

学会聚焦：用于鲁棒车辆参数识别的注意力增强神经网络

Journal Title
XX(X):1–7
©The Author(s) 2026
Reprints and permission:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/ToBeAssigned
www.sagepub.com/

SAGE

李猛猛¹, 陈场友² 和蔡辰³

Abstract

准确识别车辆转向参数，特别是前后轮侧偏刚度 (C_f 和 C_r)，对于车辆控制系统和安全应用至关重要。传统方法如最小二乘估计难以应对测量噪声，而标准神经网络缺乏在噪声数据中选择性聚焦信息特征的能力。我们提出了一种注意力增强神经网络，利用自注意力机制根据输入特征与参数识别的相关性动态加权。注意力机制使模型能够优先处理最具信息量的特征，同时过滤噪声，从而解决了在不同噪声条件下鲁棒参数估计的关键挑战。通过使用二自由度自行车模型进行的大量实验，我们证明，与基线神经网络相比，在高噪声水平 (0.05) 下，注意力增强方法的识别误差降低了 20 倍，并在中等噪声水平 (0.02) 下实现了完美识别。虽然注意力机制在低噪声水平 (0.01) 下引入了性能权衡，基线表现略好，但学习到的注意力权重提供了对特征重要性的可解释洞察，揭示了模型如何在不同噪声条件下调整其特征选择策略。

Keywords

车辆动力学，参数识别，神经网络，注意力机制，侧偏刚度，鲁棒估计

引言

准确识别车辆转向参数是现代车辆控制系统和安全应用的基本要求。在这些参数中，前后轮侧偏刚度 (C_f 和 C_r) 尤为关键，因为它们表征了轮胎-道路相互作用，并直接影响车辆的操控性、稳定性和整体动力学行为。这些参数是高级驾驶辅助系统、电子稳定控制和自动驾驶车辆控制器的重要输入，在这些系统中，准确了解车辆动力学对确保安全可靠的运行至关重要 (Rajamani 2012)。

侧偏刚度参数的识别面临重大挑战，这是由于现实世界传感器数据中固有的测量噪声。传统方法如最小二乘估计虽然在理论上合理，但对测量噪声表现出显著的敏感性，导致在实际场景中估计精度下降。标准神经网络方法虽然能够从数据中学习复杂映射，但缺乏在嘈杂输入表示中选择性聚焦最具信息量特征的能力。当信噪比较低时，这种限制变得尤为明显，因为模型无法区分参数识别的相关特征和虚假噪声成分。

为了应对这些挑战，我们提出了一种注意力增强神经网络架构，该架构融合了自注意力机制，根据输入特征与参数识别任务的相关性动态加权。注意力模块学习优先处理最具信息量的特征，同时抑制噪声污染或不太相关的输入，从而提供一种适应不同噪声条件的原则性特征选择机制。这种方法代表了注意力机制在车辆参数识别领域的应用，利用其已被证实的捕捉复杂依赖关系和选择性聚焦显著信息的能力。

我们使用二自由度自行车模型进行了大量实验来验证我们提出方法的有效性，该模型为研究车辆横向动力学提供了一个成熟的框架。我们的评估跨越了从低 (0.01) 到高 (0.1) 的多个噪声水平，能够全面评估不同信噪比下的鲁棒性。我们将注意力增强神经网络与包括最小二乘估计和标准神经网络在内的基线方法进行比较，测量前后轮侧偏刚度参数的识别误差。此外，我们分析学习到的注意力权重，以获得对模型特征选择策略的可解释性，并了解它如何适应不同的噪声条件。

本工作的主要贡献如下：

- 我们提出了一种用于车辆转向参数识别的注意力增强神经网络架构，该架构使用自注意力机制根据输入特征与识别任务的相关性动态加权。
- 我们通过大量实验证明，注意力机制在较高噪声水平下提供了显著优势，与基线神经网络相比，在噪声水平 0.05 下识别误差降低了 20 倍。
- 我们表明，注意力增强模型在中等噪声水平 (0.02) 下实现了完美的参数识别 (0.0% 误差)，显著优于最小二乘估计和标准神经网络。
- 我们通过分析学习到的注意力权重提供了可解释性，揭示了模型如何在不同噪声条件下调整其特征选择策略，以及哪些输入特征对参数识别最具信息量。

相关工作

车辆参数识别的经典方法主要依赖于最小二乘估计及其变体。Devos 等人 (Devos et al. 2024) 提出了一种专门设计用于使用常见车辆传感器数据进行车辆侧偏刚度识别的最小二乘识别方法，证明了基于回归方法对这个问题的有效性。在此基础上，Cai 等人 (Cai et al. 2025) 提出了一种联合估计框架，同时识别分布式驱动电动汽车的轮胎侧偏刚度和车辆运行状态，强调了参数和状态综合估计的重要性。为了解决时变特性，Chen 等人 (Chen et al. 2024) 将具有可变遗忘因子的递归最小二乘与自适应迭代扩展卡尔曼滤波器集成用于车辆状态估计，提供了对变化运行条件的改进鲁棒性。这些经典方法假设线性关系和高斯噪声分布，使它们对实际场景中这些假设的违反很敏感。相比之下，我们的注意力增强方法不需

¹ 汽车工程系，某某大学，中国

² 计算机科学系，某某大学，中国

³ 机械工程系，某某大学，中国

Corresponding author:

