

基于深度卷积神经网络的小样本车型分类方法

吕磊, 李文彬, 王晓鸣, 胡隆基

引用本文:

吕磊, 李文彬, 王晓鸣, 等. 基于深度卷积神经网络的小样本车型分类方法[J]. 兵器装备工程学报, 2020, 41(8): 193-200+221.

您可能感兴趣的其他文章

1. 基于序列特征的2D CNN的动态手势识别

引用本文: 李振宇, 张立民, 邓向阳, 等. 基于序列特征的2D CNN的动态手势识别[J]. 兵器装备工程学报, 2019, 40(2): 141-146.

2. 基于深度CNN和ELM的滚动轴承故障诊断研究

引用本文: 顾鑫, 唐向红, 陆见光, 等. 基于深度CNN和ELM的滚动轴承故障诊断研究[J]. 兵器装备工程学报, 2020, 41(3): 154-158.

3. 基于DLPCC和ELM的装甲车辆声识别

引用本文: 樊新海, 石文雷, 张传清. 基于DLPCC和ELM的装甲车辆声识别[J]. 兵器装备工程学报, 2018, 39(7): 20-25.

4. 基于噪声信号EMD的装甲车辆识别研究

引用本文: 樊新海, 孙国强, 张传清. 基于噪声信号EMD的装甲车辆识别研究[J]. 兵器装备工程学报, 2017, 38(7): 111-115.

5. 改进的Cifar-10模型在装甲目标二分类中的应用

引用本文: 谢晓竹, 薛帅. 改进的Cifar-10模型在装甲目标二分类中的应用[J]. 兵器装备工程学报, 2019, 40(8): 141-144.

基于深度卷积神经网络的小样本车型分类方法

吕磊¹, 李文彬¹, 王晓鸣¹, 胡隆基²

(1. 南京理工大学 智能弹药技术国防重点学科实验室, 南京 210094;

2. 32381 部队, 北京 100072)

摘要:针对战场军事车辆分类等任务的可供训练样本较少的车型分类任务,在 AlexNet 网络的基础上引入多尺度分支的思想,设计了多尺度特征提取卷积神经网络。仿照战场军用车辆数据集的特点,构建了复杂背景下多角度、多尺度、类间差异小、类内差异大的 10 类较细粒度车型分类数据集。所提出的方法在实验平台上达到了每张图片 0.003 s 的分类识别速度,依据算力计算具有嵌入式实时性应用的可能。结合自适应学习率等方法,在小样本车型分类数据集上实现了最高 92% 的分类准确率,同等实验条件下分类性能和训练速度均优于主流卷积神经网络 AlexNet。提出的多尺度特征提取卷积神经网络在智能弹药或无人机进行军用车辆识别和部分民用场景等小样本车型分类场景中具有应用价值。

关键词:卷积神经网络;车型识别;小样本数据集;细粒度;智能弹药

本文引用格式:吕磊,李文彬,王晓鸣,等. 基于深度卷积神经网络的小样本车型分类方法[J]. 兵器装备工程学报, 2020,41(08):193-200,221.

Citation format:LYU Lei, Li Wenbin, WANG Xiaoming, et al. Vehicle Type Recognition Method with a Small Dataset Based on Deep Convolutional Neural Network[J]. Journal of Ordnance Equipment Engineering, 2020,41(08):193-200,221.

中图分类号:TJ413.6;E919;TP391.41;TP183 **文献标识码:**A **文章编号:**2096-2304(2020)08-0193-08

Vehicle Type Recognition Method with a Small Dataset Based on Deep Convolutional Neural Network

LYU Lei¹, Li Wenbin¹, WANG Xiaoming¹, HU Longji²

(1. Ministerial Key Laboratory of ZNDY, Nanjing University of Science and Technology, Nanjing 210094, China;

2. The No. 32381st Troop of PLA, Beijing 100072, China)

Abstract: The training samples for classification tasks such as that the vehicle type recognition has few training samples, the idea of multi-scale branching was introduced on the basis of AlexNet network. A multi-scale feature extraction convolutional neural network was designed. According to the characteristics of military vehicle data on battlefield, the multi angles and scales 10 classes fine-grained vehicles dataset with small inter-class differences and large intra-class differences under complex background was established. The proposed method achieved a recognition speed of 0.003 seconds per image on the experimental platform, which has the possibility of real-time application on embedded platform according to the compute capability. Combined with adaptive learning rate and other methods, the classification accuracy on the established dataset is up to 92%. The classification performance and training speed in this paper are better than AlexNet under the same experimental conditions. The proposed multi-scale feature

收稿日期:2019-09-24; **修回日期:**2019-11-13

基金项目:装备预研兵器工业联合基金项目(6141B012858)

作者简介:吕磊(1994—),男,硕士研究生,主要从事车型识别、深度学习与智能弹药技术研究。

通讯作者:李文彬(1965—),男,博士,教授,主要从事弹药总体技术、毁伤机理与终点效应和智能弹药技术研究, E-mail: lwb2000cn@njust.edu.cn。

extraction convolutional neural network has application value in military vehicle classification by low speed intelligent ammunition or unmanned aircraft, civil vehicle classification scenes and other small sample vehicle classification scenarios.

