

文章编号: 2095-2163(2020)01-0067-04

中图分类号: TP391.41

文献标志码: A

基于卷积神经网络车型分类的研究

晏世武, 罗金良, 严庆

(南华大学机械工程学院, 湖南衡阳 421001)

摘要: 自 ILSVRC 大赛以来, 卷积神经网络(Convolutional Neural Network, CNN)得到了迅速发展, 众多学者将该技术用于图像的分类领域。车型分类是图像分类任务之一, 在交通安全中具有很大的作用, 因此使用 CNN 构建高效车型分类模型也越来越重要。为快速训练出高效的车型分类模型, 首先使用迁移学习来训练本文的原始车型数据, 其中 InceptionV3 模型精度最高, 约 85.91%。然而这些模型结构网络层次大多太深, 且无法直接用于工程实践。因此, 为了训练出精简且高效的模型结构, 本文从 CNN 的基本概念出发, 构建由卷积层、批规范化层(Batch Normalization, BN)、池化层、Dropout 层、全连接层和 softmax 层所组成的模型结构, 使用带有约束权重的 L2 作为损失函数, 并通过 Adam 优化算法对模型参数进行更新, 又通过逐步增加卷积层数和调整全连接层神经元个数的方法, 对数据增强过的车型数据进行训练和测试, 结果表明卷积层数为 4 和全连接层神经元个数为 256 的模型结构的精度最好, 约 85.15%, 较浅层次的网络达到了深层网络的性能。

关键词: 卷积神经网络; 车型分类; 迁移学习; 数据增强

Research on vehicle classification based on Convolutional Neural Network

YAN Shiwu, LUO Jinliang, YAN Qing

(School of Mechanical Engineering, University of South China, Hengyang Hunan 421001, China)

[Abstract] Since the ILSVRC competition, the Convolutional Neural Network (CNN) has developed rapidly, and many scholars have used this technology in the field of image classification. Vehicle classification is one of the image classification tasks and plays an important role in traffic safety. Therefore, it is more and more important to use CNN to build an efficient vehicle classification model. In order to train an efficient vehicle classification model quickly, the paper uses the transfer learning to train the original data of this vehicle at first. As a result, the InceptionV3 of transfer learning model has the highest accuracy. It's about 85.91%. However, the network structure of these models are too deep to be directly used in engineering practice. Therefore, in order to train a simple and efficient model structure, the paper starts the study based on the basic concept of CNN and builds the model by combining the convolution layer, the batch normalization (BN), the pooling layer, the dropout layer, the fully connected layer and the softmax layer. Then the paper also uses the constraint weight with L2 as the loss function and update the model parameters by using the Adam optimization algorithm. Finally the paper trains the model by gradually increasing the number of convolution layers and adjusting the number of neurons in the fully connected layer with the data which is augmented. The results show that the model structure in which the number of convolution layers is four and the fully connected layers' neurons are 256 is the best. It's about 85.15%. The ability is almost equal to the ability of the deeper networks.

[Key words] Convolutional Neural Network; vehicle classification; transfer learning; data augmentation

0 引言

车型分类^[1]主要研究车辆的结构化信息, 能够根据车型、车身颜色、车身标志性文字或字母等特征对车辆进行分类, 是智能交通系统中的重要功能。但车型分类易受到光照、天气、拍摄角度等因素的影响, 面临着很大的挑战, 长期以来即已成为了国内外学者研究的热点。

自从 AlexNet^[2] 赢得了 2012 的 ILSVRC (Large Visual Recognition Challenge) 竞赛以来, CNN^[3-5] 现已广泛用于视觉图像的分类任务中, 并取得了较传统人工选取特征基础上的更加准确分类效果。目前研究中, CNN 网络层次在不断增加, 参数规模也越

来越大, 这就极大地增加了人工成本, 并且参数规模过大还将影响该模型在实践中的可行性。因此, 为了使用较少的数据和快速训练更深的模型, 学者们开始使用迁移学习来搭建分类器, 然而, 迁移学习却会降低特定分类任务准确率, 而重新训练深层次的网络模型的所有参数则需要高性能计算机的技术支持。基于此, 为了训练出高精度的结构简单的车型分类模型, 本文拟从 CNN 的基本概念出发, 设计研发最符合车型分类的结构简单的分类器。有针对性地, 本文的研究工作可简述如下。

