Computer Science and Application 计算机科学与应用, 2025, 15(6), 231-243
Published Online June 2025 in Hans. https://www.hanspub.org/journal/csa
https://doi.org/10.12677/csa.2025.156173

Hans汉斯

基于YOLOv8n的道路目标检测方法研究

张佳宠，于　霞

沈阳工业大学信息与工程学院，辽宁 沈阳

收稿日期：2025年5月24日；录用日期：2025年6月23日；发布日期：2025年6月30日

摘　要

自动驾驶中的感知系统主要使用目标检测算法来获取障道路上目标的分布，以便进行识别和分析。当前的目标检测算法发展迅速，但在实际应用场景中平衡实时检测和高检测精度的要求具有挑战性。为了解决上述问题，本文使用YOLOv8n作为原始模型，并提出了一个名为YOLOv8n-CSS的目标检测网络。首先，引入CBAM混合注意力机制增强关键特征提取并去除冗余，从而提高网络对物体和背景的识别能力。然后，使用SPPCSPC模块替换原始模型骨干网络中的SPPF模块，能够更好地融合来自不同层次和尺度的特征信息，可以有效地捕捉不同尺度物体的特征，提高模型识别物体的准确性。最后，引入SPD-Conv模块替换原始的交错卷积层，进行下采样操作，保留了更多的特征信息，从而提高了不同尺度目标的检测能力。在KITTI数据集和BDD100K数据集上的实验结果表明，改进的网络模型的平均准确率分别达到96.1%和48.0%，比基线模型分别高出3.9%和7.9%，明显优于基线模型。该模型在保证高检测精度的基础上，可以实现一般场景中的实时图像处理。

关键词

自动驾驶，目标检测，YOLOv8n

Research on Road Target Detection Method Based on YOLOv8n

Jiachong Zhang, Xia Yu

School of Information and Engineering, Shenyang University of Technology, Shenyang Liaoning

Received: May 24th, 2025; accepted: Jun. 23rd, 2025; published: Jun. 30th, 2025

Abstract

The perception system in autonomous driving mainly uses object detection algorithms to obtain the distribution of objects on the road for identification and analysis. Although current object detection algorithms are developing rapidly, it remains challenging to balance the requirements of real-time

文章引用: 张佳宠, 于霞. 基于 YOLOv8n 的道路目标检测方法研究[J]. 计算机科学与应用, 2025, 15(6): 231-243.
DOI: 10.12677/csa.2025.156173

detection and high detection accuracy in practical application scenarios. To address the above issues, this paper uses YOLOv8n as the original model and proposes an object detection network named YOLOv8n-CSS. First the CBAM hybrid attention mechanism is introduced to enhance the extraction of key features and remove redundancy, thereby improving the network's ability to distinguish objects from the background. Then, the SPPF module in the backbone network of the original model is replaced with the SPPCSPC module. This allows for better integration of feature information from different levels and scales, effectively capturing the features of objects of various scales and improving the accuracy of object recognition by the model. Finally, the SPD-Conv module is introduced to replace the original staggered convolution layer for downsampling operations, which retains more feature information and thus enhances the detection ability for objects of different scales. Experimental results on the KITTI dataset and the BDD100K dataset show that the average accuracy of the improved network model reaches 96.1% and 48.0% respectively, which is 3.9% and 7.9% higher than that of the baseline model, significantly outperforming the baseline model. This model can achieve real-time image processing in general scenarios while ensuring high detection accuracy.

Keywords

Automatic Driving, Object Detection, YOLOv8n

1. 引言

在科技水平不断进步以及人工智能技术飞速发展的当下，自动驾驶不再仅仅是一个设想，而是逐步迈向实际应用阶段[1]。在这一技术变革进程里，目标检测技术起着极为关键的作用。道路环境复杂多样，其中有行人、机动车辆以及形形色色的障碍物。怎样让自动驾驶系统精确识别这些目标，已然成为保障行车安全的核心问题[2]。所以，搭建一个高效且精准的目标识别系统，是实现自动驾驶安全、稳定运行的重要技术保障。

随着泛化能力的增强以及精度的提升，基于深度学习的方法已逐渐替代传统算法，成为目标检测领域的主流手段[3]。不过，现有的技术仍旧面临着诸多难题。在目前的道路目标检测任务当中，诸如各类天气状况、光照变化、物体遮挡，以及真实交通场景里小物体的存在等因素，带来了大量的不确定性，进而导致道路目标检测的精度有所下降。

目前，基于深度学习的目标检测算法分为两类：两阶段目标检测和一阶段目标检测。两阶段算法以其精度闻名，将目标检测任务分为两个步骤：首先生成候选区域，然后对每个区域进行分类和回归。擅长处理复杂场景和小目标检测，典型代表为 R-CNN [4]。尽管两阶段算法越来越准确，但它们也存在某些缺点。首先，这类方法需要依次执行区域生成和目标分类等多个步骤，导致计算复杂度显著提升，对硬件资源的需求也随之提升，这在很大程度上制约了在实时系统中的部署与应用。其次，一些两阶段算法在背景复杂或目标尺度较大的场景中会遇到不稳定，导致区域选择不一致，从而导致误检或漏检的问题[5]。此外，这类算法在检测小尺寸物体方面可能表现不佳，可能会导致漏检或定位不准确，这在检测小尺寸目标时构成了挑战[6]。相比之下，单阶段算法直接预测图像中目标的位置和类别，而不生成候选区域。提供了更高的速度和实时性能，使其非常适合需要快速响应的应用，特别是在道路目标检测中[7]。典型的单阶段算法主要包括 YOLO 系列[8]、SSD [9]、RetinaNet [10]。

作为目标检测领域的重要突破，YOLO 系列算法凭借其优异的检测精度和高效的计算性能，在学术界和工业界获得了广泛认可。其中，YOLOv8 作为该系列的最新版本，在多个基准测试中展现了卓越的性能。本研究致力于对该网络架构进行改进与优化，旨在进一步提升其目标检测能力。通过在 KITTI 和 BDD100K 数据集上进行实验验证，改进后的模型在检测准确性和鲁棒性方面均取得了显著提升，为复杂场景下的目标检测任务提供了新的解决方案。为了实现这一目标，我们做了如下改进：