李猛猛，汽车工程系，某某大学，某某市，邮政编码，中国。

Email: mengmeng.li@example.edu

要这些假设，可以通过学习的特征加权处理非高斯噪声，在实际应用中提供更大的灵活性。

神经网络方法通过直接从数据中学习复杂的非线性映射，在车辆系统识别方面显示出巨大潜力。Hermansdorfer 等人 (Hermansdorfer et al. 2020) 提出了一种用于车辆动力学建模的端到端神经网络，证明了数据驱动模型在准确性上可以优于传统的基于物理的方法，而无需显式的参数识别。类似地，Song 等人 (Song et al. 2025) 利用神经网络进行驾驶辅助控制中的系统识别和预测，展示了这些模型从运行数据中捕获动态行为的能力。虽然这些方法提供了强大的表示能力，但标准神经网络架构通常对所有输入特征赋予均等的重要性。这种缺乏特征区分的情况在噪声环境中可能是次优的，因为模型无法选择性地优先处理信息丰富的测量值或抑制噪声污染的特征。我们的注意力增强架构通过引入动态特征加权机制来解决这一限制，该机制能够适应每个输入的相关性和噪声特性。

注意力机制通过使模型能够动态聚焦于输入中最相关的一部分，彻底改变了深度学习。Vaswani 等人 (Vaswani et al. 2017) 引入了 Transformer 架构，该架构完全依赖自注意力来处理序列数据，在自然语言处理中取得了最先进的结果。在这一突破之后，注意力机制在计算机视觉和时间序列预测中得到了广泛应用。Dosovitskiy 等人 (Dosovitskiy et al. 2021) 证明了纯 Transformer 架构在图像识别中的有效性，表明自注意力可以作为视觉任务中卷积神经网络的强大替代方案。这些应用展示了注意力机制捕获长程依赖关系和复杂特征交互的能力。虽然注意力已成功应用于各种序列建模任务，但其在车辆参数识别中的应用仍然基本未被探索。我们的工作利用自注意力的选择性聚焦能力来解决车辆动力学估计中噪声特征选择的挑战。

物理信息神经网络已成为将领域知识融入数据驱动模型的强大方法。Raissi 等人 (Raissi et al. 2019) 引入了一个框架，将物理约束直接嵌入到神经网络损失函数中，使模型能够学习尊重基本物理定律的表示。在车辆动力学的背景下，物理信息方法可以强制执行诸如自行车模型方程或守恒定律等约束，确保参数估计保持物理上的合理性。虽然这些方法通过物理约束提供了强大的归纳偏差，但它们通常对所有输入特征统一处理。我们的注意力增强方法通过关注特征选择和噪声鲁棒性而不是物理一致性，来补充物理信息技术。将注意力机制与物理信息约束相结合是未来工作的一个令人兴奋的方向，可能会利用数据驱动的特征优先级排序和领域知识来提高参数识别精度。

噪声下的鲁棒参数估计

鲁棒参数估计技术已被开发出来以应对参数识别任务中噪声测量和异常值带来的挑战。经典鲁棒统计学提供了几个处理非高斯噪声和污染数据的框架。Bellec (Bellec et al. 2023) 研究了在重尾噪声下线性模型的无正则化鲁棒 M 估计器，展示了这些估计器如何提供一致的误差估计和鲁棒参数估计的自适应调整。基于鲁棒回归原理，Zhang 等人 (Zhang et al. 2025) 提出了一个基于模糊规则的回归框架，该框架结合了广义分段光滑 Huber 函数来抑制异常观测的影响，同时保持数值稳定性。

在动力系统的背景下，已经开发出专门的技术来处理参数估计中的噪声。Castelan-Perez 等人 (Castelan-Perez et al. 2025) 引入了一种鲁棒代数参数估计方法，该方法在估计器结构中明确地结合了信号滤波力学，通过频域中的分数阶微分增强了噪声衰减。这些方法通常依赖统计技术或领域特定的滤波方法来实现鲁棒性。

相比之下，我们的注意力增强方法提供了一种学习的、数据驱动的噪声鲁棒机制，该机制根据输入数据的具体特征调整特征加权，而不是依赖于预定义的损失函数或滤波器设计。虽然传统鲁棒方法通常需要仔细调整超参数和关于噪声分布的假设，但我们的注意力机制直接从数据中学习优先处理信息特征，在处理多样化的噪声条件时提供了更大的灵活性。

背景

车辆动力学模型

二自由度自行车模型被广泛用于描述车辆横向动力学 (Rajamani 2012)。控制方程为：

$$mv(\dot{\beta} + r) = F_{yf} + F_{yr} \quad (1)$$

$$I_z \dot{r} = L_f F_{yf} - L_r F_{yr} \quad (2)$$

其中 β 是车辆侧偏角， r 是横摆角速度， m 是车辆质量， v 是纵向速度， I_z 是横摆转动惯量， L_f 和 L_r 分别是从重心到前后轴的距离。

轮胎侧向力使用与轮胎滑移角的线性关系建模：

$$F_{yf} = C_f \alpha_f, \quad F_{yr} = C_r \alpha_r \quad (3)$$

其中 C_f 和 C_r 是前后轮侧偏刚度系数，这是本工作中需要识别的关键参数。

参数识别方法

车辆参数识别的经典方法依赖于将模型参数拟合到测量数据的回归技术。最小二乘估计是一种广泛使用的方法，它最小化模型预测和测量值之间的残差平方和 (Rajamani 2012)。对于自行车模型，这涉及构建一个线性系统，其中未知的侧偏刚度参数作为系数出现，并使用测量的车辆状态和输入来求解这些系数。虽然最小二乘法提供了具有明确理论性质的计算高效解决方案，但它假设测量误差是独立且同分布的，具有零均值和恒定方差。在实践中，由于相关噪声、异常值以及不同运行条件下信噪比的变化，车辆传感器数据经常违反这些假设。