(1) 为了测试收集数据是否可以用于训练, 使用迁移模型, 再用 ImageNet^[6] 数据集训练好的深层

作者简介: 晏世武(1994-), 男, 硕士研究生, 主要研究方向: 深度学习、机器学习、图像算法; 罗金良(1968-), 男, 博士, 教授, 主要研究方向: 机械设计及理论; 严庆(1993-), 女, 硕士研究生, 主要研究方向: 机械设计及理论。

收稿日期: 2019-09-19

次的模型进行训练,观察测试集的准确率。

(2)为了解决数据样本过少的问题,对原始数据进行数据增强研究。

(3)构建了网络的卷积层、BN层、池化层和全连接层,通过调整卷积层数和全连接层神经元个数来构造出针对本文车型数据集最好的网络结构。

(4)使用增强的数据去训练设计的网络结构,并检测数据的准确率。

1 实验

本次研究中,采用迁移学习来检测经典图像分类模型在本文数据集上的仿真性能,并设计构建卷积网络结构进行训练,旨在达到使用较浅层次的网络结构也能获得较深网络层次的结果性能的目的。

1.1 基本的网络结构

为了使得较浅层次的车型分类 CNN 模型的精度也能媲美深层次的 CNN 迁移模型的车型分类精度,本文从 CNN 的概念出发,重新设计 CNN 车型分类模型。将指定图像作为 CNN 模型的输入,依次经过卷积层、BN^[7]、Relu、max_pool、dropout^[8]、Fc,综上所述处理后输出经过 Softmax 激活进行概率空间映射,最终输出每种车型的置信度,概率越大说明该输入的图像数据属于对应车型的可能性越大。本文则以 4 层卷积层为例,研发得出的基本 CNN 结构如图 1 所示。

1.2 实验数据

本文原始数据来自于 ImageNet 大型数据库中 11 种车型数据图像,由于 ImageNet 数据收集的年份较早,导致一些车型图像模糊,迹近无法辨认,本文剔除一些无法用于训练的车型图像,又考虑到 CNN 需要高清的图片作为训练的依据,因此本文从网络上下载原始数据五分之一的高清图片加入到网络模型的训练和测试中。各种车型数量见表 1。由表 1 可知,总共有 7 194 张图片。

根据 CNN 对数据数量的要求,本文搜集的图像数量是不够的,因此就要用到数据增强技术。数据增强是一种能够对原始图像进行翻转、旋转、缩放、裁剪、平移等图像操作使图像数量增多的技术。考虑车辆在道路上行驶的实际状态,本文采用平移、小角度旋转、水平翻转三种操作来进行数据增强,操作示例如图 2 所示。使用数据增强技术后图片的数目即从 7 194 张增加到 37 864 张,本文选取每一类的前 11 张图片进行展示,如图 3 所示。

在 CNN 模型中,通常将原始数据划分为训练集和测试集。其中,训练集用于模型的训练,测试集用

于模型性能的检测。对于百万级别的数据量,CNN 采用 99:1 的方式划分训练集和测试集,而对于万级别的数据量,采用 4:1 的方式划分,即每一类的 80% 的图片作为训练集训练模型,每一类的 20% 的图片作为测试集检测模型的性能。

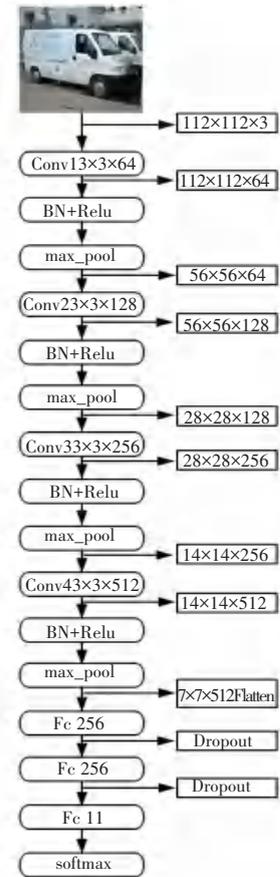


图 1 四层卷积层的 CNN 结构

Fig. 1 The CNN structure of four layers

表 1 各种车辆的图片数量

Tab. 1 The number of kinds of pictures

车型	数量/张
ambulance	756
bus	939
convertible	798
fireEngine	710
jeep	762
pickup	707
policeCar	544
rubbishCar	449
smallCar	514
taxi	361
truck	654