Key words: convolutional neural network; vehicle type recognition; small sample data set; fine-grained; intelligent ammunition

车型分类识别在民用与军用方向都有广阔应用前景。民用方面,车型分类能够帮助智能交通自动化系统完成车辆的自动判别,在电子收费、车辆监管等情景下^[1]实现交通服务和管理的智能化,比如路口自动锁定套牌车,自动驾驶中判别警车、消防车等需让行的车型,无人收费站自动收费等场景均有应用。军用方面,利用无人机、弹箭平台和地面装甲车辆的图像采集设备,通过对军事车辆的类型进行识别,可以制定针对不同类型目标的攻击策略,从而实现了对目标关键部位的精确打击,对提升智能化与无人化作战具有重要意义。其中,在军用场景中,真实敌对军用车辆数据难以获得,那么如何利用小样本数据集完成车型分类具有极高的研究价值。

可见光图像信息丰富,传感器成本低,作为车辆信息获取方式具有较高的实用性。利用可见光图像进行车型分类识别属于图像分类技术的范畴,通常利用机器学习方法进行。有两类方法,第一类是利用 SIFT、HOG 等人工浅层特征提取方法结合机器学习分类器进行车型分类识别,这类方法依赖于先验知识以及特征提取与描述质量,泛化性差。第二类是利用深度卷积神经网络(DCNN)进行分类,此类方法不需要过多的预处理,不需要人工设计特征,而是通过自主学习方式得到特征。近年来,DCNN 在图片分类和车型分类取得了非常多的成果^[2-4],但这些方法在小样本数据集中进行反复训练后均会陷入过拟合的情况,即识别方法在训练集中达到很高的识别率,但是在未经训练的数据中却识别率很低。

针对以上问题,本文提出一种基于改进的深度卷积神经网络的小样本多尺度车型分类方法,旨在为军用车辆实时性识别和部分民用场景提供思路。为智能弹药和中低速弹箭的目标识别以及关键部位识别提供了基础。

1 深度卷积神经网络

卷积神经网络起源于 20 世纪 90 年代末,一开始被用于手写字母的识别。2006 年,Hinton^[5]提出了深度学习(Deep Learning)的概念,2012 年 Krizhevsky 等^[6]构建了深度卷积神经网络 AlexNet 并获得了 ILSVRC-2012 图片分类竞赛的冠军,从此深度卷积神经网络被广泛地应用到语音识别和图像识别中,并且与其他方法相比,成为了图像识别领域最有效的方法。

对于目标图像而言,需要检测图片中的特征来决定图片的类别,通常情况下这些特征不是由整张图片决定,而是由

一些局部的区域决定的,如图 1 所示的坦克目标,其可用来识别类别的特征存在于图片的局部中。



图 1 坦克图像局部性

同类车辆具有相同的特征,虽然可能位于图片中不同的位置,但是检测这些特征的模式是一致的。并且对于车辆目标图片,进行下采样操作完成图片缩放后图片的性质基本保持不变。基于这样的识别特性,AlexNet 的网络结构如图 2 所示。

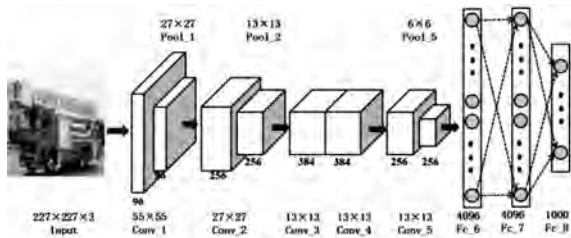


图 2 AlexNet 的网络结构示意图

输入图像为大小 227×227 像素的三通道数字图像,经过卷积层(Convolutional Layer, Conv),池化层(Pooling Layer, Pool)和全连接层(Full Connected Layer, FC)构成的前向传播通道,最终输出出各类别的得分,依据得分高低来判断图像的种类。

卷积层是对每个通道的二维图像数据进行卷积操作,卷积操作由卷积核完成,如图 3 所示。

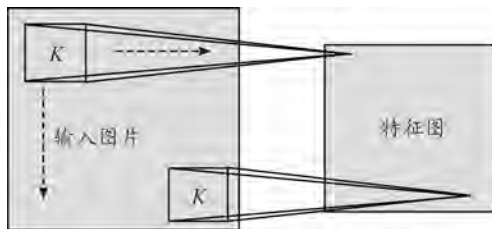


图 3 卷积操作示意图

卷积核是由可学习的参数集合的滤波器,通过卷积核在输入图像上按照一定步长进行滑动,依次通过卷积计算出特

征映射神经元并组合成特征图(feature maps)。二维卷积公式为

$$f(i, j) = \sum_{m=0}^M \sum_{n=0}^N K(m, n) A(i - m, j - n) \quad (1)$$

式中: $K(m, n)$ 为卷积核各点数值; $A(x, y)$ 为输入图各点数值; $f(i, j)$ 为输出的特征图坐标为 (i, j) 点的数值。卷积层具有参数共享和局部稀疏连接的特点,利用卷积核可以使得图像中特定模式能够得到检测。

池化层实现了图像识别的不变形,同时也降低了图像的尺寸,实现方法由池化窗口在输入图像上按照一定步长计算出的结果组合成。采取最大池化的方式,其表达式为:

$$f(i, j) = \max_{0 \leq m, n < k} \{A(i \times s + m, j \times s + n)\} \quad (2)$$

式中: s 为池化窗口移动的步长; m, n 分别为池化窗口的宽和高; $A(x, y)$ 为输入图各点数值; $f(i, j)$ 为输出的特征图坐标为 (i, j) 点的数值。

全连接层由多层神经元组成,每层具有若干神经元,层内神经元所有神经元相互独立,但层间神经元相互连接。

2 小样本车型分类网络结构与训练流程

2.1 激活函数选取和判别损失函数

车型分类是一个非线性问题,在神经网络中需要引入非线性激活函数实现整个网络的非线性。常用激活函数有 sigmoid 函数, tanh 函数和 ReLU 函数,其中 ReLU 函数会使深度卷积神经网络训练加速数倍,因此选用 ReLU 作为激活函数。其公式为

$$\text{ReLU}(x) = \max(0, x) \quad (3)$$

判别损失函数是反向传播的依据,决定了模型能否收敛并达到预期目标,模型训练的评判标准。理想状况是通过判别损失误差函数可将样本所属类别对应的输出节点的输出为 1,而其他节点的输出为 0,即 $[0, 0, 1, 0, \dots, 0, 0]$,但神经

