

DOI:10.3969/j.issn.1003-5060.2023.03.007

改进的 SSD-ResNet 算法

孟 婧, 江 平, 王 凯, 蒋鑫宇

(合肥工业大学 数学学院, 安徽 合肥 230601)

摘 要: 单次多边界框检测器 (single shot multibox detector, SSD) 算法因其性能优良已被应用于许多场景中, 但该算法对小目标物体的检测精度偏低, 主要原因是高层的语义信息没有被充分利用。为解决该问题, 文章将其基础网络替换为残差网络 (residual network, ResNet), 同时融合深浅层的特征信息来增强浅层特征图的语义信息, 此外还引入注意力模块, 保留更多的目标特征信息, 抑制无关信息, 进一步提升对小目标物体的检测效果。在 PASCAL VOC2007 数据集上进行实验测试, 平均精度均值为 80.2%, 优于其他 SSD 改进算法。由于增加了特征融合和注意力模块, 所提算法检测速度有所下降, 但相比于 SSD 改进算法, 检测速度仍有明显的优势。

关键词: 目标检测; 单次多边界框检测器 (SSD); 残差网络 (ResNet); 特征融合; 注意力机制

中图分类号: TP391.413 **文献标志码:** A **文章编号:** 1003-5060(2023)03-0326-07

An improved SSD-ResNet algorithm

MENG Jing, JIANG Ping, WANG Kai, JIANG Xinyu

(School of Mathematics, Hefei University of Technology, Hefei 230601, China)

Abstract: As an algorithm with better detection accuracy and speed, single shot multibox detector (SSD) has made great progress in many aspects. However, it cannot achieve a good detection effect for small objects because it does not make full use of high-level semantic information. Aiming at this problem, VGG is replaced with residual network (ResNet) as the backbone network, the feature fusion method is used to enhance the semantic information of the shallow feature map, at the same time, the attention module is introduced, which retains more object feature information and suppresses irrelevant information. Through the above methods, the detection effect of small objects can be enhanced. By experimenting on the PASCAL VOC2007 datasets, the validity of the proposed algorithm is proven, the mean average precision value of the algorithm is 80.2%, which is better than those of other improved SSD algorithms. Although the addition of feature fusion and attention modules to the algorithm can cause a decrease in the detection speed, it is still better than deconvolutional single shot detector (DSSD) algorithm.

Key words: object detection; single shot multibox detector (SSD); residual network (ResNet); feature fusion; attention mechanism

目标检测作为计算机视觉领域重要的研究热点之一, 已在自动驾驶、车辆检测、视频检索等方面^[1-3]取得了较好的检测效果。基于深度学习的目标检测算法可以分为两大类: 一类是基于区

域建议的目标检测算法, 主要有 R-CNN^[4]、SPP-net^[5]、Fast R-CNN^[6]、Faster R-CNN^[7]等常见网络, 该类算法先生成一些候选框, 再将这些候选框送入深度神经网络, 提取相应的特征来对目标进

收稿日期: 2021-03-26; 修回日期: 2021-04-16

基金项目: 国家自然科学基金资助项目 (12172115)

作者简介: 孟 婧 (1997—), 女, 山西忻州人, 合肥工业大学硕士生;

江 平 (1972—), 女, 安徽安庆人, 博士, 合肥工业大学教授, 硕士生导师, 通信作者, E-mail: hfutus@hfut.edu.cn.

行分类和定位,优点是精度高,但缺点是速度慢;另一类是基于回归的目标检测算法,主要有 YOLO^[8]和单步多边界框检测器(single shot multi-box detector, SSD)^[9],此类算法直接按照预定方法去划分默认框,省去了生成候选框的时间,可保证检测速度大大提升。YOLO 采取回归的方式直接检测整张图片中的目标位置和类别,是一种全新的检测方式,但在小目标或目标之间位置接近的情况下检测效果不好。SSD 在 YOLO 的基础上增加了区域建议的机制,并使用多尺度特征图预测,因此效果优于 YOLO,但是该算法对小目标物体的检测精度偏低,主要是由于高层的语义信息没有被充分利用。

