

基于卷积神经网络的道路障碍物检测算法^{*}

权鸿斌 姚剑敏 福州大学物理与信息工程学院

摘要：针对传统道路障碍物检测算法准确性和鲁棒性不强等问题，本文提出一种基于卷积神经网络的道路障碍物检测算法。该算法首先对车载图像预处理生成障碍物候选区域，再将障碍物候选区域输入到改进的卷积神经网络中，进行精确识别和剔除，区分道路障碍物和非障碍物。改进的卷积神经网络在原有网络的基础上，调整了卷积核的大小和个数、池化层的空间尺寸和神经网络的深度，并且在卷积层后选择性的加入池化层，提高障碍物的识别率。在不同场景中进行了测试，实验结果表明：本文提出的道路障碍物检测算法有效的提高了障碍物的识别率，识别率达到98.2%，并且拥有较高的鲁棒性。

关键词：障碍物 感兴趣区域 阈值分割 候选区域 卷积神经网络

1 引言

随着时代的发展与社会的进步，人们生活水平逐渐提高，家家户户都有了汽车，并且成为他们出行的主要交通工具。由于车辆增多，交通事故的发生率也随之上升，因此车辆辅助驾驶系统中对于道路中障碍物的检测成为重要的研究方向。在道路中的障碍物，占比例最大的是汽车，其次就是行人，这里把骑着自行车、摩托车的人也归为行人这一类，因为大部分对这些障碍物的检测算法都是以人为目标进行检测的，这样两类障碍物所占道路障碍物的比例将超过90%，剩余的障碍物为一些路面静止的异物，如坑洞等。由于车辆和行人所占道路中的比率最高，并且在道路障碍物中威胁度最高。因此，本文主要检测道路中的车辆和行人。

传统的障碍物检测算法主要有帧间差分法、运动光流检测法、边缘特征检测法、浅层机器学习算法。帧间差分法是检测视频中相邻两帧或多帧图像之间发生变化的区域，计算差分图像并进行二值化，然后提取运动目标，但是该方法不适合检测与摄像头相对静止的障碍物。光流法用于检测运动的物体，并且光流的计算繁琐，实时性较差。单纯基于边缘特征检测算法的鲁棒性和准确性较差，容易出现漏检和误检的情况。浅层机器学习算法的思想是让机器通过大量样本的训练与学习，最后将待检测图像输入到训练好的模型中分类识别。浅层机器学习算法主要有BP算法、AdaBoost算法、SVM算法、最大熵方法（如LR，逻辑回归算法），这些算法掀起了基于统计模型的机器学习热潮。文献[8]利用图像差分的方法分割背景图

像，再结合BP神经网络对车辆进行分类，适用于复杂道路环境。文献[9]用类Haar特征作为车辆检测的提取特征，采用AdaBoost分类器检测车辆，车辆识别率达到84.1%。文献[10]利用SVM分类器结合HoG特征，采用主动学习的方法对车辆进行检测。

浅层的机器学习算法均是人工提取图像特征，如HOG特征、LBP特征等，人工选取的特征不能很好的适合样本的多样性，所以，近年来，涌现了一股以深度学习为主要框架的研究浪潮。通过深层次、模仿人类大脑学习的方式学习，让机器从大量样本中去提取图像的特征，这样的好处在于提取的图像特征更适合于样本本质的特征，从而提高目标的识别率。在Hinton教授提出深度置信网络(DBN)后，各个学者相继提出了基于深度学习框架的算法。其中卷积神经网络算

^{*}本文所属基金项目：国家重点研发计划资助(No:2016YFB0401503)，福建省科技重大专项(No:2014HZ0003-1)，广东省科技重大专项(No:2016B090906001)，福建省资助省属高校专项课题(No:JK2014002)。

法 (CNN) 被应用到图像处理的各个领域, 如人脸识别、字符识别。CNN 算法的模型主要有 LeNet-5 模型、AlexNet 模型、VGG 模型、GoogleNet 模型、ResNet 模型等, 后三者的模型计算较为复杂。本文选取 LeNet-5 模型作为卷积神经网络的核心并对其改进, 提高模型的识别率。采用假设加验证的检测方法, 对产生的障碍物候选区域进行分类验证。

