

基于卷积神经网络的轿车车型精细识别方法

陈宏彩^{1,2}, 程煜^{1,2}, 张常有³

(1.河北省科学院应用数学研究所,河北石家庄 050081; 2.河北省信息安全认证工程技术研究中心,河北石家庄 050081; 3.中国科学院软件研究所并行软件与计算科学实验室,北京 100190)

摘要:在复杂交通场景中,公安和交管部门对车型识别的实时性和精度提出了更高要求。针对当前假牌、套牌、无牌车辆处理占用大量警力、检索效率低下、非智能化等一系列问题,提出了一种基于 GoogleNet 深度卷积神经网络的车型精细识别方法,设计了合理的卷积神经网络滤波器大小和数目,优选了激活函数和车型识别分类器,构建了一个新的卷积神经网络轿车车型精细识别模型框架。实验结果表明,在车型精细识别测试中,所提出模型的识别率达到了 97%,较原始 GoogleNet 模型有较大提升,而且,新模型有效地减少了训练参数的数量,降低了模型的存储空间。车型精细识别技术可应用于智能交通管理领域,具有重要的理论研究价值与实践意义。

关键词: 神经网络; 车型识别; 卷积神经网络; 精细识别; 深度学习

中图分类号: TP183 文献标志码: A

Fine-grained vehicle type recognition based on deep convolution neural networks

CHEN Hongcai^{1,2}, CHENG Yu^{1,2}, ZHANG Changyou³

(1.Hebei Academy of Sciences Institute of Applied Mathematics, Shijiazhuang, Hebei 050081, China; 2. Hebei Authentication Technology Engineering Research Center, Shijiazhuang, Hebei 050081, China; 3. State Key Laboratory of Computing Science, The Chinese Academy of Sciences, Beijing 100190, China)

Abstract: Public security and traffic department put forward higher requirements for real-time performance and accuracy of vehicle type recognition in complex traffic scenes. Aiming at the problems of great police forces occupation, low retrieval efficiency, and lacking of intelligence for dealing with false license, fake plate vehicles and vehicles without plates, this paper proposes a vehicle type fine-grained recognition method based GoogleNet deep convolution neural networks. The filter size and numbers of convolution neural network are designed, the activation function and vehicle type classifier are optimally selected, and a new network framework is constructed for vehicle type fine-grained recognition. The experimental results show that the proposed method has 97% accuracy for vehicle type fine-grained recognition and has greater improvement than the original GoogleNet model. Moreover, the new model effectively reduces the number of training parameters, and saves computer memory. Fine-grained vehicle type recognition can be used in intelligent traffic management area, and has important theoretical research

收稿日期:2017-08-28;修回日期:2017-10-14;责任编辑:陈书欣

基金项目:国家自然科学基金(61379048,61672508);河北省重点研发计划项目(17395602D);河北省三三三人才工程项目(2016022577-7)

第一作者简介:陈宏彩(1981—),女,河北滦南人,助理研究员,主要从事模式识别及机器视觉方面的研究。

E-mail:chenhc88@163.com

陈宏彩,程煜,张常有.基于卷积神经网络的轿车车型精细识别方法[J].河北科技大学学报,2017,38(6):564-569.

CHEN Hongcai, CHENG Yu, ZHANG Changyou. Fine-grained vehicle type recognition based on deep convolution neural networks[J]. Journal of Hebei University of Science and Technology, 2017, 38(6): 564-569.

value and practical significance.

Keywords: computer neural network; vehicle recognition; convolution neural networks; fine-grained recognition; deep learning

中国机动车保有量的稳步增长,特别是私人汽车的迅猛增长,导致城市交通超负荷运转。车辆被盗、汽车超速、超载、闯红灯以及驾驶汽车犯罪等事件与日俱增。日益严重的交通问题给社会和人们的生活带来了很大的隐患。交通卡口电子警察的快速发展在很大程度上解决了大部分涉车事件。然而,面对海量视频图像,执法部门在处理假牌、套牌、无牌等车辆案件时,仍然无法在短时间内找到相关车辆及线索。传统的车型识别系统主要按照车辆的大中小对车型进行粗分类,因此,建立车辆精细识别系统在智能交通管理领域显得尤为迫切。

车型识别是智能交通管理系统中的一项关键技术。虽然已有研究机构在此上投入了不少人力、物力,但是,由于车辆外观复杂多样,同一品牌不同车系、不同年款的车在外观上又很相近,再加上车辆容易受到背景、光照、视角等因素的影响,所以在实际应用中一直难以找到稳定、有效的视觉特征以准确识别不同车型。

