

基于轻量型卷积神经网络的交通标志识别方法^①



程 越, 刘志刚

(东北石油大学 计算机与信息技术学院, 大庆 163318)

通讯作者: 刘志刚, E-mail: dqpilzg@163.com

摘 要: 交通标志识别设备的功耗和硬件性能较低, 而现有卷积神经网络模型内存占用高、训练速度慢、计算开销大, 无法应用于识别设备. 针对此问题, 为降低模型存储, 提升训练速度, 引入深度可分离卷积和混洗分组卷积并与极限学习机相结合, 提出两种轻量型卷积神经网络模型: DSC-ELM 模型和 SGC-ELM 模型. 模型使用轻量化卷积神经网络提取特征后, 将特征送入极限学习机进行分类, 解决了卷积神经网络全连接层参数训练慢的问题. 新模型结合了轻量型卷积神经网络模型内存占用低、提取特征质量好以及 ELM 的泛化性好、训练速度快的优点. 实验结果表明, 与其他模型相比, 该混合模型能够更加快速准确地完成交通标志识别任务.

关键词: 轻量型卷积神经网络; 交通标志识别; VGG16 网络; 极限学习机

引用格式: 程越, 刘志刚. 基于轻量型卷积神经网络的交通标志识别方法. 计算机系统应用, 2020, 29(2): 198-204. <http://www.c-s-a.org.cn/1003-3254/7264.html>

Traffic Sign Recognition Method Based on Lightweight Convolutional Neural Network

CHENG Yue, LIU Zhi-Gang

(School of Computer and Information Technology, Northeast Petroleum University, Daqing 163318, China)

Abstract: Traffic sign recognition equipment has low power consumption and hardware performance, while the existing convolutional neural network model has high memory footprint, slow training speed, and high computational overhead, which cannot be applied to the recognition equipment. To solve this problem, in order to reduce model storage and improve training speed, deep separation convolution and mixed wash grouping convolution are introduced and combined with the ultimate learning machine. Two lightweight convolutional neural network models are proposed: DSC-ELM model and SGC-ELM model. The proposed models use the lightweight convolutional neural network to extract the features, and then send the features to the extreme learning machine for classification, which solve the problem of slow parameter training in the full connection layer of the convolutional neural network. The new models combine the advantages of lightweight convolutional neural network model with low memory footprint, good feature extraction quality, good generalization of ELM, and fast training and classification. Experimental results show that compared with other models, the hybrid model can accomplish traffic sign recognition tasks more quickly and accurately.

Key words: lightweight convolutional neural network; traffic sign recognition; VGG16 network; extreme learning machine

随着城市现代化建设和人工智能技术的崛起, 现代城市智能化交通系统得到了进一步发展. 道路交通

标志识别技术作为无人驾驶和智能辅助驾驶系统的重要组成部分, 在智能化交通系统的建设中有重要的战

① 基金项目: 黑龙江省自然科学基金 (LH2019F042); 大庆市科技计划 (zd-2019-08); 东北石油大学创新基金 (2017PYZL-06, 2018YDL-22)

Foundation item: Natural Science Foundation of Heilongjiang Province (LH2019F042); Daqing Science and Technology Project (zd-2019-08); Innovation Fund of Northeast Petroleum University (2017PYZL-06, 2018YDL-22)

收稿时间: 2019-07-09; 修改时间: 2019-07-26; 采用时间: 2019-08-07; csa 在线出版时间: 2020-01-16

略意义^[1],如何实时高效地完成交通标志识别任务成为国内外研究的热点。

对于交通标志的识别方法可分为两类,基于手工选取典型特征的方法和深度学习中的卷积神经网络方法 CNN^[2]。基于手工典型特征的方法如方向梯度直方图 HOG^[3]、尺度不变特征变换 SIFT^[4]、局部二值模式 LBP^[5]等,这类方法基于人工设计获取特征,数据预处理工作繁琐,提取的特征局限性大。而 CNN 无需对图像做复杂预处理,直接以像素作为输入,在输入到识别的过程中提取图像特征,具有很强的自适应性和学习能力,成为了图像识别领域的核心模型。