(1) 引入 CBAM 混合注意力机制增强关键特征提取并去除冗余，从而提高网络对物体和背景的识别能力。

(2) 引入 SPPCSPC 模块替换原始模型骨干网络中的 SPPF 模块，能够更好地融合来自不同层次和尺度的特征信息，可以有效地捕捉不同尺度物体的特征，提高模型识别物体的准确性。

(3) 引入 SPD-Conv 模块替换原始的交错卷积层，进行下采样操作，保留了更多的特征信息，从而提高了不同尺度目标的检测能力。

2. 改进的 YOLOv8 网络

2.1. YOLOv8 的网络结构

YOLOv8 的主要网络结构[11]分为五种类型，分别为 YOLOv8n、YOLOv8s、YOLOv8m、YOLOv8l 和 YOLOv8h。YOLOv8n 具有体积小、速度快的特点，更适合嵌入式设备。因此，本文选择 YOLOv8n 作为网络的基本模型。YOLOv8n 网络结构主要由主干、颈部和头部组成，如图 1 所示。

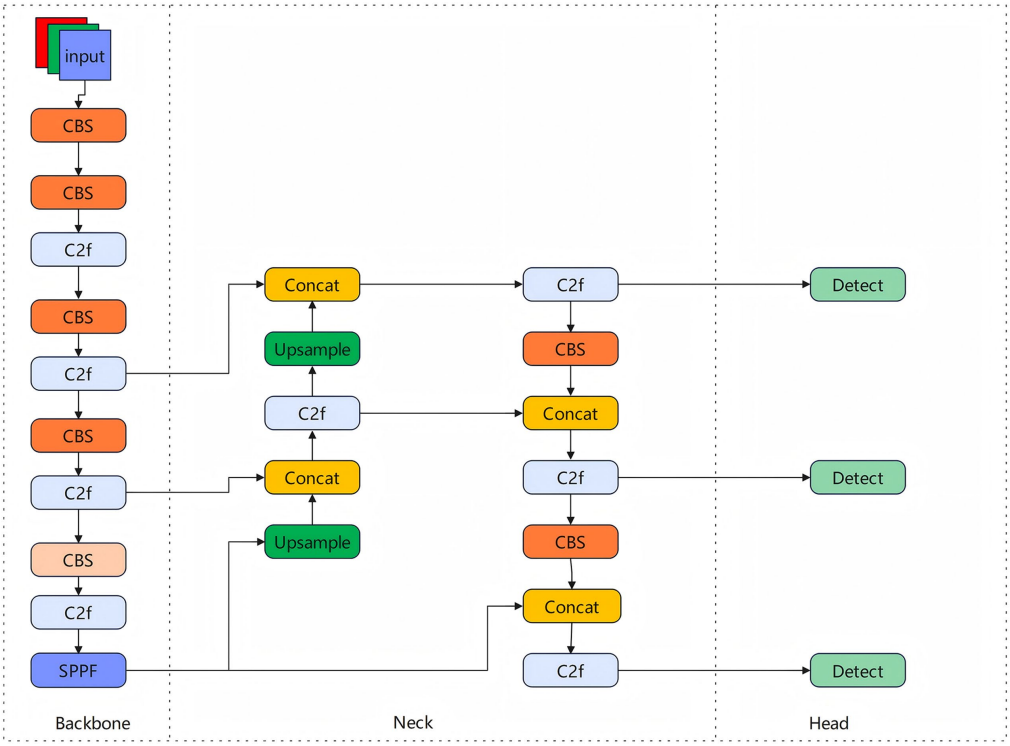


Figure 1. Architecture of YOLOv8n network
图 1. YOLOv8n 网络结构图

2.1.1. 骨干网络

YOLOv8 使用修改后的 CSPDarknet [12]作为骨干网络，包括 CBS 模块、SPPF 模块和 C2f 模块。CBS

模块由 3 × 3 卷积层、BN 归一化层和 SiLU 激活函数组成，有助于有效地提取特征并保留原始信息，降低梯度消失的风险。SiLU 激活函数有助于增强模型的非线性表达能力。C2f 模块的灵感来自 C3 模块和 ELAN 思想，采用梯度分流连接，在保持轻量级的同时丰富了特征提取网络的信息流。SPPF 模块旨在通过多个最大池化操作融合从不同尺度提取的高级语义特征来提高分类准确性。

2.1.2. 颈部

受 PANet [13]的启发，YOLOv8 在颈部设计了 PAN-FPN 结构。与 YOLOv5 和 YOLOv7 模型的颈部结构相比，YOLOv8 移除了 PAN 结构中上采样后的卷积操作，在保持原有性能的同时实现了模型的轻量化。传统的 FPN 采用自上而下的方式传递深层语义信息，但会丢失一些目标定位信息。为了缓解这一问题，PAN-FPN 在 FPN 的基础上增加了 PAN。PAN 通过自上而下的方式增强位置信息，从而实现路径增强。PAN-FPN 构建了一个自上而下和自下而上的网络结构，通过特征融合实现了浅层位置信息和深层语义信息的互补，从而实现了特征的多样性和完整性。

2.1.3. 头部

YOLOv8 的检测部分采用了解耦头结构。解耦头结构使用两个独立的分支分别进行目标分类和预测边界框回归，并且针对这两类任务采用了不同的损失函数。对于分类任务，使用二元交叉熵损失(BCE Loss)；对于预测边界框回归任务，则采用了分布焦点损失(DFL) [14]和 CLoU [15]。这种检测结构能够提高检测速度并加速模型收敛。YOLOv8 是一种无锚框的检测模型，简洁地定义了正负样本。它使用了任务对齐分配器(Task-Aligned Assigner) [16]来动态分配样本，从而提高了模型的检测精度和鲁棒性。