2 障碍物候选区域生成

为了减少数据的计算量, 选取图像三分之一以下的道路部分作为检测域。对检测域进行灰度化和中值滤波处理, 接下来对预处理后的灰度图进行灰度直方图分析和阈值分割。为了解决传统阈值分割算法在分割道路与障碍物时存在的误分割问题, 采用改进的基于道路灰度直方图的自适应阈值分割算法分割障碍物与道路背景, 选取分割效果最好的参数 $m-3s$ 作为分割阈值, 对比图如图 1 所示。

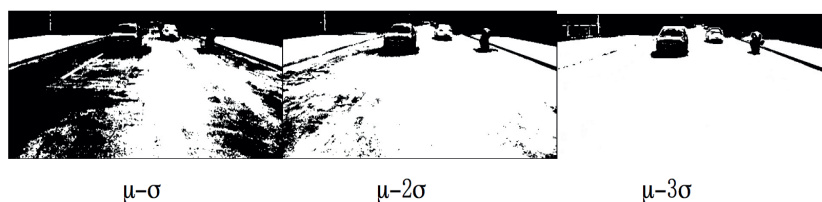


图 1 不同参数的分割情况

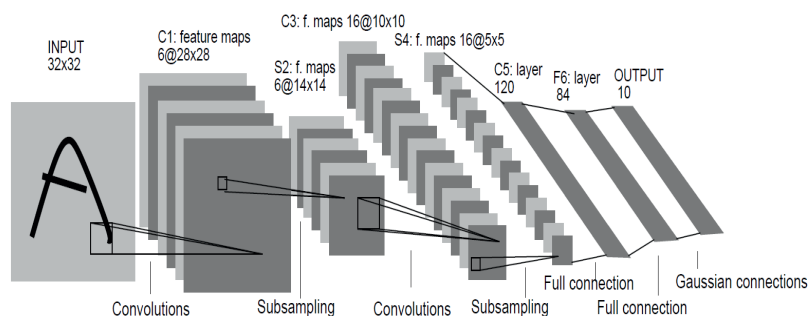


图 2 LeNet-5 网络架构

本文所采用的图片大小为 640×480 , 对于道路中的车辆、行人等障碍物, 它们均有一定的尺寸范围, 候选区域的大小需要根据障碍物反映在二值图像中的宽度和高度所占像素的个数来确定。障碍物距车体远近不同, 其宽度所占的像素个数也会不同, 但是会有一定的范围。因此本文采用基于障碍物宽度的候选区域提取方法, 根据不同宽度的障碍物生成不同大小的候选区域。

3 障碍物检测

3.1 卷积神经网络介绍

卷积神经网络 (CNN) 是深度学习中一种多层次的网络架构, 它由卷积层、池化层、全连接层三部分组成。卷积层的作用对前一层输出的特征图进行卷积运算, 进一步提取特征。池化层用于压缩特征图, 一方面降低计算量, 另一方面为了提取主要特征。最后全连接层用于连接提取的全部特

征, 送入分类器进行分类。一般池化层设在卷积层之后, 两者交替组成, 最后加上全连接层构成卷积神经网络。卷积神经网络的核心在于其通过局部连接和权值共享等方式大大降低了计算量。LeNet-5 是卷积神经网络中比较经典的网络架构, 最初用于手写体字符的识别, 后来学者们将其应用于生活的各个领域。图 2 为 LeNet-5 的网络架构。

LeNet-5 网络模型输入图片大小采用 32×32 , 包括 3 个卷积层、2 个池化层、1 个全连接层。卷积层都采用的是 5×5 大小的卷积核, C1 层有 6 个特征图, 经过卷积后大小为 28×28 。经过下采样提取后, S2 层输出 6 个 14×14 大小的特征图。C3 层将特征图的数量扩大为 16 个, 经卷积后输出特征图大小为 10×10 。再经下采样提取后, S4 层输出 16 个 5×5 大小的特征图。C5 层由 120 个 1×1 大小的特征图组成, 最后由 F6 层全连接层将 C5 层输出的所有特征进行连接, 输入到分类器中进行分类。

