

基于胶囊网络的智能交通标志识别方法

陈立潮¹, 郑佳敏¹, 曹建芳^{1,2*}, 潘理虎¹, 张 睿¹

(1. 太原科技大学 计算机科学与技术学院, 太原 030024; 2. 忻州师范学院 计算机系, 山西 忻州 034000)

(* 通信作者电子邮箱 kcxjdj122@126.com)

摘 要: 针对卷积神经网络的标量神经元无法表达特征位置信息, 对复杂的车辆行驶环境适应性差, 导致交通标志识别率低的问题, 提出一种基于胶囊网络的智能交通标志识别方法。首先采用超深度卷积神经网络改进特征提取部分, 然后在主胶囊层引入池化层, 最后采用移动指数平均法改进了动态路由算法。在 GTSRB 数据集上的测试结果表明, 改进后的胶囊网络方法在特殊场景下的识别精度提高了 10.02 个百分点, 相对于传统的卷积神经网络, 该方法的单张图片的识别时间缩短了 2.09 ms。实验结果表明, 改进后的胶囊网络方法能满足准确、实时的交通标志识别要求。

关键词: 交通标志识别; 胶囊网络; 超深度卷积; 动态路由算法; 深度学习

中图分类号: TP391.3 **文献标志码:** A

Intelligent traffic sign recognition method based on capsule network

CHEN Lichao¹, ZHENG Jiamin¹, CAO Jianfang^{1,2*}, PAN Lihu¹, ZHANG Rui¹

(1. College of Computer Science and Technology, Taiyuan University of Science and Technology, Taiyuan Shanxi 030024, China;

2. Computer Department, Xinzhou Teachers University, Xinzhou Shanxi 034000, China)

Abstract: The scalar neurons of convolutional neural networks cannot express the feature location information, and have poor adaptability to the complex vehicle driving environment, resulting in low traffic sign recognition rate. Therefore, an intelligent traffic sign recognition method based on capsule network was proposed. Firstly, the very deep convolutional neural network was used to improve the feature extraction part. Then, a pooling layer was introduced in the main capsule layer. Finally, the movement index average method was used for improving the dynamic routing algorithm. The test results on the GTSRB dataset show that the improved capsule network method improves the recognition accuracy in special scenes by 10.02 percentage points. Compared with the traditional convolutional neural network, the proposed method has the recognition time for single image decreased by 2.09 ms. Experimental results show that the improved capsule network method can meet the requirement of accurate and real-time traffic sign recognition.

Key words: traffic sign recognition; capsule network; very deep convolution; dynamic routing algorithm; deep learning

0 引言

交通标志识别是先进驾驶员辅助系统和无人驾驶系统重要组成部分。由于车辆在各种环境、光照条件、速度和地理位置下行驶, 深度学习算法必须始终保持鲁棒性和可靠性。交通标志识别领域面临的主要问题有: 环境、天气的影响; 图像采集角度的问题; 遮挡物影响; 实时性问题。目前交通标志识别常用的算法有方向梯度直方图^[1]、支持向量机^[2]、卷积神经网络(Convolutional Neural Network, CNN)等, 在正常行驶环境下, 基于卷积神经网络的 AlexNet^[3]、LeNet-5^[4]对于特殊场景识别精度和适应性存在不足。

传统的图像识别算法对于环境的适应性较差, 识别速度缓慢, 已经无法满足高精度实时性的要求。胶囊网络(Capsule Network, CapsNet)^[5]是一种新型卷积神经网络模型, 使用向量神经元克服了 CNN 对物体之间的空间辨识度差及

物体大幅度旋转之后识别能力低下的两个缺陷, 有效地弥补了 CNN 模型的不足; 动态路由算法的快速收敛及简单的网络结构, 不仅保证了较高的识别精度, 而且收敛速度得到了极大的提升。