2.2. 改进后的 YOLOv8n-CSS 网络

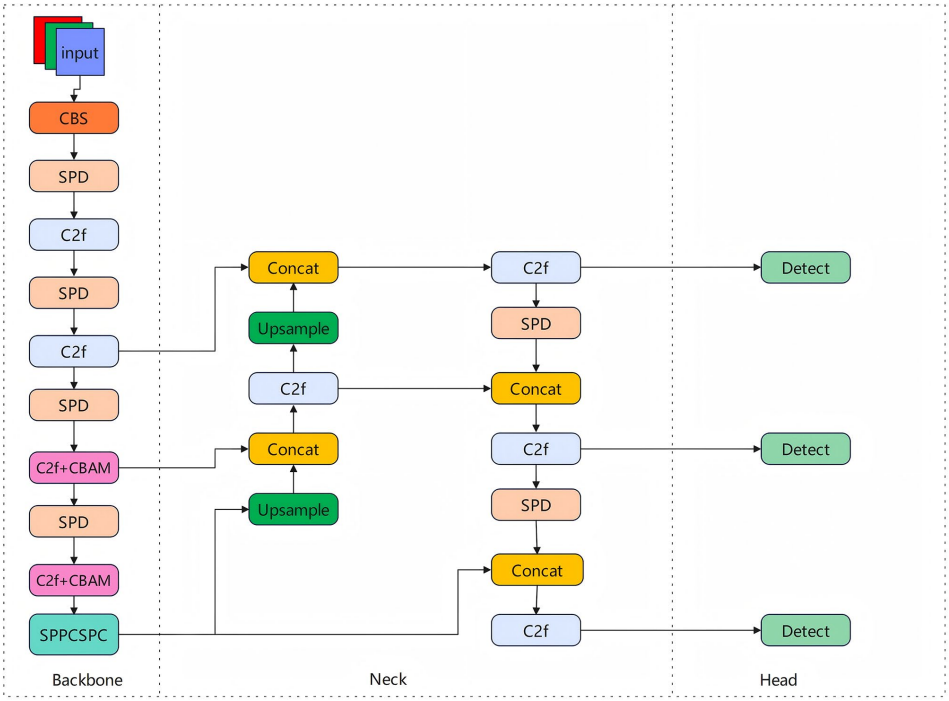


Figure 2. Architecture of YOLOv8n-CSS network
图 2. YOLOv8n-CSS 网络结构图

随着智能交通技术的持续发展，道路交通目标检测在实际应用中的重要性愈发凸显。然而，现有道

路交通目标检测算法在复杂场景下仍存在显著缺陷。实际交通场景不仅具有高度的复杂性，包含多样化的环境因素与动态变化，而且检测目标普遍呈现出尺寸小、分布密集的特点。这些因素致使现有模型检测性能欠佳，漏检、误检问题频发，难以切实满足实际应用的严格要求。因此，本文选择 YOLOv8n 算法作为基准改进，改进后的网络模型如图 2 所示。

2.2.1. CBAM 模块

复杂多变的交通场景为目标检测任务带来了诸多挑战。在实际应用中，检测算法需要应对光照变化、目标遮挡以及背景干扰等多重因素干扰，这对模型的鲁棒性提出了更高的要求。为提升网络对目标特征的提取能力，本文在特征提取模块引入了 CBAM 注意力机制[17]。该机制能够有效强化目标的关键特征表达，同时抑制无关背景因素信息的干扰，从而提升模型对目标与背景的判别能力。通过这种方式，网络能够更好地捕捉图像中的空间结构信息和深层语义信息，显著提高了检测性能。

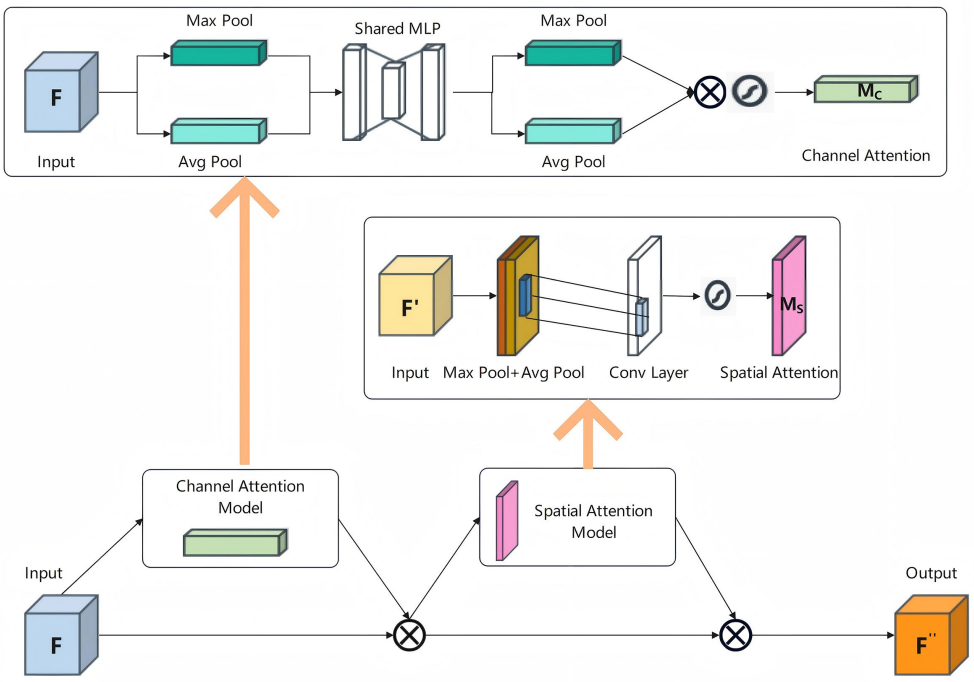


Figure 3. Architecture of CBAM network
图 3. CBAM 网络结构图

CBAM 由通道注意力模块(CAM) [18]和空间注意力模块(SAM) [19]组成，如图 3 所示。这种混合注意力结构巧妙地考虑了通道和空间两个维度，有助于网络获取目标区域的精确位置和详细信息。CAM 模块首先对输入特征图 F 进行全局最大池化和平均池化操作。随后，两种池化操作的结果通过共享的多层感知器(MLP)神经网络进行处理。MLP 输出的特征被合并，并通过 Sigmoid 函数的激活操作生成最终的特征图。这一过程有效地突出了特征图中的目标区域，增强了模型识别相关细节的能力。通道注意力的表达式如式(1)所示。

$$M_C(F)=\sigma\left(W_1\left(W_0\left(F_{avg}^C\right)\right)+W_1\left(W_0\left(F_{max}^C\right)\right)\right) \tag{1}$$

其中，σ 表示 Sigmoid 激活函数，W_0 和 W_1 分别对应多层感知器的权重($W_0 \in R^{C/r}$，$W_1 \in R^{C\times C/r}$)，参数 r 表示缩减比例。