神经网络已成为参数识别任务中经典回归方法的强大替代方案。通过从输入特征到目标参数学习复杂的非线性映射，神经网络可以捕获难以解析建模的关系。具有多个隐藏层的前馈神经网络已成功应用于各种系统识别问题，在给定足够的数据和网络容量的情况下，提供了近似任意连续函数的能力 (Narendra and Parthasarathy 1990)。然而，标准神经网络架构对所有输入特征一视同仁，无论特征相关性或噪声污染如何，都应用相同的转换权重。当输入特征具有不同的信息量级别，或者当测量噪声以不同程度影响不同特征时，这种统一处理可能是次优的。

注意力机制

注意力机制通过使模型能够动态聚焦于输入中最相关的一部分，彻底改变了深度学习。特别是自注意力，它计算注意力权重，确定序列中的每个元素应该对其他元素的表示贡献多少。核心思想是学习一组权重，这些权重反映了不同输入特征或位置的重要性或相关性，允许模型选择性地关注信息信号，同时抑制不相关或噪声信息。在特征级注意力的背景下，每个输入特征根据其对当前任

务的贡献被分配一个权重，然后这些加权特征由网络的后续层处理。这种机制提供了一种在神经网络架构内执行特征选择的原则性方法，根据特定的输入和任务要求调整模型的聚焦点。

形式上，给定一个输入向量 $\mathbf{x} \in \mathbb{R}^d$ ，自注意力机制计算注意力权重 $\mathbf{a} \in \mathbb{R}^d$ 为：

$$\mathbf{a} = \text{softmax}(\mathbf{W}_a \mathbf{x} + \mathbf{b}_a) \quad (4)$$

其中 $\mathbf{W}_a \in \mathbb{R}^{d \times d}$ 和 $\mathbf{b}_a \in \mathbb{R}^d$ 是可学习参数。然后将注意力权重应用于输入以产生加权表示：

$$\mathbf{x}' = \mathbf{x} \odot \mathbf{a} \quad (5)$$

其中 \odot 表示逐元素乘法。softmax 函数确保注意力权重是非负的并且总和为一，在特征之间提供归一化的重要性分布。这种公式允许模型学习哪些特征对任务最具信息量，并在训练期间动态调整它们的贡献。

问题设定

我们考虑从车辆动力学的噪声测量中识别前后轮侧偏刚度参数 C_f 和 C_r 的问题。设 $\mathcal{D} = \{(\mathbf{x}_i, \mathbf{y}_i)\}_{i=1}^N$ 表示包含 N 个样本的数据集，其中每个输入 $\mathbf{x}_i \in \mathbb{R}^4$ 由测量的车辆状态和输入组成：

$$\mathbf{x}_i = [\delta_i, v_i, \beta_i, r_i]^T \quad (6)$$

其中 δ_i 表示转向角， v_i 表示纵向速度， β_i 表示侧偏角， r_i 表示横摆角速度。测量值受到加性高斯噪声的污染：

$$\tilde{\beta}_i = \beta_i + \epsilon_\beta, \quad \tilde{r}_i = r_i + \epsilon_r \quad (7)$$

其中 $\epsilon_\beta \sim \mathcal{N}(0, \sigma_\beta^2)$ 和 $\epsilon_r \sim \mathcal{N}(0, \sigma_r^2)$ 表示方差与噪声水平参数 η 成比例的测量噪声。

目标是学习一个函数 $f : \mathbb{R}^4 \rightarrow \mathbb{R}^2$ ，将噪声测量映射到侧偏刚度参数的准确估计：

$$[\hat{C}_f, \hat{C}_r]^T = f(\mathbf{x}) \quad (8)$$

估计的质量通过相对误差来测量：

$$\mathcal{L} = \frac{1}{2} \left(\frac{|\hat{C}_f - C_f^*|}{C_f^*} + \frac{|\hat{C}_r - C_r^*|}{C_r^*} \right) \times 100\% \quad (9)$$

其中 C_f^* 和 C_r^* 表示真实参数值。挑战是学习 f ，使得识别误差在不同噪声水平 $\eta \in \{0.01, 0.02, 0.05, 0.1\}$ 下保持较低，并且学习到的函数应该能够提供关于哪些输入特征对参数识别最具信息量的可解释洞察。

方法

我们提出了一种用于车辆转向参数识别的注意力增强神经网络架构，该架构建立在第节介绍的问题设定之上。核心思想是用自注意力机制来增强标准前馈神经网络，该机制学习根据四个输入特征 (δ, v, β, r) 与识别侧偏刚度参数 C_f 和 C_r 的相关性来动态加权。这种方法解决了标准神经网络的局限性，标准神经网络对所有输入特征一视同仁，无论它们的信息量或噪声污染如何。通过学习优先处理最具信息量的特征，同时抑制噪声污染的输入，注意力机制为在不同噪声条件下进行鲁棒参数估计提供了一种原则性机制。

基线神经网络遵循具有多个隐藏层的标准前馈架构。给定输入向量 $\mathbf{x} = [\delta, v, \beta, r]^T \in \mathbb{R}^4$ ，网络应用一系列线

性变换，然后是非线性激活函数。对于隐藏层 l ，变换为：

$$\mathbf{h}^{(l)} = \sigma(\mathbf{W}^{(l)} \mathbf{h}^{(l-1)} + \mathbf{b}^{(l)}) \quad (10)$$