在 ILSVRC 2012 竞赛中 AlexNet^[6]取得了冠军,之后 GoogleNet^[7], VGGNet^[8], ResNet^[9]等模型又被相继提出, CNN 在图像分类任务上的精度不断提高。然而,现有 CNN 应用于嵌入式交通标志识别设备有以下不足: (1) 模型内存占用高,无法嵌入交通标志识别设备。以 AlexNet 为例,其参数一共 6000 万之多,在 Imagenet-1000 上的训练模型达到 200 MB,而 VGG16 模型参数更是达到了 1.38 亿。(2) 模型计算开销大,小型嵌入式设备无法支持运算。现有 CNN 模型在训练和使用中需要在高性能设备上进行,如高性能 CPU 和 GPU 等,这些设备往往体积大、价格高、计算开销大,而交通标志识别设备的功耗和硬件性能较低,无法支持计算。(3) 现有 CNN 模型的全连接层参数训练慢,无法满足识别设备的实时性。

针对以上问题,本文引入两种轻量化卷积方法:深度可分离卷积^[10]和混洗分组卷积^[11]并与极限学习机^[12]相结合,提出两种轻量化卷积神经网络模型: DSC-ELM 模型和 SGC-ELM 模型。新模型首先利用轻量化卷积方法构建卷积层,减少卷积层的参数使模型轻量化,然后以 ELM 作为分类器,接受经卷积层提取的特征进行分类。设计实验,在德国交通标志数据集^[13]和比利时交通标志数据集^[14]上,通过 VGG16 模型、基于深度分离卷积的 MobileNet 模型、基于混洗分组卷积的 ShuffleNet 模型与新模型 DSC-ELM 和 SGC-ELM 进行对比实验。实验结果表明:新模型在识别精度上高于其他模型,并且训练速度比 MobileNet 和 ShuffleNet 模型更快。新模型有着模型轻量化、可嵌入性高、训练速度快以及识别精度高的优点,使得新模型在分类识别交通标志的总体性能较现有 CNN 模型更加优秀。

1 轻量型卷积神经网络

1.1 深度可分离卷积

深度可分离卷积 (Depthwise Separable Convolution, DSC) 的创新点在于将卷积操作分为两部分:深度卷积 (depthwise convolution) 和 1×1 卷积 (pointwise convolution)。如图 1 所示,深度可分离卷积先保持通道间的分离,进行深度卷积,然后再进行 1×1 卷积。

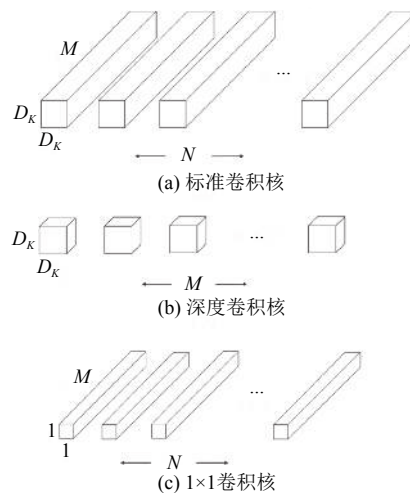


图 1 卷积核

普通卷积与深度可分离卷积的参数对比: 设 M 为通道数,卷积核的高度为 D_K ,输入的宽和高为 D_F ,卷积核的个数为 N 。普通卷积的计算量为:

$$D_K \cdot D_K \cdot N \cdot M \cdot D_F \cdot D_F$$

深度可分离卷积的计算量为:

$$D_K \cdot D_K \cdot M \cdot D_F \cdot D_F + N \cdot M \cdot D_F \cdot D_F$$

深度可分离卷积与普通卷积计算量之比为:

$$\frac{D_K \cdot D_K \cdot M \cdot D_F \cdot D_F + N \cdot M \cdot D_F \cdot D_F}{D_K \cdot D_K \cdot N \cdot M \cdot D_F \cdot D_F} = \frac{1}{N} + \frac{1}{D_K \cdot D_K}$$