为了解决 SSD 对小目标检测效果不佳的问题,一些学者提出了关于 SSD 的改进算法。文献[10]提出了 DSSD,引入了反卷积层传递上下文信息。在对小目标的检测精度上,DSSD 相比 SSD 提升了许多,但是其检测速度却下降不少。RSSD^[11]采用池化加反卷积的方式将不同层的特征进行融合,增加不同层之间特征图关系和特征图个数,提升了对小目标的检测精度。FSSD^[12]以 FPN^[13]的思想为基本依据,通过上采样的方式将金字塔中特征图不同层统一到同一尺寸,明显提升了算法的精度,同时速度下降也不是很大。ASSD^[14]引入了一个基于全局特征关系的注意单元,能够抑制无关信息,更有效地学习对象特性,同时精度和速度均有提升。

本文结合 SSD 算法的特点,利用残差结构、特征融合以及注意力模块,以提高检测精度。

1 相关知识

1.1 SSD 算法

1.1.1 SSD 网络结构

SSD 网络结构如图 1 所示。

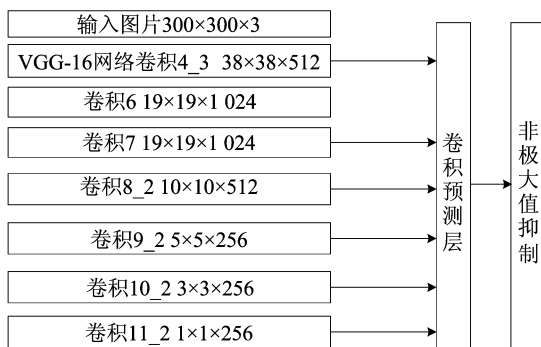


图 1 SSD 网络结构

第 1 次初步提取特征是通过 VGG-16^[15]的前 5 组卷积层进行的,随后再连接几层额外的卷积层进一步提取图像更深层次的特征,然后通过卷积预测层对不同层提取出来的特征信息进行类别和位置的预测,最终的检测结果利用非极大值抑制方式(non-maximum suppression, NMS)进行筛选后得到。

1.1.2 先验框的计算

网络中每个特征层的分辨率不同,即每个单元映射到原图的大小不同,因此先验框需要有针对性地设置。先验框的设置由尺度大小比例 s_k 、长宽比 r 、位置 $(x_{\min}, y_{\min}, x_{\max}, y_{\max})$ 3 个方面组成。

对于先验框的尺度,假设选取 m 个特征图进行预测, s_k 为第 k 个特征图对应的先验框大小与原图片的比值,则计算公式为:

$$s_k = s_{\min} + \left(\frac{s_{\max} - s_{\min}}{m - 1} \right) (k - 2), k \in [2, m] \quad (1)$$

其中: s_{\max} 、 s_{\min} 分别为该比例所取的最大值和最小值,一般 $s_{\max} = 0.9$, $s_{\min} = 0.2$, 此处 $m = 6$, $s_1 = 0.1$, 则可以计算出 $s_2 \sim s_6$ 相应的值,而计算真实值则需乘以 300, 即各个特征图对应的先验框尺度为 $f_k = 300s_k$ 。

选框的大小是根据目标的大小和不同的长宽比来设计的,长宽比参数为 $r_n \in \{1, 2, 3, 1/2, 1/3\}$, 根据这组参数来计算先验框的宽(w_k)和高(h_k),计算公式为:

$$w_k^n = s_k \sqrt{r_n}, n \in \{1, 2, 3, 4, 5\} \quad (2)$$

$$h_k^n = s_k / \sqrt{r_n}, n \in \{1, 2, 3, 4, 5\} \quad (3)$$

当长宽比 $r_n = 1$ 时,公式为:

$$w_k = h_k = \sqrt{s_k s_{k+1}} \quad (4)$$

对于先验框的位置,设定先验框的中心为 $((a+0.5)/|f_k|, (b+0.5)/|f_k|)$, 其中: $|f_k|$ 为第 k 个特征图的尺寸大小; $a, b \in \{0, 1, 2, \dots, |f_k - 1|\}$, 并归一化先验框的坐标使其在 $[0, 1)$ 内。特征图上先验框坐标与原始图像坐标的映射关系为:

$$\begin{aligned} x_{\min} &= \left(\frac{a + 0.5}{|f_k|} - \frac{w_k}{2} \right) \frac{w_i}{w_f}, \\ y_{\min} &= \left(\frac{b + 0.5}{|f_k|} - \frac{h_k}{2} \right) \frac{h_i}{h_f} \end{aligned} \quad (5)$$

$$\begin{aligned} x_{\max} &= \left(\frac{a + 0.5}{|f_k|} + \frac{w_k}{2} \right) \frac{w_i}{w_f}, \\ y_{\max} &= \left(\frac{b + 0.5}{|f_k|} + \frac{h_k}{2} \right) \frac{h_i}{h_f} \end{aligned} \quad (6)$$

