 (FCNs) [14]，结合空间上下文，自然平滑输出。另一项相关工作是[23]，它使用类似于 FCN 的编解码器网络体系结构。他们探索 RGB 以外的模式，并取得令人印象深刻的分割结果。然而，他们没有构建度量映射或演示闭环导航。



3 方法

概述。 该系统体系结构如图所示。 2. 有两个主要的传感器流，RGB 图像和 LiDAR 点



无花果。 2 我们的语义映射系统概述。

4. LiDAR 点云用于更新语义映射模块中的图5D网格映射。语义映射模块也重新

将语义分割模块中的像素级预测图像投影到 2.5D 网格映射上，随着时间的推移融合语义预测。其结果是一个以车辆为中心的 2.5D 网格地图编码连续更新的估计相关的几何和语义信息的越野导航。最后，该映射用于语义感知路径规划。对于我们的初始测试，我们使用了一个简单的后退视界路径规划器，它为每个语义类分配遍历成本，并连续选择一条路径以最小化成本。整个系统以 10Hz 运行，速度取决于语义映射模块处理图像的速度。

硬件平台。我们的车辆如图所示。对 LItisa 商用全地列车进行了改装和仪器化，用于自主越野驾驶实验。传感器套件包括一个 INS/GPS 系统，一个 64 线 Velodyne LIDAR 和一个 RGB 立体相机与 21 厘米基线由卡内基机器人制造。请注意，本文中的系统目前没有使用立体声深度信息。所有计算都是在船上进行的，有两台 COTS 笔记本电脑，通过高速以太网连接。用于语义映射的笔记本电脑包括一个 NVIDIA GT980M GPU，用于实现 CNN 分类器的实时执行。

软件平台。所有电脑都运行 UbuntuLinux。不同的系统模块作为 ROS 节点并发运行，并通过 ROS 消息进行通信。节点在 C++ 和 Python 中实现，使用 CUDA (通过 Theano 库 [3] 生成) 有效地使用 GPU。

### 3.1 语义分割

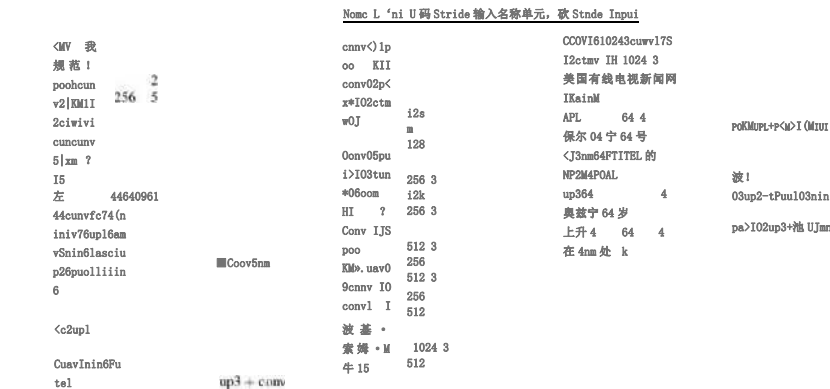
二维语义分割的目标是为图像中的每个像素分配一个 K 预定类。像计算机视觉中的许多任务一样，这项任务的艺术状态最近被深度学习，特别是卷积神经网络所彻底改变。

对于这项任务，最成功的神经网络体系结构是全卷积网络 (FCNs) [14]。这些网络的关键思想是利用卷积结构，用一个非常相似的网络与更传统的 CNN 同时标记所有像素。由于池，这导致低分辨率输出；为了扭转这种情况，将所谓的“反褶积”层添加到输出的上采样中。为了保持高频细节，增加了连接早期层到上采样特征映射的跳过层。Encoder-Decoder 体系结构 [17, 1]，其中 Up Net [23] 是一个例子，它们是相似的，但省略跳过层。

在项目开始时，我们发现艺术架构的状态相对较慢，因为它们被优化为精度超过速度。因此，我们使用 Theano [3] 和 Lasagne [6] 库来实现和训练我们自己的体系结构。目前，我们已经发现了各种可能的体系结构，以显示非常相似的准确性，我们的越野语义分割任务，主要不同的时间成本，这反过来主要是由详细的体系结构和输入/输出分辨率。