由此可得,使用深度可分离卷积代替普通卷积来构建模型可使模型轻量化。标准卷积核和深度可分离卷积核如图 1 所示。

1.2 混洗分组卷积

混洗分组卷积 (Shuffled Grouped Convolution, SGC) 是将分组卷积法和通道混洗机制结合起来。分组卷积法如图 2 所示,将卷积核拆分为两组进行分组卷积,每组包含 $D_{out}/2$ 个卷积核,分别与对应部分的输入进行卷积。每组卷积核卷积完成后都会得到 $D_{out}/2$ 通道的输出,然后将通道堆叠最终得到所需的 D_{out} 通道输

出. 此时卷积的计算量为:

$$(H \times W \times D_{in}/2 \times D_{out}/2) \times 2$$

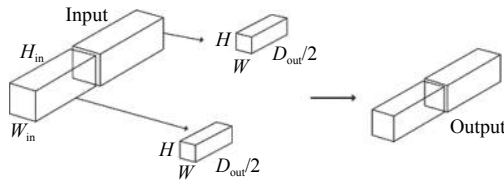


图2 分组卷积法

相比普通卷积的计算量减少一半, 当分组数增加时, 模型的参数也就随之减少.

分组卷积虽然可以减少模型参数, 但是每组卷积核仅对输入的固定部分向后传递的信息进行处理, 这样就局限了卷积核学习的特征, 阻碍了训练过程中通道组之间的信息流动, 减弱了特征表示. 解决以上问题的方法就是在分组卷积的基础上引入通道混洗机制.

通道混洗机制如图3所示, 将卷积核分为三组的分组卷积, 在执行一次卷积操作后, 将得到的特征拆分为若干组, 混洗后再输入下一次分组卷积. 利用分组卷积法和通道混洗机制相结合的方法构建卷积层, 可以减少模型参数量, 加强通道组间的信息流通, 增强特征表示能力.

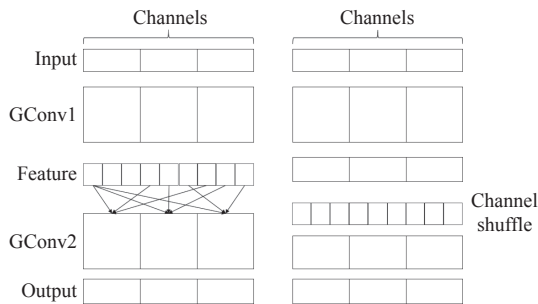


图3 通道混洗机制

2 极限学习机

现有 CNN 网络利用全连接层进行分类, 需要依靠梯度下降训练算法更新参数, 导致模型学习慢, 而且泛化能力较差, 易陷入局部最小^[15]. 极限学习机 (Extreme Learning Machines, ELM) 是由南洋理工大学黄广斌教授提出的求解单隐层前馈神经网络的学习算法. 极限学习机只需要人为确定隐含层中神经元的个数, 输入层到隐含层的权重为一次随机确定并且在算法执行过程中无需再调整, 隐含层到输出层的权重则是通过求

广义逆的方法得到, 无需经过迭代. 由于极限学习机独特的求参机制, 将 CNN 全连接层替换为 ELM 可以加快模型训练速度, 提高泛化性.

极限学习机结构如图4所示. 假设有 N 个任意样本 (X_i, t_i) , 其中:

$$\begin{cases} X_i = [x_{i1}, x_{i2}, \dots, x_{in}]^T \in R^n \\ t_i = [t_{i1}, t_{i2}, \dots, t_{im}]^T \in R^m \end{cases}$$