随后，将 M_C 与输入特征图 F 相乘，得到 SAM 模块所需的输入特征 F'，公式如式(2)所示。

$$F' = M_C(F) \otimes F \tag{2}$$

基于 CAM 输出的特征图 F'，该过程首先对通道进行全局最大池化和平均池化操作。随后，通过 7 × 7 卷积操作生成空间特征，突出目标位置和详细信息。空间注意力的表达式如式(3)所示。

$$M_S(F) = \sigma\left(f^{7\times7}\left(\left[F^S_{avg}; F^S_{max}\right]\right)\right) \tag{3}$$

最终，将空间特征图 M_S 与输入特征图 F' 相乘，得到最终的特征图 F''，如式(4)所示。

$$F'' = M_S(F') \otimes F' \tag{4}$$

2.2.2. SPPCSPC 模块

在 YOLOv5 的初始版本中，SPP 模块被用于处理多尺度特征。该模块利用 13 × 13、9 × 9 以及 5 × 5 三种不同尺度的池化核，分别设置 6、4、2 的填充值，并以固定步长 1 对输入特征进行池化处理，最终生成统一维度的输出特征。这种多尺度池化机制有效解决了目标检测中的尺度变化问题，显著提升了检测精度。

SPPF 模块作为 SPP 的优化版本，其结构如图 4(a)所示。该模块将原有的多尺度池化窗口统一替换为三个 5 × 5 的池化核，这种设计基于级联池化的等效原理：两个 5 × 5 池化层的叠加效果等同于 9 × 9 池化窗口，而三个 5 × 5 池化层的串联则与 13 × 13 池化窗口等效。通过采用小尺寸池化核的级联策略，SPPF 在保证输出特征一致性的同时显著降低了计算复杂度。实验数据表明，在输入条件相同的情况下，SPPF 与 SPP 的输出结果完全一致，但前者的处理效率提升了 100%。

SPPF 是 YOLOv8n 中用于提取多尺度特征信息的重要组件，通过池化操作整合不同尺度的特征，提高模型对各种尺度目标的检测能力。但它主要依赖于池化操作，无法充分捕捉全局与局部信息之间的语义关系。因此，在某些复杂场景(如目标尺度变化较大或背景干扰较强)下，SPPF 模块的性能可能受到限制。这对这一问题，本文引入了 SPPCSPC 模块。该模块使用三种不同的最大池化操作(5 × 5，9 × 9，13 × 13)能够更好地融合来自不同层次和尺度的特征信息，从而提高模型识别物体的准确性，结构图如图 4(b)所示。

设输入特征图为 $F \in R^{C\times H\times W}$，其中 C 为通道数，H 和 W 分别为特征图的高度和宽度。首先，将输入的特征图 F 分为两个部分，如式(5)所示。

$$F_1, F_2 = Split(F) \tag{5}$$

其中，$F_1 \in R^{\frac{C}{2}\times H\times W}$ 直接传递，$F_2 \in R^{\frac{C}{2}\times H\times W}$ 经过卷积和池化操作。然后，对 F_2 进行多尺度池化操作，得到多尺度特征，如式(6)所示。

