

文章编号: 2095-2163(2020)11-0061-05

中图分类号: TP391

文献标志码: A

基于改进 Unet 的斑马线分割

黄生鹏, 柳海南, 周克帅, 刘建宇

(上海工程技术大学 机械与汽车工程学院, 上海 201620)

摘要: 斑马线是道路交叉口的重要安全标志, 准确的检测出斑马线能有效的提高交叉路口行驶的安全性。本文在 Unet 网络的基础上, 将其主干特征提取网络替换为 MobileNetv2, 减少模型的参数, 降低模型的计算量。并在网络结构中嵌入了 SENet 模块, 对有效特征信息进行加权, 增强特征提取能力。实验结果表明, 改进后的 SEMobileNetv2 网络结构与之前相比在分类精度上提高了 1.6%; 相对于 Unet 网络, 改进后的模型在分割精度上提高了 1.8%, 验证了本文改进后分割模型的有效性。
关键词: Unet; MobileNetv2; SENet; 图像分割

Zebra Crossing Segmentation Based on Improved Unet

HUANG Shengpeng, LIU Hainan, ZHOU Keshuai, LIU Jianyu

(School of Mechanical and Automotive Engineering, Shanghai University of Engineering Science, Shanghai 201620, China)

[Abstract] The zebra crossings are important safety signs at road intersections, and accurate detection of zebra crossings can effectively improve the safety of driving at intersections. Based on the Unet network, this paper replaces its backbone feature extraction network with MobileNetv2 to reduce the parameters of the model and reduce the amount of model calculations. And the SENet module is embedded in the network structure to weight the effective feature information and enhance the feature extraction capability. The experimental results show that the improved SEMobileNetv2 network structure has a 1.6% increase in classification accuracy compared with the original network. Compared with the Unet network, the improved model has a 1.8% increase in segmentation accuracy. This verifies that the improved segmentation model is Effectiveness.

[Key words] Unet; MobileNetv2; SENet; image segmentation

0 引言

近年来,随着无人驾驶、智能小车等智能交通工具的发展,如何有效的检测出道路标识线,保障行人的安全成为热门研究方向。斑马线是道路交叉口的重要安全标志,准确的识别出斑马线能有效提高交叉路口的行驶安全性。传统的斑马线识别方法主要有以下几种:

(1) 灭点法^[1]。根据斑马线的平行特点,利用两次霍夫变换找到直线的灭点从而确定斑马线所在的位置。该方法运算速度快,但是当斑马线较多时区分性差。

(2) 频域法^[2]。该方法从斑马线黑白交替出现的纹理规律出发,根据现实中斑马线的宽度找到特定的频率,并以该频率为中心频率在一定带宽内筛选,此方法识别率很高,但是计算成本高,实时效果差。

(3) 双极系数法。该方法从斑马线黑白颜色分明特点出发,利用均值、方差获取某区域的黑白分明

程度。双极系数法计算速度快,然而易受环境的影响。

随着深度学习的发展,特别是卷积神经网络^[3]在图像领域的广泛应用,大量基于深度学习的图像分割算法被提出。如 FCN^[4]、Unet^[5]、Mask R-CNN^[6]、Deeplab^[7]等。这些算法在分割性能上相比于传统分割算法,具有更高的鲁棒性,而且速度也很快。因此,本文在深入研究 Unet 网络的前提下,提出一种基于改进 Unet 网络的斑马线分割算法。对 Unet 特征提取网络进行改进,采用轻量级 MobileNetv2^[8]网络,代替原有模型的主干网络,以减小参数量;利用 SENet 模块^[9]增强对特征图中重要信息的提取。

1 Unet 网络原理

Unet 是 Ronneberger 等人提出的一种医学影像分割网络,其网络结构如图 1 所示。整个网络可分成二部分,左侧为特征提取网络,右侧为上采样网络。其中特征提取网络就是 CNN 卷积网络,每个尺

