

DOI: 10.13382/j.jemi.B2306854

# 融合停车线方向的端到端停车位检测网络\*

汪家州 桑海峰

(沈阳工业大学信息科学与工程学院 沈阳 110870)

**摘要:**智能汽车自动泊车功能的基本要求之一是能够快速准确地检测出未被占用的停车位。针对这一问题,设计了一种融合停车线方向与全局特征的端到端可训练的停车位检测网络。首先提取出车位关键点的坐标以及关键点对应停车线的方向,并从图像全局特征中使用关键点的坐标提取出局部特征。将关键点信息、局部特征、全局特征使用交叉注意力机制融合,利用入口线鉴别器推断出关键点的停车位构成关系。根据关键点的停车位构成关系以及停车线方向,裁剪出停车位的区域图像并送入定制的车位占用分类网络进行分类,得到车位的占用信息。本文提出的方法在公共基准数据集 PS2.0 上进行了实验,其中该方法对矩形停车位的检测精度为 99.65%,对倾斜停车位的检测精度为 99.04%,在单块 GPU 上能够达到 80 fps 的检测速率。经验证,所提出的方法可以实时高精度的检测停车位位置、方向以及占用情况。

**关键词:**停车位检测;停车线方向;端到端;关键点;交叉注意力;车位占用

**中图分类号:** TP391;TN98 **文献标识码:** A **国家标准学科分类代码:** 520.20

## End to end parking slot detection network integrated with parking line direction

Wang Jiazhou Sang Haifeng

(School of Information Science and Engineering, Shenyang University of Technology, Shenyang 110870, China)

**Abstract:** One of the basic requirements for automatic parking of smart cars is to quickly and accurately detect unoccupied parking slots. To address this issue, an end-to-end train detection network that integrates the direction of the parking line with global features was designed. First, the coordinates of key parking spots and the direction of the corresponding parking line are extracted, and local features are extracted from the global features using the coordinates of the key spots. Integrate key point information, local features, and global features using the cross-attention mechanism, and use the entrance line discriminator to infer the composition relationship of key points' parking slots. Based on the composition relationship of key points and the direction of the parking line, the regional image of the parking slot is cropped and sent to a customized parking slot occupancy classification network for classification, resulting in the occupancy information of the parking slot. The proposed method was tested on the public benchmark dataset PS2.0, where the detection accuracy of the method for rectangular parking slots was 99.65%, and for tilted parking slots was 99.04%. The detection rate of 80 frames per second was achieved on a single GPU. It has been verified that the proposed method can detect the location, direction, and occupancy of parking slots in real time with high accuracy.

**Keywords:** parking slot detection; direction of parking line; end-to-end; key point; cross attention; parking slot occupancy

## 0 引言

目前,越来越多的车辆配备了环视监视器(around view monito,AVM)<sup>[1]</sup>系统,以便更好地观察周围的道路

状况。基于AVM的泊车辅助系统(parking assist system,PAS)已经成为汽车的一个重要功能。作为PAS的核心算法,停车位检测任务需要兼顾速度与精度,对停车位检测算法的实时性、准确性要求较高。

停车位检测方法主要分为基于激光雷达的方法<sup>[2]</sup>与

收稿日期:2023-08-29 Received Date: 2023-08-29

\* 基金项目:国家自然科学基金(62173078)、辽宁省自然科学基金(2022-MS-268)项目资助

纯视觉方案。基于视觉的停车位检测算法主要可以分为传统的车位检测算法<sup>[3-4]</sup>和基于深度学习的车位检测算法<sup>[5-9]</sup>。传统的车位检测算法对环境变化敏感,不适用于复杂的现实环境。近年来,基于深度卷积神经网络(convolutional neural networks, CNN)<sup>[10]</sup>的停车位检测方法取得了良好的准确性。基于深度学习的方法多采用二阶段处理,例如 Deep-PS<sup>[5]</sup>与 DMPR-PS<sup>[6]</sup>。Deep-PS 首先通过卷积神经网络检测关键点(停车位靠近当前车辆一侧的顶点),人工设计了一套规则判断停车位的关键点互相之间是否构成车位入口线。对构成的停车位入口线区域图像进行分类得出当前车位的类型。这种方法只对入口线附近的区域进行检测,忽略了全局的图像信息。DMPR-PS 将关键点分成了两类,T型和L型,并规定了关键点的方向。人工设计了一套关键点的匹配规则用来判断停车位的关键点互相之间是否构成车位入口线,这种人工设计后处理规则的方法对不同的数据集需要重新设计规则,对环境较为敏感。

基于图注意力网络的停车位检测算法 Gcn-PS<sup>[7]</sup>有效地解决了在检测到标记点后还需要人工设计后处理规则的难题。在检测到关键点后,率先使用全连接图网络(graph neural network, GNN)<sup>[11]</sup>来模拟关键点之间的关系,并通过入口线鉴别器来判断出关键点之间的车位构成关系,实现了端到端可训练<sup>[8]</sup>的停车位检测方法。然而这种方法在推断车位构成关系时,只考虑了关键点的局部信息,忽略了全局特征信息。这种方法只能检测矩形停车位,不能检测倾斜车位以及像 VPS-Net<sup>[9]</sup>一样检测车位占用,检测速度较慢。