其中: w_f, h_f 为特征层的宽和高; w_i, h_i 为原始图像的宽和高; $(x_{\min}, y_{\min}, x_{\max}, y_{\max})$ 为第 k 层特征图上以 $((a+0.5)/f_k, (b+0.5)/f_k)$ 中心、大小为 (w_k, h_k) 的先验框映射到原始图像的坐标。

1.1.3 损失函数

损失函数计算公式为:

$$L(x, c, l, g) = \frac{(L_{\text{conf}}(x, c) + \alpha L_{\text{loc}}(x, l, g))}{N} \quad (7)$$

其中: L_{conf} 为置信度损失; L_{loc} 为位置损失; N 为与真实目标框相匹配的区域候选框的个数; $x=1$ 表示预测框与标注框相匹配, $x=0$ 表示不匹配; c 为类别置信度; l 为预测框; g 为标注框; α 为权重参数, 用来决定置信度损失和位置损失所占的权重关系, 一般两者取相同的权重, 值设为 1。

位置损失函数 L_{loc} 的定义如下:

$$L_{\text{loc}}(x, l, g) = \sum_{i \in P_{\text{os}}} \sum_{m \in \{cx, cy, w, h\}} x_{ij}^p S_{\text{LI}}(l_i^m - g_j^{m'}) \quad (8)$$

$$S_{\text{LI}}(x) = \begin{cases} 0.5x^2, & |x| < 1; \\ |x| - 0.5, & \text{其他} \end{cases} \quad (9)$$

其中: P_{os} 为正样本(成功匹配的先验框)集合; $x_{ij}^p = \{1, 0\}$, 是一个取值集合, $x_{ij}^p = 1$ 表示第 i 个先验框匹配到第 j 个类别为 P 的标注框, $x_{ij}^p = 0$ 则表示没有匹配; $l_i^m - g_j^{m'}$ 表示预测窗口 l 的中心坐标 (cx, cy) 和宽高 (w, h) 相对于真实框 g 的偏移量。

对于多类别的置信度判定, 采用交叉熵损失函数, 由于本文共有 21 类目标(包括目标 20 类和背景), 因此使用 softmax 回归方法来计算。 L_{conf} 具体形式如下所示:

$$L_{\text{conf}}(x, c) = - \sum_{i \in P_{\text{os}}} x_{ij}^p \lg(c_i^{p'}) - \sum_{i \in N_{\text{eg}}} \lg(c_i^{0'}) \quad (10)$$

$$c_i^{p'} = \frac{\exp(c_i^p)}{\sum_p \exp(c_i^p)} \quad (11)$$

其中: $c_i^{p'}$ 表示第 i 个正样本(成功匹配的先验框)关于类别 P 的 softmax 置信回归; N_{eg} 为负样本集合; $c_i^{0'}$ 为第 i 个负样本(被判定为背景类的候选窗口)的置信度。

1.2 深度残差网络

从理论上讲, 网络层数的增加会使性能变得更好, 但实验发现深层网络模型很难学习到这种结果, 这是由于网络深度增加后会出现梯度消失

从而导致反向传播失效。深度残差网络(residual network, ResNet)^[16]的出现解决了此问题, 其结构如图 2 所示, 其中: z 为神经网络的输入; $H(z)$ 为期望输出, 目的是使输入 z 尽可能接近输出 $H(z)$, 使残差 $F(z)$ 接近于 0, $F(z) = H(z) - z$, 可以保证网络层数增加且检测精度不会降低。

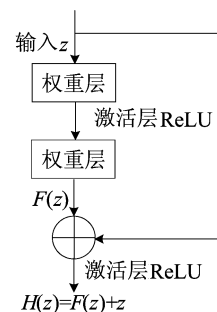


图 2 深度残差网络结构

1.3 注意力模块

注意力模块^[17]主要由通道注意力(channel attention)和空间注意力(spatial attention)2 个部分组成, 其结构如图 3 所示。