基金项目: 上海工程技术大学研究生科研创新项目(19KY0143)。

作者简介: 黄生鹏(1993-),男,硕士研究生,主要研究方向:图像处理;柳海南(1996-),女,硕士研究生,主要研究方向:超声电机。

通讯作者: 柳海南 Email: 1312894260@qq.com

收稿日期: 2020-08-15

度结构包含两个 3×3 卷积、一个 $\text{RELU}^{[10]}$ 层和一个 2×2 的最大池化层。经池化层变换后,特征通道数扩大一倍;上采样网络每个尺度结构利用 2×2 的上采样层,特征通道数减少一半,其后与对应的特征提取网络的特征图进行拼接操作,得到新的特征图,再采用 2 个 3×3 卷积层加 RELU 激活层。最后一个拼接尺度用 1×1 的卷积层将 64 维通道特征图映射到 2 维图像上,获得最终的分割结果。

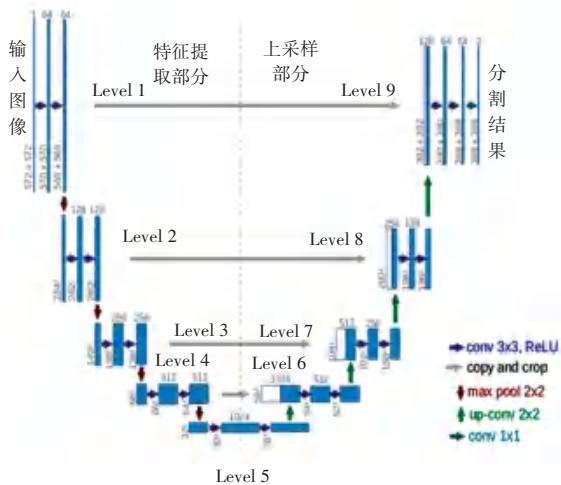


图 1 Unet 网络结构

Fig. 1 Unet network structure

2 网络结构改进

由上节可知,Unet 采用传统的卷积进行特征提取。由于斑马线分割场景往往处于复杂的环境中,这就需要更深层的卷积操作,才能有效的提取特征。随着卷积层的增加,参数量也会越来越大。为了达到既能高效提取特征,又能降低网络参数,本文采用 Mobilenet2 网络替换其特征提取网络,同时在网络结构中嵌入注意力机制——SENet 模块,进一步加强网络对有效特征信息的提取,抑制无效特征,进而提高分割精度和鲁棒性。

2.1 MoblieNetV2 网络

MobileNet 网络是一种轻量级网络,非常适用于嵌入式平台。其主要结构是基于流线型架构,结构中采用了大量的深度可分离卷积,以此来降低模型的参数量和计算量。

MobileNet1 主要采用堆叠大量的深度可分离卷积结构,构建神经网络,在保证精度的情况下,极大的压缩了模型。深度可分离卷积结构如图 2 所示。

MobileNet2 网络是 Mobilenet1 的升级版,其主要特点在于反转残差模块 (Inverted Residual Block) 和瓶颈模块 (Linear Bottleneck),是带有线性瓶颈层

的反转残差结构,如图 3 所示。首先,该结构输入的低维空间特征,通过 Expansion layer 映射到高维空间特征,应用深度可分离卷积进行处理后,使用线性瓶颈层再将特征投影回低维特征。MobileNet2 采用线性瓶颈 (Linear bottleneck) 代替非线性激活变换,保留更多的特征信息。通常认为神经网络由 n 个 L_i 层构成,每层经过激活函数最终的输出张量为 $h_i \times w_i \times d_i$,一系列卷积与激活层形成了一个兴趣流 (manifold of interest)。在神经网络中,兴趣流可以嵌入到低维子空间,就是卷积层中所有单个像素通道 d ,这些值包含多种编码信息。深度神经网络的层具有 ReLU 非线性激活函数,当完整度较高的兴趣流经过非线性激活函数时,可能会产生空间坍塌。