目前的停车位检测方法提供了较好的结果,但它们仍有 3 个主要缺点。1) 这些方法在检测关键点时,只对关键点位置进行检测,忽略了与关键点连接的停车线所指示的大量信息,因此对倾斜停车位检测效果不佳。2) 这些方法无法兼顾实时性与准确性。3) 这些方法未对停车位占用进行分类,无法得到准确的空置停车位信息。

整体算法使用轻量级特征提取网络,提高了停车位检测的速度。为了能够检测矩形与倾斜的停车位,本文算法检测关键点位置坐标以及与关键点连接的停车线的方向信息。使用在多种任务中表现良好的交叉注意力机制<sup>[12-13]</sup>让模型去学习关键点的局部特征与全局特征之间的联系,从而直接推断出停车位的位置与方向。最后裁剪出停车位并使用停车位分类网络区分出了停车位是否被占用,最终实现端到端的输出未被占用的停车位信息。

论文的贡献如下:

1) 提出了一种端到端的可训练停车位检测方法,可以同时检测停车位位置,停车位方向以及停车位是否被占用。

2) 使用交叉注意力机制融合了关键点的局部特征与

全局特征,学习关键点互相之间的相邻信息,提升了车位构成的判别准确率。

3) 所提出的方法同时检测停车位关键点的位置信息以及关键点所连接的停车线方向信息,因此无需判断关键点类型,可以直接根据停车线方向信息检测矩形与倾斜停车位,降低了因为关键点分类所导致的车位推断错误率。

## 1 融合停车线方向的端到端停车位检测网络

### 1.1 整体网络结构

沿用了 Gcn-PS 对入口线的定义,每个车位由逆时针方向的 4 个点  $P_1, P_2, P_3, P_4$  构成,如图 1 所示。每个车位入口线定义为由靠近车辆的两个关键点  $P_1, P_2$  构成。由于  $P_3$  与  $P_4$  点经常由于车辆遮挡或者镜头畸变的原因无法观测,所以在靠近车辆的  $P_1, P_2$  点的附近额外标注了  $P_{11}$  点与  $P_{22}$  点用来计算与  $P_1, P_2$  点所连接的停车线的方向。

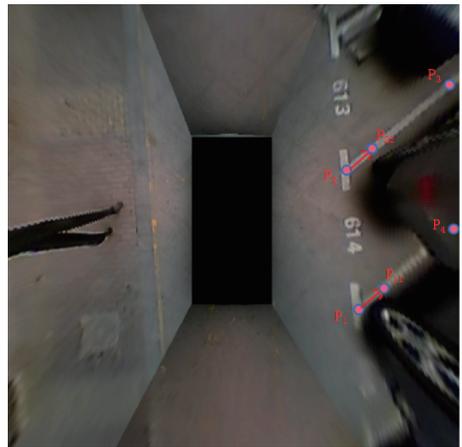


图 1 关键点定义图

Fig. 1 Key point definition

本文首先使用初始特征提取网络提取图片的原始全局特征,然后由关键点检测网络检测关键点位置信息以及停车线的方向信息。全局特征提取网络将初始图像特征进一步提取成较小的全局特征。通过关键点位置从全局特征中提取出与关键点位置对应的局部特征,然后使用图像特征编码器融合位置信息、局部特征、全局特征,使用交叉注意力机制<sup>[14]</sup>让局部特征与全局特征互相学习<sup>[15]</sup>,调整关键点所对应的特征,最终得到融合后的关键点特征向量。关键点两两匹配的编码判别特征由入口线鉴别器网络处理,以判断关键点互相之间是否可以构成停车位入口线。在得到车位构成信息后,结合停车线的