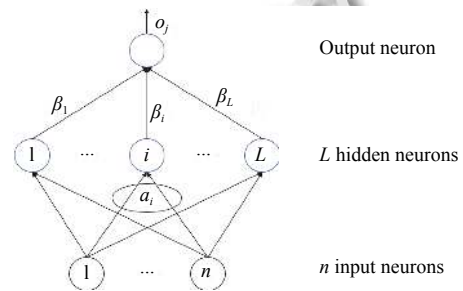


图4 极限学习机

具有 L 个隐层节点的单隐层神经网络可以表示为:

$$\sum_{i=1}^L \beta_i g(W_i \cdot X_j + b_i) = o_j, j = 1, \dots, N \quad (1)$$

其中, $g(x)$ 为激活函数, $W_i = [w_{i,1}, w_{i,2}, \dots, w_{i,n}]^T$ 为输入权重, β_i 为输出权重, b_i 是第 i 个隐层单元的偏置, $W_i \cdot X_j$ 是 W_i 和 X_j 的内积, $o_i = [o_{i1}, o_{i2}, \dots, o_{im}]^T$ 表示网络输出值. ELM 成本函数 E 为:

$$E = \sum_{j=1}^N \|o_j - t_j\| \quad (2)$$

学习目标是使得 E 最小, 即存在 β_i, W_i 和 b_i 使得:

$$\sum_{i=1}^L \beta_i g(W_i \cdot X_j + b_i) = t_j, j = 1, \dots, N \quad (3)$$

矩阵表示为:

$$H\beta = T \quad (4)$$

其中, H 为神经网络隐层节点的输出矩阵, β 为输出权重, T 为期望输出.

$$H(W_1, \dots, W_L, b_1, \dots, b_L, X_1, \dots, X_L) = \begin{bmatrix} g(W_1 \cdot X_1 + b_1) & \dots & g(W_L \cdot X_1 + b_L) \\ \vdots & \dots & \vdots \\ g(W_1 \cdot X_N + b_1) & \dots & g(W_L \cdot X_N + b_L) \end{bmatrix}_{N \times L} \quad (5)$$

$$\beta = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_L^T \end{bmatrix}_{L \times m} \quad T = \begin{bmatrix} T_1^T \\ \vdots \\ T_N^T \end{bmatrix}_{N \times m}$$

式中, W 为输入层到隐含层的权值; b 为隐含层偏置; $g()$ 为激活函数. 极限学习机的思想为: 随机确定 W_i 和 b_i 进而可得到输出矩阵 H , 再通过式 (6) 求得输出权重 β .

$$\hat{\beta} = H^+ T \quad (6)$$

3 轻量级卷积神经网络与 ELM 相结合

现有 CNN 模型内存占用高、计算开销大、训练速度慢, 无法适用于移动交通标志识别设备. 模型采用轻量化卷积代替普通卷积的策略, 可实现在牺牲合理精度的基础上极大减少卷积层的参数, 提升训练速度. 然而, 模型的全连接层在训练过程中需要耗费大量时间确定自身参数, 所以要想使模型训练更加快速, 除修改卷积层外, 还应对分类器进行改进. ELM 有着独特的求参机制, 将其作为分类器有训练速度快、泛化性强的优点. 但 ELM 对输入特征的质量有较高的要求.

根据以上情况, 将轻量级卷积神经网络与 ELM 结合起来可以充分发挥各自的优势: 使用轻量化卷积方法构建卷积层可使模型轻量化, 并且提取的特征可靠, 满足 ELM 对特征的高要求^[16]; ELM 分类器独特的求参机制, 使得模型分类更加快速准确.

3.1 DSC-ELM 模型

DSC-ELM (Depthwise Separable Convolution-ELM)

模型结构可分为两部分:

(1) 特征提取部分: 特征提取部分采用深度可分离卷积构建卷积层. 此部分包含两个普通 3×3 卷积和 17 个 bottleneck. 每个 bottleneck 包含以下 3 个操作: 首先经过 pointwise convolution 扩展通道, 扩展倍数除第一个 bottleneck 为 1 倍外均为 6 倍; 再进行 depthwise convolution; 最后再利用 pointwise convolution 压缩通道. 这种“先扩张, 后压缩”的设计可以通过将通道数扩大, 从中间层获得更多特征. 图像输入模型后首先进行 3×3 卷积, 然后将得到的特征图送入 17 个 bottleneck 中, 经过 17 个 bottleneck 后, 再进行一次 3×3 卷积, 特征提取部分结束. 表 1 为特征提取部分结构.

