基于深度学习的无人驾驶车辆环境感知与决策系统研究

# 第一章 引言

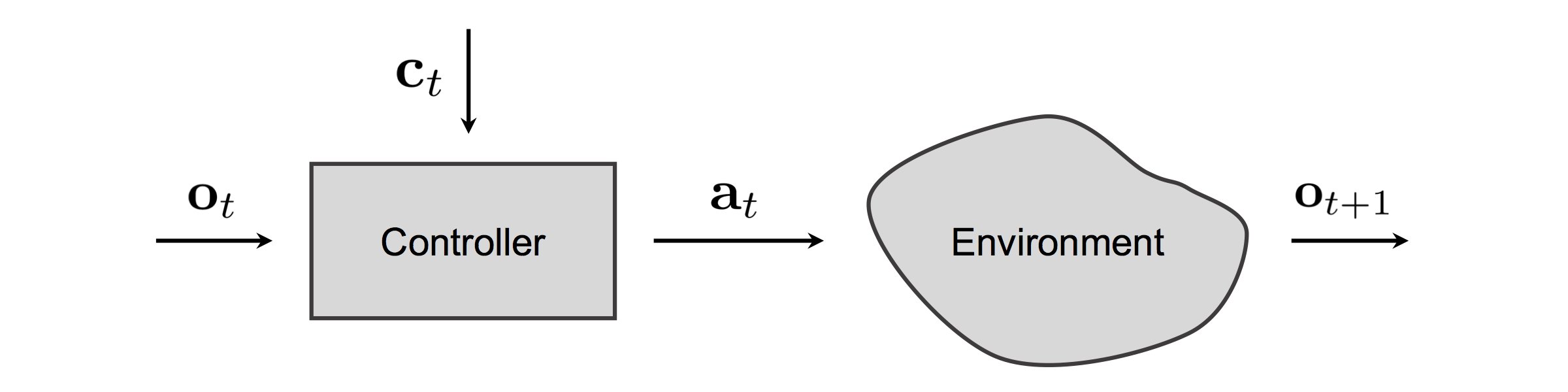
无人驾驶车辆作为人工智能与机器人技术的交叉前沿，其核心挑战在于实现高可靠性的环境感知与实时决策系统，以应对复杂动态交通场景。这一挑战源于开放道路环境的极端复杂性，包括动态变化的交通参与者、多样化的道路拓扑结构、以及不断波动的光照与天气条件。传统上，解决此问题依赖于模块化系统架构，即通过独立的感知、规划与控制模块协同工作。然而，随着深度学习技术的兴起，一种数据驱动的端到端学习范式展现出巨大潜力，它能够直接从原始传感器数据中学习感知与控制的映射关系，从而简化系统设计并提升泛化能力。本研究旨在探索并改进基于深度学习的端到端驾驶方法，特别是通过引入条件模仿学习框架来解决其在复杂场景中面临的核心难题。

传统模块化方法依赖于手工特征提取与规则系统，存在泛化能力有限和系统集成复杂度高的问题，难以适应开放道路的多样性。在模块化范式中，系统通常被分解为感知、环境建模、路径规划与控制等多个独立组件。感知模块负责从传感器数据（如摄像头、激光雷达）中提取车道线、障碍物等特征；规划模块则基于这些特征和预设规则（如交通法规）生成安全轨迹；控制模块最终执行轨迹跟踪。尽管这种方法在结构化环境中取得了成功，但其局限性日益凸显。首先，手工设计的特征提取器（如边缘检测、Hough变换）对光照变化、天气条件和道路类型（如无车道线的乡村道路）的鲁棒性不足。其次，基于规则的决策系统难以穷举所有可能的交通场景，导致在面对罕见或“长尾”事件时失效。此外，多个模块的级联设计使得误差易于累积，且系统集成与调试过程复杂而昂贵。研究表明，这种分解式优化往往导致次优的整体性能，因为各模块优化的中间目标（如检测精度）并不直接对应于最终的驾驶质量指标 `[4]`。因此，寻求一种能够联合优化所有处理步骤的替代方案成为研究的重要方向。

深度学习技术通过端到端学习范式，能够直接从原始传感器数据中学习感知与控制的映射关系，为无人驾驶提供了数据驱动的解决方案。该范式的核心思想是使用一个深度神经网络（通常是卷积神经网络，CNN）将原始输入（如单目摄像头图像）直接映射到控制输出（如转向角和加速度）。这种方法避免了手工特征工程和显式的中间表示，允许网络从数据中自动学习对任务最有用的特征表示。早期工作如Pomerleau的ALVINN系统已经证明了神经网络可以从摄像头输入直接映射到转向控制 `[27]`。近年来，随着计算能力的提升和大规模数据集的可用性，端到端方法取得了显著进展。例如，Bojarski等人 `[4]` 构建了一个包含约2700万连接的CNN，仅使用人类驾驶员的转向角作为训练信号，成功实现了在多样化道路条件下的车道跟随。该网络通过数据增强技术（如模拟车辆偏移）学习从扰动中恢复的能力，并在真实道路测试中实现了高比例的自主驾驶时间。这种端到端优化使得系统内部组件能够自我优化以最大化整体性能，而非局限于人类选择的、易于解释的中间标准。

模仿学习作为深度学习的重要分支，通过模仿人类驾驶员的示范数据，可以有效训练车辆执行车道跟随和避障等基础任务。在模仿学习框架中，智能体（即驾驶策略）的目标是学习一个从环境观测到动作的映射函数 ，使其输出尽可能接近专家（人类驾驶员）在相同观测下执行的动作 。其监督学习目标通常表示为最小化预测动作与专家动作之间的损失函数： `[3]`。这种方法的关键优势在于能够利用大量易于收集的人类驾驶数据。为了提高学习策略的鲁棒性，避免其仅在专家数据分布上表现良好而在遇到分布外状态时失效，研究者开发了诸如DAgger `[29]` 的算法，通过迭代地让策略与环境交互并由专家重新标注数据来收集恢复轨迹。此外，Bojarski等人 `[4]` 采用的三摄像头设置通过模拟车辆偏移和旋转，生成了丰富的纠正数据，有效教导了网络如何从偏离车道中心的状态中恢复航向。这些技术使得模仿学习在车道保持、简单道路跟随等反应式任务上取得了显著成功。

然而，标准模仿学习在复杂城市驾驶中面临感知-动作映射的歧义性问题，例如在交叉路口无法根据视觉输入唯一确定转向决策。这一根本性挑战源于一个隐含假设：专家的动作可以完全由观测解释，即存在一个函数  使得  `[3]`。在城市驾驶场景中，尤其是在交叉路口，这一假设不再成立。当车辆接近一个十字路口时，仅凭当前的视觉观测（图像）无法唯一确定驾驶员是打算直行、左转还是右转。驾驶员的决策还依赖于其内部状态，如预定的目的地或导航指令。因此，从观测到动作的映射不再是一个函数，而是多值的。Pomerleau早在ALVINN的研究中就观察到了这一问题，指出网络在遇到岔路时可能输出两个差异巨大的行驶方向，导致振荡 `[27]`。Bojarski等人的系统也存在类似局限，当需要变道或转弯时，必须由人类驾驶员接管控制 `[4]`。这种歧义性不仅使得函数逼近器难以拟合，更重要的是，即使网络通过某种方式（如取平均）解决了歧义并做出了决策，该决策也可能不符合乘客的意图，因为乘客缺乏控制网络行为的通信渠道。这严重限制了标准模仿学习在需要高层导航的复杂城市驾驶中的实用性。



*Figure (来源: 1710.02410v2.pdf, 第 3 页)*

本研究旨在构建一个基于条件模仿学习的无人驾驶系统，通过引入高层命令输入来解析歧义，实现可控制的端到端驾驶策略。为了克服上述歧义性问题，本研究采用条件模仿学习框架。该框架的核心思想是在训练时，除了向网络提供观测  和专家动作  外，还额外提供一个代表专家意图的高层命令 ，例如“在下一个路口左转”、“直行”或“右转”。这样，学习目标转变为最小化条件损失： `[3]`。命令  作为额外信息，揭示了专家决策所依赖的潜在状态，从而将多值映射分解为一系列条件函数，使学习变得可行。在测试时，该命令可以由乘客或一个拓扑规划器提供，从而实现对车辆导航行为的主动控制。如图2所示，控制器接收环境观测和高层命令，并产生相应的动作，形成一个可交互的驾驶系统 `[3]`。这种方法将网络从高层路径规划的任务中解放出来，使其能够专注于低层的传感器运动协调（如保持车道、避障），同时响应导航指令，类似于听从乘客指示的司机。本研究将在模拟环境和真实物理系统上验证该框架的有效性，并探究不同的网络架构（如分支设计）对性能的影响，旨在为无人驾驶提供一种兼具高性能与可控性的端到端解决方案。

# 第二章 相关技术与研究综述

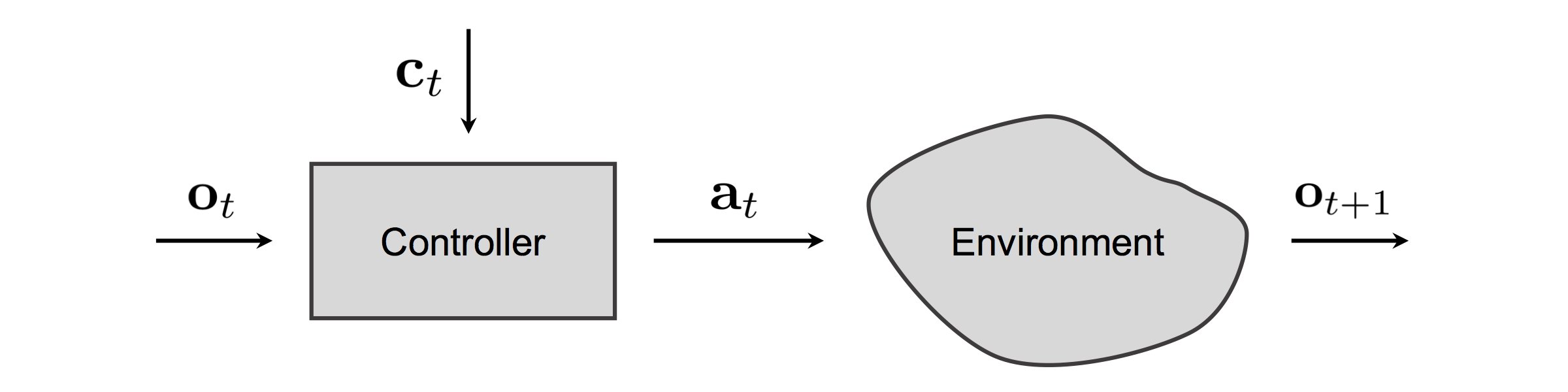
## 早期模仿学习与端到端驾驶的奠基性工作

模仿学习在机器人控制领域的应用由来已久，其核心思想是通过学习专家（如人类驾驶员）的示范数据，构建一个从环境观测到控制动作的映射策略。在无人驾驶领域，这一范式的早期代表性工作是 Pomerleau 提出的 ALVINN 系统 `[27]`。该系统使用一个全连接神经网络，直接从单目摄像头输入的图像像素映射到车辆的转向控制指令，首次在真实道路上验证了端到端学习范式在车辆控制中的可行性。然而，受限于当时的计算能力与数据规模，ALVINN 的网络容量有限，其泛化能力和鲁棒性难以应对复杂多变的真实道路环境，例如在道路分叉口容易出现决策振荡 `[27]`。尽管如此，ALVINN 为后续基于深度学习的端到端驾驶研究奠定了重要的思想基础。

随着深度学习，特别是卷积神经网络在计算机视觉领域的突破，以及大规模数据采集与 GPU 并行计算能力的提升，模仿学习在自动驾驶中的应用迎来了新的发展。Bojarski 等人的工作 `[4]` 标志着这一方向的重要进展。他们构建了一个包含约 27 万个参数的卷积神经网络，直接从原始的前视摄像头图像（YUV 三通道，分辨率 ）预测转向指令（表示为曲率 ）。该网络架构包含 5 个卷积层和 3 个全连接层，通过最小化网络预测转向角与人类驾驶员转向角之间的均方误差进行端到端训练 `[4]`。其关键创新在于通过数据增强技术，特别是利用三摄像头设置模拟车辆偏移与旋转，生成了大量“恢复轨迹”数据。具体而言，左右摄像头的图像经过视角变换，模拟车辆偏离车道中心线的情况，并相应地调整其转向标签，从而教会网络如何从扰动中纠正航向 `[4]`。这一方法显著提升了模型在多样化道路条件（包括有/无车道线、高速公路、居民区乃至未铺装路面）下的车道跟随能力与鲁棒性，实现了高达 98% 的自主驾驶时间 `[4]`。然而，此类标准模仿学习系统存在根本性局限：其策略在测试时是开环且不可控的。当车辆接近交叉路口时，相同的视觉观测可能对应多种合理的转向动作（直行、左转、右转），导致感知-动作映射存在歧义 `[27]`。网络无法获知驾驶员的潜在意图，因此无法根据乘客或导航系统的指令做出相应决策，限制了其在复杂城市驾驶场景中的实际应用价值 `[27]`。