我们认为这是由于使用的数据集相对较小，但值得进一步调查。

到目前为止，我们主要使用了两种体系结构。第一个，CNNs-FCN，是基于我们对来自[4]的VGG-CNN的“卷积”，具有227x227输入大小和109x109输出大小。第二个，dark-fcn，是基于我们自己的Darknet架构[19]的卷积，这又是类似的，但比VGG16[20]更有效。对于暗fcnn，输入和输出都是300x300，以便于与UpNet进行比较。尽管分辨率较高，但在GT980M上，暗fcnn比cnnn-fcnn更快：21ms，而不是37ms。UpNet的作者[23]描述了一个50ms与Caffe在GTX泰坦X，这在我们的经验中有类似的速度GT980M。这使我们相信我们的模型应该更快，或者至少是可比的。无花果。



3 展示了我们的两种架构。 还将提供编码和经过培训的模型。

无花果。3 我们的两个网络架构。“conv”表示卷积层；“pool”表示池层；以“nin”结尾的层是1x1卷积层；“fuse”表示元素和层；“norm”表示局部响应归一化层。除非另有规定，否则假定输入是上面的层。对于卷积层，“大小”是内核大小；对于池层，它是池接受字段。请注意，对于暗fcnn，由于大小限制，我们将表分开。

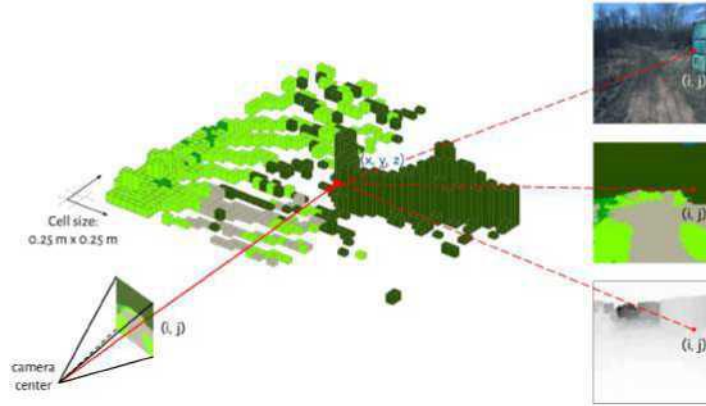
### 3.2 语义映射

语义映射步骤的输出是在二维图像空间中，但车辆在三维度量空间中规划要自然得多。在我们的例子中，为了简单起见，采用2.5D网格或高度图表示。这对大多数环境都足够，但可能会对悬垂的树木或隧道产生问题。

为了保持车辆周围环境的最新高度图，我们使用滚动网格数据结构，这在文献中已经被重新发明了多次。这种结构是环缓冲器向二维的推广，其主要特点是可以不复制其数据的情况下移动（翻译），而不是更新指示其极限的变量。这是一个速度优化；从逻辑上讲，网格的行为就像一个以车辆为中心的有限二维阵列，每个网格单元包含关于地形的各种属性。在我们的论文中，网格单元为0.25米

$\times 0$ 。每个 25 米，地图有  $400 \times 400$  个单元格。每个网格单元维护该网格单元中最小和最大高度的运行估计，该估计是通过使用来自 LiDAR 射线的占用率和自由空间约束来计算的，类似于 [8, 28]。对于点云中的每个点，我们使用 Bresenham 在 3D 中的算法在我们的网格上进行射线跟踪；通过的单元格和上面的单元格被认为是空的，光束停止的单元格和下面的单元格被认为被占用。

语义映射还集成了语义度量，如其名称所示。为了将二维语义分割的输出投影到高度图表示中，我们遵循图中描述的一个简单的过程。4.

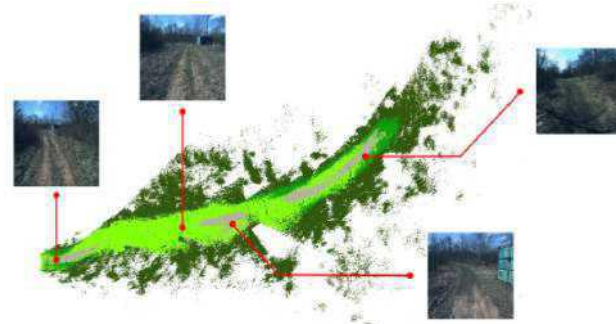


无花果。4 将语义分割映射到 2.5D 映射。

鉴于我们知道相机和 LiDAR 的相对位置，以及相机的本质，我们可以使用简单的几何图形将二维语义预测投影到 2.5D 网格单元上。然而，为了增加鲁棒性，我们将测量随时间融合。为此，我们采用了一种由占用映射 [16] 的顺序过滤过程启发的方案，但推广到 K 类。