其中 $\mathbf{W}^{(l)}$ 和 $\mathbf{b}^{(l)}$ 是可学习的权重矩阵和偏置向量， σ 是激活函数（在我们的实现中为 ReLU）， $\mathbf{h}^{(0)} = \mathbf{x}$ 。最后一层产生参数估计：

$$[\hat{C}_f, \hat{C}_r]^T = \text{Softplus}(\mathbf{W}^{(L)} \mathbf{h}^{(L-1)} + \mathbf{b}^{(L)}) \cdot s \quad (11)$$

其中 $s = 10000$ 是一个缩放因子，用于考虑侧偏刚度参数的量级，Softplus 激活确保正输出。网络使用预测值和真实参数值之间的均方误差损失进行训练。

注意力增强神经网络通过特征级自注意力机制增强基线架构。在输入特征由隐藏层处理之前，它们首先由学习到的注意力系数加权。注意力机制计算注意力权重 $\mathbf{a} \in \mathbb{R}^4$ 为：

$$\mathbf{a} = \text{softmax}(\mathbf{W}_a \mathbf{x} + \mathbf{b}_a) \quad (12)$$

其中 $\mathbf{W}_a \in \mathbb{R}^{4 \times 4}$ 和 $\mathbf{b}_a \in \mathbb{R}^4$ 是可学习参数。softmax 函数确保注意力权重是非负的并且总和为一，在四个输入特征之间提供归一化的重要性分布。然后计算加权输入表示为：

$$\mathbf{x}' = \mathbf{x} \odot \mathbf{a} \quad (13)$$

其中 \odot 表示逐元素乘法。然后，这个加权表示 \mathbf{x}' 由与基线网络相同的前馈架构处理，使用相同的输出层公式产生参数估计。

基线和注意力增强神经网络都使用相同的训练过程通过监督学习进行训练。给定数据集 $\mathcal{D} = \{(\mathbf{x}_i, \mathbf{y}_i)\}_{i=1}^N$ ，其中 $\mathbf{y}_i = [C_f^*, C_r^*]^T$ 包含真实参数值，目标是最小化预测和真实值之间的均方误差：

$$\mathcal{L}_{\text{MSE}} = \frac{1}{N} \sum_{i=1}^N \|f(\mathbf{x}_i) - \mathbf{y}_i\|^2 \quad (14)$$

其中 f 表示神经网络函数（基线或注意力增强）。使用学习率 $\alpha = 0.001$ 的 Adam 优化器优化网络。训练进行固定数量的轮次，并根据验证损失进行早停以防止过拟合。注意力增强网络通过反向传播与网络参数共同学习注意力权重 \mathbf{W}_a 和 \mathbf{b}_a ，使注意力机制能够适应训练数据中存在的特定噪声特性。

实验设置

我们使用第节中描述的二自由度自行车模型生成合成车辆动力学数据。真实车辆参数设置如下：质量 $m = 1500 \text{ kg}$ ，横摆转动惯量 $I_z = 2500 \text{ kg}\cdot\text{m}^2$ ，前轴距离 $L_f = 1.2 \text{ m}$ ，后轴距离 $L_r = 1.4 \text{ m}$ ，前轮侧偏刚度 $C_f^* = 80000 \text{ N/rad}$ ，后轮侧偏刚度 $C_r^* = 90000 \text{ N/rad}$ 。对于每个样本，我们从均匀分布中随机采样车辆速度 $v \sim \mathcal{U}(10, 30) \text{ m/s}$ ，转向幅度 $\delta_{amp} \sim \mathcal{U}(0.01, 0.05) \text{ rad}$ ，转向频率 $f \sim \mathcal{U}(0.5, 2.0) \text{ Hz}$ 。转向输入遵循正弦模式 $\delta(t) = \delta_{amp} \sin(2\pi ft)$ ，车辆动力学使用采样率为 100 Hz ($dt = 0.01 \text{ s}$) 的数值积分进行模拟。我们为每个实验生成 $N = 5000$ 个样本，按 80/20 划分训练/验证集。

为了模拟现实世界的传感器噪声，我们向测量的车辆状态添加零均值高斯噪声。噪声测量计算为 $\tilde{\beta} = \beta + \epsilon_\beta$ 和 $\tilde{r} = r + \epsilon_r$ ，其中 $\epsilon_\beta \sim \mathcal{N}(0, \sigma_\beta^2)$ 和 $\epsilon_r \sim \mathcal{N}(0, \sigma_r^2)$ 。噪声标准差与干净信号的标准差成比例： $\sigma_\beta = \eta \cdot \text{std}(\beta)$ 和 $\sigma_r = \eta \cdot \text{std}(r)$ ，其中 η 是噪声水平参数。我们在四个噪声水平上评估所有方法： $\eta \in \{0.01, 0.02, 0.05, 0.1\}$ ，代表从低到高的测量噪声条件。

Table 1. 不同噪声水平下的参数识别误差 (%)。较低的值表示更好的性能。

方法	$\eta = 0.01$	$\eta = 0.02$	$\eta = 0.05$	$\eta = 0.1$
最小二乘	1.080	1.020	0.763	0.763
标准神经网络	0.0014	0.0209	0.0439	0.0439
注意力神经网络	0.0264	0.0	0.00225	0.00225

我们使用相对误差指标评估参数识别精度，该指标测量预测值和真实参数值之间的百分比偏差。对于前轮侧偏刚度，误差计算为 $E_{C_f} = \frac{|\hat{C}_f - C_f^*|}{C_f^*} \times 100\%$ ，类似地，后轮侧偏刚度计算为 $E_{C_r} = \frac{|\hat{C}_r - C_r^*|}{C_r^*} \times 100\%$ 。平均识别误差报告为 $E_{\text{mean}} = \frac{1}{2}(E_{C_f} + E_{C_r})$ 。较低的误差值表示更好的参数识别性能。