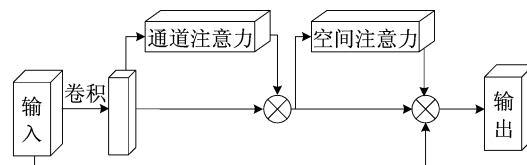


图 3 注意力模块

注意力模块的步骤如下:

(1) 先将输入特征图分别经过最大池化和平均池化操作后送入多层感知机(multilayer perceptron, MLP), 并对其输出的特征进行对应元素逐个相乘(elementwise)的运算, 再经过 sigmoid 激活函数, 生成最终的通道特征图。

(2) 将该通道特征图和输入特征图通过 elementwise 方式进行计算, 生成空间注意力模块需要的输入特征, 分别经过最大池化和平均池化操作后再进行合并(concatenation)运算, 使用卷积将通道变成 1, 经过 sigmoid 激活函数生成空间特征图, 最后将该特征和该模块的输入特征做 elementwise 乘法, 得到最终生成的特征。

2 改进的 SSD-ResNet 算法

2.1 算法框架总体结构

由于 SSD 网络由多层卷积层构成, 训练会产生梯度消失的问题。根据深度残差结构对其普通卷积层进行改进, 能有效促进网络的训练和收敛。

同时,SSD 网络结构与图像金字塔结构相似,浅层特征图尺寸大、感受野小、细节特征信息丰富,深层特征图与之相反,且语义特征信息比较丰富,因此将浅层信息与深层信息进行特征融合能够增强浅层网络的语义信息,提升浅层网络对小目标的检测能力。基于 SSD 网络模型的上述特点,本文先将基础网络换成 ResNet,通过特征融合的方式,提高对小目标的检测精度。此外,引入注意力通道,保留更多的目标特征信息,抑制无关信息,进一步提升对小目标的检测效果。

2.2 改进的 SSD-ResNet 网络结构

2.2.1 SSD-ResNet 网络结构

本文的 ResNet 有 2 种残差结构(residual block),如图 4 所示。图 4a 为恒等模块,其输入特征图先经过 1×1 卷积层、批量标准化(BN)层和激活函数(ReLU)层后,又进行 1 次 3×3 卷积层、BN 层和 ReLU 层,再进行 1 次 1×1 卷积层以及 BN 层,最后与原特征图相加。在此过程中,特征图的大小保持不变。图 4b 为卷积模块,其与恒等模块的区别是在捷径连接(shortcut path)部分增加 1 次 1×1 卷积层以及 BN 层,这种结构改变特征图大小卷积方式。

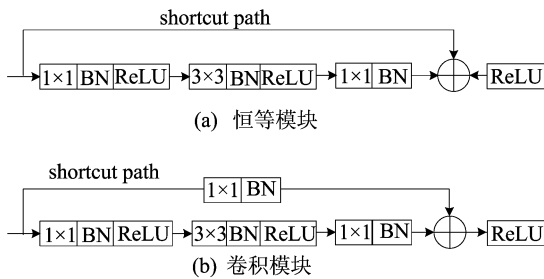


图 4 残差结构

将 SSD-ResNet 网络改进后的结构如图 5 所示。输入图片大小为 300×300 ,经过 1 个卷积层和 1 个池化层后,接入 ResNet 网络部分进行特征提取,之后进行普通的卷积操作,在特征图卷积 3-3 和卷积 4-6 后加入注意力模块来增强低层网络的信息表征能力,使感受野更关注目标特征,并进行特征融合来增强对目标的定位能力,最后将这些特征图通过卷积预测层进行预测。选出置信度最大值所属类别作为该框的类别,并过滤掉属于背景的预测框,设置置信度阈值为 0.5,将低于该阈值的预测框舍弃,根据先验框对其余预测框进行解码得到真实的位置参数,并对置信度进行降序排列,仅保留前 100 个值,利用 NMS 方法选出最终的预测结果。