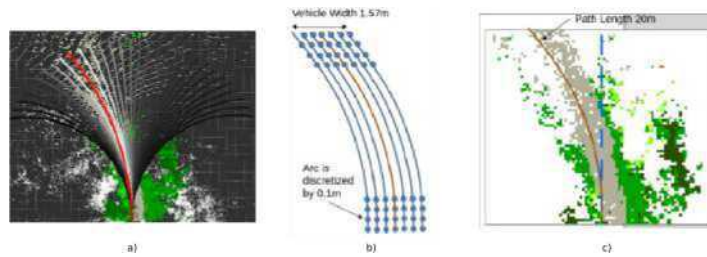
为此，我们使用分类器的概率 (Softmax) 像素输出。我们维护一个向每个网格单元投影的 K 类的日志赔率的运行总和。虽然这种软多类表示可以直接使用，但为了简单起见，当与其他系统接口时，我们使用 K 类的  $\text{argmax}$  作为我们当前对每个网格单元的语义类的最佳估计。请注意，此表示假设每个单元格有一个类，这在某些环境中可能是一个限制。

在现场运行中，语义映射的示例累积输出如图所示。5.



无花果。5 字段运行中语义映射的示例输出。

### 3.3 路径规划



无花果。6 路径规划。a) 候选路径库，覆盖在语义映射之上。红色表示可行路径。b) 说明我们如何解释车辆宽度。对于每个轨迹，我们计算的成（或奖励）超过七个移动版本的轨迹，涵盖车辆足迹。c) 一个例子选择轨迹，根据它所涵盖的语义类的可遍历性评分来选择。

为了演示自主操作，我们实现了一个非常简单的后退视界路径规划器。规划师有一个 30 条轨迹库，对应于  $-15^\circ$  /秒到  $15^\circ$  /秒的偏航率，以  $1^\circ$  /秒离散，以 9kmh 的恒定速度离散<sup>1</sup>见图。6a)。

每次地图被更新，发生在 10 赫兹，从库中选择一个轨迹。轨迹的选择最大化了从语义映射导出的奖励函数，如下所示。标记为“平滑”或“粗糙”痕迹的细胞奖励为 1，标记为“草”的细胞奖励为 0.1。其他班级均为零奖励。轨迹的总奖励是来自车辆的 20 米轨迹长度上的奖励之和。为了考虑车辆宽度，我们稍微修改了这个计算，如图 6b 所示)。

这个规划器的优点是，在其极端简单的情况下，它的性能在很大程度上取决于我们的语义映射的输出，没有来自其他因素的干扰，这些因素将是一个更复杂、多层次的系统中的存在。然而，我们的系统也被用作一个更仔细的规划师的额外输入，其中主要的表示是一个用 LiDAR 构建的几何地图。在这个规划中，我们的语义预测主要是为了避免把草当作障碍。



## 4 实验

概述。我们用两种方式来评价我们的系统。首先，我们在两个数据集中运行语义分离模块的离线基准。第二，在现场实验中演示了整个系统的自主操作。

### 4.1 离线基准

为了评估我们的语义分割模块，我们使用了两个数据集，来自 Valada 等人的深度场景数据集。[23]和我们自己的数据集，雅马哈-CMU 离线数据集。

深度场景数据集。该数据集由 233 幅训练图像和 139 幅验证图像组成，这些图像被密集地标记为六个语义类别：空洞、道路、草、植被、树木、天空和障碍。虽然这个数据集显示了一些有趣的变化，由于一天的时间，它是相当小的，似乎缺乏多样性的天气和地点。该数据集的一个关键特征是各种模式（深度、近红外），但我们目前没有使用它们。

雅马哈-CMU 越野数据集。为了训练和评估我们的方法，我们收集了我们自己的数据集，我们称之为雅马哈-CMU-Off-Road，或 YCOR。它由 1076 幅图像组成，收集在宾夕法尼亚州西部和俄亥俄州的四个不同地点（8），跨越三个不同的季节（图. 7）。数据集使用基于多边形的接口标记，有八个类：天空、粗糙的小径、光滑的小径、可穿越的草地、高植被、不可穿越的低植被、障碍物。使用密集的 CRF[13]对多边形标签进行后处理，以使标签致密化；手工检查 CRF 的输出，并在某些情况下纠正，以确保没有创建错误的标签。