## 条件模仿学习：解析歧义与实现可控驾驶

为了克服标准模仿学习在歧义场景下的局限性，条件模仿学习应运而生。该框架的核心思想是在训练和测试阶段，为策略网络引入一个额外的高层命令输入，用以显式地表示专家在示范时的潜在意图或任务目标 `[27]`。在无人驾驶场景中，这些命令通常是离散的导航指令，例如“在下一个路口左转”、“直行”或“右转”，并通过独热编码进行表示 `[27]`。训练数据因此扩展为三元组 ，其中  为观测（如图像）， 为命令， 为专家动作。优化目标相应地变为最小化条件损失函数： `[27]`。通过这种方式，命令输入  解析了映射  中的歧义，使网络能够学习到以命令为条件的策略。



*Figure (来源: 1710.02410v2.pdf, 第 3 页)*

在架构设计上，研究探索了两种主要范式。第一种是“命令输入”架构，将命令向量与图像特征、低维测量值（如车速）在早期进行拼接，然后通过共享的全连接层生成控制输出 `[27]`。然而，这种架构无法强制网络充分利用命令信息。第二种更为有效的“分支”架构则针对每个离散命令设计一个专用的输出分支（子策略模块）`[27]`。所有分支共享前端的卷积网络用于特征提取，而命令则作为一个开关，在推理时选择激活对应的分支进行计算： `[27]`。这种设计迫使不同分支专业化学习与特定命令相关的驾驶行为（如左转专精、直行专精），从而在测试时通过外部提供的命令实现精确可控的驾驶。实验表明，在复杂的模拟城市环境中，分支架构的条件模仿学习模型能够达到 88% 的导航任务成功率，显著优于非条件模仿学习（20%）和另一种基于目标向量输入的条件方法（24%）`[27]`。

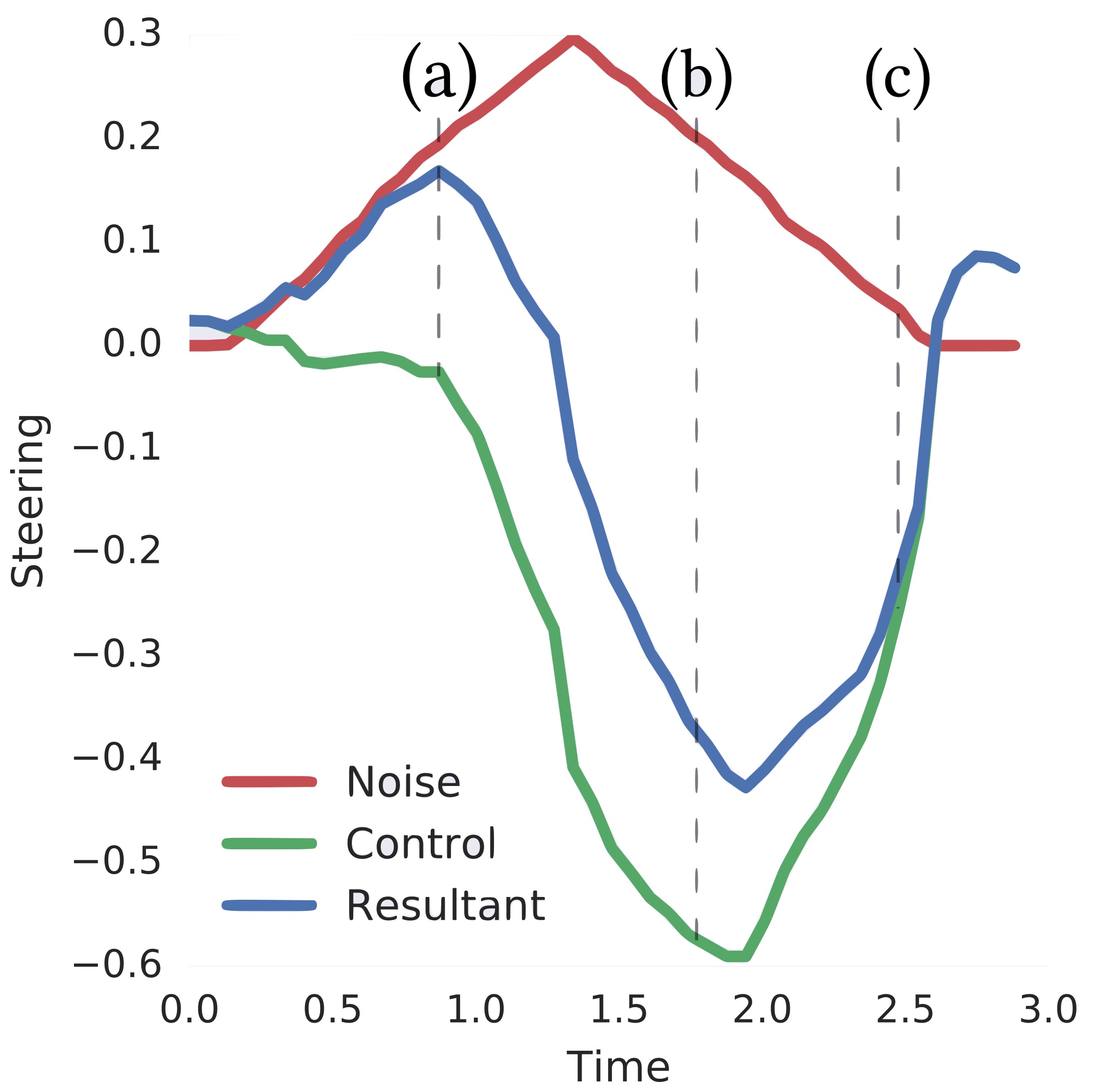
## 模块化系统与端到端系统的范式对比

纵观无人驾驶技术的发展，其系统架构主要可分为模块化与端到端两大类，二者在哲学和方法论上存在显著差异。模块化系统遵循传统机器人学中的“感知-规划-执行”范式 `[25]`。这类系统首先利用一系列精心设计的计算机视觉算法（如车道线检测、障碍物识别与跟踪、语义分割）从传感器数据中构建出对环境的结构化表示（如占据栅格图、矢量地图）`[12]`。随后，基于此环境模型，路径规划模块（如基于搜索或优化的方法）计算出安全、舒适且符合交通规则的轨迹。最后，控制模块（如 PID、模型预测控制）跟踪该轨迹，生成底层的油门、刹车和转向指令 `[25]`。模块化方法的优势在于其可解释性和可调试性，每个子模块的功能相对独立且明确。然而，其缺点也显而易见：系统性能严重依赖于各子模块的精度，错误会在流水线中累积；手工设计的特征和规则难以覆盖开放道路中无限的“长尾”场景；且整个系统集成复杂度高，开发和维护成本巨大 `[25]`。

与之相对，端到端学习系统旨在用一个统一的深度神经网络，直接从高维原始传感器数据（如图像）映射到低维控制指令 `[4, 27]`。这种数据驱动的方法避免了手工设计中间表示和规则，通过大量数据让网络自动学习完成任务所需的一切特征和决策逻辑。例如，Bojarski 等人的工作表明，CNN 在仅以转向角为监督信号的情况下，能够自主学习到对道路轮廓有用的特征表示 `[4]`。端到端范式有望通过联合优化所有处理步骤，最终实现超越模块化系统的整体性能，并减少系统复杂性 `[4]`。然而，纯粹的端到端系统长期面临“黑箱”质疑，其决策过程缺乏可解释性，且在遇到训练数据分布外的场景时，其行为难以预测和保证安全。条件模仿学习在某种程度上是对这两种范式的折中：它保留了端到端从像素到控制的学习能力，同时通过引入高层命令这一抽象接口，实现了某种程度的分层和可控性，使系统能够像听从指令的司机一样工作 `[27]`。

## 数据收集与增强技术的关键作用

模仿学习策略的性能高度依赖于训练数据的质量、多样性和规模。一个关键挑战是“分布偏移”问题：如果策略仅在与专家轨迹相同的状态分布上训练，那么当其在执行过程中因微小误差而偏离该分布时，性能会急剧下降，因为网络从未学习过如何从这些“非理想”状态中恢复 `[23, 29]`。因此，主动构建包含恢复行为的数据集至关重要。



*Figure (来源: 1710.02410v2.pdf, 第 5 页)*

DAgger 算法是一种经典的解决方案，它通过迭代地让当前策略与环境交互，并将产生的状态交由专家重新标注动作，从而将策略访问到的状态分布逐步纳入训练集 `[29]`。在自动驾驶的视觉模仿学习中，两种更直接的数据增强技术被证明极为有效。其一是源自 Bojarski 等人的三摄像头设置 `[4]`。通过同时录制左、中、右三个摄像头的图像，并对左右图像进行视角变换，可以低成本地生成大量模拟车辆横向偏移的图像，并通过几何关系计算出将其纠正回车道中心所需的转向标签。这实质上是在数据层面注入了恢复演示。



*Figure (来源: 1710.02410v2.pdf, 第 6 页)*

另一种更主动的方法是噪声注入 `[27]`。在数据收集过程中，向人类驾驶员发送的控制信号中注入时序相关的噪声（如三角波），模拟车辆逐渐漂移的过程。驾驶员会本能地进行纠正以维持正常行驶。记录时，仅将驾驶员真实的纠正响应（而非含噪声的原始信号）作为训练标签。如图4所示，这种方法提供了从显著扰动中恢复的明确示范，极大地增强了学习策略的稳定性。实验表明，即使在总时长2小时的训练数据中仅包含约12分钟的噪声注入数据，也能将模拟环境中的任务成功率从56%大幅提升至88% `[27]`。此外，在线数据增强，如随机调整对比度、亮度、添加高斯模糊或噪声、区域丢弃等，对于提升模型对不同光照、天气条件的泛化能力也至关重要，缺乏数据增强的模型在未见过的测试环境中性能会急剧恶化至接近零 `[27]`。

# 第三章 基于深度学习的无人驾驶环境感知系统设计

环境感知是无人驾驶系统实现安全、可靠导航的先决条件，其核心任务是从原始传感器数据中提取并理解车辆周围环境的语义信息。传统方法依赖于手工设计的特征提取器与独立的感知模块，如车道线检测、障碍物识别等，这不仅引入了复杂的工程集成问题，也限制了系统在多样化、非结构化场景下的泛化能力[4]。近年来，基于深度学习的端到端感知范式通过卷积神经网络直接从高维视觉输入中学习到控制输出的映射关系，避免了中间表示的手工设计，展现出强大的潜力[4, 27]。本章将详细阐述一个以摄像头为主要传感器，基于深度卷积神经网络构建的无人驾驶环境感知系统设计，涵盖传感器配置、网络架构、数据增强策略、状态表示以及实时性保障等关键方面。

## 3.1 视觉传感器配置与数据预处理

本系统采用摄像头作为核心的环境感知传感器，其优势在于能够提供丰富的纹理、颜色和上下文信息，且成本相对较低。系统配置了三个前向摄像头，分别位于车辆中心线以及左右两侧，形成特定的视角偏移[4]。这种三摄像头设置并非为了在运行时进行多视角融合，而是作为一种高效的数据增强手段。在训练阶段，来自左右摄像头的图像及其对应的、经过几何校正的转向标签被纳入训练集，用以模拟车辆偏离车道中心线时的视觉场景。例如，当车辆向右偏移时，左侧摄像头捕捉到的图像更接近于车辆需要向左纠正以回归车道中心时所“看到”的景象。通过这种方式，网络能够学习到从各种偏移状态中恢复至正常行驶轨迹的鲁棒策略，而无需在真实道路上反复进行危险的漂移测试[4]。在测试或实际部署时，仅需使用中央摄像头的单目视觉输入即可。



*Figure (来源: 1710.02410v2.pdf, 第 6 页)*

原始图像输入首先经过预处理。图像被裁剪以移除天空等无关区域，并将分辨率统一调整至网络输入尺寸，例如  像素[1]。随后，图像通常被转换为YUV色彩空间，并送入网络的第一层进行归一化处理。归一化层是硬编码的，其作用是将像素值缩放到一个固定的范围（如 ），这一操作在GPU上加速进行，有助于提升训练的稳定性和收敛速度[4]。

## 3.2 卷积神经网络架构设计

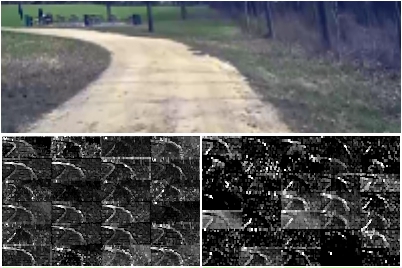
感知系统的核心是一个深度卷积神经网络，其设计目标是将原始像素映射为一个低维的环境状态表示，或直接映射为控制指令。网络架构遵循从特征提取到决策回归的经典模式，共包含9层：1个归一化层、5个卷积层和3个全连接层[4]。