表 1 DSC-ELM 特征提取部分结构

操作类型	扩展倍数	输出通道	重复次数	步长
Conv2d_bn_relu	None	32	1	2
Bottleneck	1	16	1	1
Bottleneck	6	24	2	1
Bottleneck	6	32	3	2
Bottleneck	6	64	4	2
Bottleneck	6	96	3	1
Bottleneck	6	160	3	2
Bottleneck	6	320	1	1
Conv2d_bn_relu	None	1280	1	1

(2) 分类部分: 模型使用 ELM 分类器替换卷积神经网络中的全连接层. 将特征提取部分提取的特征展平后作为 ELM 分类器的输入, 然后进行分类.

DSC-ELM 模型总体结构如图 5 所示.

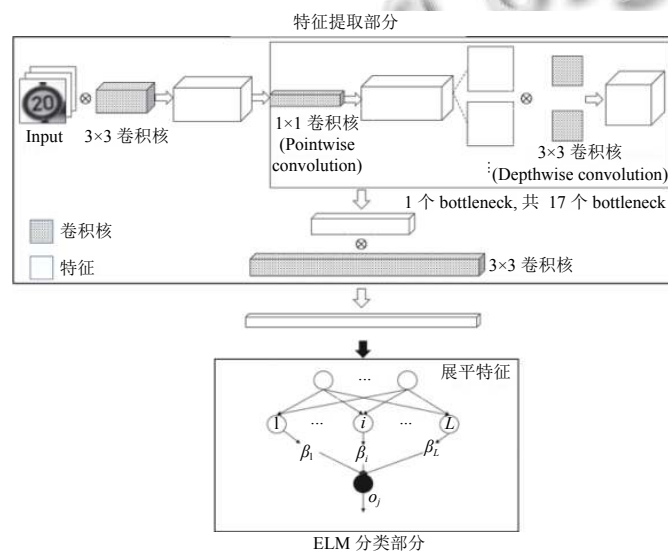


图 5 DSC-ELM 模型结构

3.2 SGC-ELM 模型

SGC-ELM (Shuffled Grouped Convolution-ELM) 模型结构可分为两部分:

(1) 特征提取部分: 特征提取部分采用混洗分组卷积构建卷积层. 此部分包括: 1个 3×3 卷积、1个 3×3 的最大池化、3个 Stage 以及 GlobalPool. 每个 Stage 包含若干个 ShuffleNet unit, ShuffleNet unit 结构如图 6 所示.

图像输入模型后先经过 3×3 卷积和 3×3 的最大池化操作, 然后将特征送入 3 个 Stage. 经过 3 个 Stage 后, 再进行 1 次 GlobalPool 操作, 特征提取部分结束. 表 2 为特征提取部分结构.

(2) 分类部分: 将特征提取部分提取的特征展平后输入 ELM 进行分类.

SGC-ELM 模型总体结构如图 7 所示.

3.3 新模型的训练

(1) 首先对模型进行预训练. 预训练模型的目的是保证网络卷积层提取的特征是有效的, 若不进行预训练, 模型提取的特征质量低, 影响最终的分类结果;

(2) 预训练结束后, 将图像输入模型提取特征;

(3) 分类器训练. 将提取的特征展平后输入 ELM 分类器, 通过公式 $\beta = H^+T$ 求得极限学习机隐含层到输出层的权值参数, 模型训练结束.

