

DMPR论文解析

2022年6月6日 10:20

概念1: 定向标记点/方向标记点

定向标记点实际上是以标记点及其邻域为特征的局部图像模式。它具有三个属性，位置、形状和方向。既然是局部图像模式，那么可以理解为一幅非常小的局部图像。

1、第一个属性：位置position，位置是什么意思，就是说**标记点是标记线的连接点**，其位置被定义为连接线与连接线重合的正方形区域，注意连接线是有一定宽度的。



(a) T-shaped



(b) L-shaped

方框的中心，注意是正方形的中心就是“位置”

2、第二个属性：形状。标记点图案有两种形状，T和L，简单易懂

3、第三个属性：方向。这个方向较难理解，

首先，T字型的标记点图案是对称的，将其方向定义为沿着对称轴的方向。

对于L型的标记点图案，定义为沿标记线的方向再顺时针旋转90°后与另一条标记线重叠，注意是重叠，这种情况是唯一的，这也是个很巧的定义方法。

综合以上三个属性，论文里将方向标记点表示为一个四维向量：

$\{x, y, s, \theta\}$

其中， x, y 表示在图像中的位置，

s 是表示图案形状的二进制值，论文中没有详细写，再继续研究

θ 表示极坐标系下图案方向的角坐标

angular coordinate (角坐标)，指的是 (r, θ) 中的 θ 的值。

检测方法和流程：

- 1、方向标记点回归
- 2、停车位推断

这两个流程是先1后2的，具体来看

第一步，方向标记点回归，做什么？

这一步是用CNN来做的，这个CNN输出的目标是方向标记点

给定一幅环视图像P，神经网络将P划分为S*S的图像网格，并使用CNN从P中提取S*S*N个**特征**

N是个特殊的东西，是个六维向量【cx, cy, s, cosθ, sinθ, 置信度C】

置信度预测标记点落入该网格单元的概率。cx和cy预测标记点在网格单元边界的位置。

s预测标记点的形状，两个三角函数是用来预测方向标记点的角度的。

论文里说的较为明白了

$$Loss = \sum_{i=1}^{S^2} \{ (i - C_i^{hat}) \}$$

CNN学的东西较多，这个block（作者称为cell）里面有没有方向标记点（L或T），这个L或T会决定一个“位置”，这个位置落在block的哪个位置（cx, cy），方向是什么？

这是个复杂工程，核心思路是这样

输入进网络的是512*512大小的图片，512/16 = 32，即全图被分为32个cell。

那么要对每一个cell做卷积？这明显是不现实的一件事情。

这一点还没能弄清楚

论文中描述了这样一段话：

The final output tensor with the size of $6 \times 16 \times 16$ is designed according to the requirement of our task. Firstly, as stated above, there are six predictions for each grid cell in the $S \times S$ grid and thus the channel dimension is 6. Secondly, a premise of our regression model is that there could be at most one marking-point falling into a cell in the $S \times S$ grid. If more than one marking-points fall into the same cell, the neural network could not predict both of them in a 6-dimensional vector. Thus, the value of S should be large enough to prevent two marking-points falling into the same cell. S should not be set too large either; otherwise, it will make DMPR computationally expensive. By examining examples in the training set of ps2.0, we set $S = 16$ in our implementation.

最终的输出张量大小为 $6 \times 16 \times 16$ ，是根据我们的任务要求设计的。首先，如上所述， $S \times S$ 网格中每个网格单元有 6 个预测，因此通道维度为 6。其次，我们的回归模型的前提是**最多有一个标记点落入 $S \times S$ 网格中的单元格**。如果多个标记点落入同一个单元格，则神经网络无法在 6 维向量中预测它们。因此，S 的值应该足够大，以防止两个标记点落入同一个单元格。S 也不宜设置过大；否则，它会使 DMPR 的计算成本很高。通过检查 ps2.0 [20] 训练集中的示例，我们在实现中设置 $S = 16$ 。

第二步

停车位推断，这一步是手写的逻辑来组合构成车位的

在检测到方向标记点，并使用非极大值抑制（找到那一坨的位置后）就要开始停车位推断包含两个步骤

①不适当的标记点过滤

首先，过滤是在干什么，一对标记点的距离应该满足真实世界的距离约束，因为车位在真实世界中的大小是有国家标准规定的范围的。例如说，正常的停车位大小是：2.5*5米，那么如果一对标记点形成的车位入口宽度过于窄了，或者过于宽了，都不是真实的车位，应该滤除。如图：

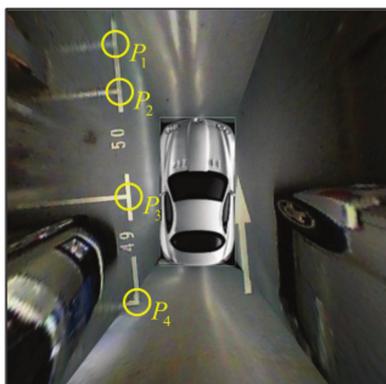
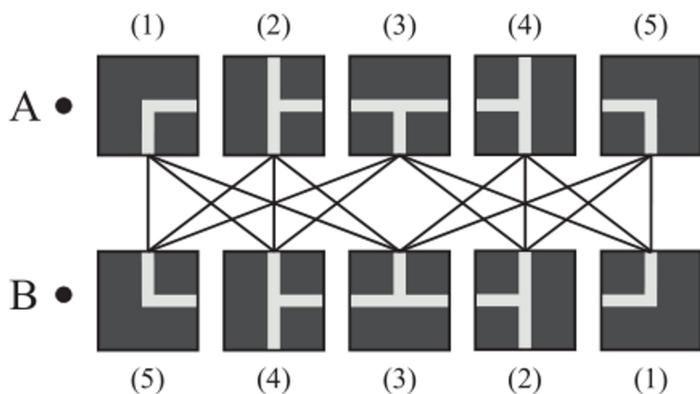


Fig. 4. Illustration of inappropriate marking-point pairs.

上图中P1和P2过于窄了，无法形成有效的入口线。P2和P4能构成一对入口，但是呢，他们根本不相邻，也构成不了停车位的入口。通过检查他们形成的入口线上是否有第三个标记点来排除这些无效的案例。在上图中，P2和P4之间夹着一个P3，所以它们不能构成有效的入口。

②定向标记点配对

对于形成入口线的一对标记点（A，B），两个标记点都可以分为5种情况。那么，对于这5种标记点，有16种组合，其中两个标记点形成有效的入口线。



对于每个标记点对，我们首先通过比较形状和方向来确定两个标记点是否属于5个标记点案例之一，然后判断这两个标记点的组合是否与入口线的16个有效案例之一匹配。如果这两个条件都满足，我们可以认为这两个标记点形成了一个有效的入口线，并且可以确定一个有序的标记点对。从而最终可以定位到这条入口线对应的停车位。

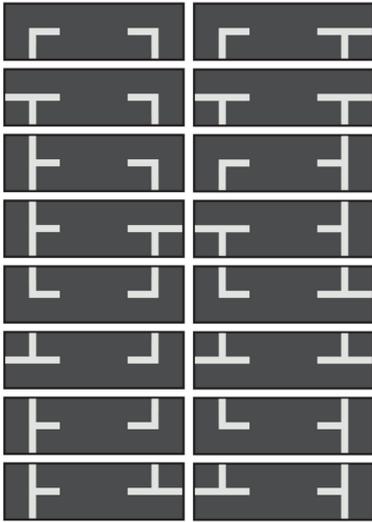


Fig. 6. 16 valid cases of entrance-lines.