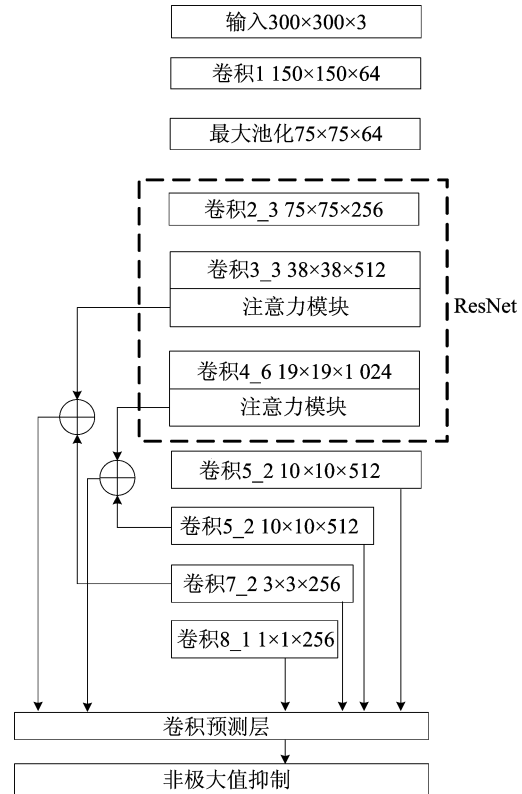


图 5 改进的 SSD-ResNet 网络结构

2.2.2 特征融合

卷积 3 和卷积 4 经过卷积 2 之后,其特征进一步被提取出来,但仍然属于浅层网络,细节特征比较丰富,而卷积 6 和卷积 7 是深层网络,语义特征丰富,实验表明该种组合的融合效果最好。

特征融合具体步骤如下:本文采用特征融合的思想,在 SSD 模型的基础框架上,首先用 1×1 的卷积核对卷积 3-3、卷积 4-6、卷积 6-2、卷积 7-2 不同尺寸的特征层进行卷积操作以确保融合时的通道统一,然后通过双线性插值的方法将特征图统一到同一尺寸,最后采用连接融合的方式得到新的不同尺寸的特征图。改进后的模型较 SSD 模型的特征层语义信息更丰富,能更好地表征目标的特性。

3 实验结果与分析

3.1 实验环境

本文所有算法基于 Tensorflow 深度学习框架,采用 Ubuntu(Linux)操作系统。实验中所使用计算机 CPU 型号为 Intel Core i7-6700K,运行内存大小为 16 GB,显卡是 TITAN X 显卡,其显存大小为 12 GB。

3.2 评价指标

由于 PASCAL VOC2007 数据集有多个类

别,本实验采用平均精度均值 (mean average precision, m- A_P) 作为模型的评价指标。其中, m- A_P 是 A_P 相加取平均得到的, 平均准确率 (A_P) 是 Precision-Recall ($P-R$) 曲线包围面积, 公式为:

$$P = \frac{T_P}{T_P + F_P}, R = \frac{T_P}{T_P + F_N} \quad (12)$$

$$A_P = \int_0^1 P(R) dR \approx \frac{1}{m} \sum_{i=1}^m P(R_i) \quad (13)$$

$$m-A_P = \frac{1}{C} \sum_{i=1}^C A_P(i) \quad (14)$$

其中: T_P 是实际为正且预测为正的样本数量; F_P 是实际为负而预测为正的样本数量; F_N 是实际为负而预测为正的样本数量; m 是取样点的数量为待检测目标的类别数。本文取 $m=11, C=20$ 。

3.3 实验和分析

为了验证本文网络模型的有效性, 共设置了以下 4 组实验: ① 将本文的特征融合方式与 FPN 以及 RetinaNet 中的特征融合方法进行对比, 探究本文特征融合方法是否有效; ② 通过 8 种不同的网络模型, 将原始 SSD 算法中的 VGG 模块与本文的 3 种模块 (ResNet 模块、特征融合模块、注意力机制模块) 作对比, 探究不同模块对网络模型性能的影响; ③ 验证本网络模型对小目标尺度的检测效果优于其他算法; ④ 在 PASCAL VOC2007 测试数据集上进行本文算法与 SSD 算法、相关改进算法的对比实验, 从精度和速度 2 个方面来比较。

3.3.1 特征融合方法的对比实验

为了探究本文特征融合方法的有效性, 通过与 FPN 以及 RetinaNet 中的特征融合方法进行对比, 并在 PASCAL VOC2007 测试数据集上对其进行测试, 结果见表 1 所列。由此可见, 本文算法的 m- A_P 值均高于方法 1 和方法 2。