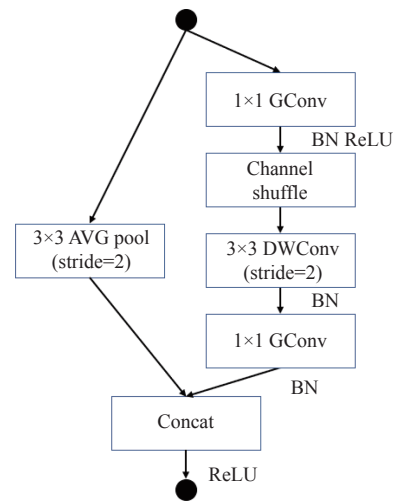


图 6 ShuffleNet unit 结构

表 2 SGC-ELM 特征提取部分结构

层	输出	内核	步数	重复次数	输出通道
Image	224×224				3
Conv1	112×112	3×3	2	1	24
MaxPool	56×56	3×3	2	1	24
Stage2	28×28		2	1	116
Stage3	14×14		2	1	232
Stage4	7×7		2	1	464
Conv5	7×7	1×1	1	1	1024
GlobalPool	1×1	7×7			

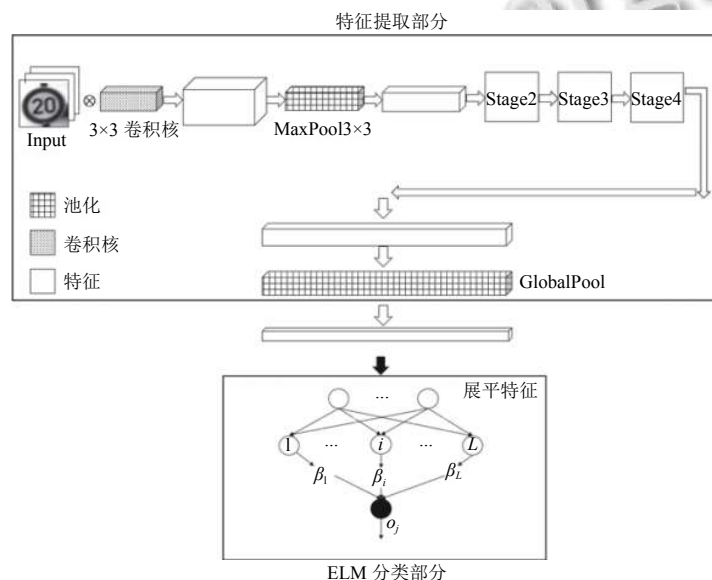


图 7 SGC-ELM 模型结构

4 实验与分析

4.1 数据集

本文采用德国交通标志数据集 (GTSRB) 和比利时交通标志数据集 (Belgium TSC) 进行实验. GTSRB 和 Belgium TSC 针对于分类问题, 数据集中的图像大部分为交通标志, 没有太多背景干扰, 比较适合本文的研究. 整个 GTSRB 数据集共包括 43 种交通标志, 51 839 张图像. 其中 39 209 张为训练集, 12 630 张为测试集, 训练集与测试集比例约为 4:1. Belgium TSC 数据集共包括 62 种交通标志, 7095 张图像, 其中 4575 张为训练集, 2520 张为测试集, 训练集与测试集比例约为 2:1.

4.2 实验设计

所有实验的硬件环境为 CPU: Intel(R) Core(TM) i5-8300H 2.30 GHz 8 G. GPU: NVIDIA GeForce GTX 1050 Ti 4G. 网络模型: 实验使用预训练 VGG16、预训练 MobileNet、预训练 ShuffleNet 与新模型 DSC-ELM 和 SGC-ELM 作对比实验.

(1) 准备数据集, 并对数据集进行预处理. 将数据集图片转换为 jpg 格式, 每张图片归一化为 48×48 , 其目的是为了节省内存开销, 加快训练速度.

(2) 准备模型, 并加载预训练参数: VGG16 加载 ImageNet 数据集预训练参数; MobileNet 加载 Cifa100 数据集预训练参数; ShuffleNet 加载 ImageNet 数据集预训练参数; 新模型 DSC-ELM 和 SGC-ELM 的特征提取部分分别加载 Cifa100 和 ImageNet 数据集预训练参数. 加载预训练参数的目的是使 ELM 接收的特征可靠, 从而达到更精确的分类结果.

(3) 设置网络超参数: 迭代次数在 GTSRB 数据集上设为 10, 在 Belgium TSC 数据集上设为 50; 批处理大小设为 32; 学习率设为 0.001, 并在每次循环后对学习率进行衰减, 衰减率为 0.98.