$$F_2^{SPP} = Concat\left(F_2^5, F_2^9, F_2^{13}\right) \tag{6}$$

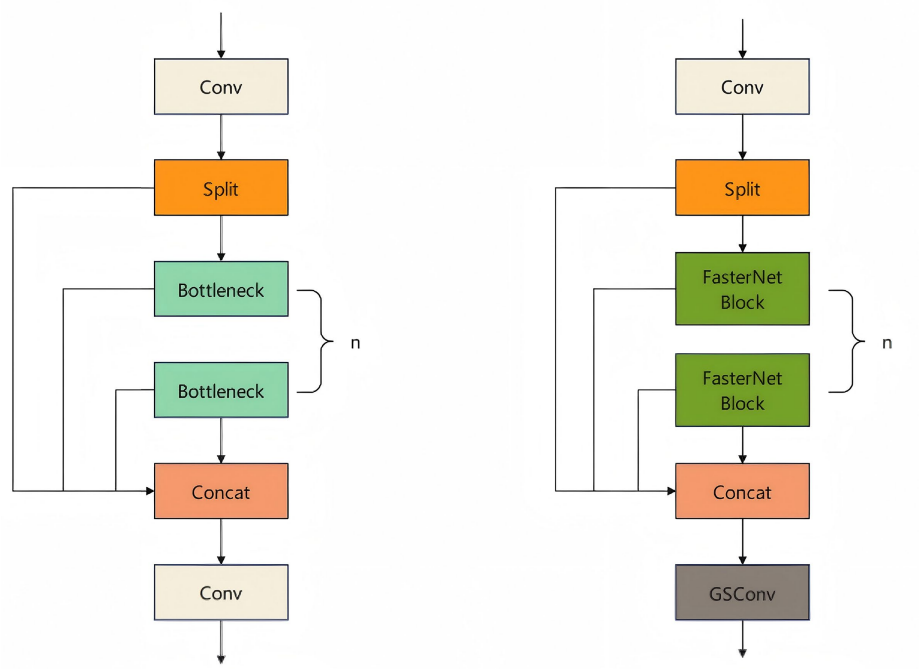
将 F_1 和多尺度池化后的 F_2^{SPP} 进行拼接，如式(7)所示。

$$F_{fused} = Concat\left(F_1, F_2^{SPP}\right) \tag{7}$$

最后，对融合的特征图进行卷积操作，以进一步提取特征，如式(8)所示。

$$F' = Conv\left(F_{fused}\right) \tag{8}$$

最终输出的特征图为 $F' \in R^{C\times H\times W}$，其通道数与输入特征图 F 相同。



（a）SPPF结构图　　（b）SPPCSPC结构图

Figure 4. Architecture of SPPF and SPPCSPC network
图 4. SPPF 和 SPPCSPC 网络结构图

2.2.3. SPD-Conv 模块

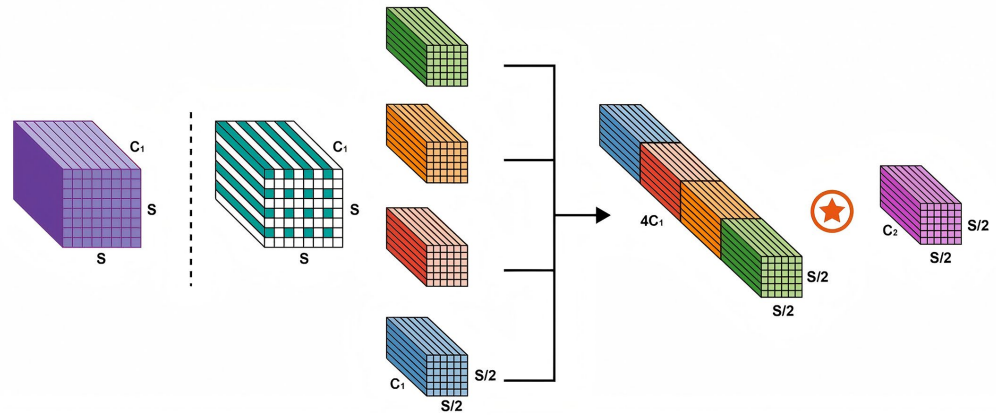


Figure 5. Architecture of SPD-Conv network
图 5. SPD-Conv 网络结构图

YOLOv8n 在进行目标检测过程中，小物体的检测精度往往低于正常大小的物体，这意味着模型学习的信息受限。此外，小尺寸物体经常与其他大型物体一起被发现，这使得模型的检测精度低，低检测精度是网络中的阶梯卷积层。虽然卷积层通过下采样扩展了接收场，但也会导致一些特征信息丢失。为了解决这一问题，本文引入了 SPD-Conv 模块[20]。它由两个部分组成：空间到深度(SPD)层和非交错卷积层，他们取代了骨干网中的交错卷积层。SPD 层对特征图进行了下采样，但保留了所有信息，使用非阶梯卷积层有助于更全面地捕捉图像的微小特征。结构图如图 5 所示。

对于初始特征图(S，S，C1)，通过以 2 的比例进行下采样，获得了四个子图，每个子图的大小(S/2，S/2，C1)与原始特征图相比减少了 2 倍。接下来，这些子特征图在信息通道维度 C1 处连接起来，形成一个新的特征图(S/2，S/2，4C1)。在通过 SPD 层后，引入非阶梯卷积层以获得最终通道维度 C2 (C2 = 4C1)的特征图(S/2，S/2，C2)，此过程消除了跨步的池化操作，同时保留了特征信息。

3. 实验及实验结果分析

3.1. 数据集介绍

本文使用 KITTI 数据集[21]和 BDD100K 数据集[22]。KITTI 数据集是自动驾驶和计算机视觉领域中著名的公开数据集之一，广泛用于目标检测、语义分割、光流分析、视觉距测等任务。该数据集由德国卡尔斯鲁厄理工学院(KITTI)和丰田工业大学芝加哥分校(TTIC)联合发布，涵盖了在不同场景中收集的图像数据，如城市、农村地区和高速公路。每幅图像都有丰富的字符、环境特征和不同程度的遮挡和截断。此数据集包含六类：car、cyclist、truck、van、pedetrain 和 tram。此外，数据集中的 7481 张图像按照 7：2：1 的比例分为训练集、验证集和测试集。BDD100K 数据集由加州伯克利大学伯克利分校 AI 实验室(BAIR)于 2018 年发布的，是目前规模最大，内容最具丰富性的公开驾驶数据集。包含 10 万张图片，涵盖美国各个地方 6 种不同的天气状况和 4 种不同时段。训练集，验证集，测试集比例为 7:2:1，共有 9 个类别。

3.2. 实验环境

实验使用 Window10 系统，基于 PyTorch 框架以 Python 作为开发语言，使用 8 核 3.2GHz AMD Ryzen 7 5800H 处理器进行模型训练和测试，并利用 NVIDIA GeForce RTX 3090 显卡加速模型运算。

3.3. 评价指标

为了评估该模型在道路目标检测中的检测性能，本文使用以下四个指标作为评估标准：精确度、召回率 mAP@0.5，以及 mAP@0.5:0.95。此外，为了确定是否满足实时要求，每秒帧数(FPS)也被用作本文的测量标准。具体公式见式(9)至式(12)所示。

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{9}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{10}$$

其中，TP 表示正确检测到的样本数量，FP 表示模型错误看到的阳性样本数量。FN 代表模型未检测到的阳性样本数量。精确度是指实际检测到的阳性样本比例，召回率是指正确看到的实际阳性样本的比例。

$$\text{AP} = \int_0^1 p(R)\,\mathrm{d}R \tag{11}$$

$$\text{mAP} = \frac{\sum_{i=1}^{M} \text{AP}_i}{M} \tag{12}$$

其中，$p(R)$表示精确度－召回率曲线。无线的 X 轴是精确的，Y 轴是可调用的。平均精度(AP)是通过计算精确度－召回率曲线下的面积获得的。AP 表示单个类别的性能。相比之下，平均精度(mAP)通过对各种 AP 值进行平均来衡量多种类型的性能，同时考虑到不同类别之间的性能差异。

3.4. 实验结果与分析

本文使用实验方法来控制变量，以验证所提出模型的检测精度，首先在 KITTI 数据集上进行实验，

各指标曲线如图 6 所示。在 KITTI 数据集上训练后，改进后的 YOLOv8n-CSS 模型所生成的 Precision-Recall (PR)曲线如图 7 所示。其中，x 轴代表召回率，y 轴代表准确率，曲线与坐标轴形成的面积越大，代表模型的精度越高。并将其与其他目标检测模型进行比较。