我们相信我们的数据集比深度场景更多样化和更具挑战性。在图中。我们显示了每个数据集的平均 RGB 图像和像素标签模式。深度场景数据集显示了一个左-右偏差和更可预测的结构比我们的；如果我们使用像素模式作为基线分类器，我们将获得 0.51。深度场景中的 30 个像素级错误率，但为 0.51。然而，我们承认，与最近的努力相比，这两个数据集都相对较小。城市景观[5]，有 25000 个标签图像。



无花果。从 YCOR 数据集 7 帧的蒙太奇。



DeepScene Mean RGB



DeepScene Label Modes



Yamaha-CMU-OR Mean RGB



Yamaha-CMU-OR Label Modes



Yamaha-CMU-OR Locations

无花果。8 前两列：数据集统计数据的比较。我们在使用的每个数据集的训练集中显示标记帧的平均 RGB 帧和像素模式。最后一列：一张地图，上面有收集 YCOR 的位置。

我们目前的分裂有 931 个训练图像和 145 个验证图像。此拆分是随机生成的，确保在培训和验证拆分中的图像之间的数据收集会话没有重叠。然而，所使用的地点存在重叠。我们将提供获取元数据的位置和时间，以便能够在这些因素的泛化方面进行进一步的评估。

定量结果。我们在这两个数据集上评估了我们的模型。在每种情况下，我们从零开始在预定义的训练集上训练我们的模型，直到与 SGD 收敛，除以初始学习速率（0.0001），乘以 10 三倍。我们使用一个标准的像素交叉熵损失与一个小的 L2 正则化因子（0.0005）。培训大约需要两天的 GT980Ti GPU。我们使用作物，旋转和颜色增强在训练时间，没有在测试时间。我们使用每类交集超过联合（IoU）作为评价度量，这是语义分割最常见的度量。

表 1 显示了深度场景的结果，表 2 显示了 YCOR 的结果。在这两种情况下，除了标准的 300x300 外，我们还包括了 448x448 分辨率的暗 fcn 模型的变体。我们报告了他们的论文[23]中的数字，其中我们用频率加权的 IoU (FW-IoU) 表示它们表

示的是 IoU，并添加了由我们自己计算的平均 IoU (M IoU)。正如我们所看到的，我们的两个模型的性能相当，黑暗 FCN 有一个轻微的优势。在深度场景数据集中，我们还可以将这两种模型与 RGB 向上网络进行比较。我们看到，我们的模型在 FW-IoU 中有一个轻微的边缘，尽管它们在障碍物上表现出明显的更差的性能，这严重扭曲了 mIoU 度量。我们注意到，数据集中的障碍物像素数比其他类少三个数量级，因此网络倾向于忽略它。尽管如此，它是一个重要的类，我们正在研究如何更准确地检测它。最后，我们看到，增加输入分辨率会稍微提高性能。

表 1 每类，平均 IoU 和频率加权 IoU 的 Up 网 (RGB) 和我们的模型在深场景数据集中。前三行使用 300x300 图像大小，如 UpNet 中的那样；最后一行使用 448x448。

		路草蔬菜/树木天空障碍 mIoUFW-IoU				
上网 (RGB) [23]	85.03+86.78	<b>90.90</b>	<b>90.39</b>	<b>45.31</b>	85.	
CNNs-FCN	85.95	87.38	90.53	1.84	58.51	87.
黑暗-FCN	<b>88.03</b>	89.21	93.17	5.03	60.35	<b>89.</b>
黑暗-FCN-448	88.80	89.46	93.35	4.61	60.61	89.