\*\*输入与归一化层\*\*：网络接收尺寸为  的三通道（YUV）图像。首个归一化层对输入进行标准化，该操作不参与训练，但作为网络的一部分实现了预处理流程的集成与加速。

\*\*卷积特征提取层\*\*：五个卷积层负责从图像中自动学习层次化的特征表示。前三个卷积层使用  的卷积核和  的步长，以逐步降低特征图的空间分辨率同时增加通道数，从而捕获更大范围的上下文信息。后两个卷积层则使用  的卷积核和步长1，进行更精细的特征提取。各层输出通道数依次为24、36、48、64、64。所有卷积层之后均使用修正线性单元作为激活函数，并在部分层后应用批归一化以加速训练并缓解内部协变量偏移[1]。为了防止过拟合，在卷积层后引入了  的随机丢弃。

\*\*全连接决策层\*\*：经过卷积层提取的特征被展平后，送入三个全连接层（神经元数量例如1164、100、50），最终输出层神经元数量取决于任务设定。在纯感知系统中，输出可能是一个低维的测量向量；而在端到端控制系统中，输出直接是控制指令，如转向角  和加速度 [1]。网络通过反向传播算法，以最小化网络预测输出与人类驾驶员真实动作之间的损失函数为目标进行优化。常用的损失函数是均方误差，对于多目标输出（如转向和加速），可采用加权求和的形式：，其中  是平衡两项权重的超参数[1]。

该架构总计包含约27万个参数和2700万个连接，能够在嵌入式平台上实现实时推理[4]。研究表明，网络在训练过程中自动学会了识别对驾驶任务至关重要的特征，如车道边界和道路轮廓，而无需任何关于这些特征的显式标注[4]。



*Figure (来源: end-to-end-dl-using-px.pdf, 第 8 页)*

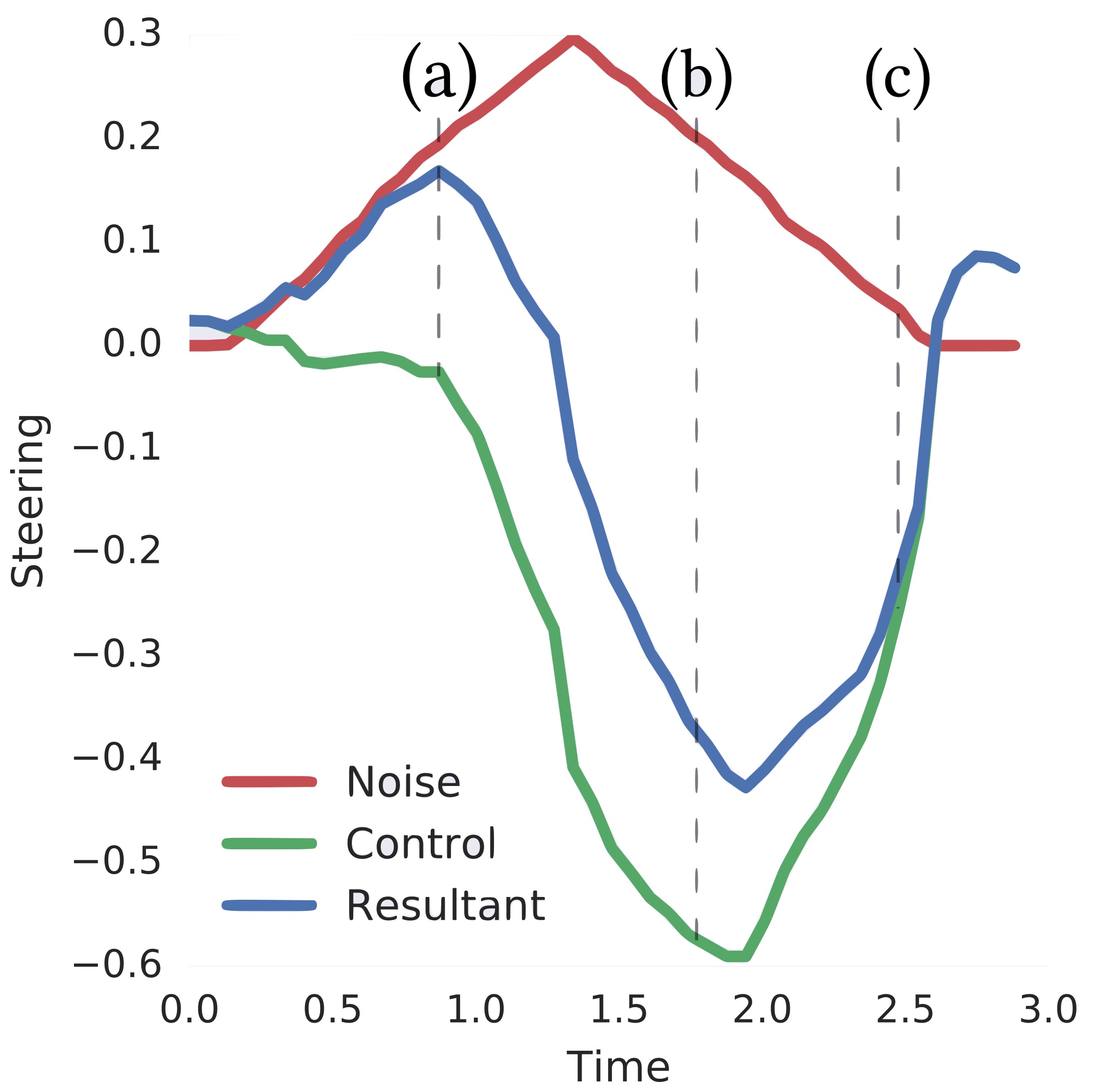
## 3.3 数据增强与鲁棒性训练策略

为了确保感知系统能够在多样化的光照、天气条件及道路环境下可靠工作，并具备从错误中恢复的能力，精心设计的数据增强与训练策略至关重要。

\*\*几何增强与三摄像头模拟\*\*：如前所述，利用左右摄像头图像进行训练是一种有效的几何增强。此外，还可以通过单应性变换在图像平面上模拟车辆更细微的平移和旋转，生成连续的“恢复”轨迹数据[4]。对于每个变换后的图像，其对应的转向标签会被重新计算，以反映将车辆在固定时间（如2秒内）纠正回期望位置和朝向所需的控制量[4]。

\*\*外观增强\*\*：在线数据增强被广泛应用于提升模型对视觉外观变化的泛化能力。增强操作包括随机调整图像的对比度、亮度、色调，以及添加高斯模糊、高斯噪声、椒盐噪声等[1]。区域丢弃是另一种有效技术，即在图像中随机掩蔽若干个约占图像面积  的矩形区域，这迫使网络不过度依赖图像的局部特定区域进行决策，从而增强其鲁棒性[1]。

\*\*噪声注入与交互式数据收集\*\*：仅使用人类驾驶员完美驾驶的数据进行训练，可能导致模型无法处理行驶中不可避免的扰动和漂移。为此，可以采用噪声注入技术。在部分数据采集过程中，向人类驾驶员的方向盘控制信号中注入一个时序相关的噪声信号（如三角波），模拟车辆逐渐偏离车道[1]。驾驶员随后做出的纠正性操作被记录为训练数据，而注入的噪声本身则被丢弃。这种方法为网络提供了从非理想状态中恢复的明确示范，显著提高了学习策略的稳定性。实验表明，即使仅占总训练数据量一小部分（如 ）的噪声注入数据，也能对最终性能带来巨大提升[1]。



*Figure (来源: 1710.02410v2.pdf, 第 5 页)*

## 3.4 环境状态表示与多模态融合

一个完整的感知系统不仅需要处理视觉信息，还需整合其他传感器数据以形成更全面的环境状态表示。在本系统设计中，卷积神经网络提取的深层视觉特征被进一步利用。

\*\*低维测量向量融合\*\*：除了图像输入 ，系统还引入一个低维测量向量  作为额外输入[1]。这个向量通常包含车辆自身的状态信息，例如当前速度 。速度信息对于决策至关重要，例如在弯道处需要根据速度调整转向幅度。在网络架构中，测量向量通过一个独立的全连接模块  进行处理，其输出与图像特征模块  的输出进行拼接，共同形成联合表征 [1]。这种多模态融合方式允许网络根据车辆动态状态调制其基于视觉的决策。

\*\*为高层决策提供特征\*\*：在更模块化的系统设计中，训练好的卷积层可以作为强大的视觉特征提取器。其输出的高级特征可以被输入到独立的决策或规划模块中，而非直接输出控制指令。这种设计在保持感知能力的同时，为引入更复杂的符号推理或长期规划提供了灵活性。

## 3.5 实时性实现与嵌入式部署

无人驾驶系统对实时性有严格要求，必须在极短的时间延迟内完成从图像采集到控制指令生成的全流程。本系统设计充分考虑了嵌入式部署的约束。

\*\*硬件平台\*\*：系统目标部署平台是诸如NVIDIA DRIVE PX或Jetson TX2等高性能车载计算单元[1, 4]。这些平台集成了强大的GPU，能够高效运行深度神经网络推理。

\*\*优化与性能\*\*：经过优化的网络模型在上述嵌入式平台上能够以每秒30帧的速度处理输入图像并生成控制信号，满足实时驾驶的需求[4]。实时性通过以下方式保障：1) 采用经过剪枝或量化的轻量级网络架构；2) 利用GPU的并行计算能力加速卷积和矩阵运算；3) 优化内存访问和数据传输流程。

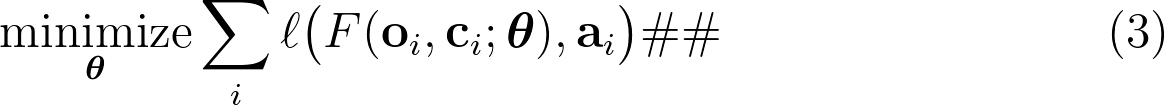
综上所述，本章所设计的基于深度学习的无人驾驶环境感知系统，通过端到端的卷积神经网络，实现了从原始视觉输入到环境理解或直接控制指令的映射。结合三摄像头设置、多层次数据增强、噪声注入训练以及多模态信息融合，该系统具备了应对复杂、动态真实世界环境的鲁棒性和泛化能力。同时，针对嵌入式平台的优化确保了其满足实时性要求，为后续章节将要讨论的、基于此感知系统的条件模仿学习决策框架奠定了坚实的基础。

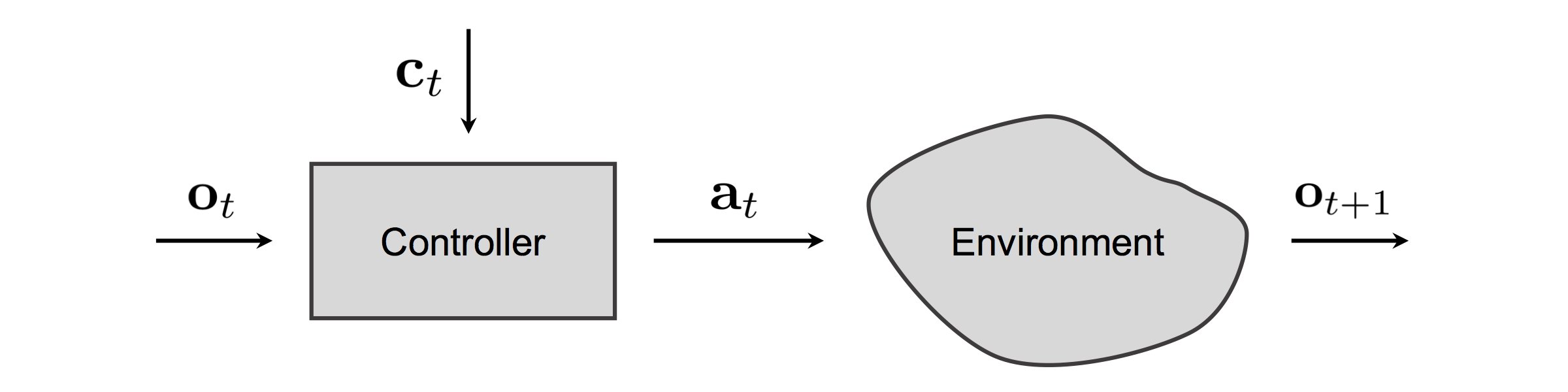
# 第四章 基于条件模仿学习的决策系统构建

为解决标准模仿学习在复杂城市驾驶场景中面临的感知-动作映射歧义性问题，本章构建了一个基于条件模仿学习的决策系统。该系统通过引入高层命令作为额外输入，解析了在交叉路口等场景中视觉输入无法唯一确定转向决策的困境，实现了可控制的端到端驾驶策略。系统采用分支网络架构，整合了感知输出与命令输入，并通过端到端训练优化整体性能，避免了模块化方法中手工设计中间表示的复杂性。