网络只能输出复杂加权和与非线性处理之后的值,所以利用判别损失误差函数计算出样本类别的概率。

本文采用交叉熵损失函数 CrossEntropyLoss,首先通过 softmax 层将神经网络原始输出的浮点数 作回归处理:

$$\text{softmax}(y_i) = y'_i = \frac{e^{y_i}}{\sum_{j=1}^n e^{y_j}} \quad (4)$$

很显然 $\sum_{j=1}^n y'_j = 1$,因此将单个节点的输出浮点数变成一个概率值。然后通过交叉熵刻画实际输出(概率)与期望输出(概率)的距离。当交叉熵值越小,概率分布就越接近,也就意味神经网络分类器更加接近任务的实际分类函数。交叉熵函数可表示为

$$H(p, q) = - \sum_x p(x) \log q(x) \quad (5)$$

其中: $H(p, q)$ 为交叉熵函数,概率分布 $p(x)$ 为期望输出,概率分布 $q(x)$ 为实际输出。期望输出除了样本所属类别概率为 1 以外,其他类别概率为 0。因此交叉熵函数越小,训练样本的输出判别结果就与样本真实结果越接近。

2.2 多尺度特征提取网络结构

军事车型分类识别相较于普适物体分类识别种类少,嵌入式需求多,所以特征空间需求较低,高维度的特征空间会使得网络陷入过拟合的可能性增加,导致训练集中识别率较高而在测试集中或实际应用中识别效果较差。因此,为了保证模型的泛化能力,以及考虑现今硬件计算能力限制,过深层次的卷积网络不适用于小样本的车型识别。

针对细粒度分类问题,文献[7]提出了双流卷积神经网络(Bilinear CNN),利用两个卷积神经网络的特征图进行特征融合,实现了多个数据集上优异的分类效果。因此在 Alex-Net 网络结构的基础上,引入多尺度初始特征提取的方法,重新设计网络结构,将多分支的特征输出进行连接,网络层数仅增加一层,特征图数量几乎不变,其结构如图 4 所示。

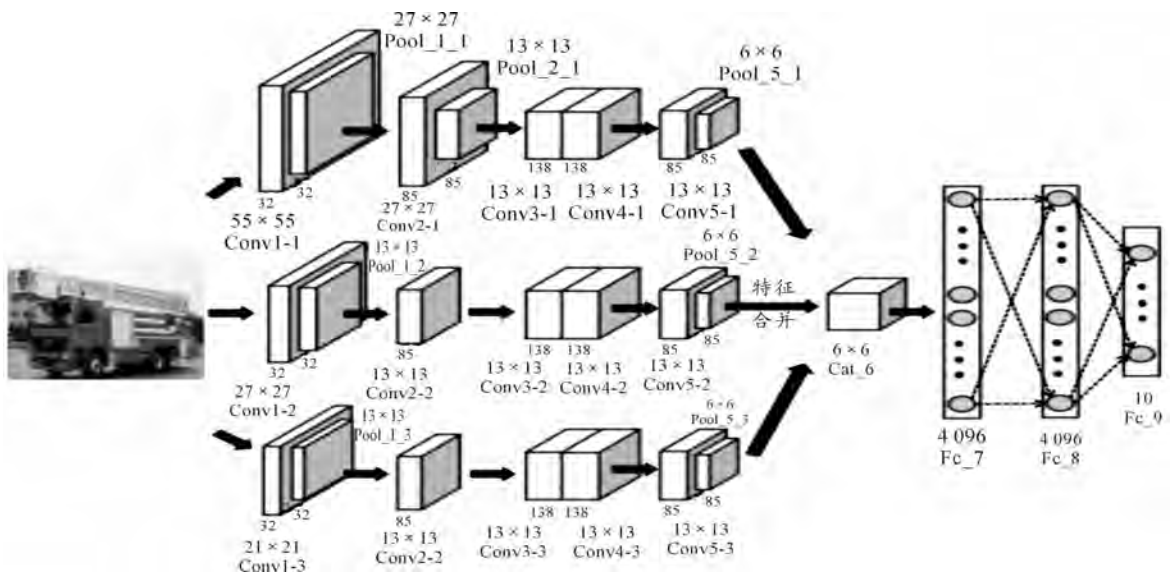


图 4 多尺度特征提取卷积神经网络结构示意图

不同尺度的目标图像中能够表征某种识别特征的区域图像大小具有差异,为了不缺失这类因尺度大小而变化的特征,设计了3个卷积特征提取通道,在最终特征图数量与大小几乎不变的情况下降低单种大小初始特征所占比重,构建多尺度的卷积核进行初始特征的提取,再进行深度特征提取后进行特征图的拼接。

利用这样的设计,特征分流改善了特征提取方法,降低了特征提取部分的复杂度和冗余度,而特征图的大小与数量没有变化,因此最终特征复杂度没变,在确保特征复杂度没变的情况下减少了卷积特征提取过程中的冗余结构,而对特征的分类效果主要集中在全连接层,此部分数目没变,保证了分类网络复杂度没有变化。这种设计能够减少过拟合现象,提升测试集准确率,主要因为:在大数据集合的情况下,特征分流可能会导致特征图的表征力不足,但在小数据集的情况,由于样本提供的信息量难以支撑极复杂并且冗余的特征空间,所以特征分流会使得过拟合现象减少。不同特征流之间减少了相互干扰,降低了无效特征对分类判断的影响。由于图片中车辆目标占据大小不同的情况(即尺度不同),相同特征具有不同的大小,例如同一车轮在不同尺度照片中尺度不同,所以不同初始尺度的卷积特征分流能够很好地适应于相同目标车辆。

这样的设计具有对尺度变化的鲁棒性,确保在图像中,目标所占图像比例不同的情况下依然能够提取出相同特征。并且在确保总特征图数量相同的情况下,减少了单分支中特征图数量,可以减少无用特征对有效特征的干扰,并且减少了训练参数的数量,从而加快训练^[8]。