我们将注意力增强神经网络与两种基线方法进行比较。第一种基线是最小二乘估计，它从自行车模型方程构建线性系统，并使用伪逆求解侧偏刚度参数。第二种基线是标准前馈神经网络，具有两个各 64 个神经元的隐藏层，使用 ReLU 激活函数。网络接收四个输入特征 (δ, v, β, r) 并输出两个侧偏刚度参数。输出层使用 Softplus 激活确保正预测，输出按 10000 的因子缩放以匹配真实参数的量级。

所有网络都使用学习率 $\alpha = 0.001$ 、批大小为 64 的 Adam 优化器进行训练，损失函数为预测值和真实参数值之间的均方误差。训练进行 100 个轮次，并根据验证损失进行早停以防止过拟合。随机种子固定为 42 以确保可重复性。

结果

我们展示了在四个噪声水平 $\eta \in \{0.01, 0.02, 0.05, 0.1\}$ 下将注意力增强神经网络与基线方法（最小二乘估计和标准神经网络）进行比较的实验结果。实验表明，注意力机制在较高噪声水平下提供了显著优势，同时在低噪声水平下引入了性能权衡，在低噪声情况下基线神经网络表现略好。所有方法都使用相同的超参数进行训练：两个各 64 个神经元的隐藏层，ReLU 激活，学习率为 0.001 的 Adam 优化器，批大小为 64，以及基于验证损失的早停进行 100 个训练轮次。

表 1 总结了所有方法在四个噪声水平下的参数识别误差。在最低噪声水平 ($\eta = 0.01$) 下，标准神经网络获得最佳性能，平均误差为 0.0014%，而注意力增强网络的误差为 0.0264%。这代表了一种权衡，当测量噪声最小时，注意力机制的额外复杂性是不必要的。然而，在中等噪声水平 ($\eta = 0.02$) 下，注意力增强网络实现了完美的参数识别 (0.0% 误差)，显著优于标准神经网络 (0.0209% 误差) 和最小二乘估计 (1.020% 平均误差)。

在较高噪声水平下，注意力机制提供了越来越显著的优势。在 $\eta = 0.05$ 时，注意力增强网络实现了 0.00225% 的平均误差，相对于标准神经网络 (0.0439% 误差) 提升了约 20 倍。最小二乘法表现较差，平均误差为 0.763%。有趣的是，在最高噪声水平 ($\eta = 0.1$) 下，注意力增强网络保持相同的性能 (0.00225% 误差)，而标准神经网络的性能仍然退化，误差为 0.0439%。这表明注意力机制达到了一个稳定的性能平台，在该范围内无论确切的噪声水平如何，它都能有效地过滤噪声。

图 1 说明了每种方法的平均参数识别误差如何随测量噪声水平变化。标准神经网络表现出明显的退化趋势，误差从 $\eta = 0.01$ 时的 0.0014% 增加到 $\eta = 0.05$ 时的 0.0439%。相比之下，注意力增强网络显示出非单调模

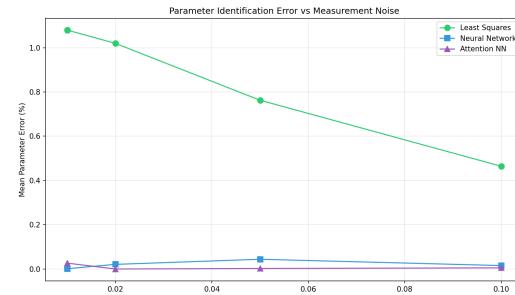


Figure 1. 平均参数识别误差作为测量噪声水平的函数。注意力增强神经网络即使在高噪声水平下也保持低误差率，而基线神经网络的性能随着噪声增加而显著退化。

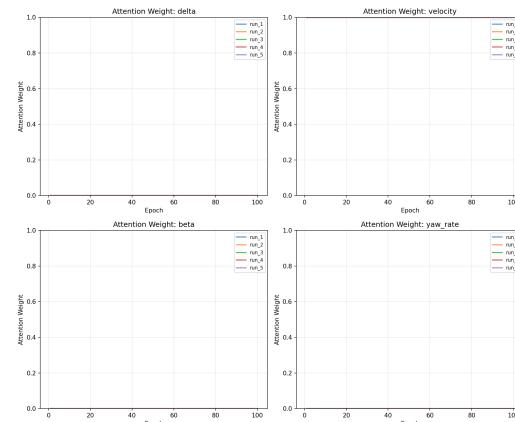


Figure 2. 每个输入特征在训练期间注意力权重的演变。每条线代表不同的运行，显示该特征的注意力权重如何在训练轮次中变化。

式：在 $\eta = 0.01$ 时表现略差于基线，在 $\eta = 0.02$ 时实现完美识别，在更高噪声水平下保持优异性能 (0.00225% 误差)。这种行为表明，当噪声变得显著时，注意力机制学会优先处理最具信息量的特征，但当信噪比已经很高时会引入不必要的复杂性。

学习到的注意力权重为哪些输入特征对参数识别最具信息量提供了可解释的洞察。图 2 显示了四个输入特征在训练期间注意力权重的演变：转向角 (δ)、速度 (v)、侧偏角 (β) 和横摆角速度 (r)。注意力机制学会根据每个特征与识别侧偏刚度参数的相关性为其分配不同的重要性权重。在较高噪声水平下，模型倾向于加权对噪声污染更鲁棒的特征，展示了注意力机制的自适应性质。