3.2 改进的卷积神经网络

本文对 LeNet-5 模型进行改进, 提高模型的识别率, 使其适合本系统的研究。首先输入的图片大小采用 64×64 , 并提高模型的深度, 网络包括 5 个卷积层、3 个下采样层和 3 个全连接层。5 个卷积层中, C4 层和 C5 层后不使用下采样层, 其余三层后均有一个下采样层。激活函数采用 ReLU, 以解决反向传播时求误差梯度计算量大的问题, 加快收敛速度。最后将最后一层的 Softmax 层改为 3 分类, 分别为汽车类、行人类和背景类, 如图 3 所示。网络各层所有参数如表 1 所示。

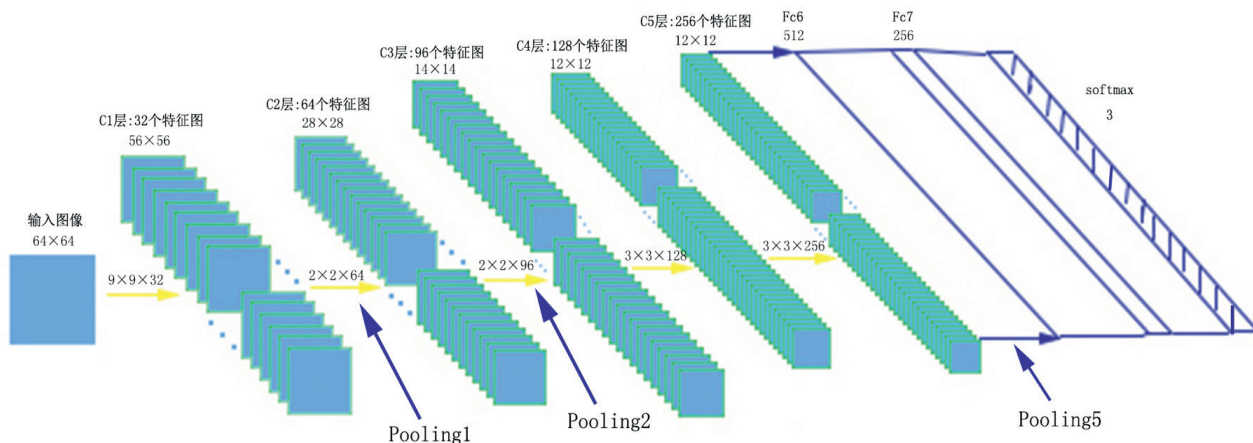


图3 卷积神经网络架构

表1 各层网络架构参数

Layer	Input Size	Kernel Size	Stride	Pad	Output Size	Feature Map
C1	64 × 64	9 × 9	1	0	56 × 56	32
S1	56 × 56	2 × 2	2	0	28 × 28	32
C2	28 × 28	3 × 3	1	1	28 × 28	64
S2	28 × 28	2 × 2	2	0	14 × 14	64
C3	14 × 14	3 × 3	1	1	14 × 14	96
C4	14 × 14	3 × 3	1	0	12 × 12	128
C5	12 × 12	3 × 3	1	1	12 × 12	256
S5	12 × 12	2 × 2	2	0	6 × 6	256

为了避免图像边缘信息在卷积时被忽略,采用 Zero Padding 的方式充分利用图片边缘信息,并在卷积时保持图片的原始尺寸。卷积核的大小采用先大后小的方式,符合人类观察事物由粗到精的特点。下采样层的选取,可以在一定程度上减少网络中需要计算的参数数量,弱化网络对目标扭曲、缩放、位移的敏感度。但下采样层计算会丢失目标的一些特征,为了弥补这些损失,采用在卷积层中增加目标特征图数量的方式,以提高识别率。