## 条件模仿学习框架

条件模仿学习框架的核心在于将高层命令作为决策过程的显式条件输入。在标准模仿学习中，控制器  的目标是学习一个从观测  到动作  的映射函数，其优化目标为最小化预测动作与专家示范之间的损失  `[1]`。然而，在城市驾驶等复杂场景中，仅凭当前观测（如单帧图像）往往无法唯一确定最优动作，例如在交叉路口，相同的视觉输入可能对应左转、直行或右转等多种合法操作。这种映射的歧义性导致标准模仿学习难以收敛或产生振荡行为 `[1]`。

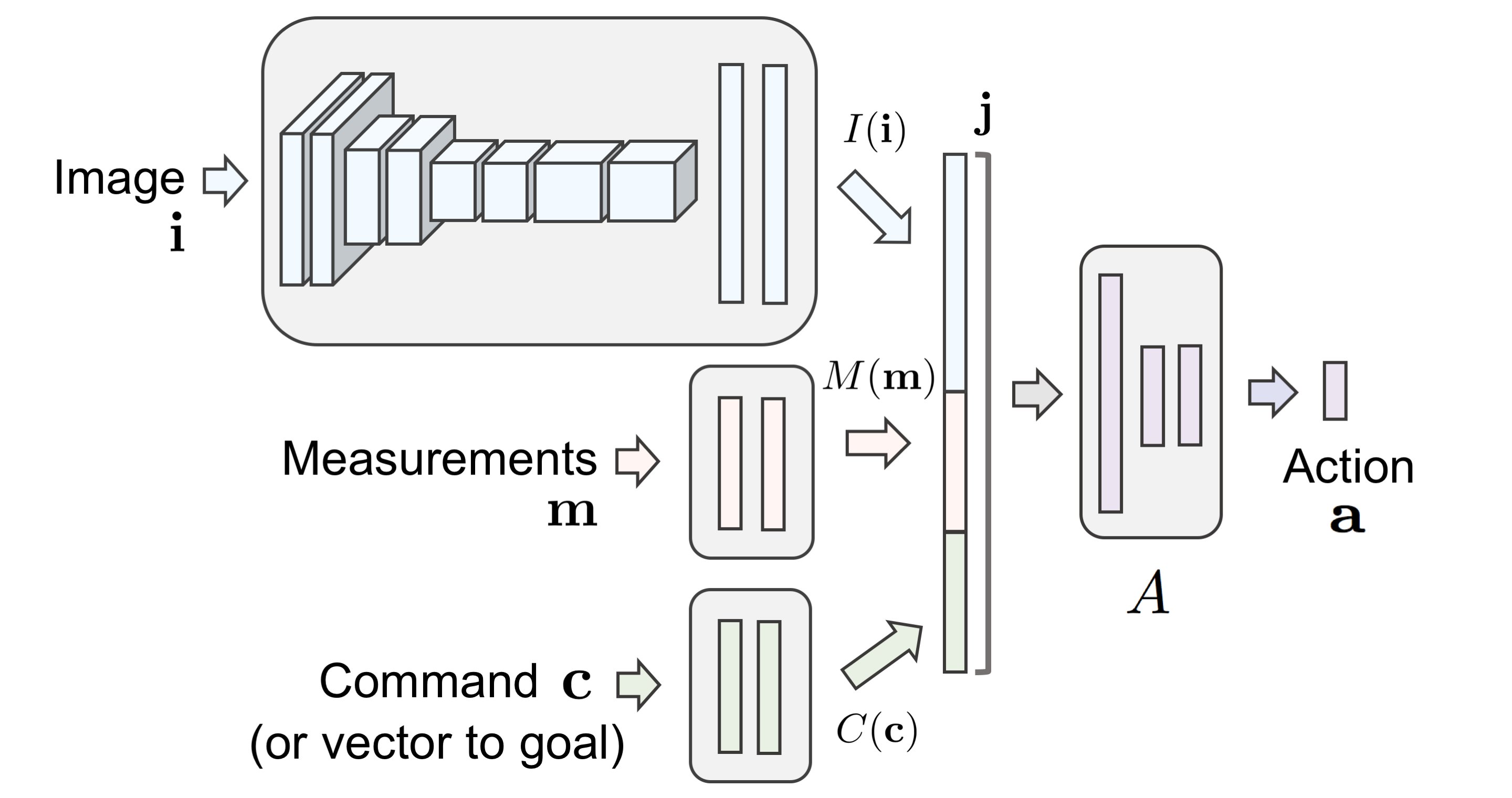
为解决此问题，条件模仿学习显式地引入一个代表专家意图或高层决策的隐状态向量 ，并假设专家动作由观测和隐状态共同决定： `[1]`。通过定义一个命令输入  来部分或完全暴露该隐状态，训练数据扩展为三元组 。相应地，控制器的优化目标转变为最小化条件损失函数：  
  
此框架在训练时利用命令解析歧义，促进学习；在测试时，命令则成为一个通信通道，允许乘客或规划模块（如拓扑导航器）通过指定“左转”、“直行”、“右转”等高层指令来实时控制车辆行为，从而将策略从无目标的模仿提升为可引导的“司机”模式 `[1]`。



*Figure (来源: 1710.02410v2.pdf, 第 3 页)*

## 分支网络架构设计

为实现命令条件化，本研究评估并采用了两种主要的网络架构：命令输入架构与分支架构 `[1]`。命令输入架构将命令  作为一个额外的输入模态，与图像  和测量值 （如车速）并行处理。图像经由卷积网络模块  处理，测量值和命令分别经由全连接网络模块  和  处理，三者的输出被拼接为联合表示 ，最终由一个控制模块  输出动作 `[1]`。尽管该架构通用性强，但网络可能无法充分利用命令信息。



*Figure (来源: 1710.02410v2.pdf, 第 4 页)*

因此，本研究主要采用了性能更优的分支架构。如 `[1]` 中图3(b)所示，该架构针对离散命令集合 （例如 `[1]` 中定义的“左转”、“直行”、“右转”及“继续行驶”）进行设计。其核心思想是移除独立的命令处理模块，转而为每个命令  配备一个专用的分支子网络 。所有分支共享同一个由图像和测量值提取的感知流 。在任一时刻，输入的命令作为一个开关，选择激活对应的分支来生成最终动作：  
  
这种设计强制每个分支学习与特定命令相关的子策略。例如，一个分支可能专门学习车道保持，另一个分支专门学习右转操作等。共享的感知流确保了环境特征提取的一致性，而独立的分支则允许策略针对不同高层意图进行精细化调整。实验表明，分支架构在模拟环境和物理系统中的成功率均显著高于命令输入架构 `[1]`。

## 训练目标与命令表示

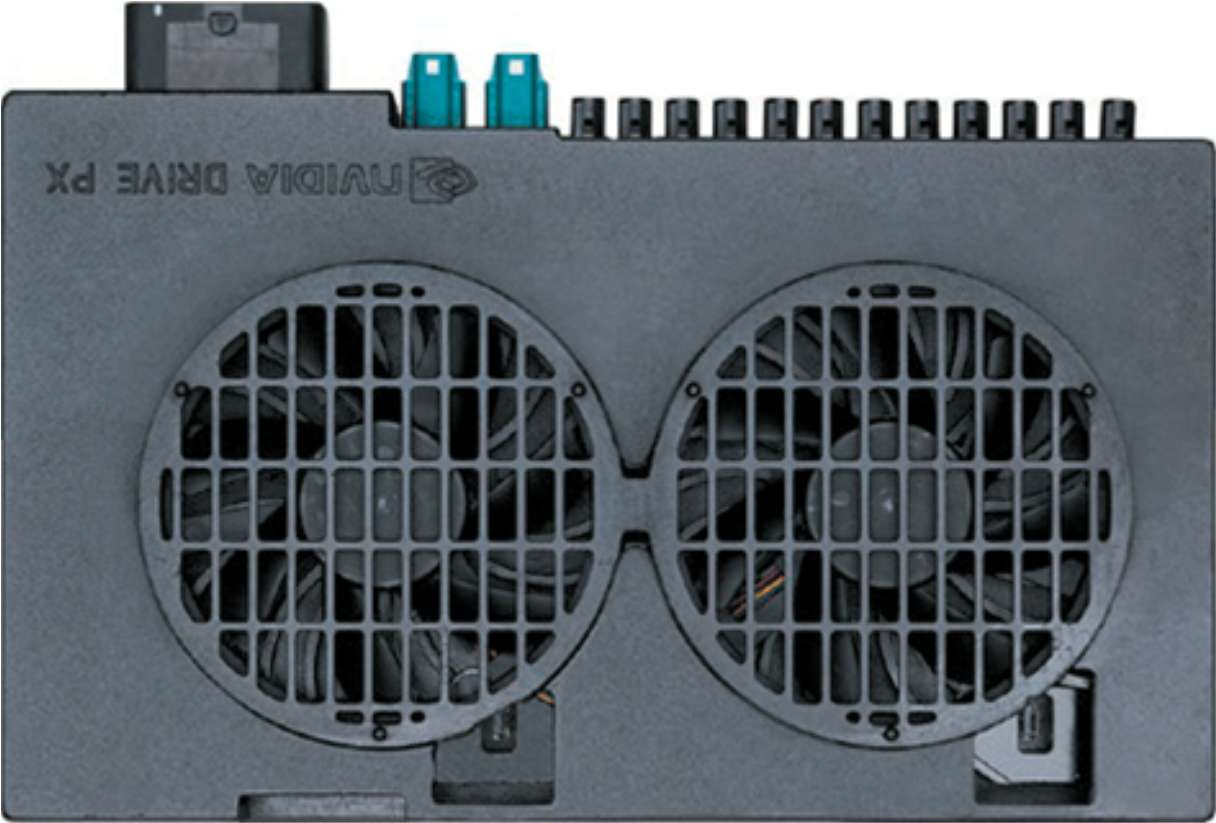
决策系统的训练目标是使网络预测的动作尽可能接近人类专家的示范动作。动作空间定义为二维连续向量 ，其中  代表转向角， 代表加速度（负值表示制动）`[1]`。损失函数采用加权均方误差，以平衡转向与加速度预测的精度：  
  
参数  用于调整两项损失的相对重要性。所有模型均使用Adam优化器进行训练，并构建包含等量各命令样本的小批量数据以平衡学习 `[1]`。

高层命令采用独热编码进行表示。在数据收集阶段，人类驾驶员在接近交叉路口时，通过方向盘按钮或遥控器指明其意图（如按下“左转”按钮），此时刻的命令与对应的观测、动作一同被记录 `[1]`。这种收集方式模拟了现实中转向灯或导航指令的发出时机，确保了数据的真实性。在测试阶段，命令可由乘客通过接口指定，或由上层拓扑规划器根据全局路径实时生成，从而实现了对驾驶策略的实时、可控引导。

## 数据收集与鲁棒性增强策略

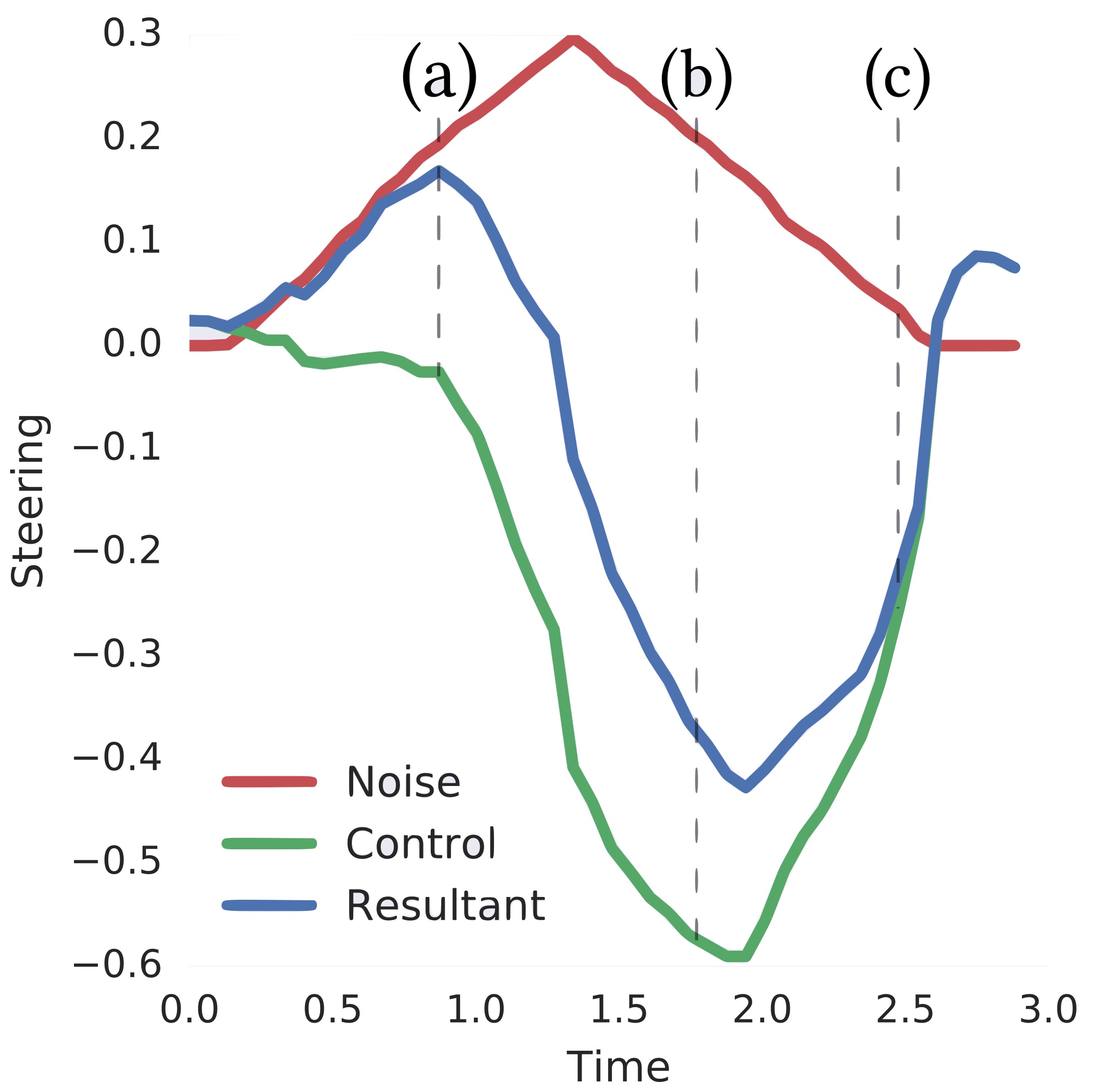
仅使用专家在理想轨迹上演示的数据进行训练，容易导致学习到的策略无法应对扰动和漂移，缺乏鲁棒性 `[1]`。为此，本研究综合采用了多种数据收集与增强技术。