表 1 不同方法中的特征融合对比实验

方法	ResNet	FPN	Focal loss(RetinaNet)	特征融合	m- A_P /%
1	✓	✓			79.2
2	✓	✓	✓		79.4
3	✓			✓	79.5

3.3.2 不同网络模型的对比实验

为了探究 SSD 算法中的 VGG 模块与本文的 3 种模块 (ResNet 模块、特征融合模块、注意力机制模块) 对网络模型性能的影响, 将这 3 种模块组合成 7 种不同的网络模型后再训练, 并在 PASCAL VOC 2007 测试数据集上对其进行测试, 实验结果见表 2 所列。从表 2 可以看出, 每种组合

模型的 m- A_P 均有提升, 其中特征融合模块在目标检测精度提升中作用最大, 这是由于它可以有效地传递上下文信息, 减少各层次特征图之间的语义差异, 且注意力机制模块对检测精度也有一定的提升, 抑制了无关信息, 保留了更多的目标特征信息。实验结果表明, 将这 3 种模块组合在一起可以更加有效地提升检测精度。

表 2 不同网络模型的对比实验

VGG	ResNet	注意力机制	特征融合	m- A_P /%
✓				77.1
✓		✓		77.6
	✓			78.0
	✓	✓		78.3
✓			✓	78.4
✓		✓	✓	78.8
	✓		✓	79.5
	✓	✓	✓	80.2

3.3.3 不同算法在小尺度目标上的检测效果

为了验证本文模型对小目标的检测效果, 在 PASCAL VOC2007 测试集中选择一些具有小尺度目标 (目标真实框面积与整张图片面积的比值小于 20%) 的图片进行测试, 实验结果见表 3 所列。从实验结果可以看出, 本文算法的 m- A_P 值达到了 51.1%, 优于其他算法。

表 3 不同类别的小尺度目标的 A_P 值 %

目标	SSD	DSSD	FSSD	ASSD	本文
aeroplane	42.5	46.8	48.2	50.1	51.7
bike	61.2	62.3	63.4	64.3	65.9
bird	36.8	43.6	47.5	52.5	53.7
boat	25.7	30.4	34.2	35.9	37.8
bottle	9.5	15.8	19.7	23.8	24.6
bus	48.9	54.6	56.2	59.2	60.1
car	44.6	46.4	48.1	48.7	51.5
cat	67.2	67.9	68.2	68.0	68.4
chair	19.8	20.5	21.4	22.3	23.6
cow	43.1	50.8	56.7	63.4	65.2
table	35.5	35.9	36.1	36.2	36.3
dog	57.0	59.5	61.3	61.9	62.7
horse	62.1	65.4	69.7	72.4	73.9
motorbike	47.6	49.6	52.1	53.7	55.6
person	31.4	34.2	35.4	38.6	40.2
plant	7.3	15.1	17.8	22.8	25.4
sheep	41.6	43.4	47.5	49.1	50.1
sofa	59.6	59.8	60.1	60.2	60.3
train	64.2	66.8	67.2	68.4	69.0
TV	30.4	35.1	38.7	43.3	45.4
m- A_P	41.8	45.2	47.5	49.7	51.1

3.3.4 不同算法对比实验

为了进一步验证本文算法的有效性, 通过与

SSD 算法及其他算法进行对比,结果见表 4 所列。本文算法的 $m-A_p$ 值达到 80.2%,相比于原始 SSD 算法、DSSD 算法、FSSD 算法、ASSD 算法,分别提升了 3.1%、2.2%、1.5%、0.7%。

表 4 不同算法各个类别的 A_p 值 %

算法	SSD	DSSD	FSSD	ASSD	本文
aeroplane	78.1	79.7	82.3	84.5	85.8
bike	83.2	84.6	85.1	85.3	85.7
bird	76.5	77.3	78.4	78.6	78.9
boat	68.3	67.5	69.3	70.2	70.8
bottle	53.4	54.8	55.2	55.4	55.7
bus	88.1	86.9	88.7	89.5	89.6
car	83.2	85.8	86.1	88.1	89.2
cat	87.5	88.1	88.9	89.2	89.7
chair	61.1	61.7	62.0	62.8	63.9
cow	79.5	81.4	82.7	84.0	85.4
table	78.3	79.2	79.1	79.6	79.8
dog	85.2	85.4	85.2	86.3	87.3
horse	86.1	86.8	87.5	88.1	88.4
motorbike	80.6	81.3	81.8	81.9	82.1
person	77.1	78.4	78.2	79.2	79.7
plant	56.4	57.1	58.0	58.3	58.6
sheep	78.0	81.2	81.3	82.4	84.7
sofa	77.5	77.6	78.7	79.1	79.6
train	85.9	86.4	85.6	86.3	87.1
TV	78.3	78.7	79.2	80.5	81.9
$m-A_p$	77.1	78.0	78.7	79.5	80.2