方向信息裁剪停车位对应区域的图像,送入到车位占用检测网络去检测当前停车位是否可用。

网络整体结构图如图 2 所示。

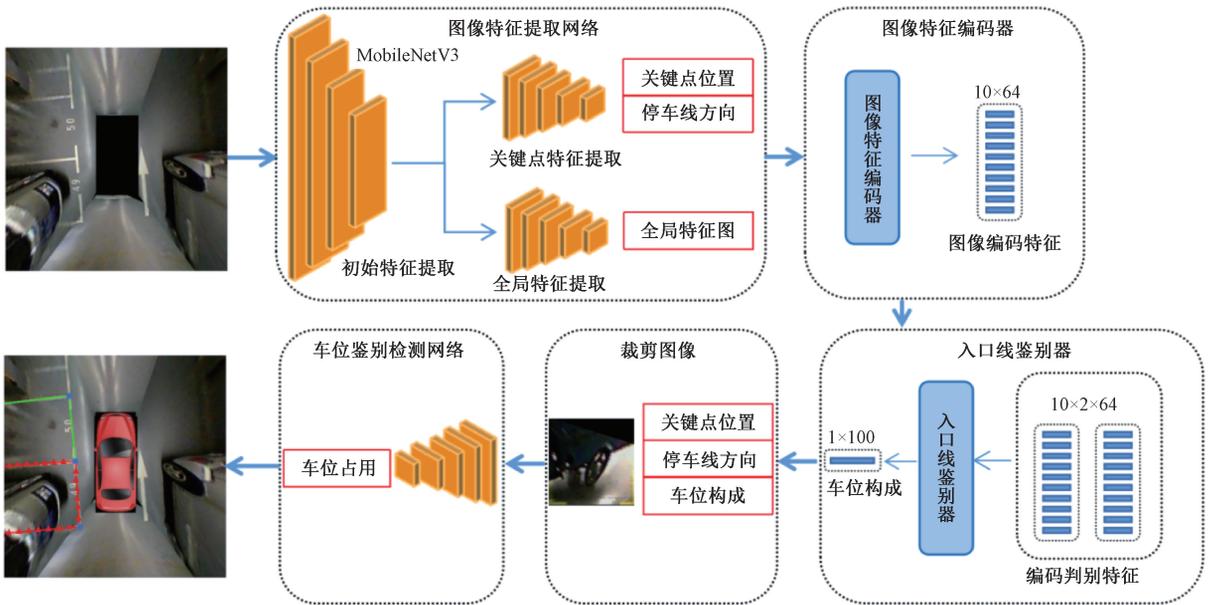


图 2 融合停车线方向的端到端停车位检测网络整体结构

Fig. 2 Overall structure of end-to-end parking slot detection network integrating parking line direction

## 1.2 图像特征提取网络

停车位检测任务需要兼顾准确性与实时性,因此轻量化停车位检测网络是十分必要的。在提取初始图像特征的部分,采用 MobileNetV3<sup>[16-17]</sup>来提取图像的初始特征。首先将一副大小为  $3 \times 512 \times 512$  的图像按照  $32 \times 32$  的大小,分成了  $S \times S$  (遵循 DMPS-PS,  $S=16$ ) 个图块,最终通过 MobileNetV3 网络提取出  $1024 \times S \times S$  的全局初始图像特征。初始特征提取网络提取的图像特征被送入关键点特征提取网络和全局特征提取网络。

关键点特征提取网络由 4 个卷积层组成,输出维度为  $5 \times S \times S$  的特征图。其中 5 个通道分别是该方格内存在关键点的置信度  $c_{\text{point}}$ 、关键点相对于方格的偏移量  $x$ 、关键点相对于方格的偏移量  $y$ 、关键点对应停车线方向的  $\cos$  值  $c$ 、关键点对应停车线方向的  $\sin$  值  $s$ 。在对置信度通道应用非极大值抑制 (non-maximum suppression, NMS) 后检测到  $N$  个关键点 (在具体实现中,  $N=10$ ), 所以最终能够得到  $10 \times 4$  的关键点位置信息与停车线方向信息  $Kp_i = (x_i, y_i, c_i, s_i), i = (1, 2, \dots, 10)$ 。在预测方向信息时,本文使用  $\tanh$  作为激活函数,从而使得回归出的方向信息的取值范围在  $[-1, 1]$  之间。全局特征提取网络也由 4 个卷积层组成,输出大小为  $128 \times S \times S$  的全局特征图  $Fg$ 。在得到关键点位置信息与停车线方向信息后,将其与全局图像特征一起送入图像特征编码器去进行编码融合。

## 1.3 图像特征编码器

利用关键点特征提取网络检测到的  $N$  个关键点的信息  $Kp_i = (x_i, y_i, c_i, s_i), i = (1, 2, \dots, 10)$ , 利用双线性插值  $BI$  (bilinear interpolation), 从全局特征图中逐点计算单点局部特征。如式(1)所示,根据关键点坐标,利用双线性插值从全局特征得到大小为  $N \times 128$  的关键点局部特征。

$$f_i = BI(Fg, (x_i, y_i)) \quad (1)$$

式中:  $Fg$  为全局特征,  $(x_i, y_i)$  为关键点坐标,  $f_i$  为关键点所对应的单点局部特征,  $i = (1, 2, \dots, 10)$ 。为了增强特征表示,使用多层感知机 (multi-layer perceptron, MLP)<sup>[18]</sup> 将点  $P_i$  的位置信息与停车线方向信息嵌入到高维向量中,并通过逐元素相加将它们与原始关键点特征  $f_i$  融合得到关键点信息增强后的局部特征  $Fl_i$ 。

$$Fl_i = f_i + MLP(x_i, y_i, c_i, s_i) \quad (2)$$

式中:  $x_i$  为第  $i$  个关键点的  $x$  坐标,  $y_i$  为第  $i$  个关键点的  $y$  坐标,  $c_i$  为第  $i$  个关键点的停车线方向余弦值,  $s_i$  为第  $i$  个关键点的停车线方向正弦值。