近年来,基于深度学习的相关理论和应用发展迅猛,与传统依赖先验知识的特征提取算法不同,深度神经网络可在训练数据驱动下自适应地构建特征描述,具有更高的灵活性和普适性^[1]。卷积神经网络(convolutional neural network, CNN)是深度学习的一项重要技术^[2],已在图像识别^[3]、语音识别^[4]、物体检测^[5]等领域取得卓越成绩。1994年BOTTOU等^[6]将CNNs用于手写数字识别取得了很好的效果。2012年,KRIZHEVSKY等^[7]利用深层的卷积神经网络在图像识别数据集ImageNet上取得最好成绩。2014年FAN等^[8]利用pyramid CNNs方法在人脸识别公共数据LFW上取得第1名的成绩。

传统图像识别方法主要依赖人工设计特征,当前常用的人工提取特征方法有SIFT^[9-10],LBP^[11],HOG^[12]等。通常,这些特征针对特定问题有效,而且运算量大、常常需要图像预处理工作。而CNNs方法直接输入原始图像,在训练数据的驱动下通过自主学习获取特征描述,简化特征模型的同时提高运算效率^[1]。另外,CNNs对平移、缩放等仿射形变具有良好的抵抗能力^[1],能有效克服车辆外观多变带来的困难。为此,利用卷积神经网络设计相应的车型特征提取算法,并结合Softmax分类器构建车型分类器,对网络结构进行优化调整。实验结果显示,该系统能显著提高车型识别的准确率,同时保持了较高的特征提取速度。

1 卷积神经网络

卷积神经网络模拟生物神经网络,其权值共享网络结构可产生生物视觉中的局部感受野效应,降低网络模型复杂度、减少权值数量的同时对平移、比例缩放、倾斜等视觉形变均具有高度的抵抗能力^[1]。一个基本的卷积神经网络结构可以分为特征提取层、特征映射层以及特征池化层^[13],通过不同层次结构的合理构建可以形成一个卷积神经网络。

1) 特征提取层 给定一张大小为 $r \times c$ 的图像 I ,利用一个大小为 $w \times w$ 可训练的过滤器组 K ,对其进行卷积操作, s 为步长,输出 Y ,大小为 $((r-w)/s+1) \times ((c-w)/s+1)$ 。

$$y_i = b_i + \sum_j k_{ij} * x_i, \quad (1)$$

式中: $*$ 表示卷积运算; x_i 是输入图像; k_{ij} 是卷积核参数; b_i 是偏值。

2) 特征映射层 考虑到数据处理的复杂性,通常利用一个非线性函数作用于过滤器层,得到特征图 f :

$$f_s = \sigma(b_i + \sum_j k_{ij} * x_i), \quad (2)$$

式中,激活函数 σ 反映的是神经元的输入与输出之间的映射关系^[6]。

3) 特征池化层 理论上,图像经过卷积操作之后得到的特征可以直接作用于分类器。然而,一幅图像经过卷积操作后的特征维度都很高,如果直接使用,在训练过程中容易出现过拟合现象。为了避免此类现象的发生,常常在卷积层后面卷积特征进行下采样(也称为池化)。经过池化后的图像特征维数会极大地降低。

卷积神经网络是一种全监督的深度学习模型^[14],在训练样本之前需要对样本做标签,也要初始化权重。CNN网络训练部分包含以下两阶段。

1) 正向传播阶段:将训练样本输入CNN网络得到实际输出;

2)反向传播阶段^[15]:计算实际输出与理想输出间的误差,通过反向传播算法逐层传播误差,并更新各层权值。

2 车型精细识别系统框架

传统的图像识别中常使用人工手动设计特征,这在识别同一物体的不同特征时可能效果显著,但对于不同物体识别时效果并不太理想。卷积神经网络可将原始数据直接作为输入,避免了传统识别算法中的数据预处理过程。本文提出基于卷积神经网络的车型精细识别模型,如图 1 所示。

2.1 卷积神经网络特征学习模型

CNN 模型训练过程包括构建 CNN 网络和训练 CNN 网络。

本文采用的基础网络结构是 22 层的 GoogleNet 网络^[16]。第 1 层为网络的输入 I ,因为 CNN 网络能够自动处理并提取图像特征,将具有单个车辆的原始图像直接作为网络模型的输入。网络的中间结构包括卷积层和下采样层。