图 2 数据增强示例
Fig. 2 Data augumentaion



图 3 数据展示
Fig. 3 Data show

本文车型分类模型是在 Google 的 Tensorflow 平台上搭建的。Tensorflow 是一个可广泛用于机器学习和深度学习,并且基于数据流编程的数学符号系统,在该平台上能够快速搭建 CNN 模型。CNN 需要 Nvidia 显卡来加速训练,本文实验中的计算机采用 Windows 10 Enterprise 64 位系统,其硬件配置见表 2。

表 2 实验用计算机配置参数

Tab. 2 The parameters of the experiment computer

硬件	型号
处理器	AMD Ryzen 5 2600X 六核
内存(16 G)	威刚 DDR4 2 667 MHz
主硬盘	NVMe SSDPEKKW25
显卡	Nvidia GeForce GTX 1060 6 GB

2 结果及分析

2.1 迁移学习

为快速训练模型,研究中使用迁移模型来训练原始数据。迁移学习通过固定已经学习好的参数和改变想要学习的网络层,快速训练出针对特定任务的网络模型。本文中的迁移学习共同参数见表 3。

本文选用 3 个较深层次的 CNN 模型,分别是 VGG16、InceptionV3 和 ResNet50。这 3 个网络在测试集上的准确率见表 4。

表 3 迁移学习共同参数

Tab. 3 The common parameters of transfer learning

参数名称	参数值
批量大小(batch_size)	32
学习率(learning_rate)	0.001
正则化参数(weight_decay)	0.000 04
训练步数(max_number_of_steps)	10 000
优化器(optimizer)	Adam

表 4 3 种迁移模型下测试集的准确率

Tab. 4 The precision of three transfer learning

网络模型名称	测试集上的准确率/%
VGG16 ^[9]	77.70
InceptionV3 ^[10]	85.91
ResNet_50 ^[11]	63.16

由表 4 分析可知,在 3 种网络上训练一万次,训练集上的损失值已经很小并趋于稳定,训练集的准确率非常高,而在测试集上的表现差强人意,说明模型已经出现了过拟合。在 3 个模型中,性能表现最好的是 InceptionV3,这与其网络模型的卷积核大小有关,InceptionV3 采用 3 种不同的卷积核,能够得到传入网络结构图片不同大小的感受野,在最后的卷积层中也拼接了不同尺度的特征;而对于车型中的警车、救火车和出租车来说,其分类关键点在于对车顶闪亮标志的识别和车身上文字与字母的不同,警车与出租车的不同在于车顶上红蓝闪烁灯,出租车上的是“TAXI”闪烁字样;警车与救护车的不同在于车身上有“POLICE”的标识,而救护车车身上有“AMBULANCE”标识,类似这种细粒度的图像识别采用不同大小的卷积核能取得更好的结果。

2.2 模型训练

迁移学习能够快速训练出针对特定任务的模型结构,但是在测试集的准确率上还不能达到工程实践的要求。为满足高准确率的网络模型,需要对特定任务重新训练所有层的参数,然而较深层次的卷积模型结构需要非常大的数据量和高性能显卡来加速训练,并且不能很好地应用于实践。为此,本文从 CNN 的基本概念出发,设计较浅不同层数的模型结构来训练模型,并且对原始数据进行增强以达到 CNN 训练的大数据量的要求,不同层数的模型结构

的共同参数详见表5。

表5 不同卷积层数的共同参数

Tab. 5 The common parameters of different convolution layers

参数名称	参数值
批量大小(batch_size)	16
学习率(learning_rate)	0.000 1
训练步数(max_number_of_steps)	100 000
优化器(optimizer)	Adam
全连接层神经元个数	1 024