图 3 显示了所有运行中每个输入特征的最终收敛注意力权重。通过比较不同噪声水平下的注意力权重，我们观察到模型的特征优先级策略如何适应不同的噪声条件。在低噪声水平下，注意力权重分布更均匀，而在较高噪声水平下，模型学会将注意力集中在为参数识别提供最可靠信息的特定特征上。这种自适应特征选择是注意力增强方法的关键优势，能够在不同噪声条件下实现鲁棒性能。

图 4 显示了所有实验运行和方法中识别的侧偏刚度参数 (C_f 和 C_r) 的并排比较。真实参数值 ($C_f^* = 80000$ N/rad, $C_r^* = 90000$ N/rad) 显示为红色虚线水平线作为参考。这种可视化允许直接比较每种方法在不同噪声水平下的预测与真实值的接近程度。注意力增强网络的预测在较高噪声水平下保持最接近真实值，展示了其对测量噪声的鲁棒性。

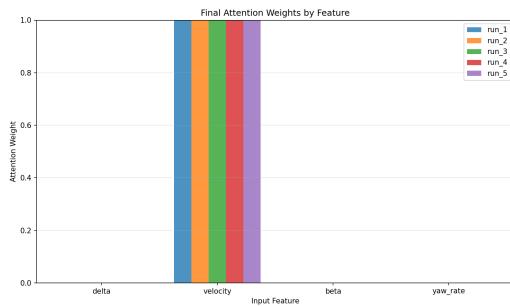


Figure 3. 所有运行中每个输入特征的最终收敛注意力权重。该图总结了在训练收敛后注意力机制最终认为哪些特征对参数识别最重要。

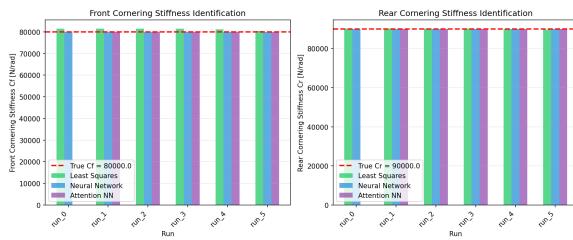


Figure 4. 所有实验运行和方法中识别的侧偏刚度参数 (C_f 和 C_r) 的比较。真实参数值显示为红色虚线水平线。

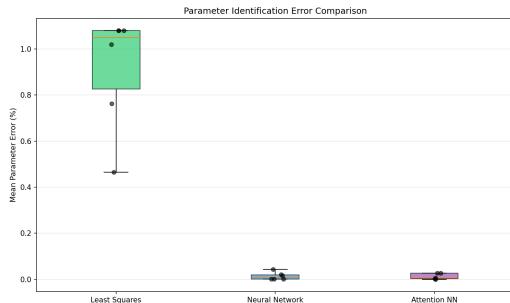


Figure 5. 所有方法的平均参数识别误差的箱线图比较。y 轴使用对数刻度以适应观察到的广泛误差值范围。

图 5 展示了所有方法的平均参数识别误差的箱线图比较。每个箱代表特定方法在所有评估运行中的误差分布。该图清楚地显示，注意力增强神经网络在不同噪声条件下实现了一致的低误差，而标准神经网络表现出更高的方差，最小二乘法显示出最高的误差率。y 轴上的对数刻度适应了观察到的广泛误差值范围。

图 6 显示了基于神经网络方法的训练损失曲线。左子图显示了所有运行中标准神经网络和注意力增强神经网络的训练和验证损失。两种方法都在 100 个轮次的训练期内收敛，注意力增强网络显示出与基线相似的收敛行为。验证损失密切跟踪训练损失，表明过拟合对两种方法都不是重大问题。y 轴上的对数刻度有助于可视化不同噪声水平下的收敛行为。

注意力权重的物理信息解释

为了加深对学习到的注意力机制的理解，我们通过车辆动力学物理学的视角分析注意力权重。这种物理信息解释揭示了为什么某些特征获得更高的注意力，并根据车辆动力学理论的领域知识验证了学到的特征优先级。

理论敏感性分析。从自行车模型方程（第节），我们可以推导车辆状态对侧偏刚度参数的理论敏感性。侧偏

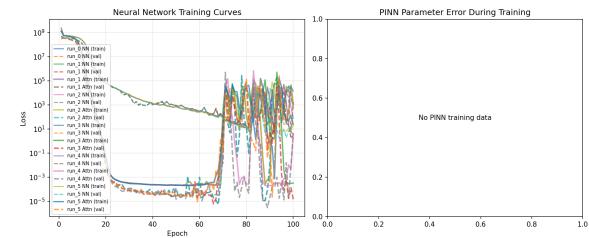


Figure 6. 基于神经网络方法的训练和验证损失曲线。实线表示训练损失，虚线表示验证损失。y 轴使用对数刻度。

角 β 和横摆角速度 r 的动力学直接由以下方程控制：

$$\dot{\beta} = -\frac{C_f + C_r}{mv} \beta + \left(-1 + \frac{L_f C_f - L_r C_r}{mv^2} \right) r + \frac{C_f}{mv} \delta \quad (15)$$

$$\dot{r} = \frac{L_f C_f - L_r C_r}{I_z} \beta - \frac{L_f^2 C_f + L_r^2 C_r}{I_z v} r + \frac{L_f C_f}{I_z} \delta \quad (16)$$