在这里,为了进一步增强关键点相邻信息表示,全局特征与局部特征在经过一层交叉注意力机制<sup>[15]</sup>后,全局特征  $Fg$  与增强后的局部特征  $Fl$  会融合在一起,从而获得  $N \times 128$  的融合特征  $Fo$ 。如式(3)、(4)所示。

$$K = Fl, K = Fg, V = Fg \quad (3)$$

$$Fo = Attention(Q, K, V) = Softmax(Q \cdot K) \cdot V \quad (4)$$

式中:  $Q, K, V$  分别为查询向量、键向量和值向量。直接将局部特征作为查询向量,全局特征作为键向量,通过计

算这两个向量的乘积,并使用 *Softmax* 函数归一化后,作为交叉权重,再与作为值向量的全局特征  $F_g$  相乘,得到最终的融合特征  $F_o$ 。

图像特征编码器整体结构图如图 3 所示。

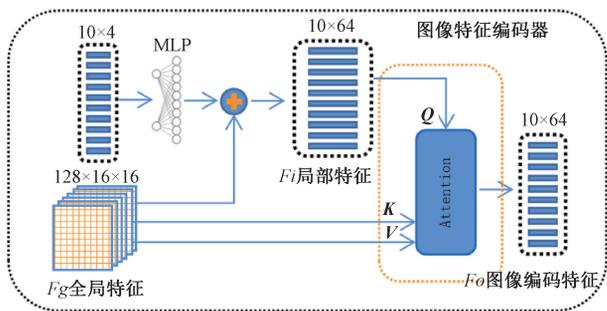


图 3 图像特征编码器结构

Fig. 3 Image feature encoder

### 1.4 入口线鉴别器

在得到融合特征之后,将 2 组相同的融合特征进行拼接,形成编码判别特征。这些编码判别特征会被送入一个由 3 层 *MLP* 网络构成的入口线鉴别器。在经过一层 *FC* 全连接层后,将得到一个  $1 \times 100$  的入口线匹配特征。这个特征表示了 10 个关键点之间两两匹配能够形成停车位的概率。

通过这个过程,得到了关键点能否形成停车位的信息,也就是车位构成信息。

入口线鉴别器结构如图 4 所示。

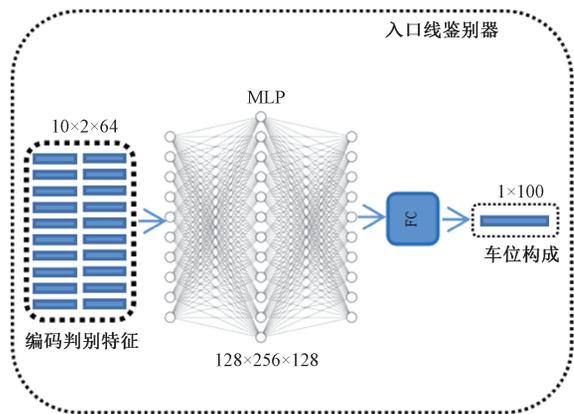


图 4 入口线鉴别器结构

Fig. 4 Entrance line discriminator

### 1.5 车位占用网络

得到车位构成信息后,根据  $P_1, P_2$  点坐标以及对应的停车线方向,可以推断出  $P_3, P_4$  点的坐标。在推断出完整的车位后,将其在图像中的位置裁剪下来,以区分是否空置。由于环视图像中的停车位大小不同,因此需要一种规则化的形式来最大化分类性能。如图 5 所示,本

文沿用 VPS-Net 的做法,将停车位根据其其在图像中的位置被切割并变形为  $128 \times 128$  pixels 的统一尺寸。

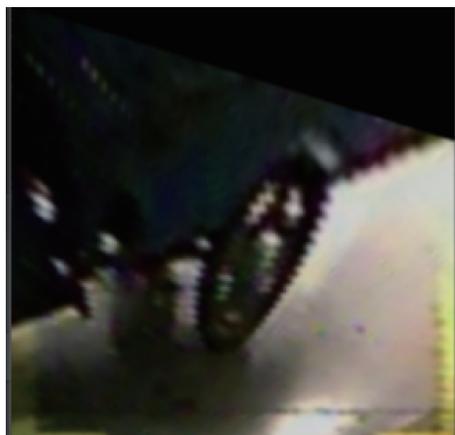


图 5 根据车位线方向裁剪的停车位

Fig. 5 Slot cut according to the direction of the parking line

将停车位按照空置、占用、当前车辆停放车位分成了 3 类,使用了定制的车位占用检测网络进行分类。定制的车位分类网络如表 1 所示。

表 1 车位占用检测网络

Table 1 Parking slot occupancy detection network

Layer Name	Output
Input	$3 \times 128 \times 128$
Conv1	$16 \times 128 \times 128$
Maxpool1	$16 \times 64 \times 64$
Conv2	$32 \times 64 \times 64$
Maxpool2	$32 \times 32 \times 32$
FC1	$1 \times 64$
FC2	$1 \times 3$