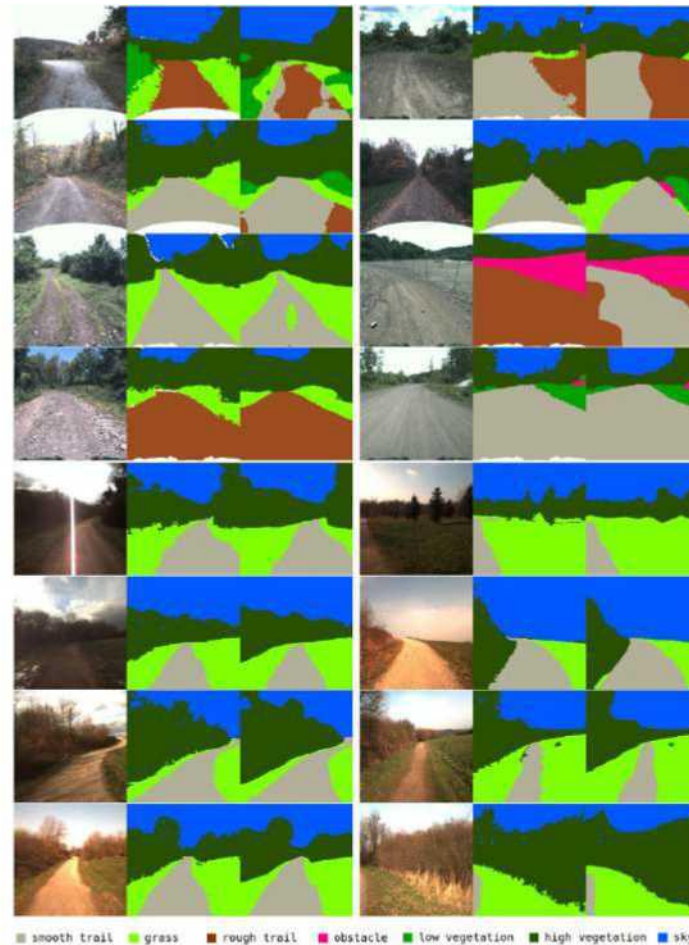
表 2 我们的模型在 YCOR 数据集中的每类、平均 IoU 和频率加权 IoU。

	光滑的草粗糙的水坑障碍低 VEG。高蔬菜。天空 mIoUFW-IoU							
CNNs-FCN	46.70	64.03	0	32.74	24.32	79.15	88.01	61.3
黑暗-FCN	<b>46.24</b>	<b>71.25</b>	0	<b>29.74</b>	<b>28.17</b>	<b>80.15</b>	<b>91.45</b>	<b>63.6</b>
黑暗-FCN-448	52.46	72.11	0	35.56	24.61	82.51	92.69	65.1

定性结果。我们展示了图中每个数据集的 CNN-FCN 体系结构的一些定性标签。9. 可以看出，结果一般是相当准确的。对于 YCOR 来说，大部分的混淆来自平滑和粗糙的轨迹，这是人类很难做出一致的区分。

## 4.2 现场实验

我们在 2017 年 3 月和 7 月在宾夕法尼亚州匹兹堡附近的测试地点周围的不同地点进行了各种自动驾驶实验。尽管很简单

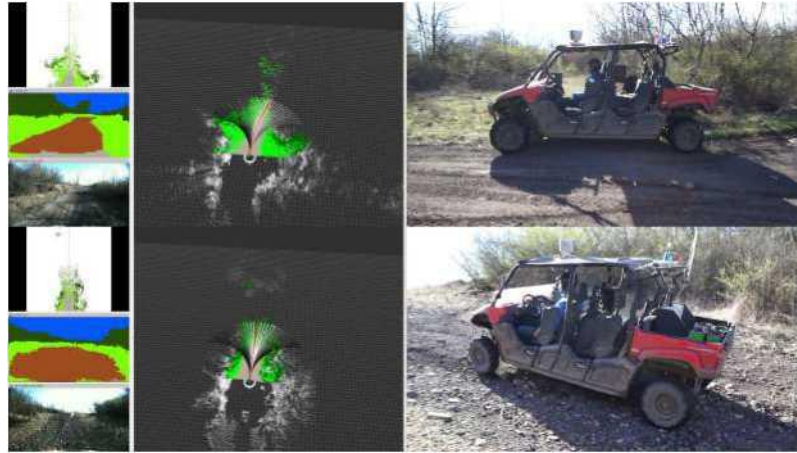


无花果。9YCOR 数据集（前四行）和深度场景（下四行）中 CNN-FCN 预测的蒙太奇）。在每种情况下，我们显示三个图像：输入、地面真相标签和预测标签。

我们的规划师，这辆车成功地穿越了各种小径，这对以前的 LiDAR 系统来说太具挑战性了。其中包括有水坑的地方，小径中间的草地，以及狭窄的小径。我们将在 <http://dimatura.net/offroad> 上传视频到项目网站 无花果。10 显示车辆处于自动运行状态。