首先，借鉴 `[2]` 的方法，系统采用三摄像头设置（左、中、右）。在训练时，不仅使用中央摄像头的图像和对应的原始转向指令，还利用左右摄像头的图像，并通过适当的几何变换模拟车辆偏离车道中心的情况，同时调整转向标签为能使车辆在约两秒内回归中心的指令，从而生成大量的“恢复轨迹”数据 `[1, 2]`。



*Figure (来源: end-to-end-dl-using-px.pdf, 第 3 页)*

其次，引入了主动噪声注入技术以进一步增强策略的稳定性。在部分数据采集过程中，向人类驾驶员的操作信号中注入时序相关的噪声（如三角波噪声），模拟车辆逐渐偏离预期轨迹的漂移 `[1]`。



*Figure (来源: 1710.02410v2.pdf, 第 5 页)*

如 `[1]` 中图4所示，注入的噪声（红色曲线）会使车辆产生偏移，驾驶员随后会施加纠正性控制（绿色曲线）以恢复稳定行驶。关键的是，用于训练的是驾驶员的纠正响应信号，而非包含噪声的原始控制信号。这种方法为网络提供了从意外扰动中恢复的明确示范，而不会让网络学习到导致偏离的错误行为。

此外，在线数据增强对于提升模型在不同环境条件下的泛化能力至关重要。在训练过程中，对输入的图像随机施加一系列变换，包括对比度与亮度调整、高斯模糊、高斯噪声、椒盐噪声以及区域随机丢弃等 `[1]`。这些增强手段有效模拟了变化的光照、天气及部分传感器失效等情况。

## 端到端系统集成与优势

最终构建的决策系统实现了从原始感知输入（图像、测量值）到控制指令（转向、加速）的端到端映射，并以高层命令为条件。该系统整合了前述所有组件：分支架构的网络接收来自感知系统的图像特征和车辆速度等信息，结合当前的高层命令，直接输出连续的控制动作。

这种端到端的集成方式具有显著优势。它避免了传统模块化自动驾驶系统中必需的、且通常需要大量手工调试的中间表示（如车道线方程、障碍物边界框、详细路径点等）`[1]`。所有处理步骤通过梯度下降法联合优化，以最终驾驶性能为统一目标，使得内部表征能够自我优化以适应整体任务，而非局限于人为设定的中间指标 `[2]`。这不仅有望获得更优的系统性能，也简化了系统设计流程。实验结果表明，该条件模仿学习系统在模拟和真实物理环境中均能实现高成功率的可控驾驶，验证了该框架在处理复杂城市驾驶场景中的有效性与实用性 `[1]`。

# 第五章 系统实现与实验验证

为全面评估所提出的基于条件模仿学习的无人驾驶系统的性能与泛化能力，本章节详细阐述了系统的实现细节、实验设置、评估指标以及对比分析结果。实验分别在高度逼真的模拟环境与物理机器人平台上展开，旨在从多个维度验证系统的有效性、鲁棒性与可控性。

## 系统实现平台与统一架构

本研究采用双平台验证策略，分别在CARLA模拟器[10]与一台1/5比例的机器人卡车（Traxxas Maxx）上实现了所提出的条件模仿学习系统。两个平台共享相同的核心网络架构与数据处理流程，确保了方法的一致性与可比性。在模拟环境中，系统接收来自CARLA渲染引擎的合成图像作为输入；在物理系统中，输入则来自安装在卡车上的三台低成本网络摄像头。控制输出均为二维连续动作向量 ，其中  代表归一化的转向角， 代表归一化的加速度（负值对应刹车或倒车）。这种统一架构设计使得在模拟环境中验证的设计决策能够直接迁移至物理世界，显著降低了实际部署的风险与成本。



*Figure (来源: 1710.02410v2.pdf, 第 1 页)*

模拟环境基于开源的CARLA驾驶模拟器构建，该模拟器利用虚幻引擎4提供了高度逼真的动态城市环境，包含建筑物、植被、交通标志以及车辆与行人交通[10]。实验使用了两个专业设计的城镇：Town 1用于模型训练，Town 2则专门用于测试，以评估模型在未见过的视觉环境中的泛化能力。物理系统的硬件配置如图6所示，核心计算单元为NVIDIA TX2嵌入式平台，负责实时运行训练好的神经网络模型。车辆的低层控制由Holybro Pixhawk飞行控制器执行，它将网络预测的高层控制指令转换为驱动转向舵机和速度控制器的PWM信号[10]。数据收集阶段，人类驾驶员通过遥控器操作卡车，并同时提供高层命令（左转、直行、右转），所有传感器数据（三摄像头图像、控制信号、GPS/IMU数据）均被同步记录。

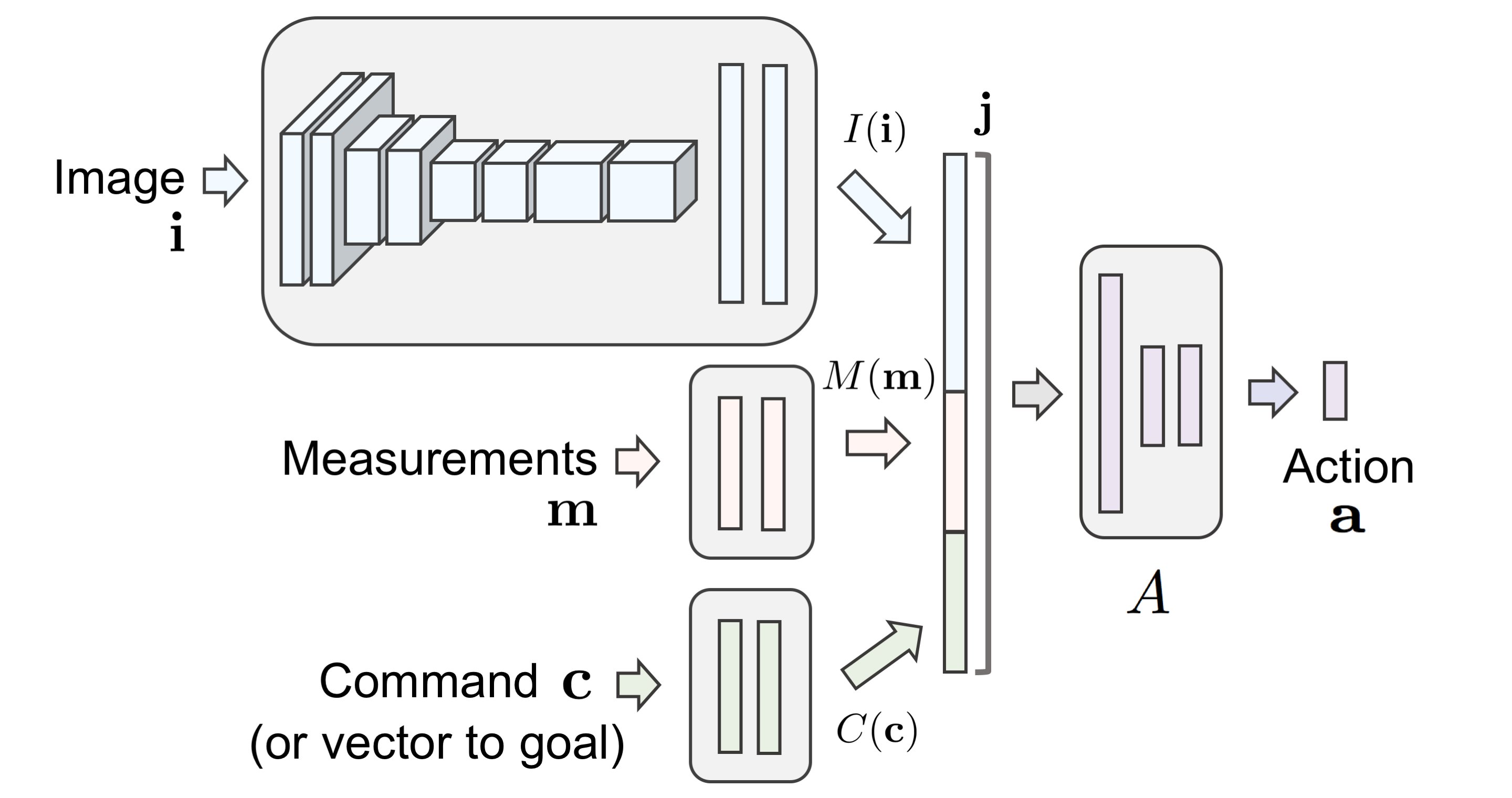
## 模拟环境实验设置与量化评估

在CARLA模拟器中的评估采用情节式（episodic）设置，以客观量化驾驶质量。在每个评估情节中，智能体被随机初始化于一个新的起始位置，并需要根据一个拓扑规划器提供的高层转向命令，驾驶至指定的目标点。一个情节被视为成功，仅当智能体在规定的时间限制内抵达目标位置。除了成功率这一核心指标，研究还记录了“每公里违规次数”（或等价地，“平均无违规行驶距离”）作为衡量驾驶安全性与平顺性的辅助指标。违规行为包括与车辆、行人发生碰撞，以及偏离车道（驶入对向车道或人行道）。这种评估方式能够综合反映系统在复杂动态城市环境中完成导航任务的能力，而不仅仅是基本的车道保持。

实验使用了50对相距至少1公里的起点-目标点组合，分别在训练镇（Town 1）和测试镇（Town 2）进行评估。训练数据总计约2小时的人类驾驶记录，其中仅包含约12分钟（占10%）的噪声注入数据。尽管噪声数据比例较小，但其对于提升策略的稳定性至关重要[10]。通过在这种可控且可重复的模拟环境中进行大规模测试，能够高效地比较不同算法变体，并进行细致的消融研究。

## 与基线方法的性能比较

为验证所提出的分支架构条件模仿学习方法的优越性，在模拟环境中将其与多个基线方法进行了系统比较。基线包括：1) \*\*标准（非条件）模仿学习\*\*：网络仅根据观测  和测量值 （如车速）预测动作，无法接收任何导航指令；2) \*\*目标条件模仿学习\*\*：网络额外接收一个指向目标点的向量作为输入（采用与图3(a)类似的“命令输入”架构），该向量在训练和测试的每一步都提供。此外，还比较了所提方法的两种架构变体：`branched`（分支架构，图3(b)）和 `command input`（命令输入架构，图3(a)）。