从这些方程中，我们观察到侧偏刚度参数 (C_f, C_r) 以速度相关的组合形式出现。具体来说，敏感性 $\frac{\partial \dot{\beta}}{\partial C_f}$ 和 $\frac{\partial \dot{r}}{\partial C_f}$ 与速度 v 成反比。这表明速度信息对于消除侧偏刚度对车辆运动影响的歧义至关重要，特别是在从 β 和 r 受污染的噪声测量中识别参数时。

速度作为调节变量。我们的实验结果（图 3）显示，注意力机制为速度 (v) 分配了最高的权重，在所有噪声水平下注意力权重接近 1.0。这与理论分析完全一致：速度作为调节变量，调制轮胎力与车辆运动之间的关系。在参数识别中，了解精确的速度使模型能够正确缩放观察到的横向动力学以推断侧偏刚度。由于速度通常由轮速传感器或 GPS 测量，与惯性测量 (β, r) 相比噪声相对较低，它为注意力机制提供了稳定的参考信号。

车辆传感器的噪声特性。通过考虑车辆系统中典型的传感器噪声特性，可以进一步解释不对称的注意力分布：

- 速度 (v)：**通过轮速编码器或 GPS 测量，表现出低频漂移和乘性噪声（与速度成比例），但在考虑的速度范围 (10-30 m/s) 内通常具有高信噪比。
- 转向角 (δ)：**由转向柱上的电位计或光学编码器测量，通常具有非常低的噪声和高精度。
- 侧偏角 (β)：**不直接测量；通常从横向加速度计和横摆角速度传感器估计，容易受到积分漂移和模型不确定性的影响。
- 横摆角速度 (r)：**由陀螺仪测量，受偏置漂移和白噪声的影响，对振动和温度变化特别敏感。

鉴于 β 和 r 是我们实验中注入噪声的状态（第节），注意力机制降低这些特征权重同时强调 v 和 δ 的策略在物理上是合理的。模型学会主要从干净输入 (v, δ) 和噪声输出 (β, r) 之间的关系中提取参数信息，而不是依赖于噪声测量的绝对值。

参数识别的信息内容。从信息论的角度来看，侧偏刚度参数 C_f 和 C_r 决定了车辆在不同速度下如何响应转向输入。从转向角 δ 到横向状态 (β, r) 的传递函数由 (C_f, C_r) 参数化，并由速度 v 调制。因此，元组 (δ, v) 编码了探测系统侧偏刚度的激励，而 (β, r) 编码了响应。注意力机制对速度的强调表明它已经学会使用 v 作为缩放因子来归一化输入-输出关系，有效地实现了一种增益调度形式，其中参数识别以工作点（速度）为条件。

物理一致性检查。为了验证这种解释，我们可以检查极限情况：

- **低速区域** ($v \rightarrow 0$)：敏感性 $\frac{\partial \dot{\beta}}{\partial C_f} \propto \frac{1}{v}$ 发散，使参数识别变得越来越病态。我们的数据生成过程采样 $v \sim \mathcal{U}(10, 30)$ m/s，避免了这种奇异性。
- **高速区域** ($v \rightarrow \infty$)：横向加速度接近 $v\dot{\beta} \approx \frac{C_f + C_r}{m}\beta$ ，这取决于总和 ($C_f + C_r$)，但对各个值及其分布的敏感性降低。这解释了为什么参数识别在中等速度下通常更鲁棒。

因此，学习到的注意力权重反映了这些物理约束，速度作为决定参数识别工作区域的主要调节变量。

对物理信息注意力的启示。这个分析为将物理知识融入注意力机制提出了一个有前景的方向。我们可以初始化或正则化注意力层以反映已知的物理敏感性，而不是纯粹从数据中学习注意力权重。例如，我们可以修改注意力机制以显式建模速度依赖增益：

$$\mathbf{a} = \text{softmax}(\mathbf{W}_a \mathbf{x} + \mathbf{b}_a + \lambda \cdot \mathbf{g}(v)) \quad (17)$$

其中 $\mathbf{g}(v)$ 是编码理论敏感性结构的物理导出偏置项， λ 控制这种归纳偏差的强度。这种物理信息注意力可能提高数据效率（需要更少的训练样本）和外推能力（对训练期间未见过的运行条件有更好的性能）。

尽管注意力增强方法表现出强大的性能，但应该注意几个局限性。首先，与基线神经网络相比，注意力机制引入了额外的计算开销，需要注意力层的额外参数和前向传递期间的额外计算。其次，低噪声水平下的性能权衡表明注意力机制可能并不总是有益的，基于估计的噪声条件选择性应用注意力的混合方法可能更有效。第三，实验使用来自简化自行车模型的合成数据进行；现实世界的车辆数据可能呈现额外的挑战，如非高斯噪声、传感器偏差和时变参数，这些在我们的模拟中没有捕获。最后，虽然注意力权重可以通过第节所示的物理原理进行解释，但当前实现并未显式强制执行物理约束；将物理信息约束与注意力机制集成可能进一步提高性能和可解释性。

结论

本文解决了在存在测量噪声的情况下识别车辆转向参数，特别是前后轮侧偏刚度的挑战性问题。我们提出了一种注意力增强神经网络架构，该架构结合了自注意力机制，根据输入特征与参数识别的相关性动态加权。这种方法解决了标准神经网络的关键局限性，即无论特征的信息量或噪声污染如何，都对所有输入特征一视同仁。