针对目前交通标志识别存在的问题, 本文提出了一种基于胶囊网络的智能交通标志识别方法。胶囊网络本身对图像旋转的适应性很好地解决了图像采集角度问题, 通过超深度卷积模型改进了胶囊网络结构中的特征提取部分, 并在主胶囊层引入池化层, 以及采用移动指数平均法改进动态路由算法, 提高了胶囊网络在交通标志识别领域的识别精度, 并降低了运算负荷, 实现了比原结构更优越的性能。

1 胶囊网络

1.1 基本结构

胶囊网络是一个非常浅的卷积神经网络, 其结构如图 1

收稿日期: 2019-09-23; 修回日期: 2019-11-04; 录用日期: 2019-11-05。 基金项目: 山西省应用基础研究项目(201801D221179)。

作者简介: 陈立潮(1961—), 男, 山西万荣人, 教授, 博士, CCF 高级会员, 主要研究方向: 大数据软件工程、智能图像信息处理; 郑佳敏(1994—), 男, 山西晋中人, 硕士研究生, 主要研究方向: 图像处理、深度学习; 曹建芳(1976—), 女, 山西忻州人, 教授, 博士, CCF 高级会员, 主要研究方向: 数字图像理解、大数据; 潘理虎(1974—), 男, 河南上蔡人, 副教授, 博士, CCF 高级会员, 主要研究方向: 智能信息处理; 张睿(1987—), 男, 山西太原人, 讲师, 博士, CCF 会员, 主要研究方向: 智能信息处理。

所示,基本结构为:输入层、卷积层、主胶囊层和数字胶囊层。卷积层采用 9×9 步长为 1 的卷积核进行特征提取,特征提取不完全,因此在交通标志识别上精度较低。

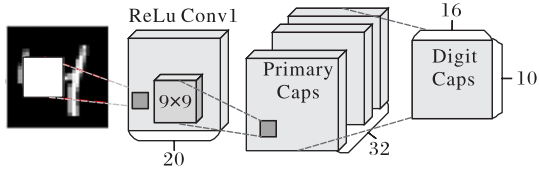


图 1 胶囊网络基本结构

Fig. 1 Basic structure of capsule network

1.2 主胶囊层

主胶囊层是胶囊网络的核心之一,实现了标量到向量的转化,原胶囊网络主胶囊层的输入维度为 $20 \times 20 \times 256$ 的特征图,通过 8 组 $9 \times 9 \times 32$ 步长为 2 的卷积核卷积得到 8 组 $6 \times 6 \times 32$ 的特征图,然后将 8 组特征图对应位置线性组合,得到 1152 个胶囊,具体过程如图 2 所示。

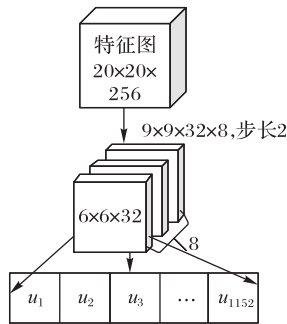


图 2 主胶囊层

Fig. 2 Main capsule layer

1.3 动态路由算法

主胶囊层实现了标量神经元到向量神经元的过渡,胶囊之间的通信无法使用传统的全连接方式,传统神经网络的反向传播 (Back Propagation, BP) 训练算法也不适用于 CapsNet 的训练,于是 Sabour 等^[5]提出了胶囊之间新的训练算法——动态路由 (dynamic routing) 算法,该算法的主要思想可概况为:低级别胶囊需要将输出传递给高级别胶囊。具体就是,路由的数据传输是由低层次胶囊的输入 (预测向量) 与高层次胶囊的输出 (激活向量) 的相似度决定,如果低层胶囊的预测向量与高层胶囊的激活向量具有较高相似度,则判定两个胶囊是高度相关的。