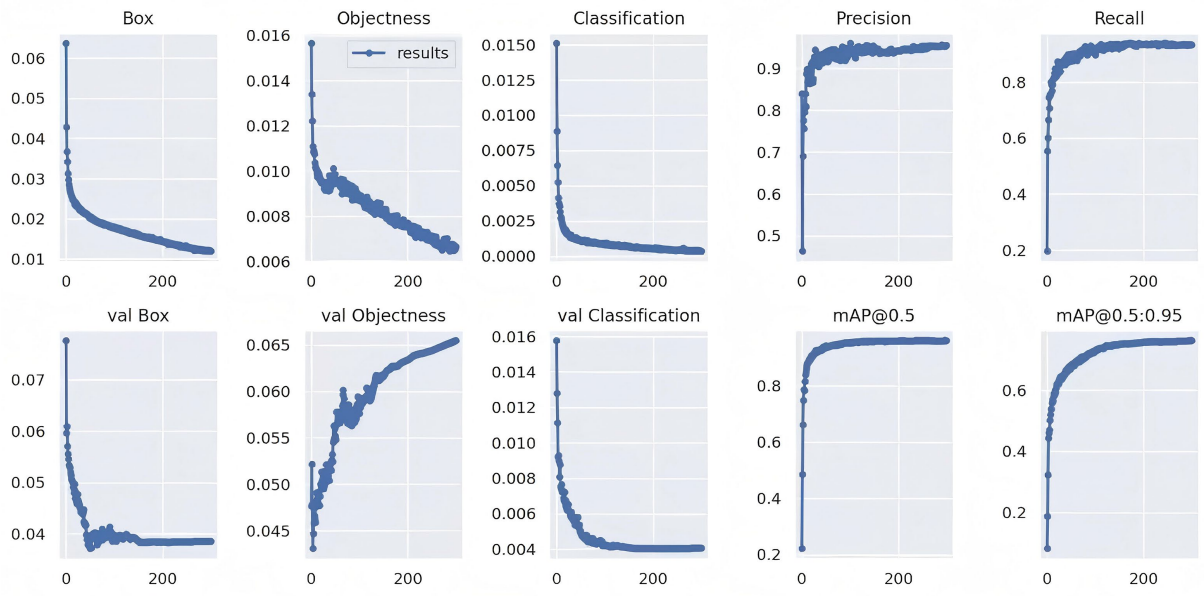


Figure 6. Performance curves of the YOLOv8n-CSS model on the KITTI dataset
图 6. YOLOv8n-CSS 模型在 KITTI 数据集上的各指标曲线

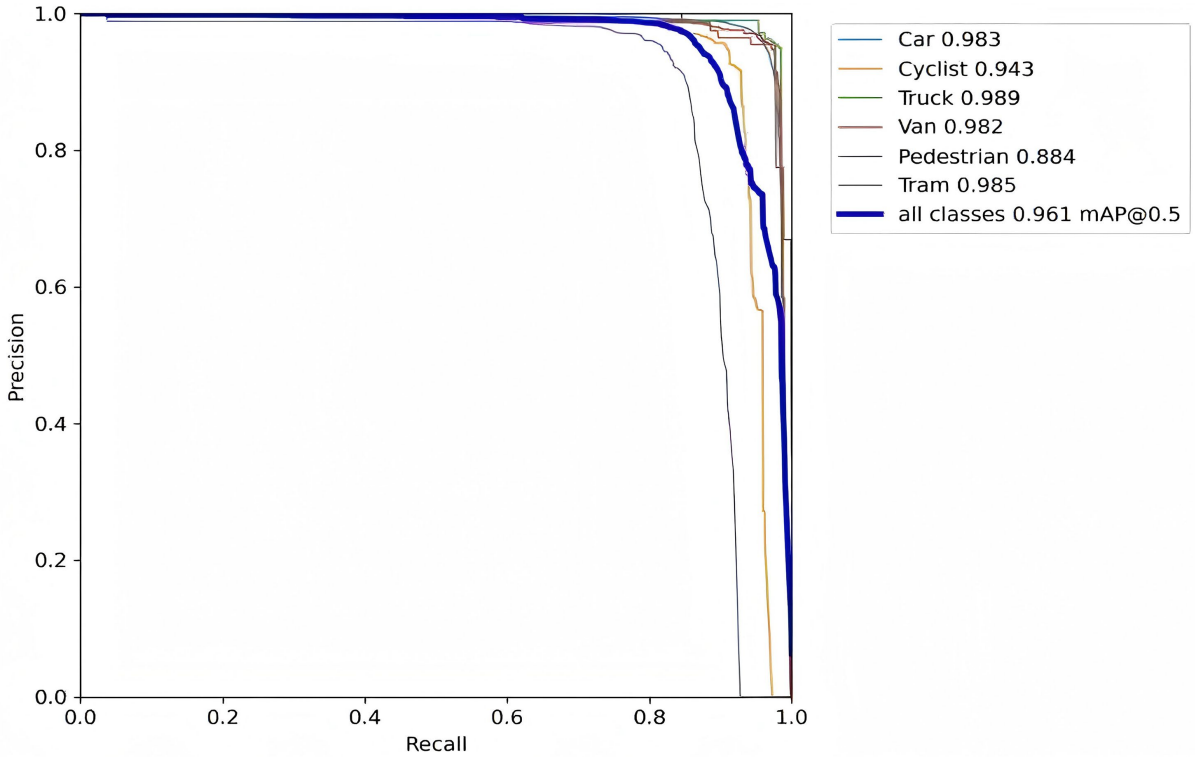


Figure 7. The Precision-Recall curve of YOLOv8n-CSS trained on the KITTI dataset
图 7. YOLOv8n-CSS 在 KITTI 数据集上训练的 PR 曲线

进一步验证本文提出的模型的性能和各模块的有效性。本文将模块单独添加到原始的 YOLOv8n 模型中进行消融实验，对实验结果进行评估和分析，参数配置与上述相同。在 KITTI 数据集和 BDD100K 数据集上分别进行消融实验，结果见表 1、表 2。

Table 1. Experimental results of the YOLOv8n-CSS on the KITTI dataset
表 1. YOLOv8n-CSS 在 KITTI 数据集上的实验结果

Group	CBAM	SPPCSPC	SPD-Conv	mAP@0.5%	mAP@0.5:0.95%
1				92.2	74.3
2	✓			94.6	75.9
3		✓		93.8	75.0
4			✓	93.5	74.9
5	✓	✓		95.8	76.1
6	✓		✓	95.6	75.9
7	✓	✓	✓	96.1	76.2

实验结果表明，在相同条件下，通过引入 CBAM 混合注意力机制、替换 SPPF 模块以及引入 SPD-Conv，检测模型的性能得到了显著改善，多项评价指标均有所提升。将各个改进的模块融入 YOLOv8n 后，与 YOLOv8n 原始算法相比，在 KITTI 数据集中，平均精度均值 mAP@0.5 提升了 3.9%，mAP@0.5:0.95% 提升了 1.9%。

Table 2. Experimental results of the YOLOv8n-CSS on the BDD100K dataset
表 2. YOLOv8n-CSS 在 BDD100K 数据集上的实验结果