2.3 参数更新与训练方法

训练方法采用监督学习(有导师学习)训练方法,如图5所示。

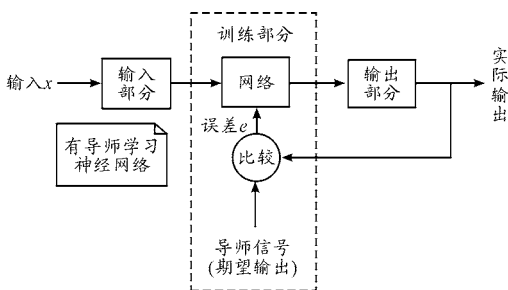


图5 有导师学习训练方法框图

训练样本输入到网络中,经过各层的数值计算与映射,前向传播得到响应,通过比较响应和标注好的样本标签,得到误差信息,利用反向传播算法反向调整网络中的参数大小,使误差信息向降低的方向发展,从而使得整个网络模型达到更好的分类效果。通过重复以上步骤,进行反复迭代使得网络模型达到收敛状态就可完成分类方法的训练深度卷积神经网络通行方法为利用反向传播方法结合梯度下降法进行,采用增加动量项的随机梯度下降法(SGD + Momentum)进行参数的训练,既可以减少梯度运算量,又可以跳出局部极小

点,其算法伪代码如表1所示。

表1 带动量项的随机梯度下降法的伪代码

参数需求:学习率 η , 动量参数 γ 。输入需求:初始梯度 θ^{i-1}
计算速度更新项: $v = \gamma \times \theta^{i-1} - \eta \times \nabla L(\theta^{i-1})$
更新参数: $\theta^i = \theta^{i-1} + v$

速度更新项

$$v = \gamma \times \theta^{i-1} - \eta \times \nabla L(\theta^{i-1}) \quad (6)$$

若假设每个时刻梯度相同,那么每次更新步长为: $\frac{\eta \times \nabla L(\theta^{i-1})}{1 - \gamma}$, 当动量参数 γ 分别设为 0.5、0.9、0.99 时,其最大速度为 SGD 算法的 2 倍,10 倍以及 100 倍。

学习率常用做法通常有两种,方法一采用固定学习率,方法二采用固定阶段衰减学习率做法。由于在模型训练达到一定精度后,过大的学习率可能会使得参数错过最优位置,需要缩小学习率达到慢慢逼近全局最优点^[9],所以方法一中如果选取较大学习率将会使得最终难以收敛到较好位置,而较小学习率则会使得训练速度变慢以及难以跳出局部最优点。方法二采取经过固定批次降低学习率,这样的做法可能会由于模型尚未收敛就降低学习率从而降低训练速度。因此引入自适应学习率变化,当损失函数值与迭代次数乘积大于设定值时,说明模型在现有学习率下收敛到较好状态,需要减小学习率使得模型进一步收敛,此时对学习率进行衰减处理:

$$\eta' = \alpha \times \eta \quad (7)$$

α 为学习率衰减系数,本文以 0.05 为间隔,分别选取了 0.6 到 0.9 的各个值作为衰减系数进行尝试,实验结果表明 0.8 的学习率衰减系数分类效果最佳,因此选取学习率衰减系数为 0.8。

相比固定学习率做法以及文献中的分段做法,在相同实验条件下,采用本文的以损失函数值与迭代次数乘积值为条件的自适应学习率在本文的实验中最佳准确率分别提升了 3% 与 1.5%。

2.4 算法总体流程

整个卷积神经网络的训练与参数学习包含前向传播和反向传播两个主要阶段。利用预处理好的图像数据,按照一定格式输入到网络中进行前向传播计算,前向传播计算过程如图6所示,激活函数为 2.1 节的 ReLU 函数。

初始输入图像大小为 $227 \times 227 \times 3$ 。

分支1:第1层初始卷积核大小为 11×11 ,数量 32 个,步长为 4,输出特征图图像大小 $(227 - 11)/4 + 1 = 55$,即 $55 \times 55 \times 32$ 。经过最大池化输出 $27 \times 27 \times 32$;第2层卷积核大小 5×5 ,数量 85,步长 2,经最大池化输出 $13 \times 13 \times 85$;

分支2:第1层初始卷积核大小为 19×19 ,数量 32 个,步长为 8,输出特征图图像大小 $(227 - 19)/8 + 1 = 27$,即 $27 \times 27 \times 32$ 。经过最大池化输出 $13 \times 13 \times 32$;第2层卷积核大小

3×3 , 数量 85, 步长 1, 输出 $13 \times 13 \times 85$;

分支 3, 第 1 层初始卷积核大小为 27×27 , 数量 32 个, 步长为 10, 输出特征图像大小 $(227 - 27) / 10 + 1 = 21$, 即 $21 \times 21 \times 32$ 。经过最大池化输出 $13 \times 13 \times 32$; 第 2 层卷积核大小 5×5 , 数量 85, 步长 1, 输出 $13 \times 13 \times 85$;

3 个分支的 3~5 层采取一样操作, 第 3 层、第 4 层输出 $13 \times 13 \times 138$, 第 5 层输出 $13 \times 13 \times 138$ 。第 6 层对特征图进行融合, 将 3 个分支各自 $13 \times 13 \times 85$ 的特征图融合为 $13 \times 13 \times 255$ 的特征图。将特征图数据转换为 $43\ 095 \times 1$ 的张量形状, 第 7 层、第 8 层各有 4 096 个神经元。最后一层有 10

个神经元, 输出数值代表了该样本在各个类别的得分情况, 得分最高的类别被判定为该样本所在类别。

计算公式如第 1 节中所示, 前向传播最终输出 $10 \times$ 批次的向量, 将最后输出的向量输入到 2.2 节的判别损失函数得到结果, 通过结果与标注好的车型信息对比, 采用 2.4 节增加动量项的随机梯度下降法将误差信息反向传播到网络中的连接权值上, 更新权值。