经过车位占用检测网络后,就得到了完整的停车位信息,包括停车位的位置与方向,停车位是否被占用。通过这些信息就可以将停车位的检测图像完整的显示出来。

### 1.6 损失函数

停车位检测算法是一个多任务算法,主要包括 3 个任务:关键点预测、车位入口线构成预测和车位占用分类预测。使用均方误差损失来衡量预测结果与真实关键点之间的差异。使用交叉熵损失来衡量预测结果与真实车位入口线构成之间的差异。使用交叉熵损失来衡量预测结果与真实车位占用分类之间的差异。

由于关键点特征提取网络的输出层被划分为  $S \times S$  个单元的网格,因此关键点预测的损失函数定义为  $S \times S$  个单元的预测与相应地面的误差平方和。

$$loss_{point} = \frac{1}{S^2} \sum_{i=1}^s \left\{ \begin{array}{l} (c_i - \hat{c}_i)^2 + \\ (x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2 + \\ (\cos_i - \widehat{\cos}_i)^2 + (\sin_i - \widehat{\sin}_i)^2 \end{array} \right. \quad (5)$$

其中,  $(x_i, y_i)$  表示关键点在网格内的预测偏移位置,  $(\hat{x}_i, \hat{y}_i)$  表示关键点在网格内的真实偏移位置。  $(\cos_i, \sin_i)$  表示与该关键点连接的停车线方向值,  $\widehat{\cos}_i, \widehat{\sin}_i$  表示根据  $P_{11}$  点与  $P_{22}$  点计算得到的真实方向值。  $c_i$  表示在当前方格内预测有关键点的概率,  $\hat{c}_i$  表示当前方格内存在关键点概率的真实值。

此外, 本文沿用 Gen-PS 对于关键点损失函数计算时所采取的掩膜处理<sup>[18]</sup>, 对于当前方格内不存在关键点的方格, 本文将  $(x_i, y_i)$  与  $(\hat{x}_i, \hat{y}_i)$  设置为 0, 减少环境因素对关键点检测的负面影响。

入口线鉴别器的损失函数定义如下:

$$loss_{edge} = \frac{1}{N^2} \left( - \sum_{i=1}^N \sum_{j=1}^N \hat{e}_{ij} \log e_{ij} \right) \quad (6)$$

其中,  $e_{ij}$  表示第  $i$  个和第  $j$  个关键点构成停车位入口线的预测概率。用  $\hat{e}$  表示的符号表示相应的真实值,  $N$  表示关键点的个数。

停车位占用的损失函数定义如下:

$$loss_{vacant} = \sum_{i=1}^M \left( - \sum_{j=1}^3 \hat{v}_{ij} \log v_{ij} \right) \quad (7)$$

其中,  $M$  为一张图片内检测到停车位的个数, 3 表示将停车位的占用情况分成了 3 类,  $\hat{v}_{ij}$  表示车位是否被占用的真实值,  $v_{ij}$  为模型的车位占用预测值。 本文将每一个车位的分类损失逐个相加得到单张图片的车位占用分类损失。

在得到 3 个任务的损失之后, 本文采用加权求和的方式计算总损失。总损失函数定义如下:

$$loss1 = \alpha_1 \cdot loss_{point} \quad (8)$$

$$loss2 = \alpha_2 \cdot loss_{edge} \quad (9)$$

$$loss3 = \alpha_3 \cdot loss_{vacant} \quad (10)$$

$$loss_{all} = loss1 + loss2 + loss3 \quad (11)$$

其中,  $\alpha_1, \alpha_2, \alpha_3$  为 3 个任务的平衡权重, 因为 3 个任务的难度不同, 所以较为困难的任务会设置较大的损失权重。  $loss_{point}$  为关键点的损失,  $loss_{edge}$  为车位入口线构成的损失,  $loss_{vacant}$  为车位占用分类的损失,  $loss_{all}$  为总损失。

## 2 实验结果与讨论

在本章中, 在 PS2.0 数据集上进行了一系列实验来验证所提出方法的性能, 并对实验得出的结果进行了

讨论。

### 2.1 实验

#### 1) 数据集

PS2.0 数据集由 9 827 张训练图像和 9 476 个停车位和 2 338 张测试图像和 2 168 个停车位组成。遵循 DMPR-PS, 选择 7 780 张图像作为训练集, 2 290 张图像作为测试图像。 PS2.0 数据集中的图像采集自各种环境条件下的典型室内外场景, 其分辨率为  $600 \times 600$  pixels, 对应一个  $10 \text{ m} \times 10 \text{ m}$  的物理平面区域。