不同算法检测速度的对比结果见表 5 所列。从表 5 可以看出,相比于 DSSD 算法,本文算法的网络层数较少,因此速度上存在明显的优势。

表 5 不同算法检测速度的对比实验

算法	SSD	DSSD	FSSD	ASSD	本文
速度/(帧·s ⁻¹)	48.0	9.5	36.5	27.5	27.0

4 结 论

本文将残差网络作为基础网络,引入了注意力机制模块和特征融合的方法,提升了算法的检测精度。通过实验验证了本文算法中特征融合方法优于 FPN 中的特征融合方法。同时,将不同模块对网络性能的影响进行对比实验,加入不同模块后效果均有提升,其中特征融合模块对检测精度的提升效果最佳。此外,对小尺度物体的检测效果较好,其 $m-A_p$ 值为 51.1%。最后,在 PASCAL VOC2007 测试数据集上进行了本文算法与 SSD 算法及其相关改进算法的对比实验,再次验证了本文算法的优越性,但速度有所下降,接下来

将会在模型压缩方面做进一步的研究,以提升算法的运行速度。

[参 考 文 献]

- [1] 阮祥伟,李华,余焯. 基于改进 YOLOv3 的快速车标检测方法[J]. 合肥工业大学学报(自然科学版), 2020, 43(12): 1608-1613.
- [2] 李涛,张代胜,张飞,等. 基于改进 YOLO 算法的 FSAC 赛车实时目标检测系统[J]. 合肥工业大学学报(自然科学版), 2020, 43(9): 1153-1158.
- [3] CHEN B, MIAO X. Distribution line pole detection and counting based on YOLO using UAV inspection line video [J]. Journal of Electrical Engineering & Technology, 2020, 15: 441-448.
- [4] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and segmentation[C]//Proceeding of 2014 IEEE Conference on Computer Vision and Pattern Recognition. [S. l. : s. n.], 2014: 580-587.
- [5] HE K, ZHANG X, REN S, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2015, 37(9): 1904-1916.
- [6] GIRSHICK R. Fast R-CNN[C]//Proceedings of 2015 IEEE International Conference on Computer Vision. [S. l.]: IEEE, 2015: 1440-1448.
- [7] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks [C]//Proceeding of the 28th International Conference on Neural Information Processing Systems. [S. l. : s. n.], 2015: 91-99.
- [8] REDMOM J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection[C]//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. [S. l.]: IEEE, 2016: 779-788.
- [9] LIU W, ANGUELOV D, ERHAN D C, et al. SSD: single shot multibox detector[C]//Proceeding of the 14th European Conference on Computer Vision. [S. l. : s. n.], 2016: 21-37.
- [10] FU C Y, LIU W, RANGA A, et al. DSSD: deconvolutional single shot detector[EB/OL]. (2017-01-23) [2019-08-28]. <https://arxiv.org/abs/1701.06659>.
- [11] JEONG J, PARK H, KWAK N. Enhancement of SSD by concatenating feature maps for object detection[EB/OL]. (2017-05-26) [2019-08-28]. <https://arxiv.org/abs/1705.09587>.
- [12] LI Z X, ZHOU F Q. FSSD: feature fusion single shot multibox detector[EB/OL]. (2018-05-17) [2019-08-28]. <https://arxiv.org/abs/1712.00960>.
- [13] LIN T Y, DOLLAR P, GIRSHICK R, et al. Feature pyramid networks for object detection [C]//Proceedings of

- 2016 IEEE conference on computer vision and pattern recognition. [S. l.]:IEEE,2016:2117-2125.
- [14] YI J R, WU P X, METAXAS D N. ASSD: attentive single shot multibox detector[J]. Computer Vision and Image Understanding, 2019, 189:102827. 1-102827. 7.
- [15] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [EB/OL]. (2014-09-04) [2019-08-28]. <https://arxiv.org/abs/1409.1556v3>.
- [16] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//Proceedings of 2016 IEEE conference on computer vision and pattern recognition. [S. l.]: IEEE, 2016:770-778.
- [17] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module[C]//Proceedings of the European Conference on Computer Vision. [S. l. : s. n.], 2018:3-19.