Softmax 输出层输出为三类,每一类会对应一个概率值。该概率值可用式(1)

进行表示。

$$p(y=i|x;\theta)=\frac{\exp(\theta_i^T x)}{\sum_{j=1}^3 \exp(\theta_j^T x)} \quad (1)$$

其中 i 表示分类的个数, θ 表示训练模型参数。

使用代价函数 J 来评价卷积神经网络的准确率,可求得网络代价函数的最小值来调节网络中各层的参数,经过反复迭代提高神经网络的识别率,如式(2)所示。

$$J(\theta)=-\frac{1}{m} \left[\sum_{i=1}^m \sum_{j=1}^3 1\{y^{(i)}=j\} \log \frac{\exp(\theta_j^T x^{(i)})}{\sum_{l=1}^3 \exp(\theta_l^T x^{(i)})} \right] \quad (2)$$

其中 $1\{\cdot\}$ 表示示性函数, $y^{(i)}$ 表示第 i 张样本属于的真实类别, $x^{(i)}$ 为

第 i 张输入样本。

4 实验结果分析与评价

4.1 网络训练

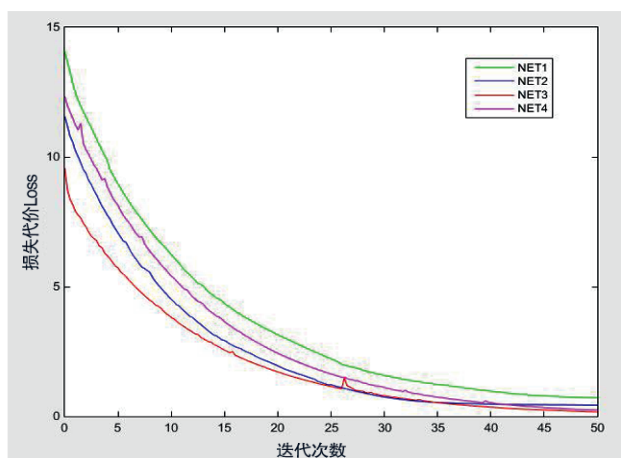
本文选用车辆样本数量为 10000 张,行人样本数量 10000 张,背景样本(无车辆、无行人)10000 张,部分训练样本如图 4 所示。其中 80% 作为训练集,20% 作为测试集。将样本归一化处理后,得到 64×64 大小像素的图像作为网络训练的输入。初始化网络中各层参数,为了避免网络进入饱和区,加快训练速度,采用较小的权值参数,因此选取区间 $[-0.015, 0.015]$ 中的随机值作为初始化参数。

4.2 实验数据对比

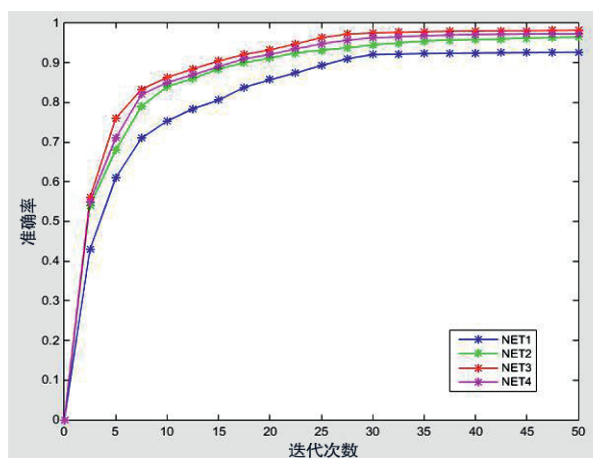
本文调整了卷积核的大小和个数、池化层的空间尺寸和网络的深度,实验探究了 4 种不同卷积网络架构,如表 2 所示。4 种网络在 C1 层采用了不同的卷积核,分别为 5×5 、 7×7 、 9×9 、 11×11 ,探究不同卷积核对卷积神经网络准确率的影响。经过 5 万次迭代,输出网络的损失代价 Loss 训练结果,并使用测试集验证每个网