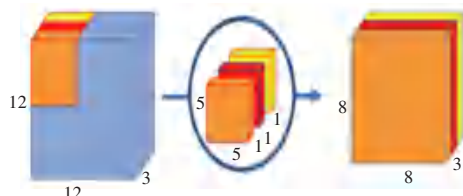


图 2 深度可分离卷积

Fig. 2 Depth separable convolution

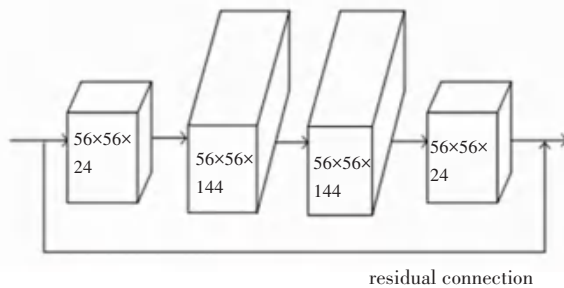


图 3 反转残差结构

Fig. 3 Reverse residual structure

2.2 基于 SENet 改进的 SE-MobileNetv2

SENet 模块是在通道维度上进行 attention 操作,这种注意力机制可以让网络更加关注有效的通道特征,抑制那些不重要的通道特征。SENet 模块结构如图 3 所示。首先 SENet 模块对输入的特征图进行 Squeeze 操作,得到 channel 级的全局特征,然后对全局特征进行 Excitation 操作。学习各个 channel 之间的关系,获得不同 channel 的权重,最后与原来的特征图相乘得到校准过的特征 SENet 模块的 3 个主要操作。

2.2.1 Squeeze 操作

一般的卷积操作仅仅在一个局部空间内进行,因此 U 中很难获得足够的信息来提取通道之间的关系。对于网络深度较潜的特征层来说,其感受野尺度都比较小,这种情况更加糟糕。 U 中的多个特

征图,可被视为局部特征表征的一个子集,对于整个特征图来说具有很好的表现力。Squeeze 操作利用全局平均池化(global average pooling)将每一个通道上的空间特征编码为一个全局特征,其在某种程度上代表了全局感受野,压缩过程如公式(1)所示。

$$z_c = F_{sq}(u_c) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W u_c(i, j). \quad (1)$$

式中: u_c 表示卷积变换后第 c 个通道的输出; z_c 表示压缩操作后第 c 个通道的输出; (W, H) 表示特征图的宽和高; F_{sq} 代表全局特征压缩变换函数。

2.2.2 Excitation 操作

图 4 中的 $F_{ex}(\cdot, W)$ 变化操作,即为 Excitation 操作。其利用两个全连接和两个激活函数来获取通道之间的依赖关系,计算如公式(2)所示。

$$s = F_{ex}(z, W) = \sigma(g(z, W)) = \sigma(W_2 \text{ReLU}(W_1 z)). \quad (2)$$

式中: σ 代表 Sigmoid 激活函数; z 是 Squeeze 操作获得的全局特征,其维度是 $1 \times 1 \times C$; W_1 代表第一个全连接层中的降维权重,其维度是 $\frac{C}{r} \times C$, 其中 r 为降维比例,在文献^[11]中该参数取值为 16; W_2 代表第二个全连接层中的升维权重,其维度是 $C \times \frac{C}{r}$, 第二个全连接层输出的维度是 $1 \times 1 \times C$,再经过 Sigmoid 激活将权重归一化到 $(0, 1)$ 范围内, s 的最终维度是 $1 \times 1 \times C$ 。

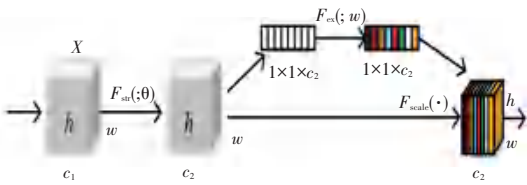


图 4 ENet 模块

Fig. 4 SENet module

Excitation 操作通过两个全连接层,获取特征 U 中的 C 个特征图的权重。Squeeze 操作是将单个特征图进行特征编码,而 Excitation 操作是对所有编码后的特征图进行信息汇总。这样做可以降低模型的复杂度,同时也提高了模型的计算效率。