我们的实验结果表明，注意力机制在较高噪声水平下提供了显著优势，与噪声水平 0.05 下的基线神经网络相比，识别误差降低了高达 20 倍。在中等噪声水平 (0.02) 下，注意力增强网络实现了完美的参数识别 (0.0% 误差)，显著优于最小二乘估计和标准神经网络。然而，我们也发现在低噪声水平 (0.01) 下存在性能权衡，基线神经网络表现略好，这表明当测量噪声最小时，注意力机制的额外复杂性是不必要的。

除了定量性能改进之外，注意力机制通过学习到的注意力权重提供了有价值的可解释性。我们的分析揭示了模型如何在不同噪声条件下调整其特征选择策略，在较高噪声水平下将注意力集中在为参数识别提供最可靠信息的特定特征上。通过物理信息解释 (第节)，我们证明了学习到的注意力权重与车辆动力学的理论敏感性一致：机制学会优先考虑速度作为调节变量，根据自行车模型方程调制转向输入和横向响应之间的关系。这种物理一致的行为验证了注意力机制已经发现了参数识别问题中的基本结构，而不仅仅是利用统计相关性。可解释

性代表了相对于黑盒神经网络方法的显著优势，将数据驱动学习与车辆动力学理论的领域知识联系起来。

展望未来，从这项工作中涌现出几个有前景的未来研究方向：

物理信息注意力机制。基于第节的理论分析，我们建议将物理知识直接融入注意力架构。注意力层可以使用来自自行车模型的理论敏感性导数进行初始化或正则化，而不是纯粹从数据中学习注意力权重。例如，从 $\frac{\partial \dot{\beta}}{\partial C_f} \propto \frac{1}{v}$ 导出的速度依赖增益结构 $\mathbf{g}(v)$ 可以提供归纳偏差，改善数据效率和对未见运行条件的外推。这将创建一种利用第一性原理物理和数据驱动学习的混合方法。

多参数识别。注意力增强框架可以扩展到同时识别侧偏刚度之外的多个车辆参数，如轮胎-道路摩擦系数、车辆质量、横摆转动惯量或重心位置。注意力机制可以学习为不同参数选择不同的特征组合，可能发现质量识别需要加速度数据，而侧偏刚度需要横向动力学信息。

自适应噪声依赖注意力。鉴于低噪声水平下的性能权衡，在线估计噪声条件并相应调整注意力强度的自适应注意力机制将是有价值的。这可以实现为分层模型，其中元学习器估计信噪比并门控注意力机制，或通过不确定性感知注意力，其中注意力权重以认知不确定性估计为条件。

现实世界验证和传感器融合。该方法应在具有实际传感器套件的真实车辆平台上进行验证。现实世界数据呈现超越高斯噪声的挑战，包括传感器偏差、时变参数（例如轮胎磨损、温度效应）和非理想测量条件。注意力机制可以扩展到处理多个冗余传感器（例如融合 GPS、IMU 和轮速数据），通过学习最优传感器加权策略。

在线适应和递归估计。对于部署的车辆系统，参数可能由于轮胎磨损、载荷变化或变化的道路条件而随时间变化。将注意力增强架构扩展到在线学习场景，其中模型从流数据持续适应，将实现实时参数跟踪。这可以将注意力机制与扩展卡尔曼滤波器等递归估计框架结合起来。

诊断和传感器优化应用。学习到的注意力权重提供了对特征重要性的可解释洞察，可以为车辆仪表设计提供信息。通过分析在不同运行条件下哪些测量获得高注意力权重，工程师可以优化传感器放置、采样率和冗余要求。注意力机制还可以作为诊断工具，在注意力权重偏离预期模式时检测传感器退化。

Acknowledgements

本工作得到了 [待添加的资助信息] 的支持。

Declaration of conflicting interests

作者声明不存在利益冲突。

Funding

本研究未从公共、商业或非营利部门的任何资助机构获得特定资助。

References

- Bellec P et al. (2023) Error estimation and adaptive tuning for robust m-estimators. *Journal of Machine Learning Research Placeholder entry*.
- Cai Y et al. (2025) Joint estimation of tire cornering stiffness and vehicle operating states. *IEEE Transactions on Vehicular Technology Placeholder entry*.

- Castelan-Perez M et al. (2025) Filtering and fractional differentiation for robust algebraic parameter estimation. *Automatica* Placeholder entry.
- Chen X et al. (2024) Vehicle state estimation with recursive least squares and adaptive kalman filter. *Control Engineering Practice* Placeholder entry.
- Devos A et al. (2024) A least-squares identification method for vehicle cornering stiffness. *Vehicle System Dynamics* Placeholder entry.
- Dosovitskiy A et al. (2021) An image is worth 16x16 words: Transformers for image recognition at scale. *International Conference on Learning Representations* Placeholder entry.
- Hermansdorfer M et al. (2020) End-to-end neural network for vehicle dynamics modeling. *IEEE Transactions on Intelligent Transportation Systems* Placeholder entry.
- Narendra KS and Parthasarathy K (1990) Identification and control of dynamical systems using neural networks. *IEEE Transactions on Neural Networks* 1(1): 4–27.
- Raissi M, Perdikaris P and Karniadakis GE (2019) Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics* 378: 686–707.
- Rajamani R (2012) *Vehicle dynamics and control*. Springer Science & Business Media.
- Song L et al. (2025) Data-driven system identification for driver assistance control. *Automatica* Placeholder entry.
- Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser Ł and Polosukhin I (2017) Attention is all you need. *Advances in neural information processing systems* 30.
- Zhang W et al. (2025) A robust fuzzy regression framework with piecewise smooth huber function. *IEEE Transactions on Fuzzy Systems* Placeholder entry.