另一方面，我们观察到了我们当前系统的一些局限性。许多限制是由于后退的地平线规划器的简单性，它经常从侧面转向更宽的小径。

一些失败也是由于我们的语义分类系统。例如，它有时无法检测到小径旁边的稀疏草，从而导致



无花果。10 自动越野驾驶在我们的测试现场的动作镜头。左侧显示从车辆上看到的传感器数据和地图的截图。视频将在项目页面上提供。



无花果。在我们的测试现场 11 自主越野驾驶的空中动作镜头。左侧显示车辆，右侧显示车辆所见语义图。视频将在项目页面上提供。

车辆偏离轨道。有一次，它还把一个大的不可移动的灌木丛和可移动的草混淆起来，迫使我们手动干预。

为了解决这些问题，我们将使用更谨慎和安全的计划器、更多的培训数据和更好的网络架构。

时机。我们使用带有 6GB GT980M GPU 的 i7 笔记本电脑在车辆上运行所有计算。该系统的瓶颈在于语义映射，语义分割每幅图像约 35ms，标签投影每幅图像约 60ms。这些步骤依次发生，导致系统大约 10Hz 的速率操作。这对于中速操作来说是足够的，我们预计即使不改变我们的体系结构，GPU 的进步将导致在不久的将来至少 2 倍的更快的操作。

## 5 结论

我们介绍了一种高效、健壮的用于越野导航的语义映射系统，该系统具有最先进的 CNN 分类器。为了训练 CNN，我们收集并标记了一个新的越野图像数据集。我们在离线基准中评估了它，其结果与较低延迟的最新技术状态相当。我们还演示了在具有挑战性的越野地形中的闭环操作。

在未来的工作中，我们感兴趣的是结合语义分割方面的最新进展，如脱毛层[27]和多种输入模式，如[23]。

在第一手验证了准确标记大量数据的困难之后，在未来的工作中，我们感兴趣的是手动标记的替代方法，例如自我监督和反向强化学习。

感谢我们感谢雅马哈汽车公司的支持，雅马哈-CMU 团队建设和维护自主 ATV，以及 Mesh 机器人提供现场测试镜头。

## 参考资料

1. Badrinarayanan, V, Handa, A., Cipolla, R.: SegNet: 一个用于健壮的语义像素-wise 标签的深度卷积编码器解码器架构 (2015 年)。网址 <http://arxiv.org/abs/1505.07293>
2. Bajracharya, M., Howard, A., Matthies, L.H., Tang, B., Turmon, M.: 自主越野导航与端到端学习的拉格朗日程序。田间机器人学杂志 26 (1), 3-25(2009)
3. Bergstra, J., Breuleux, O., Bastien, F., Lamblin, P., Pascanu, R., Desjardins, G., Turian, J., Warde-Farley, D., Bengio, Y.: Theano: CPU 和 GPU 数学表达式编译器。In: Sci Py (2010 年)
4. Chatfield, K., Simonyan, K., Vedaldi, A., Zisserman, A.: 在细节上返回魔鬼: 深入到卷积网中。共同 RRabs/1405.3531 (2014 年)。网址 <http://arxiv.org/abs/1405.3531>
5. Cordts, M., Omran, M., Ramos, S., Rehfeld, T., Enzweiler, M., Benenson, R., Franke, U., Roth, S., Schiele, B.: 用于理解城市场景语义的城市场景数据集。In: 检察官。IEEE 计算机视觉和模式识别会议 (CVPR) (2016)
6. Dieleman, S., Schluter, J., Raffel, C., Olson, E., Sonderby, S.K., Nouri, D., et al. 激光: 第一次释放。(2015)。DOI10.5281/zenodo.27878。网址 <http://dx.doi.org/10.5281/zenodo.27878>
7. 藤森, T, 卡纳德, T: 一种基于知识的户外自然彩色道路场景解释方法。In: 视觉与导航: 卡内基梅隆纳瓦拉, pp. 39-81 (1990)
8. 哈塞尔, R, Bagnell, J. A., Huber, D., Hebert, M.: 用于精确粗略地形估计的非静止空间携带核。ijrr29 (8), 981-996(2010)
9. 哈塞尔, R, Erkan, A, Sermanet, P, Scoffier, M, Muller, U: 深度信念网络学习在一个远程视觉系统的自主越野驾驶。2008IEEE/RSJ 智能机器人和系统国际会议 1 (1)、628-633(2008)
10. Jackel, L.D., Krotkov, E., Perschbacher, M., Pippine, J., Sullivan, C.: DARPA LAGR 计划: 目标、挑战、方法和第一阶段结果。野外机器人学杂志 23 (1112), 945-973(2006)
- 14 丹尼尔·马图拉娜, 周波伟, 青山正树和塞巴斯蒂安·谢尔
11. Kelly, A., Stentz, A.T., Amidi, O., Bode, M.W., Bradley, D., Diaz-Calderon, A.,