图4 部分训练样本示例



(a) 损失代价 loss 图



(b) 准确率验证图

图5 四种网络训练结果

表2 不同网络结构参数

网络	C1	S1	C2	S2	C3	C4	C5	S5
Net1	5×5	2×2	3×3	2×2	3×3	3×3	3×3	3×3
Net2	7×7	2×2	3×3	3×3	3×3	5×5	3×3	2×2
Net3	9×9	2×2	3×3	2×2	3×3	3×3	3×3	2×2
Net4	11×11	2×2	3×3	3×3	3×3	3×3	3×3	2×2

表3 四种网络准确率

网络	NET1	NET2	NET3	NET4
准确率 %	92	96.5	98.2	97.3

表4 不同障碍物检测算法对比

方法	本文算法	类 Haar+AdaBoost	HOG+SVM
准确率 %	98.2	90.6	92.3

网络的准确率,结果如图5所示。在图5(a)中可看出在迭代5万次后,4种网络均达到收敛,Net1网络Loss最高收敛于0.5,Net3网络Loss最低收敛于0.02。为了检测网络是否存在过拟合现象,需要对网络进行准确率的验证。如图5(b)准确率验证中,Net3网络准确率最高,其次为Net4、Net2、Net1网络测量的准确率最低,四种卷积神经网络准确率的具体数值如表3所示。由此可得出结论,采用9×9大小的卷积核的网络性能最好,识别率最高,也是本文所采用的神经网络架构。卷积核的大小如同人类获取外界信息的视野大小,需要选择合适的感受野。感受野过大,提取的信息过于复杂,

冗余信息过多,不能很好的表示目标的特征;感受野过小,提取的信息不够全面,导致网络的准确率过低。

将本文的算法同传统障碍物检测算法进行对比,对比的传统障碍物检测算法包括类 Haar+AdaBoost 算法、HOG+SVM 算法,使用本文的样本集训练、验证,验证结果如表 4 所示。

上述对比结果表明,本文提出的障碍物检测算法优于传统障碍物检测算法,识别率可达 98.2%。由于假设区域中已经包含障碍物,并且假设区域内包含极少的背景区域,不需要使用滑窗机制在图中选择大量的窗口,这就使得检测的时间大大降低,系统处理每幅图像的时间大约为 100ms 到 120ms,基本满足实时性的要求。

最后采取了高速公路、城镇道路等场景图像进行测试,测试结果如图

6 所示。结果中分别展示了分割后的二值图、障碍物宽度提取图、候选区域生成图、最终确认图。在图 6(a)(b)(c)(d) 中可看出行人和车辆均被正确检测出,伪车辆区域全部被剔除,车辆使用绿色的矩形框框出,行人使用红色的矩形框框出,说明本文设计的卷积神经网络具有较高的识别精度。

5 结论

提出一种改进的基于卷积神经网络的道路障碍物检测算法,将图像预处理生成障碍物候选区域后,再将车辆候选区域输入到具有较高精度的卷积神经网络中验证,剔除伪障碍物区域,最后实现了道路车辆、行人的识别。经实验验证,本文提出的道路障碍物检测算法,实现了复杂道路情况下的多目标障碍物检测,障碍物识别率可