首先,分析卷积神经网络结构中卷积核规模对网络模型精度的影响。如果设置的卷积核数量较多,那么提取的车辆图像特征也多。但是,卷积核数量并不是越多越好。设置过多的卷积核数量,一方面提取的特征可能会有重叠,延长训练和计算时间;另一方面,在训练网络模型时,很可能产生过拟合现象^[17]。如果设置卷积核个数为 64 个,可以算得仅一层网络就需要内存 4.19 MB,若卷积神经网络深度为 7 层,前向传播就需要 30 MB 左右,加上反向传播则需要 60 MB,处理一张图像需要内存为 60 MB,若设定批处理的图像个数为 50,则需要 3 GB 内存,所以卷积核个数的设定要结合所使用的计算机硬件设施。在本文中,笔者探索了几种不同的卷积核配置对模型识别精度的影响。由于卷积核数量较多,仅列出模型前几层和后几层卷积核数目,实验结果如表 1 所示。根据测试情况,在车型识别模型中,选择表 1 中的第 3 个卷积核数目设置方案。

表 1 不同卷积核数量在 CAR-240 数据库上的训练结果

Tab.1 Train results of different filter numbers in CAR-240 database

卷积核数目	训练识别率/%	测试识别率/%
32-64-32-...-160-32-128	90	94
64-128-16-...-256-160-320	87	91
64-192-64-...-384-128-128	93	96

在卷积神经网络结构中,不仅卷积核个数对识别精度有影响,而且卷积核大小的设置也很关键。理论上,卷积核越小,提取的特征越精细。但是,实际采集的图像可能会受到不同程度的噪声污染,如果卷积核设置过小,容易提取到一些杂乱无章的信息。通过实验验证,本文车型识别模型中卷积核大小设置为 3×3 , 5×5 , 7×7 。图 2 是第 1 个卷积层滤波器的权重可视图。从图 2 中可以看出,车辆图像经过卷积滤波后能够提取目标的边缘、角点等底层信息。

在网络的输出层中,用于训练的车型图片经过标注后有 240 类,标注的号码从 0 到 239,每一号码与具体的车辆型号是一一对应的。因此,最后基于深度卷积神经网络的模型输出节点为 240。

在卷积神经网络结构中,激活函数对模型的收敛性影响很大。设置合理的激活函数不仅可以提高目标识别的精度,而且也会加快网络模型训练的收敛速度。常用的激活函数有 Sigmoid^[18], Tanh^[19], ReLU (Rectified Linear Units)^[20-21]等。传统神经网络中常使用 Sigmoid 和 Tanh 作为激活函数。然而,这两种函数的输出不具备稀疏性^[18],而且在 2010 年 GLOROT 等^[22]发现 Sigmoid 函数在训练过程中会聚集在 0 区域,导致训练过程比较长。激活函数 ReLU 是线性的、非饱和的,对于随机梯度下降的收敛有巨大的加速作用^[7],而且只需要一个阈值就可以得到激活值,其函数图像如图 3 所示,函数表达式为

$$g(x) = \max(0, x)。$$

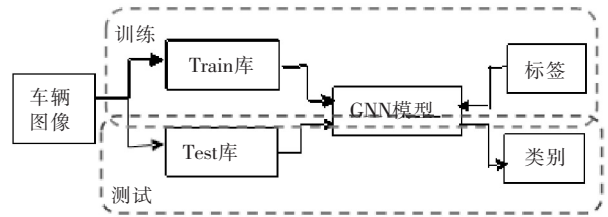


图 1 基于学习的车型识别模型

Fig.1 Model of vehicle type recognition based on learning

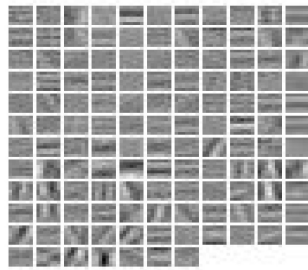


图 2 卷积层响应图

Fig.2 Convolution layer response

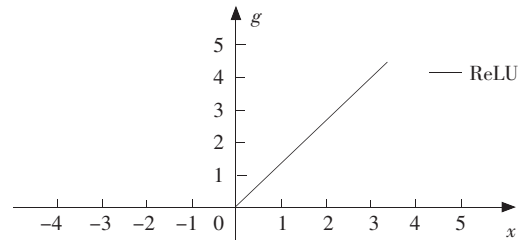


图 3 ReLU 函数图

Fig.3 ReLU function

从 ReLU 函数图(图 3)可以看出,当输入值小于 0 时,输出值等于 0,使用这样的激活函数会带来网络的稀疏性。而当网络层的输入值大于 0 时,输出值保持原值。这样在训练网络模型时会加快其收敛速度。ReLU 函数的导函数如式(4)所示:

$$g'(x) = \begin{cases} 1, & x > 0, \\ 0, & x < 0. \end{cases} \quad (4)$$

当输入值大于 0 时,ReLU 的导数输出值为 1。而输入值小于 0 时,输出结果为 0。在网络模型反向传播数据时,这个激活函数能够避免梯度消失的问题。综合 ReLU 激活函数的特点,选择 ReLU 函数作为车型识别网络模型的激活函数。

2.2 车型识别

在卷积神经网络应用于车型图像分类任务时,在神经网络最后一层全连接层后接一个 Softmax 分类器用于计算车辆属于哪个具体车型的概率,计算公式为

$$d_j^{(i)} = \frac{\exp(W_j^T x^{(i)} + a_j)}{\sum_{j=0}^{239} \exp(W_j^T x^{(i)} + a_j)} \quad (5)$$

式(5)分类器参数 $W = [W_0, W_1, W_2, \dots, W_{239}]$, $a = [a_0, a_1, a_2, \dots, a_{239}]$, 其中 $d_j^{(i)}$ 是对样本 $x^{(i)}$ 属于第 j 类的预测。

3 实验

3.1 数据集

为了验证算法的有效性,首先构建一个用于精细车型识别的图像数据库。以交通卡中的车辆图像为识别对象。为确保实验的准确和可靠性,采用 ImageNet^[19] 标注方式,自建了一个可供车辆识别研究及应用的标准数据库,该数据库包含 240 类不同品牌、车系、年款的小轿车车型图像,共采集不同尺度、光照、角度的车辆图片 58 265 张,本文将此车辆库命名为 CAR-240。每类车型样本数量在几十到几百张不等。由于每张图片的大小不一,为了满足卷积神经网络的输入要求,在本文实验中将车型图像大小归一化为 256×256 。在每种车型样本中,随机选取本车型样本数量的 75% 作为训练样本,其余的作为测试样本。部分车型图片见图 4,图 4 左侧列出了几款车辆的具体品牌、车系、年款,右侧是每款车型的样本图。



图 4 部分车型样本示例

Fig.4 Samples of some vehicle types

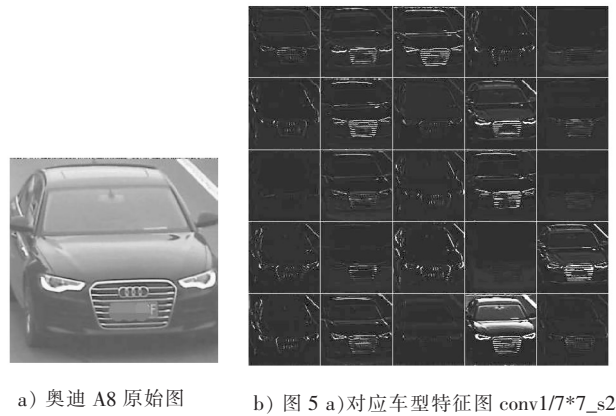
3.2 结果及分析

车型识别实验测试是在装有一块 NVIDIA-K20c GPU 的计算机上完成的。以样本车型中的奥迪 A8 车型为例(如图 5 a)所示),经过卷积神经网络特征提取后的第 1 层卷积特征如图 5 b)所示。从提取的车辆特

征图可以看出,经过第1层的卷积层特征提取,车辆的边缘轮廓清晰地被呈现出来。

模型训练过程中采用随机梯度下降法,设置初始化的学习率等于0.01,每进行10 000次迭代将学习率下降,采用步进衰减策略,衰减因子设置为0.000 5,冲量值等于0.9。车型识别模型的训练结果如图6所示,最终得到平均验证准确率为97.8%。