图 3 为动态路由算法向量运算的形象表示,特征分量加权之后相加,得到预测向量 s_j ,采用非线性激活函数得到 v_j :

$$v_j = \frac{\|s_j\|^2}{1 + \|s_j\|^2} \times \frac{s_j}{\|s_j\|} \quad (1)$$

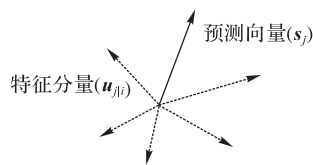


图 3 动态路由算法

Fig. 3 Dynamic routing algorithm

用 v_j 与特征分量的向量点积表示高层胶囊 v_j 与低层胶囊 v_{j-1} 的相关程度,调整特征分量的权重,使特征分量的加权求

和结果越来越接近真实向量。

2 结构改进的胶囊网络

本文所提出的网络结构如图 4 所示,图中 Conv3-128 表示 128 个 3×3 的卷积核,首先采用深度特征提取结构提取深度特征,然后通过引入最大池化层 (MaxPool) 的主胶囊层实现标量神经元到向量神经元的转换,最后用自适应路由算法实现胶囊间的参数更新,由数字胶囊层 (Digital Capsule Layer, DigitCaps) 得到识别结果。

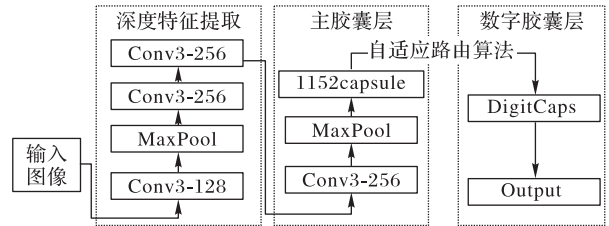


图 4 深度特征提取网络

Fig. 4 Depth feature extraction network

2.1 特征提取结构的改进

深度卷积网络有助于提取图像深层特征,从而提高识别精度^[6]。超深度卷积神经网络由牛津大学计算机视觉组 (Visual Geometry Group, VGG) 与 Google DeepMind 公司共同研发,该网络结构说明了深度有益于提高分类的正确率,通过在传统的卷积网络框架中使用更深的层能够在同样的数据集上取得优异的结果^[7]。VGG 通过使用大量 3×3 的小卷积核进行堆叠,稳步增加网络深度,用更少的参数实现了与大卷积核相同的感受野,不但减少了网络参数,而且提高了训练速度。因此,本文基于 VGG 模型改进胶囊网络特征提取部分,实现深度特征提取,达到提高识别精度的目的。

如图 5 所示,本文采用多个 3×3 卷积核改进原结构,在基于 VGG 结构上,提出三种深度特征提取结构,特征提取层数分别为 2 层、3 层、5 层。训练时,输入图片均调整为 28×28 的灰度图,激活函数选取线性整流函数 (Rectified Linear Unit, ReLU)^[8],采用最大池化层进行数据降维,得到 $14 \times 14 \times 256$ 的特征图。

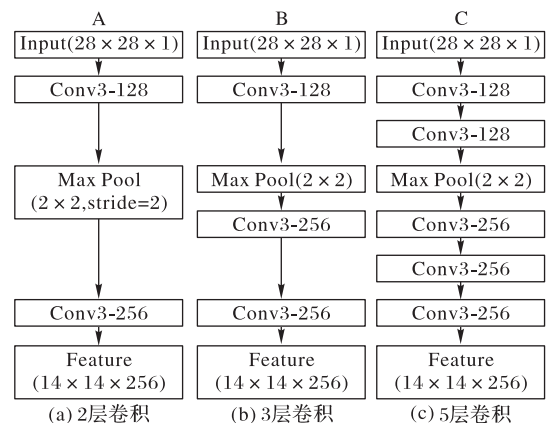


图 5 三种深度特征提取结构

Fig. 5 Three depth feature extraction structures

2.2 主胶囊层的改进

传统卷积神经网络的最大创新就是共享卷积参数与池化压缩参数,提高网络表达能力,减小过拟合风险^[9]。Hinton 认为 CNN 中的池化操作是一个灾难,因为池化过程在一定程度

上弱化了神经网络结构对于位置信息的准确把握,并不能精准地学习到不同物体的位置关联^[10]。但是数据降维获得抽象特征是每一个卷积神经网络的重点,胶囊网络也不例外,主胶囊层为了数据降维,原胶囊网络采用步长为2的卷积操作,这种方式必定会丢失特征信息,尽管最大池化也丢失特征信息,但是有选择地保留了主要特征^[11]。