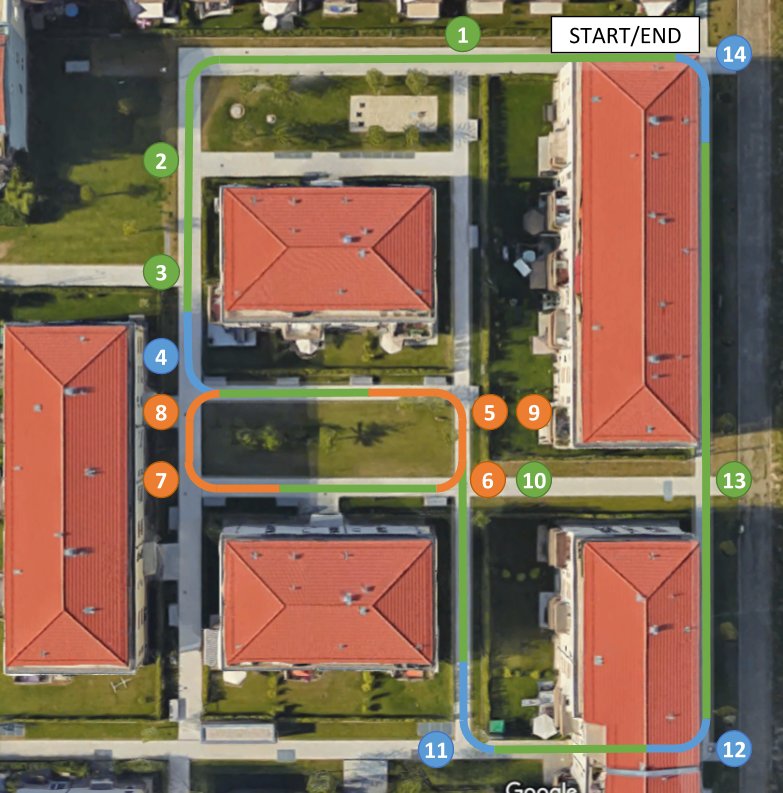


*Figure (来源: 1710.02410v2.pdf, 第 4 页)*

实验结果（表1）清晰地表明，分支架构的条件模仿学习控制器在成功率上显著优于所有基线方法。在训练镇（Town 1）达到了88%的成功率，在未见过的测试镇（Town 2）也达到了64%的成功率。相比之下，标准模仿学习模型由于完全不知道目标，成功率仅为20%-26%。更具启示意义的是，目标条件方法虽然每一步都获得了精确的目标方向向量，但其成功率（24%-30%）仅略高于非条件模型，且定性观察发现，该控制器常常试图“抄近路”直接冲向目标，导致过早偏离道路。这表明，简单的全连接网络难以自动地将一个空间目标向量转换为一系列正确的转弯决策序列，凸显了引入高层语义命令（如“在下一个路口左转”）的必要性。分支架构在成功率上也优于命令输入架构，这表明强制网络为不同命令学习专用子策略（分支）的设计，比让网络自行决定是否使用命令输入更为有效。

## 物理系统实验与真实世界验证

物理系统的实验在一个真实的住宅区进行，沿着一条包含14个交叉路口的固定路线（图8）评估系统的性能。评估指标包括：\*\*错过转弯次数\*\*（车辆未能按指令在路口转弯的比例）、\*\*人工干预次数\*\*（当车辆偏离道路超过5秒或可能发生碰撞时，人类驾驶员接管控制）以及\*\*完成路线所需时间\*\*。这些指标直接反映了系统在真实、不可控环境中的鲁棒性、对命令的响应能力以及整体驾驶效率。



*Figure (来源: 1710.02410v2.pdf, 第 7 页)*

物理实验结果（表2）与模拟实验结论一致，并进一步揭示了数据增强与噪声注入技术的关键作用。分支架构模型在三次测试运行中平均实现了0%的错过转弯率、仅0.67次干预，并以2分19秒顺利完成路线，表现最佳。消融实验表明，\*\*噪声注入\*\*对于策略稳定性至关重要：未使用噪声注入数据训练的模型，其干预次数激增至8.67次，错过转弯率达24.4%，完成时间翻倍。\*\*数据增强\*\*的影响更为显著：未使用数据增强的模型完全失效，错过了73%的转弯，引发了近40次干预，完成时间超过10分钟。这极端地证明了，在光照、天气、背景不断变化的真实世界中，通过数据增强（如随机调整对比度、添加高斯噪声、区域丢弃等）来提升模型的泛化能力是必不可少的[10]。这些结果强有力地证实，即使训练数据有限（2小时），通过精心设计的数据收集与增强策略，基于视觉的端到端驾驶系统能够在真实城市环境中实现可靠、可控的驾驶。

## 消融研究与组件重要性分析

为了深入理解系统中各个组件的贡献，在模拟和物理实验中都进行了系统的消融研究。除了上述关于噪声注入和数据增强的消融实验外，还分析了\*\*网络深度\*\*的影响。如表1所示，使用一个更浅的网络架构会导致性能急剧下降（Town 2成功率从64%降至14%），这表明处理丰富、复杂的城市视觉场景需要足够深度的网络来学习有效的感知-动作映射。综合所有消融实验结果，可以得出结论：分支架构、噪声注入的数据收集策略、广泛的数据增强以及足够深度的网络，这四个核心组件对于实现高成功率、强泛化能力和稳定驾驶行为都至关重要，缺一不可。

## 系统泛化能力展示

系统的泛化能力通过两种方式得到验证。首先，在模拟环境中，模型在视觉风格迥异的Town 2（专门用于测试，未在训练中出现）中取得了64%的成功率，这证明了其对新环境的适应性。其次，在物理系统中，除了在主要测试路线上评估，还将训练好的卡车部署到另外三个外观截然不同的新环境中进行定性测试（图9）。在这些新环境中，卡车能够持续保持车道跟随，并对高层转向命令做出正确响应。这进一步表明，系统学习到的驾驶策略并非过拟合于训练数据的特定视觉特征，而是捕捉到了更具普适性的道路结构与驾驶逻辑。



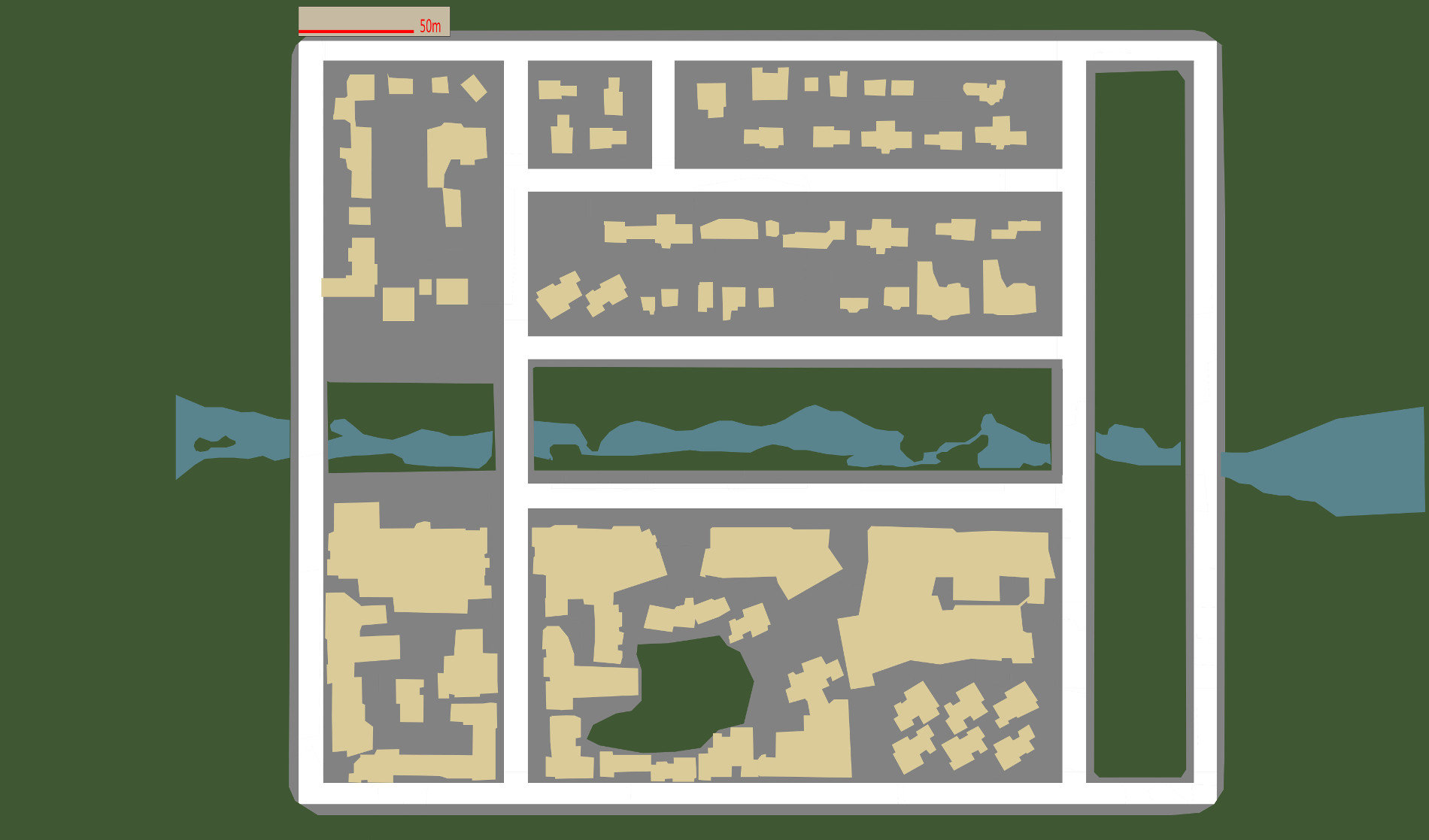
*Figure (来源: 1710.02410v2.pdf, 第 7 页)*

综上所述，本章通过在多平台、多指标下的系统实验与对比分析，全面验证了基于条件模仿学习的无人驾驶系统在解决动作歧义、实现可控驾驶方面的有效性。实验数据表明，该系统不仅在受控的模拟环境中表现优异，更能成功迁移至复杂多变的真实世界，展现出良好的鲁棒性、可控性与泛化能力，为端到端自动驾驶的实际应用提供了有力的实证依据。

# 第六章 结论与展望

本研究成功构建并验证了一个基于条件模仿学习的无人驾驶系统。该系统通过引入高层命令作为额外输入，有效解决了传统端到端模仿学习中固有的感知-动作映射歧义问题。具体而言，在交叉路口等场景中，仅凭视觉输入无法唯一确定转向决策，而高层命令（如“左转”、“直行”、“右转”）的引入为网络提供了明确的意图信息，从而将原本多对一的映射关系转化为可学习的确定性函数 `` [3]。这一设计使得训练后的策略在测试时能够响应外部指令，实现了从被动模仿到可控驾驶的转变。系统的核心架构采用了分支设计（`branched`），其中每个命令对应一个专用子模块，共享感知流但独立学习控制策略，从而强制网络学习与命令相关的子策略，如车道跟随、左转或右转 [4]。这种设计相较于将命令作为普通输入的架构（`command input`），在性能上表现出显著优势。

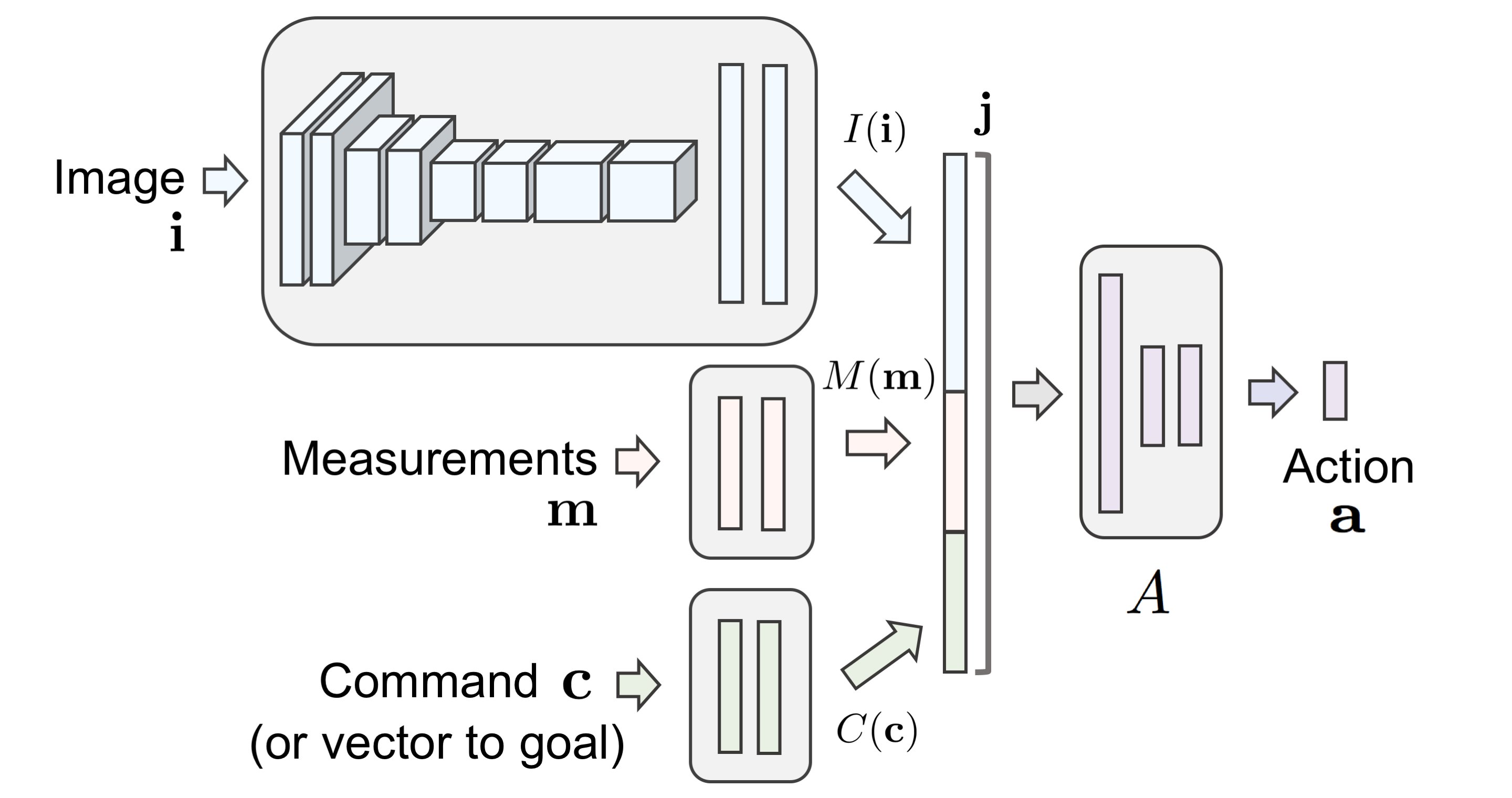
实验结果表明，该系统在模拟环境和物理环境中均实现了高成功率和可控性，充分证明了条件模仿学习在复杂城市驾驶场景中的适用性。在CARLA模拟器的评估中，分支架构的条件模仿学习控制器在训练集所在城镇（Town 1）达到了88%的成功率，在未见过的测试城镇（Town 2）也达到了64%的成功率，显著优于非条件模仿学习（20%和26%）和目标条件模仿学习（24%和30%）等基线方法 [6]。同时，系统在物理世界中的表现同样稳健。在一辆1/5比例的机器人卡车上进行的实地测试显示，完整的`branched`模型能够以零错过转弯率和平均仅0.67次人工干预完成预设路线，而移除噪声注入或数据增强的消融模型则分别导致了24.4%的错过转弯率、8.67次干预以及73%的错过转弯率、39次干预的严重性能退化 [7]。这些对比实验凸显了噪声注入与数据增强对于策略鲁棒性和泛化能力的关键作用。此外，系统还展示了在新环境中的泛化能力，能够在视觉外观迥异的不同环境中保持车道跟随并正确响应命令 [7]。