本文对 PS2.0 数据集进行了停车线方向的标注, 如图 6 所示。其中车位空置标注为 0, 车位占用标注为 1, 当前车辆占用车位标注为 2。

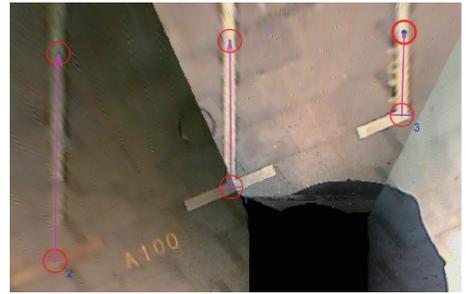


图 6 停车位方向标注

Fig. 6 Parking slot direction annotation

#### 2) 实验设置

将所提出的方法与以下最先进的停车位检测方法进行比较: 3 种基于深度学习的停车位检测方法: DeepPS、DMPR-PS 和 Gen-PS。精确率 *Precision* 为正确预测的正样本数量与预测为正样本的全部样本数量的比值, 召回率 *Recall* 为正确预测的正样本数量与所有正样本数量的比值:

$$Precision = TP / (TP + FP) \quad (12)$$

$$Recall = TP / (TP + FN) \quad (13)$$

式中:  $TP$  为预测为正的样本数量;  $FP$  为预测为正的负样本数量;  $FN$  为预测为负的正样本数量。

#### 3) 实施细节

使用 Pytorch 框架来完成具体实现, 并使用 Adam 优化器<sup>[20]</sup> 在 Nvidia RTX3090 上对模型进行了 200 个 epoch 的训练。初始学习率为 0.003, 批量大小设置为 16。损失函数的平衡权重参数  $\alpha_1, \alpha_2, \alpha_3$  分别设置为 100.0、1.0 和 1.0。选择 MobileNetV3 作为图像特征提取的主干网络。在训练时, 使用标注的关键点信息进行车位构成推断和裁剪对应的停车位区域, 从而减少关键点检测、停车位构成推断、停车位占用分类的训练时的互相影响。

### 2.2 结果与讨论

#### 1) 评估

表 2 显示了在 PS2.0 测试集上的量化结果。对比了目前较为主流的停车位检测方法,分别对比矩形停车位与倾斜停车位的检测精确率与召回率。由表 2 可知,所提出的方法在倾斜停车位的检测精度上则具有显著的优势,这表明基于端到端学习的方法可以学习更多的判别特征,对于复杂的视觉条件下可以有有效的检测出停车位。

表 2 PS2.0 数据集下的有效性对比结果

Table 2 Effectiveness comparison under PS2.0 dataset

(%)				
算法	矩形精确率	矩形召回率	倾斜精确率	倾斜召回率
DeepPS	98.99	99.13	96.15	92.59
DMPR-PS	99.42	99.37		
VPS-Net	99.63	99.31	90.12	90.12
Gcn-PS	99.56	99.42		
Ours	99.65	99.44	99.04	98.85

本文的方法在精度和召回率方面都达到了与最先进的基于学习的方法 Gcn-PS 相当的性能,值得注意的是, Gcn-PS 只能检测矩形停车位,并不能检测倾斜停车位以及检测车位占用情况,而 VPS-Net 不能进行端到端的训练,不方便实际应用。

图 7 为在 PS2.0 数据集上测试的效果图,在最终的推理阶段,采用不同的线型来区分车位的不同类别,空置停车位用光滑线条表示,已占用停车位使用带三角形图块的线条表示,当前车辆所占据的停车位使用带矩形图块的线条表示。

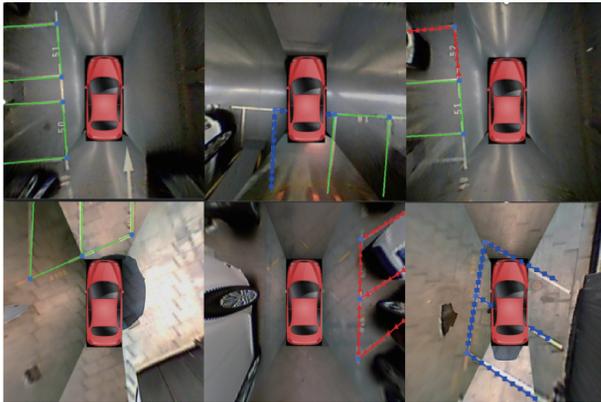


图 7 PS2.0 数据集测试效果

Fig. 7 PS2.0 dataset test results

此外本文还对模型的运行效率进行了测试,如表 3 所示,相比于优秀的车位检测算法 Gcn-PS,算法的模型大小为它的 56.6%,每帧图像的运行时间与其相比缩短 10 ms。需要注意的是,本算法为全功能、端到端的算法,能够同时检测倾斜停车位以及车位是否被占用,这是其他算法所不具备的。

表 3 PS2.0 数据集下的推理效率对比结果

Table 3 Inference efficiency under the PS2.0 dataset

算法	模型大小/MB	推理时间/ms	检测精度/%
DeepPS	255	10	98.99
DMPR-PS	121.3	9	99.42
VPS-Net	226	12	99.63
Gcn-PS	465	22.5	99.56
Ours	263	12.5	99.65

2) 消融实验

为了探究轻量特征提取网络与基于交叉注意力机制的图像特征编码器如何影响停车位检测的性能。以 Gcn-PS 所公开的代码为参考标准,使用 VGG-16<sup>[21]</sup> 作为特征提取网络,图注意力网络 GAT<sup>[22]</sup> 作为图形特征编码部分。表 4 显示的是依次添加轻量级特征提取网络 MobileNetV3 以及使用交叉注意力机制学习全局特征的情况。