如图6所示,如果使用步长为1的卷积核加最大池化操作,将会得到4×4区域中最显著的特征,如果采用步长为2的卷积操作,将会受图片中数字位置的影响,得到四个卷积结果的其中之一,影响识别精度,因此,通过引入池化层实现数据降维可以提高模型性能^[12]。

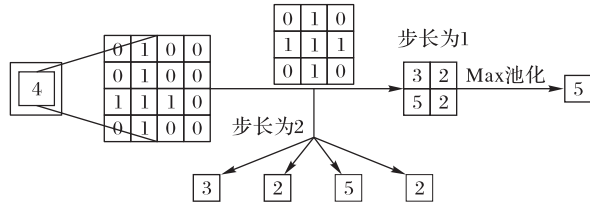


图6 池化层的作用

Fig. 6 Role of pooling layer

本文深度特征提取得到的特征图为 $14 \times 14 \times 256$,根据上述分析,主胶囊层采用8组 $3 \times 3 \times 32$ 步长为1的卷积核,对比实验采用8组 $3 \times 3 \times 32$ 步长为2的卷积核,用VALID方式填充,进行对比实验,验证池化层引入的效果。

2.3 动态路由算法的改进

主胶囊层通过引入池化层实现了数据降维,但是训练后期无法稳定收敛,精度出现2%的上下浮动,分析发现,原动态路由算法更新参数 b_{ij} 时直接使用误差更新,未设置学习率,无法适应深度特征提取网络的训练过程,为了解决这个问题,本文采用随机优化方法Adam(Adaptive moment estimation)^[13],对胶囊之间的误差进行移动指数平均,实现自适应矩估计,动态调整路由算法的学习率,使其平稳收敛。

算法1 改进后的动态路由算法。

输入 第 l 层胶囊,迭代次数 r 。

输出 b_{ij} 。

1) 初始化参数。

第 l 层胶囊 i 路由到 $l+1$ 层胶囊 j 的先验概率: $b_{ij} \leftarrow 0$

一阶矩: $m_i = 0$

二阶矩: $n_i = 0$

指数衰减率: $a = 0.9, b = 0.999$

步数: $t = 0$

稳定数值的常数: $\varepsilon = 10^{-8}$

2) for r do

3) $c_{ij} \leftarrow \text{softmax}(b_{ij})$

4) $s_j \leftarrow \sum_i c_{ij} u_{ji}$

5) $v_j \leftarrow \text{squash}(s_j)$

6) $m_i \leftarrow a m_{i-1} + (1-a) u_{ji} \cdot v_j$

7) $n_i \leftarrow b n_{i-1} + (1-b) (u_{ji} \cdot v_j)^2$

8) $\hat{m}_i \leftarrow \frac{m_i}{1-a^t}$

9) $\hat{n}_i \leftarrow \frac{n_i}{1-b^t}$

10) $b_{ij} \leftarrow b_{ij} + \frac{\hat{m}_i}{\sqrt{\hat{n}_i} + \varepsilon}$

11) end for

3 实验与结果分析

3.1 实验环境与数据集

硬件环境:i7-8700k, GTX 1060-6G, 16 GB内存。

软件环境: Windows 10, Python 3.6.5, TensorFlow-GPU

1.11.0, CUDA-9.2, cuDNN-v7.5.0

本文实验数据集采用德国交通标志数据,该数据集包括43类交通标志,训练集39 209张,测试集12 630张,其中有大量模糊不清的图像,部分遮挡、倾斜、运动模糊等不利条件下的图像,基本可以达到实际情况下的识别难度。部分交通标志如图7所示。



图7 数据集示例

Fig. 7 Dataset examples

本文网络训练测试均采用灰度图,网络训练参数定义如下:每批128张,路由更新次数为3,一阶矩修正指数 a 为0.9,二阶矩修正指数 b 为0.999,稳定常数 ε 为 10^{-8} ,迭代50次。