(责任编辑 李 凯)

(上接第 302 页)

由试验结果可知,整个试验过程共停车 11 次,开阀 4 次,试验前液压支路压力差为 0 MPa,试验过程中最大压力差为 0.76 MPa,试验结束时支路压力差为 0.18 MPa,结果表明该控制方法能够适时打开电磁阀平衡压力差,并避免电磁阀误开启造成压力差加剧,有效降低 HIS 系统液压缸内泄影响。

4 结 论

本研究 HIS 电磁平衡阀的控制策略,以减小液压缸内泄影响,提高车辆性能。首先,利用 AMEsim 建立了 HIS 与电磁阀模型,并通过实验分析对比了电磁阀的流量特性与平衡效果;然后通过对实验数据的分析,依据电磁阀与车辆特性设计了模糊控制器对电磁阀进行控制,通过 MATLAB/Simulink 验证模糊控制策略的可行性;最后以某 SUV 对设计的控制器进行试验分析。结果表明,基于模糊控制的电磁平衡阀能有效降低液压缸内泄对 HIS 系统的影响。

[参 考 文 献]

- [1] 丁飞,张农,韩旭. 安装液压互联悬架货车的机械液压多体系统建模及模态分析[J]. 机械工程学报, 2012, 48(6): 116-123.
- [2] 赵贺雪,张邦基,张农,等. 高度可调式抗侧倾液压互联悬架建模及控制策略研究[J]. 振动与冲击 2018, 37(3): 202-209.
- [3] 朱海燕,苏校,陈齐平,等. 车辆减振器油液微小内泄漏分析[J]. 润滑与密封, 2019, 44(8): 81-86.
- [4] ZHANG N, SMITH W A, JEYAKUMARAN J. Hydraulically interconnected vehicle suspension: background and modelling[J]. Vehicle System Dynamics, 2010, 48(1): 17-40.
- [5] SMITH W A, ZHANG N, JEYAKUMARAN J. Hydraulically interconnected vehicle suspension: theoretical and experimental ride analysis[J]. Vehicle System Dynamics, 2010, 48(1): 41-64.
- [6] 曾繁鑫,郑敏毅,刘军,等. 半主动液压互联悬架客车的平顺性研究[J]. 合肥工业大学学报(自然科学版), 2020, 43(3): 307-315.
- [7] ZHU S, DU H, ZHANG N. Development and implementation of fuzzy, fuzzy PID and LQR controllers for an roll-plane active Hydraulically Interconnected Suspension[C]//2014 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE). [S. l.]: IEEE, 2014: 2017-2024.
- [8] 吴晓建,周兵,文桂林. 液压互联悬架抗侧倾控制研究[J]. 中国公路学报, 2018, 31(3): 123-132.
- [9] 江治东,郑敏毅,张农. 半主动抗俯仰液压互联悬架俯仰力学的研究[J]. 振动与冲击, 2020, 39(14): 272-278.
- [10] 陈盛钊,钟义旭,张邦基,等. 液压互联悬架系统关键参数对车辆动力学响应影响及试验验证[J]. 机械工程学报, 2017, 53(14): 39-48.
- [11] 杨扬,张翼. 基于 AMESim 的锥阀和球阀式液压阀动态特性分析[J]. 煤矿机械, 2015, 36(4): 154-155.
- [12] 朱佳鑫. 桥梁变形监测中基于 MEMS 加速度倾角传感器的设计与研究[D]. 广州:华南理工大学 2020.
- [13] 徐小东,张冰战. 基于模糊逻辑的混合动力汽车控制策略研究[J]. 合肥工业大学学报(自然科学版), 2012, 35(6): 725-728.
- [14] 陈龙,马瑞,王寿静,等. 车辆半主动悬架阻尼多模式切换控制研究[J]. 振动与冲击, 2020, 39(13): 148-155.

(责任编辑 胡亚敏)