表 4 不同模块消融实验结果

Table 4 Module ablation experiment

轻量特征提取网络	交叉注意力	精确率/%	模型大小/MB	速度/ms
×	×	99.56	465	22.5
√	×	99.54	289	15.3
√	√	99.65	263	12.5

从表 4 可以看出,采用了轻量特征提取网络后,模型的大小显著下降,运行速度得到明显提升。在采用交叉注意力机制后,模型的停车位检测精确率也得到了提升。

同时,为了验证局部特征  $F_l$  与全局特征  $F_g$  做交叉注意力机制时, $Q$ 、 $K$ 、 $V$  的取值不同对模型性能的影响,本文做了对比实验如表 5 所示。

表 5 交叉注意力取值影响对比结果

Table 5 Comparison of different cross attention values

$Q$	$K$	$V$	精确率/%	召回率/%
$F_l$	$F_g$	$F_l$	99.48	99.30
$F_l$	$F_g$	$F_g$	99.65	99.44

从表 5 可以看出,当局部特征作为  $Q$ ,全局特征作为  $K$  与  $V$  时,模型的停车位检测精确率与召回率均有提升。

在进行损失函数权重设置时,不同的损失函数权重系数设置对模型的最终性能具有影响。当损失函数选取不同的权重系数时,得到的停车位检测结果如表 6 所示。

表 6 不同的损失权重系数对比结果

Table 6 Comparison of different loss weight coefficients

$\alpha_1$	$\alpha_2$	$\alpha_3$	精确率/%	召回率/%
1	1	1	98.40	98.24
100	100	1	98.33	98.20
100	1	100	99.28	99.15
100	1	1	99.65	99.44

由表 6 可知,当关键点损失函数权重系数设置为 100,入口线损失函数权重系数设置为 1,车位占用分类损失函数权重系数设置为 1 时,模型性能能够达到最优。

本算法经过消融试验、对比实验后,通过对算法的不断优化,在实时性方面,每帧图像的处理时间为 12.5 ms,模型大小缩减至 263 MB。在有效性方面,本文所提出的算法达到了 99.65% 的精确率、99.44% 的召回率,在能够检测倾斜车位和车位是否被占用的同时,有效地平衡了算法的实时性和准确性。

### 3 结 论

本文提出了一种融合停车线方向的端到端可训练停车位检测网络。方法的核心贡献是融合了停车线的方向,并利用交叉注意力机制融合了图片的全局特征与关键点的局部特征,从而提高了关键点相邻信息特征的信息表达,提升了模型的准确率。PS2.0 数据集的实验结果表明,本文所提出的方法是一种端到端全功能的停车位检测算法,达到了与最先进的基于 CNN 的方法 Gcn-PS 相当的精度。并且本文所提出的方法相较于目前最先进的方法参数量更小,功能更多,更适合嵌入式设备部署。