Group	CBAM	SPPCSPC	SPD-Conv	mAP@0.5%	mAP@0.5:0.95%
1				40.1	19.8
2	✓			44.5	21.0
3		✓		43.1	20.4
4			✓	43.7	20.7
5	✓	✓		46.9	22.1
6	✓		✓	46.2	21.8
7	✓	✓	✓	48.0	22.3

在 BDD100K 数据集中，mAP@0.5 提升了 7.9%，mAP@0.5:0.95%提升了 2.5%。由此可见，本文提出的改进算法能够助力网络更好地学习道路目标特征。

为验证本文所提出改进算法的有效性，本文选取 BDD100K 和 KITTI 这两个广泛使用的公开数据集开展实验。在实验过程中，会将改进算法的检测结果与当前主流算法模型的相应结果进行全面的比较与深入分析，以此探究改进算法的性能表现。

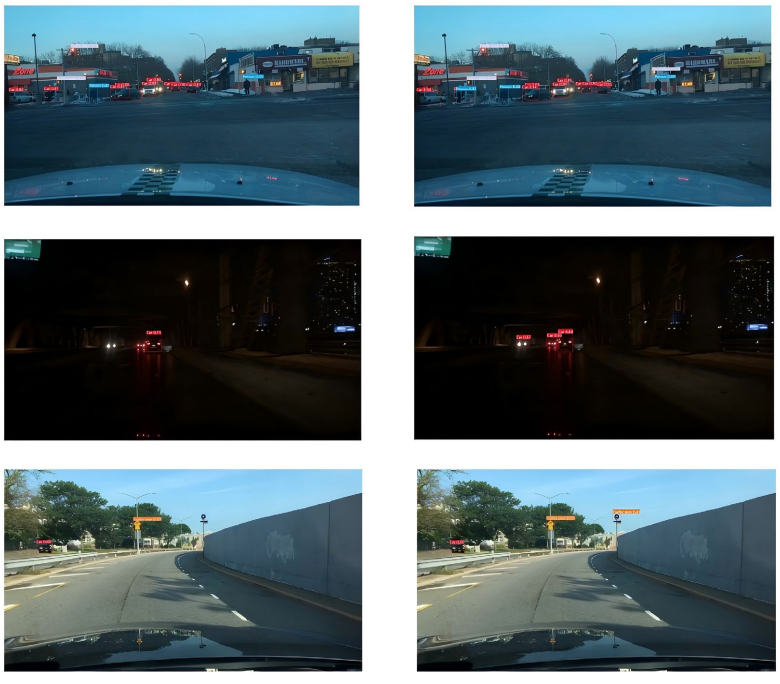
本文改进的模型 YOLOv8n-CSS 与 Faster R-CNN、SSD、YOLOv5s、YOLOv7-tiny、Z-YOLOv8s 和 SES-YOLOv8n 模型在 BDD100K 数据集上进行对比实验，实验结果见表 3。实验结果表明，本文提出的改进模型在检测效果方面表现最为优异。其平均精度均值 mAP@0.5 达到了 48.0%，mAP@0.5:0.95 达到了 22.3%。与其他模型相比，优势显著。

Table 3. Performance comparison results of different algorithms on the BDD100K
表 3. 不同算法在 BDD100K 上的性能对比结果

算法	输入尺寸	mAP@0.5%	mAP@0.5:0.95%
SSD	640 × 640	46.1	—
Faster R-CNN	640 × 640	41	19.4
YOLOv5s	640 × 640	32.2	—
YOLOv7-tiny	640 × 640	44.8	39.9
Z-YOLOv8S	640 × 640	39.5	—
SES-YOLOv8n	640 × 640	41.9	22.3
YOLOv8n-CSS	640 × 640	48.0	22.3

Table 4. Performance comparison results of different algorithms on the KITTI
表 4. 不同算法在 KITTI 上的性能对比结果

算法	输入尺寸	mAP@0.5%	mAP@0.5:0.95%
SSD	640 × 640	68.2	37.8
Faster R-CNN	640 × 640	87.5	58.1
YOLOv5s	640 × 640	91.3	62.7
YOLOv7-tiny	640 × 640	87.5	56.4
BFIDet	640 × 640	93.9	—
SES-YOLOv8n	640 × 640	92.7	69.2
YOLOv8n-CSS	640 × 640	86.1	76.2



(a) (b)

Figure 8. Visualization of results
图 8. 结果可视化图

接下来，本文在 KITTI 数据集上开展对比实验。实验结果汇总于表 4。从结果可以看出，与其他主流算法相较而言，本文所提出的改进算法展现出显著优势。在所有参与对比的算法中，本文改进算法的平均精度均值 mAP@0.5%以及 mAP@0.5:0.95%均达到最优水平，且相较于其他算法存在明显领先。由此可见，本文的改进算法性能更为卓越，在一定程度上缓解了小目标漏检与误检的问题。

为了更有效地验证本章提出的基于 YOLOv8n-CSS 的道理小目标检测算法的性能，本文从 BDD100K 数据集的测试集中随机抽取 3 张不同交通场景的图片，结果如图 8 所示。图中，最左侧的为 YOLOv8n 原算法的检测结果，最右侧则为本文改进算法的输出。通过对几组对比图的细致观察，可清晰发现，相较于 YOLOv8n 原算法，本文改进算法的检测精度在整体上有所提升。以最后一幅图为例，YOLOv8n 原算法未能识别出右侧的交通标志，而本文改进算法成功检测出，这一实验结果表明，本文所改进的算法有效缓解了小目标的误检与漏检问题。综上所述，基于 YOLOv8n 网络所改进的交通目标检测算法 YOLOv8n-CSS，在实际应用中展现出更为出色的性能表现。

4. 结论

本文提出了一种基于 YOLOv8n 的改进目标检测算法 YOLOv8n-CSS。解决了传统算法在目标检测方面的挑战，实现了更精确、更稳定的检测。为了确保实时性能，网络中引入了 CBAM 混合注意力机制和 SPD-Conv 模块，SPPCSPC 模块取代了 SPPF 模块。这些模块增强了对关键特征的感知和不同尺度特征图的融合能力，提高了目标检测的准确性。在 KITTI 和 BDD100K 数据集上的实验表明，YOLOv8-CSS 在检测不同尺度物体时的准确性明显优于其他模型。我们提出的方法在自动驾驶的检测任务中表现良好。