- Hapold, M., Herman, H., Mandelbaum, R., Pilarski, T., Rander, P., Thayer, S., Vallidis, N.M., Warner, r. \*在具有挑战性的环境中运行的可靠的越野自主车辆。 国际机器人研究杂志 25 (5-6), 449-483(2006)
12. Kim, D., Oh, S.M., Rehg, J.M.: UGV 导航的遍历性分类: 补丁和超像素表示的比较。 IEEE 智能机器人和系统国际会议 pp. 3166-3173 (2007)
  13. Krahenbuhl, P., Koltun, V.: 具有高斯边缘电位的全连接 CRFS 的有效推理。 共同 RRabs/1210.5644 (2012 年)
  14. 龙, J, Shelhamer, E, Darrell, T: 用于语义分割的完全卷积网络。 In: IEEE 计算机视觉和模式识别会议(CVPR) (2015)
  15. Manduchi, R., Castano, A., Talukder, A., Matthies, L.: 用于自主越野导航的障碍物检测和地形分类。 自主机器人 18 (1), 81-102(2005)
  16. Moravec, H., Elfes, A.: 从广角声纳获得的高分辨率地图。 In: 机器人和自动化国际会议 (1985 年)
  17. Noh, H., Hong, S., Han, B.: 学习反褶积网络进行语义分割。 共同 RRabs/1505.04366 (2015)
  18. Procopio, M.J., Kegelmeyer, W.P., Grudic, G.: 用于自主机器人导航的专家在线混合物的地形分割。 图像(罗切斯特, N. y.) pp. 385-397 (2009)
  19. Redmon, J.: Darknet: <http://pjreddie.com/darknet/>中的开源神经网络 (2013-2016)
  20. Simonyan, K., Zisserman, A.: 用于大规模图像识别的非常深的卷积网络。 共同 RRabs/1409.1556 (2014 年)
  21. Suleymanov, T., Paz, L.M., Pinis, P., Hester, G., Newman, P.: 较少采取的路径: 用于闭环控制的场景分割的快速变分方法。 In: 2016 年 IEEE/RSJ 智能机器人和系统国际会议 (IROS), pp. 3620-3626 (2016)
  22. Thrun, S., Montemerlo, M., Dahlkamp, H., Stavens, D., Aron, A., Diebel, J., Fong, P., Gale, J., Halpenny, M., Hoffmann, G., Lau, K., Oakley, C., Palatucci, M., Pratt, V., Stang, P., Strohband, s. Dupont, C., Jendrossek, L.E., Koelen, C., Markey, C., Rummel, C., van Niekerc, J., Jensen, E., Alessandrini, P., Bradski, G., Davies, B., Ettinger, S., Kaehler, A., Nefian, A., Mahoney, P.: 赢得 DARPA 的巨大挑战。 野外机器人学杂志 (2006 年)
  23. Valada, A., Olivera, G.L., Brox, T., Burgard, W.: 森林环境的深度多光谱语义场景理解。 In: 2016 年实验机器人国际研讨会论文集 (ISER) (2016)
  24. Vernaza, P., Taskar, B., Lee, D.D.: 在线, 通过判别训练的亚模马尔可夫随机场自监督地形分类。 In: 2008 年 IEEE 机器人和自动化国际会议, pp. 2750-2757. IEEE (2008)
  25. Wellington, C., Courville, A., Stentz, A.: 植被中自主导航的地形生成模型。 国际机器人研究杂志 25 (12), 1287-1304(2006)
  26. Wolf, D.F., Sukhatme, G.S.: 使用移动机器人进行语义映射。 IEEE 机器人交易 24 (2), 245-258(2008)
  27. Yu, F., Koltun, V.: 通过扩张卷积进行多尺度上下文聚合。 共同 RRabs/1511.07122 (2015)
  28. 周, S, Xi, J, McDaniel, M.W., Nishihata, T., Salesses, P, Iagnemma, K.: 自我监督学习, 以视觉检测在森林地形中操作的自主机器人的地形表面。 田间机器人学杂志 29 (2), 277-297(2012)