3.2 实验结果分析

3.2.1 确定特征提取深度

用图5给出的三种深度特征提取结构进行对比实验,确定特征提取的深度。表1中A-CapsNet、B-CapsNet、C-CapsNet分别采用2、3、5层卷积层进行特征提取,之后与胶囊网络原结构进行对比。

表1 原结构模型与深度特征提取模型的精度对比

Tab. 1 Accuracy comparison of original structure model and depth feature extraction models

算法	精度/%	
	无池化	引入池化层
A-CapsNet	96.52	97.13
B-CapsNet	96.89	97.84
C-CapsNet	96.67	97.57
CapsNet	93.17	95.56

表1显示,原胶囊网络的识别精度为93.17%,三种深度特征提取模型精度均比原模型高,因此深度特征提取结构很好地提高了模型性能。从表1可以看出:深度特征提取参数层为3层时精度达到最高;A结构深度较浅,特征提取不完全,导致精度较低;C结构可能在训练后期出现梯度消失问题,导致精度无法再提高。另外池化层的对比显示,主胶囊层引入池化层可以一定程度上提高精度,因为池化层所选择的是局部主要特征,而步长为2的卷积操作得到的特征会受到图像拍摄角度的影响,而交通标志识别算法适应车辆行驶过程中

的拍摄角度是非常重要的,因此主胶囊层引入池化层提高了网络性能,而步长为2的卷积操作严重影响网络性能。本文接下来的对比实验均采用在主胶囊层引入池化层的 B-CapsNet 结构。

3.2.2 验证动态路由算法改进效果

图8是采用自适应路由算法与动态路由算法的损失对比,是训练批次达到6 000以后的损失变化情况,黑色线条为本文算法损失变化,灰色线条为采用动态路由算法的损失变化,此时网络已经基本收敛。从图8可以看出,灰色线条的波动程度远大于黑色线条,原动态路由算法在训练后期损失振荡严重,不利于网络收敛,而改进后的波动程度很小,更加稳定。这是因为使用移动指数平均法会使近几次的损失占更大的比重。训练初期损失较大,所以参数更新速度很快,训练后期损失较小,即使某一次损失较大,也会因为近几次的平均而减弱波动程度,因此不会出现太大的震荡,既保证了训练前期的快速更新,又保证了训练后期的平稳收敛。

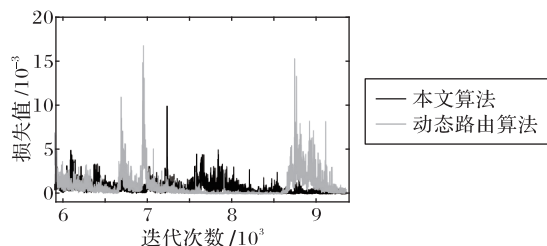


图8 本文算法与动态路由算法的损失对比

Fig. 8 Loss comparison of the proposed algorithm and dynamic routing algorithm

3.2.3 本文算法性能验证

为了验证本文改进后的胶囊网络的性能,本文将43类交通标志的分类结果分为三大类以方便展示:禁令标志(红色圆形)、警告标志(红色三角形)和指示标志(蓝色圆形)。分别计算三大类的识别精度及总精度,评价模型性能,实验结果如表2所示。

表2 本文算法性能评价

Tab. 2 Performance evaluation of the proposed algorithm

标志类别	精度/%	召回率/%	F1 Score
禁令	97.8	99.0	0.984
警告	98.3	95.3	0.968
指示	97.6	97.8	0.977

从表2可看出,禁令标志的识别效果最好,警告标志识别效果最差,这是因为警告标志为红色三角形,其中有多种相似的图案,区分度较弱,在识别过程中有较大的难度,因此识别效果较差。

为了验证本文算法在特殊场景中的识别效果,采用3.2.1节中的B-CapsNet训练得到的模型和训练好的AlexNet、Let5-Net模型,并从原数据集中挑选图像模糊、变形和缺失等难以识别的图片,测试三个模型,以验证本文算法在特殊场景下的识别效果,测试结果如图9所示。为了直观地表示识别