参考文献

[1] Wang, F., Wang, P., Zhang, X., Li, H. and Himed, B. (2021) An Overview of Parametric Modeling and Methods for Radar Target Detection with Limited Data. *IEEE Access*, **9**, 60459-60469. https://doi.org/10.1109/access.2021.3074063

[2] Zhang, Y., Zhang, W. and Bi, J. (2017) Recent Advances in Driverless Car. *Recent Patents on Mechanical Engineering*, **10**, 30-38. https://doi.org/10.2174/2212797610666170215144809

[3] Zhao, R., Tang, S.H., Bin Supeni, E.E., Rahim, S.A. and Fan, L. (2024) Z-YOLOv8s-Based Approach for Road Object Recognition in Complex Traffic Scenarios. *Alexandria Engineering Journal*, **106**, 298-311. https://doi.org/10.1016/j.aej.2024.07.011

[4] Girshick, R., Donahue, J., Darrell, T. and Malik, J. (2016) Region-Based Convolutional Networks for Accurate Object Detection and Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **38**, 142-158. https://doi.org/10.1109/tpami.2015.2437384

[5] He, D., Qiu, Y., Miao, J., Zou, Z., Li, K., Ren, C., *et al.* (2022) Improved Mask R-CNN for Obstacle Detection of Rail Transit. *Measurement*, **190**, Article ID: 110728. https://doi.org/10.1016/j.measurement.2022.110728

[6] Qiu, Z., Bai, H. and Chen, T. (2023) Special Vehicle Detection from UAV Perspective via YOLO-GNS Based Deep Learning Network. *Drones*, **7**, Article 117. https://doi.org/10.3390/drones7020117

[7] Soylu, E. and Soylu, T. (2023) A Performance Comparison of YOLOv8 Models for Traffic Sign Detection in the Robotaxi-Full Scale Autonomous Vehicle Competition. *Multimedia Tools and Applications*, **83**, 25005-25035. https://doi.org/10.1007/s11042-023-16451-1

[8] Redmon, J. and Farhadi, A. (2018) YOLOv3: An Incremental Improvement. arXiv: 1804.02767.

[9] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C., *et al.* (2016) SSD: Single Shot Multibox Detector. In: Leibe, B., Matas, J., Sebe, N. and Welling, M., Eds., *Computer Vision—ECCV* 2016, Springer, 21-37. https://doi.org/10.1007/978-3-319-46448-0_2

[10] Lin, T., Goyal, P., Girshick, R., He, K. and Dollar, P. (2017) Focal Loss for Dense Object Detection. 2017 *IEEE International Conference on Computer Vision* (*ICCV*), Venice, 22-29 October 2017, 2999-3007. https://doi.org/10.1109/iccv.2017.324

[11] Wang, G., Chen, Y., An, P., Hong, H., Hu, J. and Huang, T. (2023) UAV-YOLOv8: A Small-Object-Detection Model Based on Improved Yolov8 for UAV Aerial Photography Scenarios. *Sensors*, **23**, Article 7190. https://doi.org/10.3390/s23167190

[12] Redmon, J. and Farhadi, A. (2018) YOLOv3: An Incremental Improvement. arXiv: 1804.02767

[13] Liu, S., Qi, L., Qin, H., Shi, J. and Jia, J. (2018) Path Aggregation Network for Instance Segmentation. 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, 18-23 June 2018, 8759-8768. https://doi.org/10.1109/cvpr.2018.00913

[14] Li, X., Wang, W., Wu, L., Chen, S., Hu, X., Li, J., Tang, J. and Yang, J. (2020) Generalized Focal Loss: Learning Qualified and Distributed Bounding Boxes for Dense Object Detection. *NeurIPS Proceedings*, **33**, 21002-21012.

[15] Zheng, Z., Wang, P., Liu, W., Li, J., Ye, R. and Ren, D. (2020) Distance-IoU Loss: Faster and Better Learning for Bounding Box Regression. *Proceedings of the AAAI Conference on Artificial Intelligence*, **34**, 12993-13000. https://doi.org/10.1609/aaai.v34i07.6999

[16] Feng, C., Zhong, Y., Gao, Y., Scott, M.R. and Huang, W. (2021. TOOD: Task-Aligned One-Stage Object Detection. 2021 *IEEE/CVF International Conference on Computer Vision* (*ICCV*), Montreal, 10-17 October 2021, 3490-3499. https://doi.org/10.1109/iccv48922.2021.00349

[17] Zhu, F., Cui, J., Zhu, B., Li, H. and Liu, Y. (2023) Semantic Segmentation of Urban Street Scene Images Based on Improved U-Net Network. *Optoelectronics Letters*, **19**, 179-185. https://doi.org/10.1007/s11801-023-2128-8

[18] Tejashwini, P., Thriveni, J. and Venugopal, K. (2023) A Novel SLCA-UNet Architecture for Automatic MRI Brain Tumor Segmentation. arXiv: 2307.08048.

[19] Wang, C., He, W., Nie, Y., Guo, J., Liu, C., Han, K. and Wang, Y. (2023) Gold-YOLO: Efficient Object Detector via Gather-and-Distribute Mechanism. *Neural Computing and Applications*, **36**, 1-15.

[20] Sunkara, R. and Luo, T. (2023) No More Strided Convolutions or Pooling: A New CNN Building Block for Low-Resolution Images and Small Objects. In: Amini, M.R., Canu, S., Fischer, A., Guns, T., Kralj Novak, P. and Tsoumakas, G., Eds., *Machine Learning and Knowledge Discovery in Databases*, Springer, 443-459. https://doi.org/10.1007/978-3-031-26409-2_27

[21] Geiger, A., Lenz, P., Stiller, C. and Urtasun, R. (2013) Vision Meets Robotics: The KITTI Dataset. *The International Journal of Robotics Research*, **32**, 1231-1237. https://doi.org/10.1177/0278364913491297

[22] Yu, F., Chen, H., Wang, X., Xian, W., Chen, Y., Liu, F., *et al.* (2020) BDD100K: A Diverse Driving Dataset for Heterogeneous Multitask Learning. 2020 *IEEE/CVF Conference on Computer Vision and Pattern Recognition* (*CVPR*), Seattle, 13-19 June 2020, 2633-2642. https://doi.org/10.1109/cvpr42600.2020.00271