(4) 如图 8 所示, 将数据集送入模型进行训练和测试.

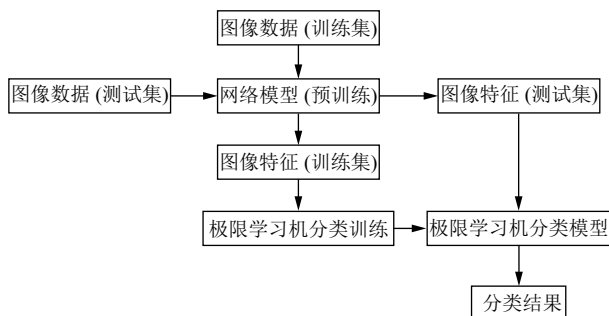


图 8 模型的训练和测试

(5) 从精度、训练时间、内存占用对比 5 种模型的实验结果.

4.3 实验结果与对比分析

5 种网络模型在 GTSRB 和 Belgium TSC 数据集上的实验结果如表 3 所示.

表 3 实验结果

数据集	网络模型	测试精度 (%)	模型耗时	内存占用 (MB)
GTSRB	VGG16	97.8	58 m 24 s	129.0
	MobileNet	94.8	41 m 51 s	27.0
	ShuffleNet	92.3	35 m 03 s	15.0
	DSC-ELM	98.9	31 m 24 s	8.9
	SGC-ELM	98.2	15 m 44 s	4.9
BelgiumTSC	VGG16	98.2	28 m 24 s	129.0
	MobileNet	94.6	20 m 32 s	27.2
	ShuffleNet	93.9	17 m 18 s	15.2
	DSC-ELM	99.3	15 m 05 s	8.9
	SGC-ELM	98.6	12 m 07 s	4.9

(1) 测试精度对比分析. 5 种模型的测试精度从高到低依次为 DSC-ELM、SGC-ELM、VGG16、MobileNet、ShuffleNet. MobileNet 和 ShuffleNet 精度低于 VGG16, 原因是采用了轻量化卷积方法构建卷积层, 卷积层参数减少, 导致损失一定的精度. DSC-ELM 与 SGC-ELM 模型精度均达到 98% 以上, 高于其他模型, 其原因在于新模型采用 ELM 作为分类器, ELM 泛化性强, 在接受到良好特征后能达到很好的分类效果, 而其他模型采用梯度下降算法更新参数, 易陷入局部极值, 损失精度.

(2) 训练时间对比分析. 从表 3 看出, 5 种模型训练速度从快到慢依次为: SGC-ELM、DSC-ELM、ShuffleNet、MobileNet、VGG16. VGG16 模型参数庞大, 导致计算缓慢, 在 5 种模型中耗时最长. ShuffleNet、MobileNet 模型参数量少, 训练速度较快, 而 ShuffleNet 相较于 MobileNet 少了 1×1 卷积, 速度快于 MobileNet. SGC-ELM 和 DSC-ELM 在训练速度上明显优于其他模型, 原因在模型的 ELM 分类器可通过广义逆一次求得权值参数, 而其他模型在训练时需要计算梯度信息, 并通过迭代进行调整, 耗费时间较长.

(3) 内存占用对比分析: 从表 3 看出, 5 种模型内存占用从高到低依次为: VGG16、MobileNet、ShuffleNet、DSC-ELM、SGC-ELM. VGG16 模型参数庞大, 导致内存占用高. MobileNet 和 ShuffleNet 采用轻量化卷积构建卷积层, 模型参数量较少, 内存占用较 VGG16 低, 而 ShuffleNet 模型结构中采用 Shuffle 替换 1×1 卷积,

相比 MobileNet 减少大量权值参数. DSC-ELM、SGC-ELM 模型无全连接层, 而是采用单隐层的 ELM 作为分类器, 相比其他模型内存占用低, SGC-ELM 采用 ShuffleNet unit 构成, 模型参数量小于 DSC-ELM, 内存占用较 DSC-ELM 低.