结果,采用清晰的能够代表类别的图片表示识别结果。

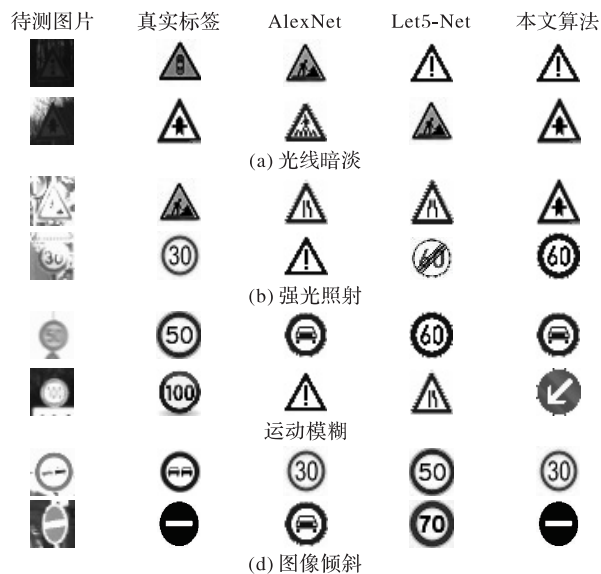


图9 特殊场景识别部分结果

Fig. 9 Some results of special scene recognition

图9中待测的交通标志图片模糊、变形和缺失非常严重,在这种极难识别的场景中本文算法仍然能够正确识别2张,而另外两种算法均无法正确识别,可见本文算法识别特殊场景中的交通标志拥有一定的优势。根据待测图片的场景分析,本文算法与经典的交通标志识别算法在光线暗淡场景下的识别错误率都比较高,图像采样的低分辨率或摄像机运动而产生的图像模糊场景下识别率也比较低,可见无法提取到图像特征是识别错误的根本原因。表3是本次实验的精度及识别耗时的数据记录。

表3 不同算法性能对比结果

Tab. 3 Performance comparison of different algorithms

分类方法	精度/%	单张耗时/ms
AlexNet	79.34	7.65
Let5-Net	76.68	5.98
本文算法	89.36	5.56

从表3结果可以看出,本文算法能够取得很好的分类结果,在特殊场景下,识别精度比AlexNet提升了10.02个百分点,比Let5-Net提升了12.68个百分点。这是因为胶囊网络采用向量神经元,保留图像特征的位置信息,因此对整体同变性有非常好的鲁棒性,能够正确识别不同拍摄角度下的交通标志;而基于卷积神经网络的经典图像识别算法使用标量神经元,没有保留特征的位置信息,对于图像旋转的鲁棒性较差,不能很好地解决拍摄角度问题,因此经典的CNN算法相对本文算法精度较低。本文算法对单张图片的识别时间相对AlexNet缩短了2.09 ms,相对Let5-Net缩短了0.42 ms,效率提升不明显。虽然本文算法的网络复杂度低于另外两种算法,但是动态路由算法部分运算复杂,占用了总识别时间的80%。经测试,训练过程主要耗时也在动态路由算法部分,因此动态路由算法是胶囊网络识别效率提升不明显的主要原因。

4 结语

本文提出的结构改进的胶囊网络,首先采用VGG深层特征提取思想进行深度特征提取,为主胶囊层提取图片的深层特征;然后在主胶囊层引入池化层,降低数据维度,并保留显著特征;最后用移动指数平均法改进动态路由算法,使学习率可以自适应调节,以适应深度特征提取网络的训练,使模型能够快速平稳地收敛,最终实现了比原结构更好的性能。结构改进的胶囊网络提高了交通标志识别的精度,解决了基于卷积神经网络的经典结构在复杂场景下识别率低的问题。通过对识别效率和误报图像的分析,本文所提出的模型采用的动态路由算法复杂度较高,识别场景受光线影响较大,下一步将针对动态路由算法进行优化,提高识别效率,并完善网络结构,提高对光线强弱的适应性,进一步提高识别精度,使其能够更加高效、准确地进行交通标志识别。

参考文献 (References)