训练完成后利用保存的网络权值, 对测试样本进行前向传播计算, 得出的向量中最大值所处维度即预测的样本类别。

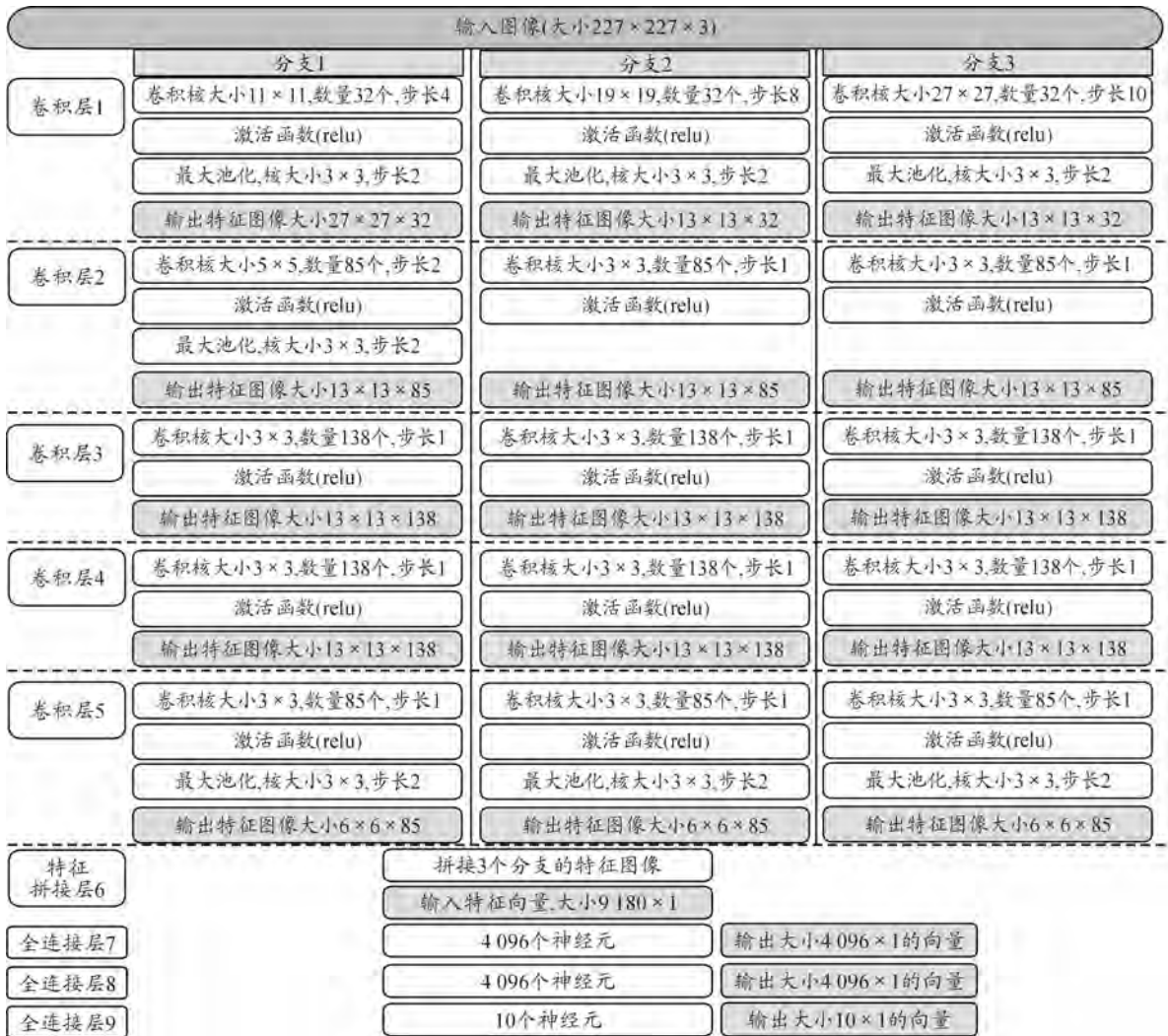


图 6 多尺度特征提取卷积神经网络前向传播计算过程框图

3 数据集

通常对于传统的深度学习分类方法, 所需训练样本量每一类需要数万张才能达到较好的识别效果, ImageNet 数据集有 30 种大类共 1 400 多万张标注好的图片数据, CIFAR-10 数据集 10 类物体具有 6 万张标注好的图片数据。为了测试

本文设计的方法在小样本数据集集中的分类效果, 选取了公交车、消防车、小卡车、吉普车、面包车、赛车、SUV、出租车、重型卡车、家庭轿车这十类较细粒度分类的车型进行车型分类实验。

3.1 数据集样本选取与质量分析

车辆图片遵照图片分类竞赛通用要求, 来源于互联网, 车型信息标签按照国际 ImageNet 分类竞赛 ILSVRC 的标准,

由人工标注类别信息,以文本形式储存。在训练时同步读取文本获取车型信息。每一类使用 140 张样本用作训练集、20 张样本用作算法测试集。图片样本囊括不同车辆在不同复杂背景下的情况,多角度与多尺度,并且类间差距有些较小,因此可以算是较细粒度车型识别。共 1 400 张训练集,规模总数不足 ImageNet 数据集的万分之一

数据集中部分类间差异较小,具有极其相近的外观(轿车与 SUV,重型卡车与小卡车),类内差异大,如图 7 所示。数据集包含了正面、侧前面、正侧面、侧后面等角度的样本,能够研究视角不同因素对车型识别的影响,车辆颜色与涂装不同,背景相对复杂并各不相同,图片中具有多辆车共存的