考虑到在训练过程中,损失函数的权重系数设置对模型性能影响较大,下一步将对损失函数权重系数的动态设置做进一步的研究。

### 参考文献

- [ 1 ] LI W, CAO H, LIAO J C, et al. Parking slot detection on around-view images using DCNN [ J ]. *Front Neurobot*, 2020, 14: 46-46.
- [ 2 ] 冯明驰,高小倩,汪静姝. 基于立体视觉与激光雷达的车辆目标外形位置融合算法研究[ J ]. *仪器仪表学报*, 2021, 42(10): 210-220.  
FENG M CH, GAO X Q, WANG J SH. Research on vehicle target shape and position fusion algorithm based on stereoscopic vision and LiDAR[ J ]. *Chinese Journal of Scientific Instrument*, 2021, 42 ( 10 ): 210-220.
- [ 3 ] SUHR J, JUNG H. Sensor fusion-based vacant parking slot detection and tracking [ J ]. *IEEE Transactions on Intelligent Transportation Systems*, 2014, 15 ( 1 ): 21-36.
- [ 4 ] UNGER C, WAHL E, ILIC S. Parking assistance using dense motion-stereo [ J ]. *Machine Vision and Applications*, 2014, 25(3): 561-581.
- [ 5 ] LIN Z, HUANG J, LI X, et al. Vision-based parking-slot detection: A DCNN-based approach and a large-scale benchmark dataset [ J ]. *IEEE Transactions on Image Processing: A Publication of the IEEE Signal Processing Society*, 2018: 5350-5364.
- [ 6 ] HUANG J, ZHANG L, SHEN Y, et al. DMPR-PS: A novel approach for parking-slot detection using directional marking-point regression [ C ]. 2019 IEEE International Conference on Multimedia and Expo ( ICME ). IEEE, 2019: 212-217.
- [ 7 ] CHEN M, XU J L, XIAO L, et al. Attentional graph neural network for parking-slot detection [ J ]. *IEEE Robotics and Automation Letters*, 2021, 6 ( 2 ): 3445-3450.
- [ 8 ] SUHR J, JUNG H. End-to-end trainable one-stage parking slot detection integrating global and local information [ J ]. *IEEE Transactions on Intelligent Transportation Systems*, 2021, 23(5): 4570-4582.
- [ 9 ] LI W, CAO L, YAN L, et al. Vacant parking slot detection in the around view image based on deep learning[ J ]. *Sensors*, 2020, 20(7): 2138-2138.
- [ 10 ] ALEX K, ILYA S, GEOFFREY E, et al. Imagenet classification with deep convolutional neural networks[ J ]. *Commun ACM*, 2017, 60(6): 84-90.
- [ 11 ] 徐冰冰,岑科廷,黄俊杰,等. 图卷积神经网络综述[ J ]. *计算机学报*, 2020, 43(5): 755-780.  
XU B B, CEN K T, HUANG J J, et al. Overview of graph convolutional neural networks [ J ]. *Journal of Computer Science*, 2020, 43(5): 755-780.
- [ 12 ] 蒯晓,丁飞,张登银. 多分支融合注意力机制的车道线检测模型[ J ]. *电子测量与仪器学报*, 2023, 37(4): 35-43.  
KUI X, DING F, ZHANG D Y. A lane detection model based on multi branch fusion attention mechanism [ J ]. *Journal of Electronic Measurement and Instrumentation*, 2023, 37(4): 35-43.
- [ 13 ] 胡昊,史天运,关则彬. 融合混合注意力和改进 YOLOX 的铁路落石检测方法[ J ]. *电子测量技术*, 2022, 45(20): 110-116.  
HU H, SHI T Y, GUAN Z B. A railway rockfall detection method combining mixed attention and improved YOLOX [ J ]. *Electronic Measurement Technology*, 2022, 45(20): 110-116.
- [ 14 ] LIN H, CHENG X, WU X, et al. CAT: Cross attention in vision transformer [ C ]. 2021 IEEE International Conference on Multimedia and Expo ( ICME ), 2021: 1-6.
- [ 15 ] GIRSHICK B, DONAHUE J, DARRELL T, et al. 2014 Rich feature hierarchies for accurate object detection and semantic segmentation[ C ]. *CVPR*, 2014: 580-587.
- [ 16 ] HOWARD A, SANDLER M, CHU G, et al. Searching for MobileNetV3 [ C ]. 2019 IEEE/CVF International

Conference on Computer Vision. ICCV, 2019: 1314-1324.

- [17] 冯婧, 顾梅花, 刘晓龙, 等. 课堂场景下改进 MobileNetV3 人脸识别算法[J]. 国外电子测量技术, 2022, 41(10): 47-55.  
FENG J, GU M H, LIU X L, et al. Improving MobileNetV3 facial recognition algorithm in classroom settings [ J ]. Foreign Electronic Measurement Technology, 2022, 41(10): 47-55.
- [18] 廖亚桢, 刘昱, 张立强, 等. 面向 PDR 定位系统的多层感知机 IMU 阵列数据融合方法[J]. 电子测量与仪器学报, 2022, 36(8): 35-42.  
LIAO Y ZH, LIU Y, ZHANG L Q, et al. A multi-layer perceptron IMU array data fusion method for PDR positioning systems [ J ]. Journal of Electronic Measurement and Instrumentation, 2022, 36 ( 8 ): 35-42.
- [19] HE K, GKI G, PIOTR D, et al. Mask R-CNN[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2019, 24(3): 23-29.
- [20] KINGMA D P, BA J. Adam: A method for stochastic optimization [ J ]. arXiv preprint arXiv: 1412.6980, 2014.
- [21] KAN B, AMAR S. Development of VGG-16 transfer learning framework for geographical landmark recognition [ J ]. Intelligent Decision Technologies, 2023, 17 ( 3 ): 799-810.
- [22] MAJID R, MOH F, MOH B. Text-based automatic personality prediction using KGrAt-Net: A knowledge

graph attention network classifier[J]. Scientific Reports, 2022, 12(1): 21453.

## 作者简介



**汪家州**, 2017 年于沈阳工业大学获得学士学位, 现为沈阳工业大学硕士研究生, 主要研究方向为深度学习与停车位检测。

E-mail: 1473665306@qq.com

**Wang Jiazhou** received his B. Sc. degree from Shenyang University of Technology in 2017. Now he is a M. Sc. candidate at Shenyang University of Technology. His main research interests include deep learning and parking slot detection.



**桑海峰** (通信作者), 2000 年于东北师范大学获得学士学位, 2003 年于东北师范大学获得硕士学位, 2006 年于东北大学获得博士学位, 现为沈阳工业大学教授, 主要研究方向为智能视频分析、机器视觉检测与图像识别、无人驾驶之环境感知技术、深度学习技术。

E-mail: sanghaif@163.com

**Sang Haifeng** (Corresponding author) received his B. Sc. degree from Northeast Normal University in 2000, M. Sc. degree from Northeast Normal University in 2003 and Ph. D. degree from Northeastern University in 2006, respectively. Now he is a professor in Shenyang University of Technology. His main research interests include intelligent video analysis, machine vision detection and image recognition, autonomous driving environment perception technology, and deep learning technology.