随机选取测试样本集中的几类车型,其中包括同一品牌、车系下不同年代的车型,每类车型选取500张,测试车型精细识别率情况。表2给出了原始GoogleNet模型和本文提出的车型识别模型在车型识别中的对比结果。实验结果表明,本文建立的卷积神经网络模型在车型精细识别上取得了较好的效果,车型精细识别正确率平均在97%,较原始GoogleNet模型提升2.3%。通过对本文方法识别错误样本进行分析得知,即便是被识别错误的车型,也是被识别为同一品牌下的不同车系或是同一车系的不同年款。这些被识别错误的车辆其外观都十分相似,通常仅在车辆的某些部件上有细微差别。



a) 奥迪 A8 原始图 b) 图 5 a) 对应车型特征图 conv1/7*7_s2

图 5 车型特征图

Fig.5 Features map of vehicle types

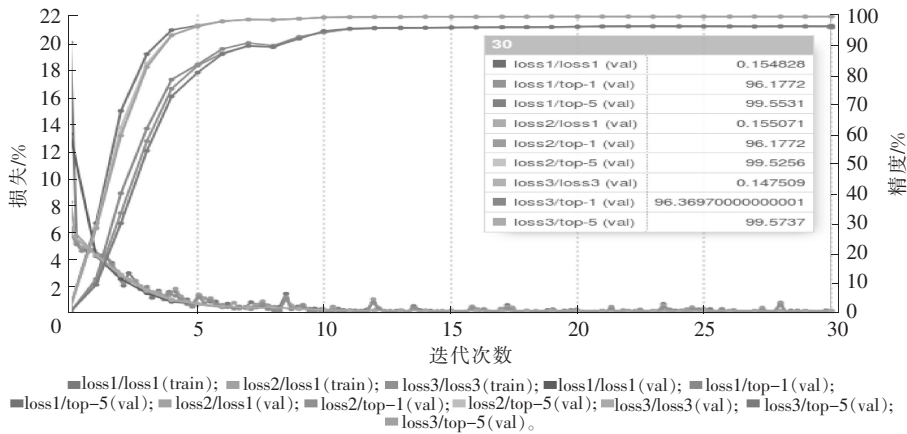


图 6 训练和测试错误率图

Fig.6 Error ratio of training and testing

表 2 部分不同车型识别的实验结果

Tab.2 Recognition result of some vehicle types

车 型	GoogleNet 模型识别率/%	本文方法识别率/%	车 型	GoogleNet 模型识别率/%	本文方法识别率/%
奥迪_A6L_2011	94.2	96.4	尼桑_逍客	97.3	99.8
奥迪_A4L_2015	95.2	97.7	丰田_凯美瑞_2006	95.8	98.4
宝马_5系	97.4	99.8	马自达_6_2011-2015	89.0	91.6
标致_207	95.6	97.7	长城_哈佛 H6_2011	94.2	96.4
大众_POLO_2013	95.6	97.7			
平均识别率	94.7	97.0			

4 结 论

针对现有车型识别方法中无法对车辆品牌、车系、年款等综合信息进行精细识别的问题,建立了一种新的卷积神经网络结构的车型精细识别模型。新模型建立在基础模型 GoogleNet 框架之上,设计了一个适度规模大小的网络结构,通过实验测试选取了最佳的滤波器大小和数目,分析并优选了用于车型识别的激活函数和分类器,最终构建了一个适用于轿车车型精细化识别的新的卷积神经网络模型。实验结果表明,建立的

车型识别模型在车型精细化识别测试中识别率达到了97%,与经典的卷积神经网络模型 GoogleNet 方法相比,不仅在识别效果上有所提高,而且训练参数大幅度减少,训练所需的时间更短。该方法能够准确识别车辆品牌、车系和年款,可帮助车管所、交管部门实现假牌车、套牌车的自动检测,准确率高,实时性强。但是,目前公开的车型数据库信息较少,车辆更新速度比较快,而手工标注车型的成本较大,导致该方法在实际应用中存在一定的局限性。因此无监督或半监督的车型识别方法将是今后的一个研究方向。

参考文献/References:

- [1] 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.
DENG Liu, WANG Zijie. Deep convolution neural networks for vehicle classification[J]. Application Research of Computers, 2016, 33(3): 930-932.
- [2] 李彦冬,郝宗波,雷航. 卷积神经网络研究综述[J]. 计算机应用, 2016, 36(9): 2508-2515.
LI Yandong, HAO Zongbo, LEI Hang. Survey of convolutional neural network[J]. Journal of Computer Applications, 2016, 36(9): 2508-2515.
- [3] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[C]//ICLR.[S.l.]:[s.n.], 2015: 1-14.
- [4] HAN K, WANG D L. Neural network based pitch tracking in very noisy speech[J]. IEEE/ACM Transactions on Audio, Speech and Language Processing(TASLP), 2014, 22(12): 2158-2168.
- [5] 刘洪公,王学军,李冰莹,等. 基于卷积神经网络的桥梁裂缝检测与识别[J]. 河北科技大学学报, 2016, 37(5): 485-490.
LIU Honggong, WANG Xuejun, LI Bingying, et al. Detection and recognition of bridge crack based on convolutional neural network[J]. Journal of Hebei University of Science and Technology, 2016, 37(5): 485-490.
- [6] BOTTOU L, CORTES C, DENKER J S, et al. Comparison of classifier methods: A case study in handwritten digit recognition[C]//Proceedings of the 12th IAPR International. [S.l.]: IEEE Press, 1994: 77-82.
- [7] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[C]//Proceedings of the 25th International Conference on Neural Information Processing Systems. [S.l.]:[s.n.], 2012: 1097-1105.
- [8] FAN Haoqiang, CAO Zhimin, JIANG Yuning, et al. Learning deep face representation[EB/OL]. <https://arxiv.org/pdf/1403.2802.pdf>, 2014-03-12.
- [9] GE Fengxiang, SHI Yishu, SUN Bo, et al. Sparse representation based classification by using PCA-SIFT descriptors[C]// 2014 4th IEEE International Conference on Information Science and Technology(ICIST). Shenzhen: IEEE, 2014: 429-432.
- [10] 冯玉朋,曾庆喜,马杉,等. 无人驾驶车辆基于角点和斑点的特征提取算法[J]. 河北科技大学学报, 2017, 38(3): 237-243.
FENG Yupeng, ZENG Qingxi, MA Shan, et al. A feature extraction algorithm based on corner and spots in self-driving vehicles[J]. Journal of Hebei University of Science and Technology, 2017, 38(3): 237-243.
- [11] RANJAN R, PATEL V M, CHELLAPPA R. A deep pyramid deformable part model for face detection[C]// 2015 IEEE 7th International Conference on Biometrics Theory, Applications and Systems (BTAS). Arlington: IEEE, 2015: 1-8.
- [12] DELORME A, MULLEN T, KOTHE C, et al. EEGLAB, SIFT, NFT, BCILAB, and ERICA: New tools for advanced EEG processing [J]. Computational Intelligence and Neuroscience, 2011, 2011:1-12.
- [13] 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.
ZHOU Feiyan, JIN Linpeng, DONG Jun. Review of convolutional neural network[J]. Chinese Journal of Computers, 2017, 40(6): 1229-1251.
- [14] 何克磊,史颖歆,高阳,等. 一种基于原型学习的多示例卷积神经网络[J]. 计算机学报, 2017, 40(6):1265-1274.
HE Kelei, SHI Yinghuan, GAO Yang, et al. A prototype learning based multi-instance convolutional neural network[J]. Chinese Journal of Computers, 2017, 40(6): 1265-1274.
- [15] RAZAVIAN A S, AZIZPOUR H, SULLIVAN J, et al. CNN features off-the-shelf: An astounding baseline for recognition[C]// Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition. Washington DC: IEEE Computer Society, 2014: 512-519.
- [16] SZEGEDY C, LIU W, JIA Y Q, et al. Going deeper with convolutions[C]// IEEE Conference on Computer Vision and Pattern Recognition(CVPR). Boston: IEEE, 2016:1-9.
- [17] ZBONTAR J, LECUN Y. Stereo matching by training a convolutional neural network to compare image patches[J]. Journal of Machine Learning Research, 2016, 17:1-32.
- [18] AGOSTINELLI F, HOFFMAN M D, SADOWSKI P J, et al. Learning Activation Functions to Improve Deep Neural Networks[EB/OL]. <https://arxiv.org/pdf/1412.6830.pdf>, 2015-12-21.
- [19] 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.
- [20] WANG Jinjun, YU Kai, LV Fengjun, et al. Locality-constrained linear coding for image classification[C]// IEEE Conference on Computer Vision and Pattern Recognition. San Francisco: IEEE, 2010: 3360-3367.
- [21] DAHL G E, SAINATH T N, HINTON G E. Improving deep neural networks for LVCSR using rectified linear units and dropout[C]// 2013 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Vancouver: IEEE, 2013: 8609-8613.
- [22] GLOROT X, BENGIO Y. Understanding the difficulty of training deep feedforward neural networks[J]. Proceedings of Machine Learning Research, 2010, 9:249-256.