2.2.3 Scale 操作

Scale 操作是通过相乘的方法,将 Excitation 操作获得归一化权重加权到标准卷积之后的特征上,经过加权后的特征图包含了先 2 前特征 U 的每个通道的重要程度,因此可以达到加强重要特征信息的目的。计算如公式(3)所示:

$$\tilde{x}_c = F_{scale}(u_c, s_c) = s_c \cdot u_c. \quad (3)$$

式中: u_c 表示单个通道的二维矩阵; s_c 表示权重; \tilde{x} 表示加权计算后得到的特征图。

在传统卷积操作中,更多的关注点在于空间特征的融合上,对于通道维度的特征融合,卷积操作基本上默认对输入特征图的所有通道进行融合。而 MboileNetv2 网络中的组卷积(Group Convolution)和深度可分离卷积是对通道进行分组,使得模型参数量减少。这种方式并没有关注通道之间的关系,从而模型无法自动学习到不同通道特征之间的重要程度。因此,本文在反转残差结构的基础上,嵌入 SENet 模块,使得模型更加关注有效通道信息,提高模型精度。结合模块结构如图 5 所示。

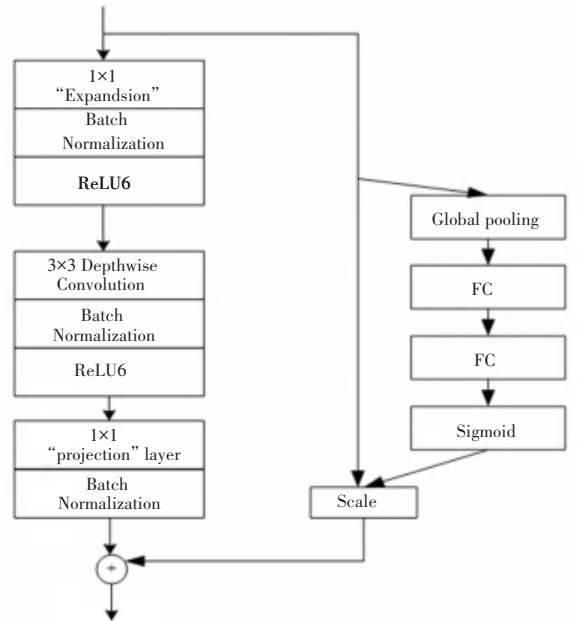


图 5 SENet 模块结合图

Fig. 5 SENet module binding diagram

3 数据集构建

深度神经网络模型的训练,需要大量优良的数据支撑。本文构建的斑马线数据集来源分为二部分,一部分是通过互联网爬取网页图片,另一部分由人工采集,并进行人工手动标注。为了使模型具有更强的鲁棒性,本文在晴天、阴天、雨天、复杂街道环境等条件下采集斑马线数据集。

卷积神经网络的数据拟合过程,需要给网络提供一个真实的标准数据。通过网络预测的结果与真实结果的损失值,自动的调整模型权重,从而使模型的预测值逐渐接近真实值,实现斑马线分割检测任务。因此,需要对数据集进行人工手动标注,本文采用 Lbaleme 图像标注工具对每张图片进行手工标