不同卷积层层数模型都采用(3,3)大小的卷积核来进行卷积,其测试效果见表6。

表6 不同卷积层数下数据集的准确率

Tab. 6 The precision of different convolution layers in dataset

卷积层数	准确率/%
3	77.69
4	84.45
5	80.16

本实验分别采用卷积层数为3、4和5的网络模型来训练增强的数据,其网络模型都出现过拟合现象,表现较好的是卷积层数为4的网络模型。为了说明对于本文车型数据来说,卷积核大小为3×3的卷积层不断堆砌易造成模型的过拟合,本文又进行了AlexNet和VGG16的实验,实验结果见表7。

表7 AlexNet和VGG16训练模型精度

Tab. 7 The precision of AlexNet and VGG16

模型	准确率/%
AlexNet	79.71
VGG16	77.70

AlexNet和VGG16都是经典的CNN结构,其中AlexNet卷积深度为5,VGG16的卷积深度为13,然而这两种模型的精度较卷积层数为4的模型低,这是由于模型过拟合造成的,因此对于本文数据来说,卷积层数为4的模型较为合适。

为进一步降低过拟合现象,需要减少卷积层数为4中的全连接层的神经元个数重新训练模型来增强模型的鲁棒性,其测试效果见表8。当全连接层的神经元个数为256时,在神经元个数为1 024的模型上其精度增加了约0.7%,说明调整神经元个数可以减轻模型的过拟合问题。

表8 不同神经元个数下测试数据的准确率

Tab. 8 The precision of dataset in different neurons

全连接层神经元个数	准确率/%
512	83.68
256	85.15
128	77.97

3 结束语

(1)从深度学习的角度出发,快速探究了CNN

在车型分类中的应用,使用迁移学习对本文的车型原始数据集进行简单的训练并测试,在使用InceptionV3、ResNet50和VGG16三种模型进行训练和测试时,发现InceptionV3模型结构对车型分类的结果表现最好,其测试集的准确率达到85.91%。

(2)从CNN基本概念出发,探究卷积层数对车型分类精度的影响,通过分析经典的CNN的模型结构,设计了本文的基本卷积神经网络结构。本文的CNN结构由卷积层、BN层、最大池化层、Dropout层、全连接层和softmax层所组成,使用带有约束权重的L2作为损失函数,并使用Adam优化算法对模型参数进行更新。测试结果表明对于本文的车型数据集,卷积层数为4且全连接层的神经元个数为256的模型结构的精度表现得较好,最终的平均精度为85.15%,能达到使用InceptionV3迁移学习训练的效果,即较浅网络结构在特定数据集能与较深结构相媲美。

参考文献

- [1] 黄跃珍,王乃洲,梁添才,等.基于改进型MobileNet网络的车型识别方法[J].电子技术与软件工程,2019(1):22-24.
- [2] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks [C]//Advances in Neural Information Processing Systems(NIPS 2012). Nevada:Nips, 2012:1097-1105.
- [3] 刘健,袁谦,吴广,等.卷积神经网络综述[J].计算机时代,2018(11):19-23.
- [4] 杨真真,匡楠,范露,等.基于卷积神经网络的图像分类算法综述[J].信号处理,2018,34(12):1474-1489.
- [5] 李策,陈海霞,汉语,等.深度学习算法中卷积神经网络的概念综述[J].电子测试,2018(23):61-62.
- [6] RUSSAKOVSKY O, DENG J, SU H, et al. ImageNet large scale visual recognition challenge [J]. International Journal of Computer Vision, 2015, 115(3):211-252.
- [7] IOFFE S, SZEGEDY C. Batch normalization: Accelerating deep network training by reducing internal covariate shift [C]//ICML'15 Proceedings of the 32nd International Conference on Machine Learning. Lille, France:ACM, 2015,37:448-456.
- [8] SRIVASTAVA N, HINTON G, KRIZHEVSKY A, et al. Dropout: A simple way to prevent neural networks from overfitting [J]. Journal of Machine Learning Research, 2014, 15(1):1929-1958.
- [9] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [C]//International Conference on Learning Representations. San Diego, USA: ICLR, 2015:1-5.
- [10] SZEGEDY C, LIU Wei, JIA Yangqing, et al. Going deeper with convolutions [C]//2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Boston, MA, USA: IEEE Computer Society, 2015:1-12.
- [11] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C]//Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: CVPR, 2016:770-778.