情况,车辆尺度也从占据图片大小的 50%~90% 均有。光照强度从明到暗均有,部分车辆具有一定遮挡,可以充分考察模型的泛化能力。部分图片只拥有局部车辆的干扰样本,以观察模型的鲁棒性。这样的十类车型数据集模拟军事车辆分类等小样本情况的样本质量,增加了训练难度,从而能更好考量分类方法在小样本细粒度分类这一情况下的表现。

根据车辆数据集的质量分析可知,该数据集属于识别难度较高的较细粒度车型分类,识别难度与难以获取的战场真实军用车辆数据识别难度相当,因此作为小样本车型分类数据集,具有很高的判定意义。



图 7 小样本车型分类数据集

3.2 图像预处理

数据集中图像数据的格式各种各样,有 png、jpg 等数字图像格式,不能够直接进行计算,需要转换为所需要的格式,并且由于图片的尺寸大小不一,所以在将图像数据输入到卷积神经网络中进行运算前需要进行相应的预处理,并且为了算法的实时性和之后的可嵌入式移植,应尽量采取简单的预处理流程,如图 8 所示。

首先,由于训练集大小尺寸不一,但真实应用中由于同一设备采集图像,所以图像大小固定,所以先将图片尺寸归一化到 $256 \times 256 \times 3$ 。

对于训练集,先采取 50% 概率随机水平翻转及随机切割,目的是进行数据增强。这样操作后理论上可以将样本数扩充 1 000 多倍。再转换为张量数据结构,数字图像的各通道像素值范围为 $[0, 255]$, 散布范围大,所以在转为张量数据结构的同时,等比例得将其转换至 $[0, 1.0]$ 。由于光照角度与强度等原因,不同目标图像的像素数据均值差异较大,不具有可比性。因此通过正则化操作将图片的像素值限制在某一固定范围内,从而更好地泛化识别,避免在后续过程中大数值区间的属性过分支配小数值区间的属性。其计算公式为

$$\begin{pmatrix} p_{N,1} \\ p_{N,2} \\ p_{N,3} \end{pmatrix} = \begin{pmatrix} 1/S_1 & 0 & 0 \\ 0 & 1/S_2 & 0 \\ 0 & 0 & 1/S_3 \end{pmatrix} \begin{pmatrix} p_1 - m_1 \\ p_2 - m_2 \\ p_3 - m_3 \end{pmatrix} \quad (8)$$

其中: p_i 表示图片中某一坐标第 i 通道的像素值; m_i 表示第 i 通道的平均值; S_i 表示第 i 通道的标准差; $p_{N,i}$ 为经过正则化后该坐标第 i 通道的像素值。为保证预处理速度,平均值序列与标准差序列在训练前通过抽样计算得出^[10]。

对于测试集省去翻转操作,其余与训练集相同。这样的预处理方法省去了传统图像处理中复杂的预处理计算,极大

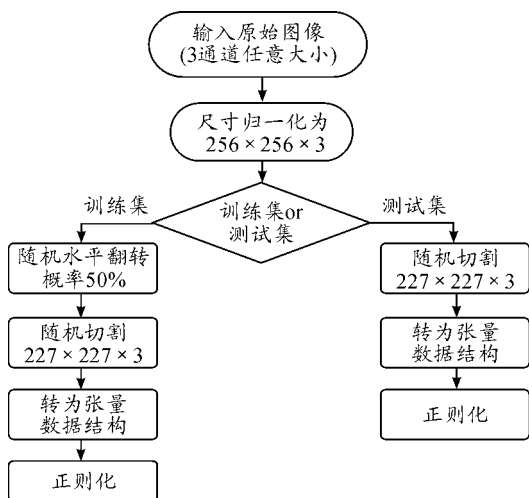


图 8 图像预处理流程框图

地提升了算法速度,降低了嵌入式应用的硬件要求。

4 车型分类实验

为了验证本文结构在小样本车型分类中的性能,采用2012年ILSVRC图片分类竞赛的冠军 AlexNet 做对比实验。实验样本与实验环境均一致,实验样本采用第3节构建的小样本数据集,旨在验证本文方法与主流分类方法在小样本情况下的性能对比。

从2015年resnet^[11]开始,极大层数的深度神经网络结构被开发并在目标分类领域取得很好的效果,但是这些算法通常都需要大量的样本集训练使模型收敛到较好状态,否则会产生 Hughes 现象^[12],即高维度的特征空间却比低维度特征空间对车型特征的描述能力差,分类效果不如浅层网络。对于只能提供少量样本的分类任务,只能通过迁移学习^[13]的方法,并且要求迁移前后的分类任务具有一定的相关性,所以对于军用车辆分类这种与其他任务相关性较低、可用样本量少的分类任务,resnet152等结构只能达到较低准确率并且运行速度大大降低,因此,选用 AlexNet 作为本文结构的对比参照。

4.1 实验配置与步骤

整个实验包括神经网络的训练与测试两大部分,利用编程语言完成数据集的预处理、定义网络结构和前向传播过程、实现反向传播与数据存储。随机顺序对1400张训练集各训练一次称为一批或一个 epoch,训练完成后利用测试集对训练好的模型进行测试。

算法中包含了大量重复运算与并行化操作,以 GPU 为核心计算器件可以加快算法的训练。为了对比本文结构与 AlexNet 在小样本车辆数据集上的分类效果,保证训练参数一致,实验环境与训练参数配置如表2所示。

表2 实验环境与计算设备配置

配置名称	配置信息
CUDA	10.0
CPU	Core i5 - 8500 @ 3.00 GHz
GPU	NVIDIA GeForce GTX 1080 Ti
GPU memory	6 GB
RAM	16 GB
训练批数	500
初始学习率	0.008
学习率衰减系数	0.8
动量参数	0.8