注,获取分割 label,标注文件以 Json 格式保存,主要记录标注点的位置、类别标签。标注过程如图 6 所示。

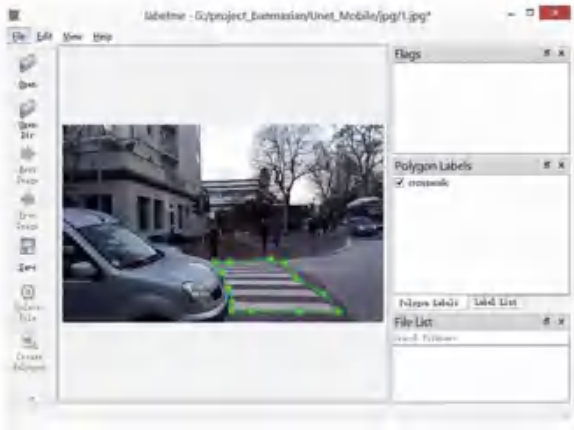


图 6 斑马线标注

Fig. 6 Zebra mark

完成标注后,将网页获取的数据划分为训练集、验证集,对模型进行预训练。然后将预训练模型在自制数据集上进行调优和测试。

4 网络模型训练与实验结果分析

4.1 MobileNetv2 改进前后的对比实验

本文首先对 MobileNetv2 与改进后的 SEMobileNetv2 进行对比实验。通常使用分类数据集对卷积神经网络进行分类精度的测试对比,本文选取 Cifar10 数据集进行实验,该数据集包含 10 个类别,其中 50 000 张为训练集,10 000 张为测试集;深度学习框架选用 Keras,优化策略为 Adam,初始学习率(learning rate)为 0.001,当学习率在连续 3 次迭代中保持不变时,学习率变为原来的 0.5 倍,batch_size 为 32,迭代次数(epoch)为 20 000 次。实验结果如图 7 所示。

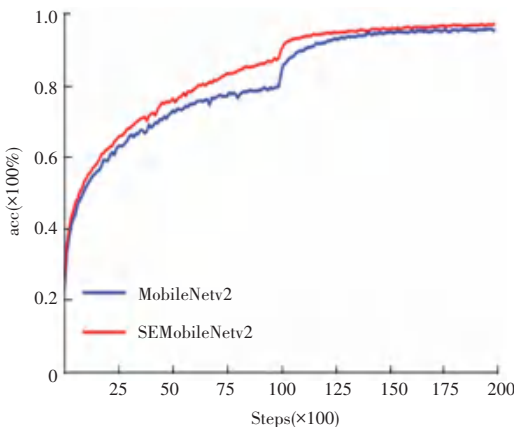


图 7 改进前后分类精度对比

Fig.7 Comparison of classification accuracy before and after improvement

由图 7 可知,通过 20 000 次的迭代,改进后的 SEMobileNetv2 分类精度比之前精度提升了 1.6%。

通过嵌入 SENet 模块对模型的特征提取能力有一定的提升。

4.2 基于 SEMobileNetv2 的斑马线分割实验

将 Unet 特征提取网络替换为 SEMobileNetv2,搭建分割模型。分别对改进前后的 Unet 分割网络进行训练,其训练集 8 000 张,测试集 500 张,训练参数见表 1。

表 1 实验环境

Tab. 1 Experimental environment

硬件环境	软件环境	初始学习率	优化策略	学习率更新
GPU 1080Ti	Keras1.8,python3.5	Lr = 0.001	Adam	分段常数衰减

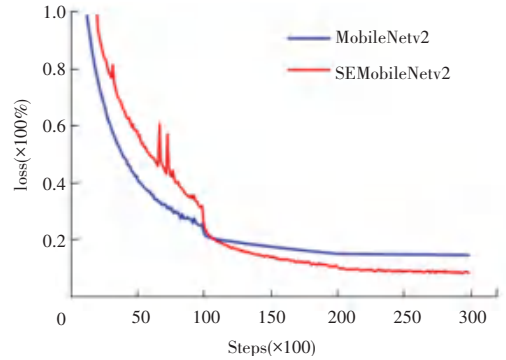


图 8 loss 值对比

Fig. 8 Comparison of loss

经过 3 万次的迭代,由图 8 可知:基于 SEMobileNetv2 的 Unet 分割网络的 loss 收敛效果更佳。

获取两种 Unet 模型后,本文分别在不同环境中对斑马线进行了测试,测试效果如图 9 所示。在有车辆行驶的条件下和无障碍物的条件下,改进后的 Unet 网络模型的分割精度均比原来的分割网络提高了 1.8%。进一步验证了改进模型的有效性。



(a) SEMobileNetv2 斑马线分割 (b) MobileNetv2 斑马线分割
(a) SEMobileNetv2 divides the zebra crossing (b) MobileNetv2 divides the zebra crossing

图 9 分割效果图对比

Fig. 9 Comparison of segmentation effect

(下转第 69 页)