5 结论

现有 CNN 为了追求高的分类精度, 模型深度加深, 复杂度提高, 导致模型内存占用高, 训练速度慢. 交通标志识别设备的功耗和硬件性能较低, 对模型的速度和精度要求较高, 使得现有 CNN 模型无法应用于这种设备. 针对以上问题, 本文通过引入两种轻量化卷积方法并与极限学习机相结合, 提出两种轻量型卷积神经网络模型: DSC-ELM 模型和 SGC-ELM 模型. 新模型通过采用轻量化卷积方法构建卷积层进行特征提取, 使模型轻量化, 并以 ELM 作为分类器, 接受提取的特征进行分类, 解决了现有 CNN 全连接层参数训练速度慢的问题. 通过实验验证, 新模型相较于对比模型训练速度快, 分类精度高, 内存占用低, 可以很好的应用于嵌入式交通标志识别设备.

参考文献

- 1 Gudigar A, Chokkadi S, Raghavendra U. A review on automatic detection and recognition of traffic sign. *Multimedia Tools and Applications*, 2016, 75(1): 333–364. [doi: 10.1007/s11042-014-2293-7]
- 2 Jin J Q, Fu K, Zhang CS. Traffic sign recognition with hinge loss trained convolutional neural networks. *IEEE Transactions on Intelligent Transportation Systems*, 2014, 15(5): 1991–2000. [doi: 10.1109/TITS.2014.2308281]
- 3 Wang GY, Ren GH, Wu ZL, *et al.* A hierarchical method for traffic sign classification with support vector machines. *Proceedings of 2013 International Joint Conference on Neural Networks*. Dallas, TX, USA. 2013. 1–6.
- 4 Zhang Q, He N, Chi Y, *et al.* Traffic sign detection and recognition based on the improved SIFT algorithm. *Journal of Beijing Union University*, 2017, 31(2): 59–65.
- 5 Guo ZH, Zhang L, Zhang D. A completed modeling of local binary pattern operator for texture classification. *IEEE Transactions on Image Processing*, 2010, 19(6): 1657–1663. [doi: 10.1109/TIP.2010.2044957]
- 6 Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, 2017, 60(6): 84–90. [doi: 10.1145/3065386]
- 7 Szegedy C, Liu W, Jia YQ, *et al.* Going deeper with convolutions. *Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition*. Boston, MA, USA. 2015. 1–9.
- 8 Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. *arXiv: 1409.1556*, 2014.
- 9 He KM, Zhang XY, Ren SQ, *et al.* Deep residual learning for image recognition. *Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas, NV, USA. 2016.
- 10 Howard AG, Zhu ML, Chen B, *et al.* MobileNets: Efficient convolutional neural networks for mobile vision applications. *arXiv: 1704.04861*, 2017.
- 11 Zhang XY, Zhou XY, Lin MX, *et al.* ShuffleNet: An extremely efficient convolutional neural network for mobile devices. *Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City, UT, USA. 2018. 6848–6856.
- 12 Huang GB, Zhou HM, Ding XJ, *et al.* Extreme learning machine for regression and multiclass classification. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 2012, 42(2): 513–529. [doi: 10.1109/TSMCB.2011.2168604]
- 13 Stallkamp J, Schlipsing M, Salmen J, *et al.* The German traffic sign recognition benchmark: A multi-class classification competition. *Proceedings of 2011 International Joint Conference on Neural Networks*. San Jose, CA, USA. 2011. 1453–1460.
- 14 Mathias M, Timofte R, Benenson R, *et al.* Traffic sign recognition—How far are we from the solution? *Proceedings of 2013 International Joint Conference on Neural Networks*. Dallas, TX, USA. 2013. 1–8.
- 15 刘志刚, 李盼池, 许少华. 基于 ELM-PNN 算法的第 24 周太阳黑子预测预报. *控制与决策*, 2017, 32(4): 642–646.
- 16 闫河, 王鹏, 董莺艳, 罗成, 等. 一种 CNN 与 ELM 相结合的船舶分类识别方法. *重庆理工大学学报 (自然科学)*, 2019, 33(1): 53–57.