4.2 实验结果与分析

本文网络结构与 AlexNet 结构大小相差不大,在特征图数量相同的情况下本文结构占用存储比 AlexNet 少 8%

左右。

判别损失函数是模型进行训练的目标函数,如果判别损失函数不收敛,代表模型不收敛,即训练失败。判别损失函数曲线如图9所示。

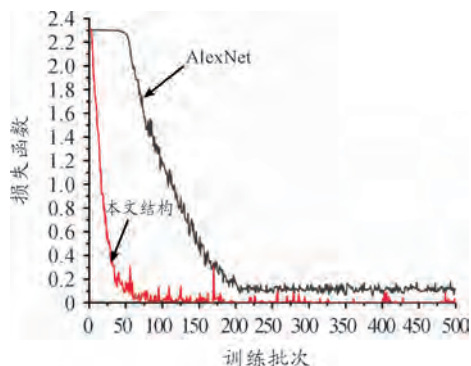


图9 不同结构的训练过程中损失函数曲线

观察图9,两种结构在训练后判别损失函数输出均达到较低水平,并且随着训练批数增加维持在一定阶段,代表模型均训练完成。在保证其他训练参数相同情况下,本文结构在训练至50个 epoch 左右,而 AlexNet 需要200个 epoch,训练速度提高了300%。与前文分析一致,实验结果表明本文结构收敛速度快于 AlexNet,原因主要是因为分支之间的特征不会相互影响,从而减少了无效特征带来的参数更新,从而加快了训练速度。

表3 AlexNet 与改进结构的对比

结构主要参数	AlexNet	本文结构
参数数量	约6000万	约5500万
节点数量	约81万	约80万
层数	8	9

训练集准确率可以表示训练集的利用程度,也可以与测试集准确率结合观察过拟合程度,训练过程中1400张训练集准确率变化如图10所示。

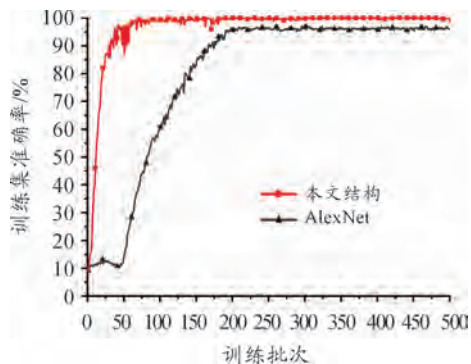


图10 训练过程中训练集准确率曲线

测试集准确率是判断模型分类性能的重要指标,训练过

程中 200 张测试集准确率变化如图 11 所示。

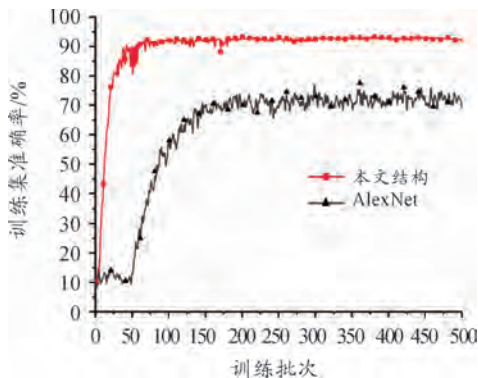


图 11 训练过程中测试集准确率曲线

观察图 10、图 11,两种结构在训练集上均达到了 95% 以上的准确率并几乎稳定,但 AlexNet 在测试集中准确率最终只能达到 70% 左右,说明过拟合程度较高。而本文结构在测试集中准确率能达到 90% 左右,与训练集准确率相差不多,说明过拟合现象较轻。

对训练后的模型分别利用 200 张测试集进行分类实验,每个模型共进行 5 次实验。实验结果如表 4 所示。

表 4 测试集模型分类实验结果

实验 次数	AlexNet		本文方法	
	准确率/%	用时/s	准确率/%	用时/s
1	71.00	0.611	91.50	0.656
2	72.50	0.596	91.50	0.649
3	72.00	0.605	91.50	0.657
4	74.00	0.604	91.00	0.666
5	71.50	0.599	92.00	0.639
最佳值	74.00	0.596	92.00	0.649
平均值	72.20	0.603	91.50	0.653

通过 5 次实验结果平均值,相同实验条件下本文方法在构建的小样本车辆测试集中准确率达到 91.5%,而 AlexNet 只达到 72.2%,存在较严重的过拟合现象。

实验结果表明:由于多尺度特征分流提取以及自适应学习率等策略的改变,减少了特征提取的冗余部分,减少了无效特征对有效分类特征的干扰,以及能够适应于多尺度情况下的相同特征提取,所以本文方法在小样本车型分类任务性能优于 AlexNet,识别速度达到 0.653 s 每 200 张,即约 0.003 s 每张,其中包括了实际应用不需要的文件存储等操作。根据 NVIDIA 公司显卡算力数据^[14],嵌入式 GPU 模块 Jetson TX1 的计算能力能够实现嵌入式识别。因此本实验平台验证的算法在嵌入式应用场景中具备实时性应用的条件。

5 结论

1) 本文仿照战场真实军用车辆数据集的特点,构建了复杂背景下多角度与多尺度,类间差异较小,类内差异大的较细粒度车型分类数据集。通过引入多分支卷积特征提取方法,设计了多尺度卷积特征提取神经网络,制定了网络的训练策略与流程,编写了自动化车型识别程序,进行了 GPU 平台上的车型分类算法训练与性能测试实验。

2) 同等实验环境,在训练样本较少的情况下,利用多分支提取不同尺寸的初始特征可以有效提高车型分类的识别效果,同时也能显著提高网络的收敛速度。本文方法分类性能优于主流卷积神经网络 AlexNet,在识别相当战场真实军用车辆数据难度的较细粒度车型分类数据集上取得了测试集 92% 的识别率。本文的方法可在嵌入式平台中达到约 0.003 s 每张图片的识别速度。本文方法具有较高的实时性实用价值,为战场车辆目标识别奠定基础。