达 98.2%,具有较强的鲁棒性和精确性。

参考文献

- [1]Ting Jin, Fugen Zhou, Xiangzhi Bai. Moving Vehicles Detection in Airborne Video[J].Institute of Electrical and Electronic Engineers.2008:316-321.
- [2]陈存祺.汽车自动驾驶中基于单目视觉的障碍物检测[D].吉林:吉林大学计算机科学与技术学院,2013.
- [3]卞建勇,高爽,徐建阔.基于帧差与统计法的运动目标背景图像提取算法[J].微电子学与计算机,2011,28(4):91-94.
- [4]Cheng Y H, Wang J. A Motion Image Detection Method Based on the Inter-Frame Difference Method[J]. Applied Mechanics & Materials,2014,490-491:1283-1286.
- [5]郭文俊,常桂然.基于光流技术的障碍物检测方法[J].软件,2016,37(4):64-68.
- [6]钟彩.边缘检测算法在图像预处理中的应用[J].软件,2013,34(1):158-159.
- [7]王蒙蒙.基于单目视觉的车辆前方通行性分析方法[D].大连理工大学,2014.
- [8]Zhi-Pan W U,Research on Vehicle Type Recognition Based on BP Neural Network[J].Modern Computer,2013.
- [9]金立生,王岩,刘景华,等.基于 Adaboost 算法的日间前方车辆检测[J].吉林大学学报:工学版,2014,44(6):1604-1608.
- [10]Sivaraman S,Trivedi M M.Active learning for on-road vehicle detection: a comparative study[J].Machine Vision and Applications, 2014,25(3):599-611.
- [11]Hinton G E,Osindero S,Teh Y W.A fast learning algorithm for deep belief nets[J].Neural Computation,2006,18(7):1527-1554.
- [12]Sri S. Gaussian Mixture Model and Deep Neural Network based Vehicle Detection and Classification[J]. International Journal of Advanced Computer Science and Applications, 2016,7(9):17-25.
- [13]Zhang X, Zou J, He K, et al. Accelerating Very Deep Convolutional Networks for Classification and Detection[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2015, 38(10):1943-1955.
- [14]Zhou T, Han G, Li B N, et al. Quantitative analysis of patients with celiac disease by video capsule endoscopy: A deep learning method[J]. Computers in Biology & Medicine, 2017, 85:1.
- [15]He K, Zhang X, Ren S, et al. Deep Residual Learning for Image Recognition[J]. 2015:770-778. CATV

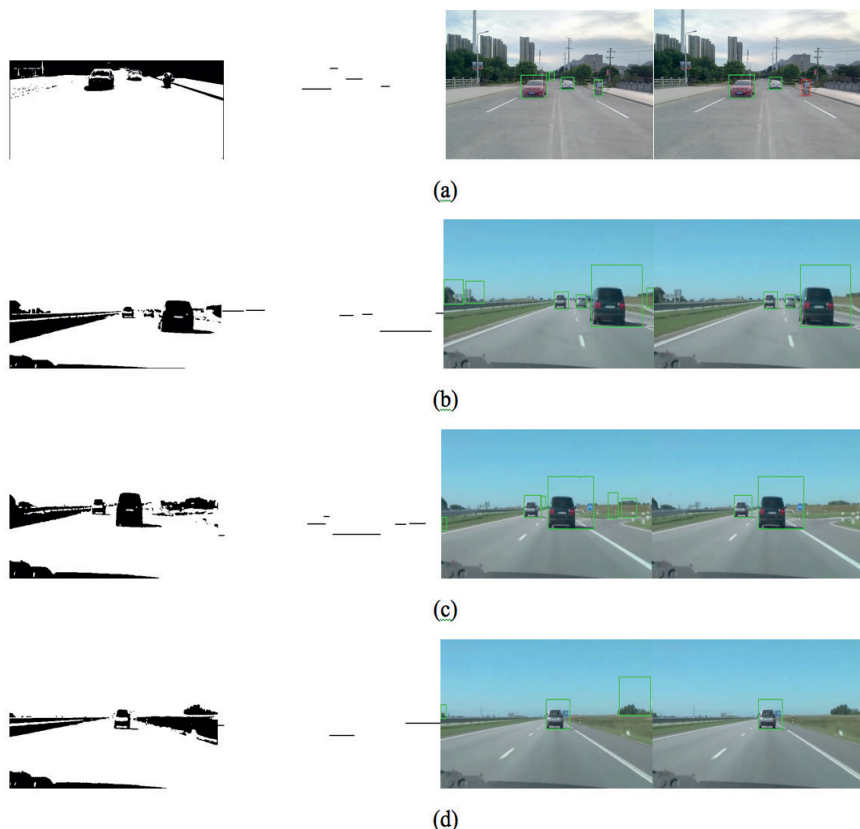


图 6 障碍物验证结果