- [1] BISWAS R, TORA M R, BHUIYAN F H. LVQ and HOG based speed limit traffic signs detection and categorization [C]// Proceedings of the 2014 International Conference on Informatics, Electronics and Vision. Piscataway: IEEE, 2014: 1-6.
- [2] HOU Y, HAO X, CHEN H. A cognitively motivated method for classification of occluded traffic signs [J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2016, 47(2): 255-262.
- [3] 张邯,罗晓曙,袁荣尚. 基于优化的卷积神经网络在交通标志识别中的应用[J]. 现代电子技术, 2018, 41(21): 132-136. (ZHANG H, LUO X S, YUAN R S. Application of optimized convolutional neural network in traffic sign recognition [J]. Modern Electronics Technique, 2018, 41(21): 132-136.)
- [4] 汪贵平,盛广峰,黄鹤,等. 基于改进LeNet-5网络的交通标志识别方法[J]. 科学技术与工程, 2018, 18(34): 78-84. (WANG G P, SHENG G F, HUANG H, et al. Traffic sign recognition method based on improved LeNet-5 network [J]. Science Technology and Engineering, 2018, 18(34): 78-84.)
- [5] SABOUR S, FROSST N, HINTON G E. Dynamic routing between capsules [C]// Proceedings of the 31st Conference on Neural Information Processing Systems. New York: Curran Associates Inc., 2017: 3859-3869.
- [6] 洪瑞,康旭东,郭军,等. 基于复杂网络描述的图像深度卷积分类方法[J]. 计算机应用, 2018, 38(12): 3399-3402. (HONG R, KANG X D, GUO J, et al. Image deep convolution classification method based on complex network description [J]. Journal of Computer Applications, 2018, 38(12): 3399-3402.)
- [7] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [EB/OL]. [2019-07-20]. <https://arxiv.org/pdf/1409.1556.pdf>.
- [8] GLOROT X, BORDES A, BENGIO Y. Deep sparse rectifier neural networks [C]// Proceedings of the 14th International Conference on Artificial Intelligence and Statistics. New York: JMLR. org, 2011: 315-323.
- [9] 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 F Y, JIN L P, DONG J. Review of convolutional neural network [J]. Chinese Journal of Computers, 2017, 40(6): 1229-1251.)
- [10] 曹荣川. 图像分类中基于显著关系的改进特征编码与池化算法研究[D]. 长春: 吉林大学, 2017: 34-41. (CAO R C. Research on improved feature coding and pooling algorithm based on saliency relationship in image classification [D]. Changchun: Jilin University, 2017: 34-41.)
- [11] 付家慧,吴晓富,张索非. 基于仿射变换的胶囊网络特征研究[J]. 信号处理, 2018, 34(12): 1508-1516. (FU J H, WU X F, ZHANG S F. Study on characteristics of capsule network based on affine transformation [J]. Journal of Signal Processing, 2018, 34(12): 1508-1516.)
- [12] 胡聪,屈瑾瑾,许川佩,等. 基于自适应池化的神经网络的服装图像识别[J]. 计算机应用, 2018, 38(8): 2211-2217. (HU C, QU J J, XU C P, et al. Garment image recognition based on adaptive pooling neural network [J]. Journal of Computer Applications, 2018, 38(8): 2211-2217.)
- [13] KINGMA D P, BA J L. Adam: a method for stochastic optimization [EB/OL]. [2019-07-20]. <https://arxiv.org/pdf/1412.6980.pdf>.

This work is partially supported by the Applied Basic Research Project of Shanxi Province (201801D221179).

CHEN Lichao, born in 1961, Ph. D., professor. His research interests include big data software engineering, intelligent image information processing.

ZHENG Jiamin, born in 1994, M. S. candidate. His research interests include image processing, deep learning.

CAO Jianfang, born in 1976, Ph. D., professor. Her research interests include digital image understanding, big data.

PAN Lihu, born in 1974, Ph. D., associate professor. His research interests include intelligent information processing.

ZHANG Rui, born in 1987, Ph. D., lecturer. His research interests include intelligent information processing.