3) 由于缺少细粒度分类的战场真实军用车辆数据集,不能实际验证算法在军用车辆分类识别的表现,只能说明识别机理以及在识别难度增加的民用车型数据集上验证算法。下一步将尝试构建军事车辆专用数据集,利用迁移学习进行军事车辆上的算法训练,并将算法移植到现有嵌入式平台中进行实时应用验证。

参考文献:

- [1] 易校石,刘念.一种基于支持向量机中分离超平面求取的算法[J].重庆工商大学(自然科学版),2018,35(3):66-69.
- [2] DONG Z, PEI M, HE Y, et al. Vehicle type classification using a semisupervised convolutional neural network [J]. IEEE Transactions on Intelligent Transportation Systems, 2015, 16(4): 2247-2256.
- [3] YANG L, LUO P, CHANGE LOY C, et al. A large-scale car dataset for fine-grained categorization and verification [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2015:3973-3981.
- [4] 邓柳,汪子杰.基于深度卷积神经网络的车型识别研究[J].计算机应用研究,2016,33(3):930-932.
- [5] HINTON G E, SALAKHUTDINOV R R. Reducing the dimensionality of data with neural networks [J]. Science, 2006, 313(5786): 504-507.
- [6] ALEX K, ILYA S, GEOFFREY E. Hinton. ImageNet classification with deep convolutional neural networks [C]//Advances in Neural Information Processing System, 2012: 1097-1105.

- [4] REMENNIKOV A, NGO T, MOHOTTI D, et al. Experimental investigation and simplified modeling of response of steel plates subjected to close-in blast loading from spherical liquid explosive charges [J]. *International Journal of Impact Engineering*, 2017, 101(09): 78 - 89.
- [5] 宋彬, 黄正祥, 翟文, 等. 聚脲弹性体夹芯防爆罐抗爆性能研究[J]. *振动与冲击*, 2016, 35(7): 138 - 144.
- [6] 廖瑜, 石少卿, 梁朝科, 等. 聚脲-编织玻璃纤维网格布复合材料加固钢板抗冲击力学性能研究[J]. *兵工学报*, 2018, 39(10): 1988 - 1996.
- [7] LIAO Yu, SHI Shaoqing, LIU Zhaohui, et al. Study on shock resistance of steel plate reinforced with polyurea-woven fiberglass mesh composite under shock wave[J]. *International Journal of Protective Structures*, 2018, 00(0): 1 - 19.
- [8] 赵宇峰, 路志超, 董永香, 等. 复合壳体对炸药抗破片冲击起爆影响研究[J]. *振动与冲击*, 2019, 38(20): 126 - 147.
- [9] PARNIANI S, TOUTANJI H. Monotonic and fatigue performance of RC beams strengthened with a polyurea coating system[J]. *Construction and Building Materials*, 2015, 101(01): 22 - 29.
- [10] TOUTANJI H A, CHOI H, WONG D, et al. Applying a polyurea coating to high-performance organic cementitious materials[J]. *Construction and Building Materials*, 2013(38): 1170 - 1179.
- [11] 蒲兴富. 弹性体增强混凝土砌体墙爆炸响应的数值分析[D]. 宁波: 宁波大学, 2009.
- [12] 蔡桂杰. 弹性体涂覆钢筋混凝土板抗爆作用设计方法研究[D]. 太原: 中北大学, 2015.
- [13] 张青艳, 靳晓庆, 郑宇轩, 等. 聚脲包覆混凝土的压缩和吸能特性研究[J]. *工程力学*, 33(04): 205 - 209.
- [14] SHI Shaoqing, LIAO Yu, PENG Xiongqi, et al. Behavior of polyurea-woven glass fiber mesh composite reinforced RC slabs under contact explosion[J]. *International Journal of Impact Engineering*, 2019, 132(07): 1 - 12.
- [15] RAMAN S N, NGO T, MENDIS P, et al. Elastomeric polymers for retrofitting of reinforced concrete structures against the explosive effects of blast[J]. *Advances in Materials Science and Engineering*, 2012(11): 1 - 8.

科学编辑 杨继森 博士(重庆理工大学教授)
责任编辑 唐定国

(上接第 200 页)

- [7] LIN T Y, ROYCHOWDHURY A, MAJI S, et al. Bilinear CNNs for Fine-grained Visual Recognition[C]//*IEEE International Conference on Computer Vision (ICCV)*. 2015: 1449 - 1457.
- [8] HUANG K, ZHANG B. Fine-grained vehicle recognition by deep Convolutional Neural Network [C]//*International Congress on Image and Signal Processing, Biomedical Engineering and Informatics. IEEE*, 2017: 465 - 470.
- [9] 郭守恒. 深度学习算法在高光谱影像分类中的应用研究[D]. 成都: 成都理工大学, 2017.
- [10] 周飞燕, 金林鹏, 董军. 卷积神经网络研究综述[J]. *计算机学报*, 2017, 40(6): 1229 - 1251.
- [11] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//*Proceedings of the IEEE conference on computer vision and pattern recognition*. Piscataway: IEEE, 2016: 770 - 778.
- [12] ZHANG J, ZHANG T, YANG Z L, et al. Vehicle model recognition method based on deep convolutional neural network [J]. *Transducer & Microsystem Technologies*, 2016, 35(11): 19 - 22.
- [13] ZHENG H, FU J, MEI T, et al. Learning Multi-attention Convolutional Neural Network for Fine-Grained Image Recognition[C]//*IEEE International Conference on Computer Vision. IEEE Computer Society*, 2017: 5219 - 5227.
- [14] 吴鑫, 张建奇, 杨琛. Jetson TK1 平台实现快速红外图像背景预测算法[J]. *红外与激光工程*, 2015, 44(09): 2615 - 2621.
- [15] 应晨露. 基于深度学习的车型识别[D]. 合肥: 中国科学技术大学, 2018.

科学编辑 王浩伟 博士(北京航空航天大学云南创新研究院特聘研究员、硕导)
责任编辑 唐定国