*Figure (来源: 1710.02410v2.pdf, 第 5 页)*

然而，本研究构建的系统仍存在一定的局限性，这些局限性也为未来的研究指明了方向。首先，系统性能的提升在很大程度上依赖于更大规模、更多样化的数据集以及更高容量的网络架构。当前的研究仅使用了约2小时的模拟驾驶数据和2小时的物理世界数据，其中包含噪声注入的数据仅占约10% [6]。虽然数据增强技术（如随机对比度调整、高斯噪声注入和区域丢弃）有效缓解了过拟合问题，但要支持更广泛的自动驾驶部署（如处理极端天气、复杂交通流或罕见长尾场景），仍需收集涵盖更广泛工况的海量数据，并探索更高容量的网络模型（如更深的卷积层或Transformer架构）以提升模型的表征能力。其次，当前的网络架构主要处理当前时刻的感知输入与命令，对于动态交通场景中普遍存在的长时序依赖（如预测其他交通参与者的意图、理解交通信号的时序变化）处理能力有限。

针对上述局限性，未来的研究工作可以从多个维度展开。在网络架构方面，可以探索引入更复杂的时序建模机制，例如注意力机制（Attention Mechanism）或递归神经网络（RNN）。注意力机制能够使网络动态聚焦于图像中的关键区域（如即将进入的交叉路口、动态障碍物），从而提升感知的效率和准确性。而递归神经网络或其变体（如长短时记忆网络LSTM）则能够显式地建模历史观测序列，使决策不仅依赖于当前帧，还能利用过去的上下文信息，这对于处理车辆轨迹的连续性、预测他车行为以及执行需要多步规划的动作（如变道、超车）至关重要。在交互方式上，可以扩展命令的词汇与表达形式。当前系统使用预定义的离散命令（如“左转”），未来可探索支持自然语言交互，使乘客或规划器能够使用更丰富、更灵活的指令（如“在前方便利店右转后靠边停车”）来控制车辆。这将极大地增强系统的人机通信能力，使其更贴近实际驾驶场景中的自然交互需求。实现这一目标需要结合自然语言处理（NLP）技术，将文本指令编码为网络可理解的语义向量，并与视觉特征进行融合。



*Figure (来源: 1710.02410v2.pdf, 第 4 页)*

最终，本研究为无人驾驶的环境感知与决策系统提供了一种可扩展的端到端框架。该框架避免了传统模块化方法中手工设计特征提取器、环境模型和规划器的复杂性，通过数据驱动的方式直接从原始传感器数据中学习驾驶策略。条件模仿学习的引入，不仅解决了动作歧义这一核心挑战，更重要的是为端到端系统开辟了一个可控的“通信通道”，使其能够与高层规划器或人类用户进行有效协作。这推动了数据驱动方法在自动驾驶领域的应用，为实现更智能、更灵活且易于交互的无人驾驶系统奠定了坚实的基础。未来的工作应致力于在保持端到端学习优势的同时，进一步提升系统的安全性、可解释性以及对开放世界无限长尾场景的适应能力。

# 附录：图片汇总



*Figure (来源: 1710.02410v2.pdf, 第 1 页)*



*Figure (来源: 1710.02410v2.pdf, 第 1 页)*



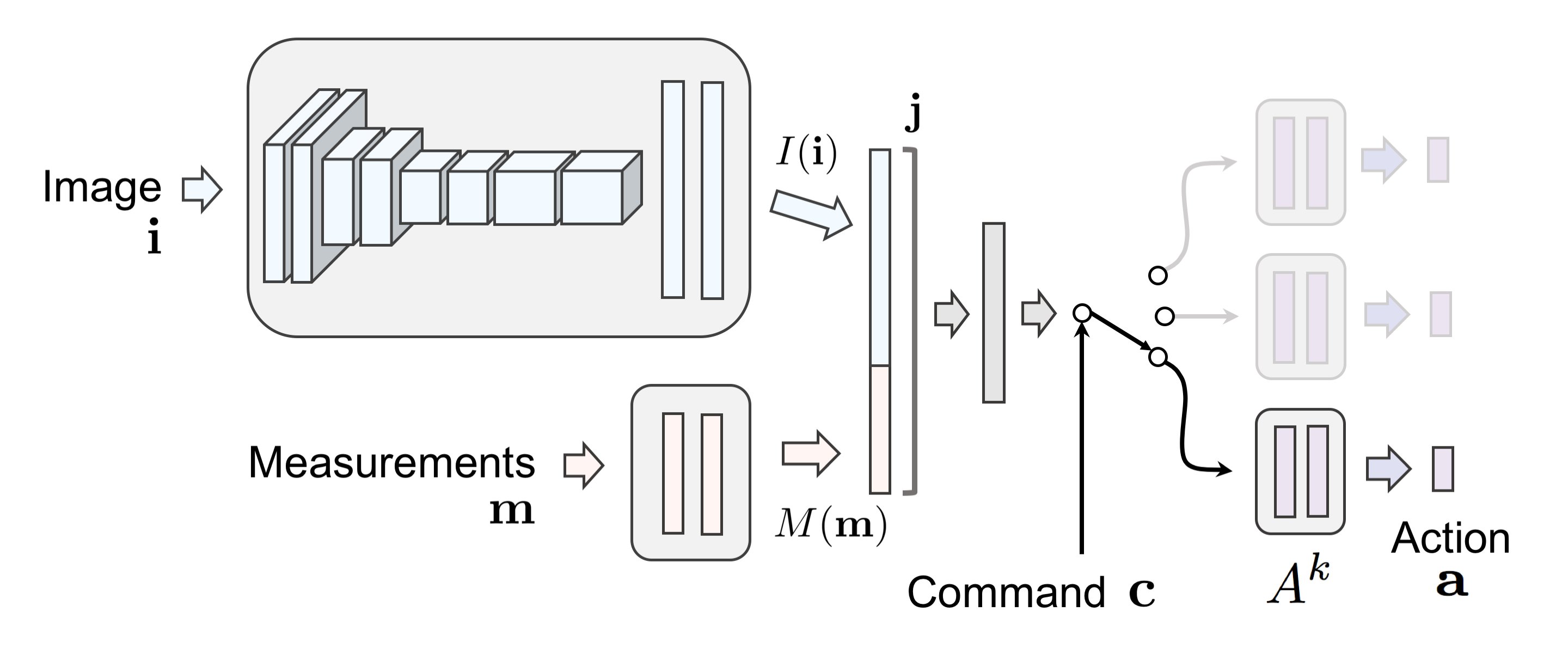
*Figure (来源: 1710.02410v2.pdf, 第 1 页)*



*Figure (来源: 1710.02410v2.pdf, 第 1 页)*



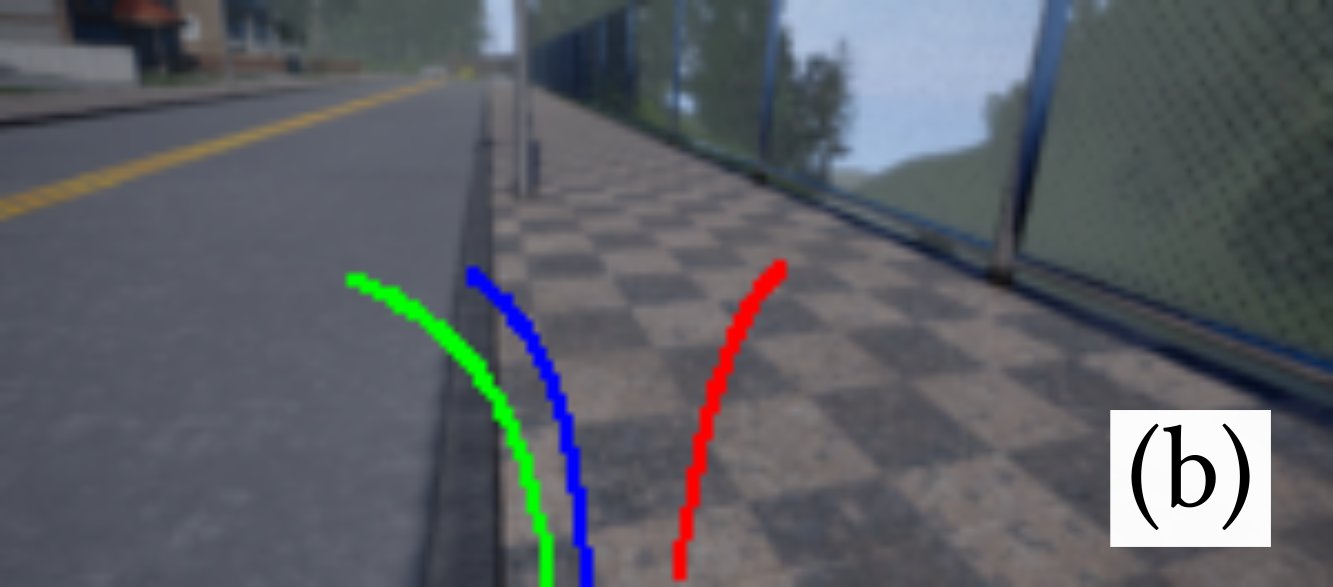
*Figure (来源: 1710.02410v2.pdf, 第 1 页)*



*Figure (来源: 1710.02410v2.pdf, 第 4 页)*



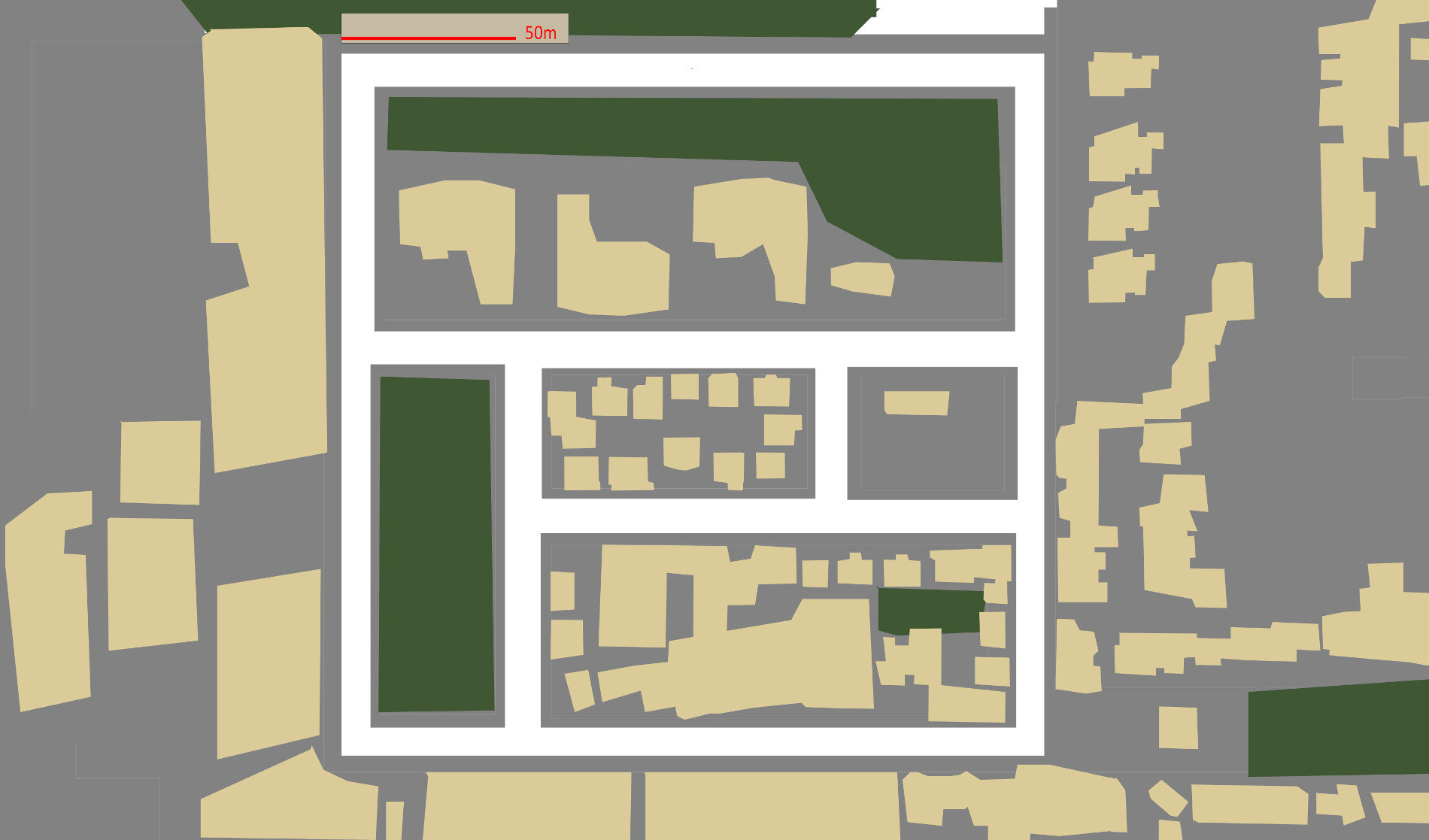
*Figure (来源: 1710.02410v2.pdf, 第 5 页)*



*Figure (来源: 1710.02410v2.pdf, 第 5 页)*



*Figure (来源: 1710.02410v2.pdf, 第 5 页)*



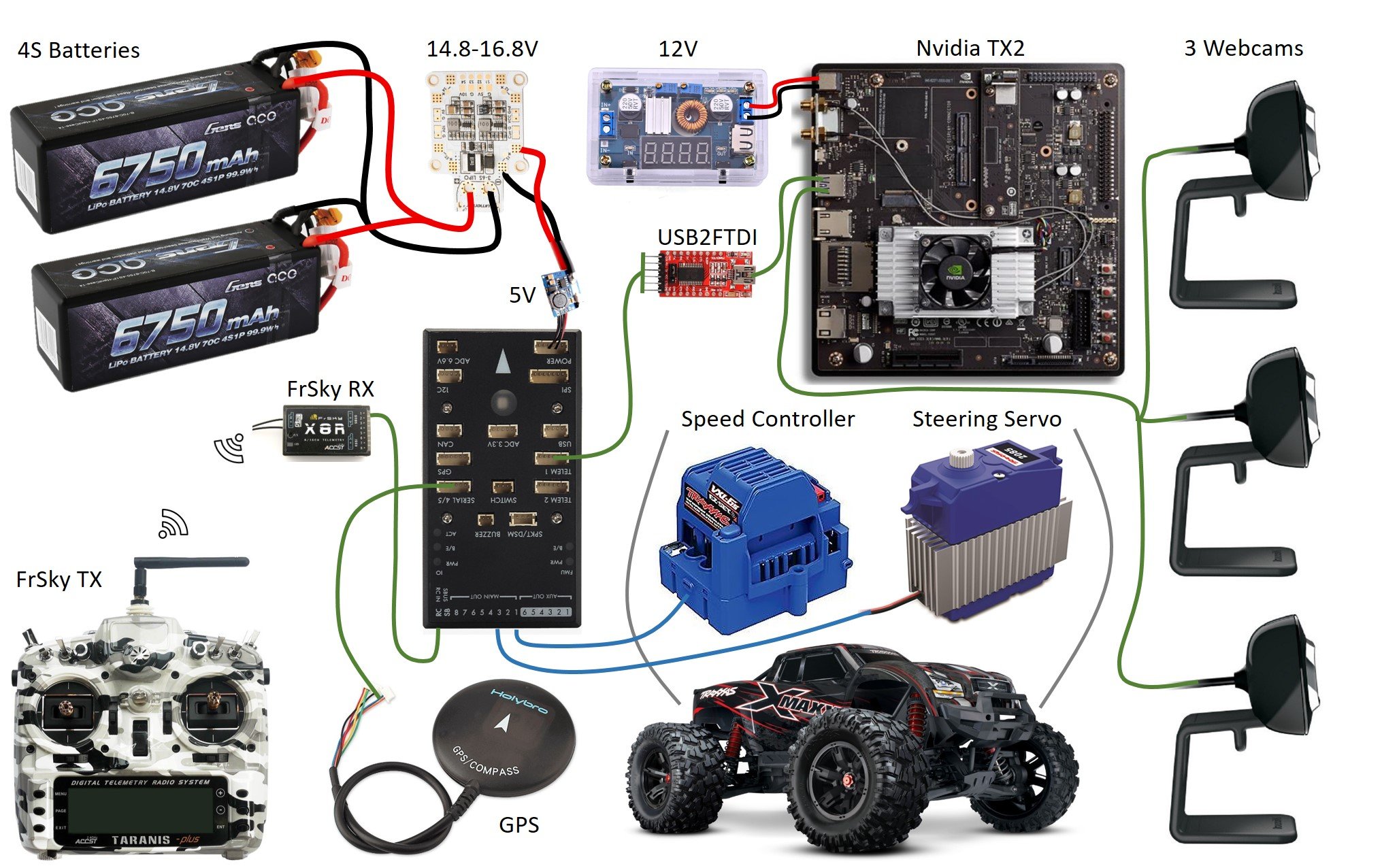
*Figure (来源: 1710.02410v2.pdf, 第 5 页)*



*Figure (来源: 1710.02410v2.pdf, 第 5 页)*



*Figure (来源: 1710.02410v2.pdf, 第 5 页)*



*Figure (来源: 1710.02410v2.pdf, 第 6 页)*



*Figure (来源: 1710.02410v2.pdf, 第 6 页)*



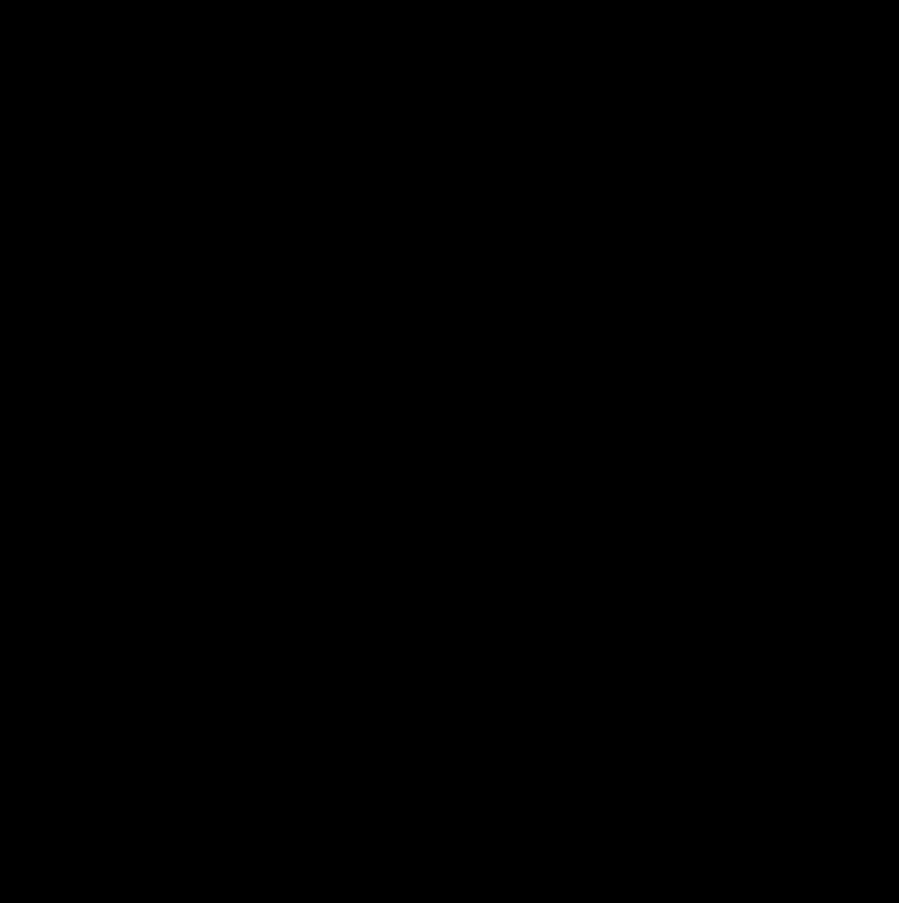
*Figure (来源: 1710.02410v2.pdf, 第 6 页)*



*Figure (来源: 1710.02410v2.pdf, 第 7 页)*



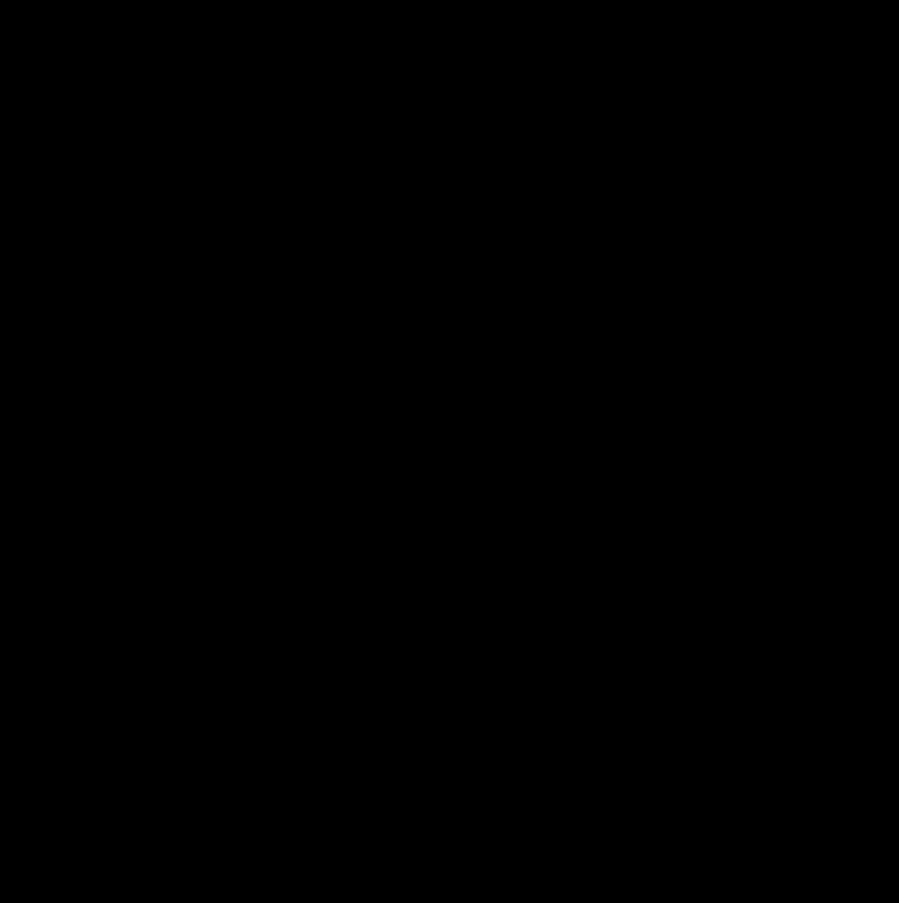
*Figure (来源: 1710.02410v2.pdf, 第 7 页)*



*Figure (来源: end-to-end-dl-using-px.pdf, 第 3 页)*



*Figure (来源: end-to-end-dl-using-px.pdf, 第 3 页)*



*Figure (来源: end-to-end-dl-using-px.pdf, 第 3 页)*



*Figure (来源: end-to-end-dl-using-px.pdf, 第 6 页)*



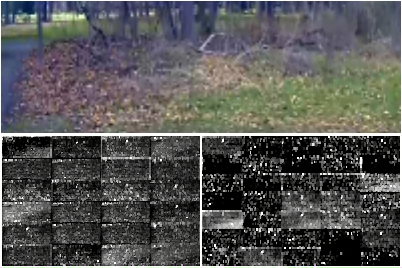
*Figure (来源: end-to-end-dl-using-px.pdf, 第 6 页)*



*Figure (来源: end-to-end-dl-using-px.pdf, 第 6 页)*



*Figure (来源: end-to-end-dl-using-px.pdf, 第 7 页)*



*Figure (来源: end-to-end-dl-using-px.pdf, 第 8 页)*