

树状卷积神经网络的车标识别应用^①

吴章辉, 李志清, 杨晓玲, 刘雨桐

(湘潭大学 信息工程学院, 湘潭 411105)

摘要: 为了提高在自然环境下车标识别率, 提出一种多通路树状结构的卷积神经网络模型. 该模型采用多通路树状结构, 在传统卷积网络单一种类卷积核的卷积层上, 使用多种类型的卷积核进行卷积操作, 并且采用树状网络结构. 通过对每个通路的顶层提取特征, 作为全连接层的输入, 进行车标的分类任务. 通过理论分析和实验表明, 与传统的卷积神经网络训练获得的分类器相比, 车标识别率提升至 98.43%.

关键词: 深度学习; 卷积神经网络; 特征图; 车标识别

引用格式: 吴章辉, 李志清, 杨晓玲, 刘雨桐. 树状卷积神经网络的车标识别应用. 计算机系统应用, 2017, 26(10): 166-171. <http://www.c-s-a.org.cn/1003-3254/6002.html>

Vehicle Logo Recognition Using Tree-Based Convolution Neural Network

WU Zhang-Hui, LI Zhi-Qing, YANG Xiao-Ling, LIU Yu-Tong

(The College of Information Engineering, Xiangtan University, Xiangtan 411105, China)

Abstract: In order to improve the recognition rate of vehicle in natural situations, this paper proposes a vehicle logo recognition modal based on a multi-path tree structure convolutional neural networks, which modal with different convolution kernel in the same convolutions, namely T-CNN. Firstly, different layer convolution features are obtained and are joined together as the input of the fully connected layer to get classifiers. Compared with the traditional method, the theoretical analysis and simulation results show that T-CNN can increase the recognition accuracy up to 98.43%.

Key words: deep leaning; convolutional neural network(CNN); feature map; vehicle logo recognition

车标自动识别系统是车辆自动识别系统的重要组成部分. 车标是车辆的重要信息, 包含了车型信息, 更重要的是包含了厂家的信息, 由于厂家信息难于更换, 有了车型的信息, 在交通事故和利用车辆犯罪案件中能迅速缩小范围, 对车辆信息识别提供了很大的作用.

对于车标识别, 国内外已经有一些研究人员使用各类方法去尝试. 如文献[1]提出在 PCA 的基础上, 利用图像的质量分析, 通过对车标图像模糊度的计算, 将车标质量相似图像进行分类. 文献[2]中提出利用 adaboost 算法进行车标图像的定位, 然后使用 Techebichef 矩进行识别. 文献[3,6]中采用方向梯度直方图 HOG 算法进行特征的提取, 然后使用支持向量机 SVM 分类. 文献

[4]采用了深度学习方法进行车标的特征自动提取和分类, 分类正确率有很大的提升, 并且网络的输入是原始的图像, 并不要人为的图像预处理. 文献[5]改进了传统的卷积神经网络, 利用最后两个池化层的特征进行分类, 车标的识别率比较传统的卷积网络有了一定的提高. 但是, 目前车标识别系统在对光照不均, 部分缺失, 形变等因素下的车标图像正确识别率存在不足. 为了降低这些因素的影响, 提出了多通路的树状结构的卷积神经网络模型. 该模型能够充分利用网络底层提取的特征, 并且能够多尺度提取图像特征. 实验证明, 在提出的 T-CNN 网络进行测试, 车标的正确识别率达到了 98.43%.

^① 收稿时间: 2017-01-12; 采用时间: 2017-02-20

1 树状卷积神经网络

1.1 卷积神经网络

传统的卷积神经网络结构一般是由输入层, 多个交替的卷积层 (C) 和池化层 (P), 全连接层 (FC) 及输出层 (O) 组成. 如图 1 所示.

一般地, 卷积层为特征提取层, 每个神经元的输入与前一层的局部感受野相连, 提取该局部的特征. 在同一个特征图上的神经元连接的权值共享. 池化层的输入一般是前一层的卷积层的输出, 主要对特征进行模糊, 从而获得平移, 尺度的不变性. 在卷积神经网络中^[12], 有四种基本的运算. 依次定义为: 内卷积, 外卷积, 下采样, 上采样.

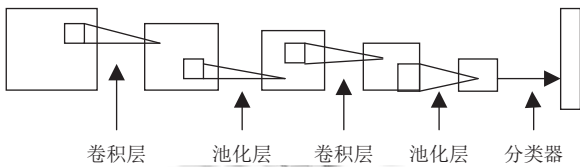


图 1 CNN 网络实现结构

假设 A, B 是矩阵, 大小分别为 $M \times N, m \times n$, 且 $M \geq m, N \geq n$, 则它们的内卷积 $C = A \otimes B$ 的所有元素定义为:

$$c_{ij} = \sum_{s=1}^m \sum_{t=1}^n a_{i+m-s, j+n-t} \cdot b_{st} \quad (1)$$

$$1 \leq i \leq M - N + 1, 1 \leq j \leq N - n + 1$$

它们的外积定义如下:

$$A \oplus B = \bar{A} \otimes B \quad (2)$$

式中 $\bar{A} = (\bar{a}_{ij})$ 是一个通过用 0 对 A 进行扩充得到的矩阵, 大小为 $(M + 2m - 2) \times (N + 2n - 2)$, 且矩阵元素

$$\bar{a}_{ij} = a_{i-m+1, j-n+1} \quad (3)$$

其中, $m \leq i \leq M + m - 1, n \leq j \leq N + n - 1$

如果对矩阵 A 进行不重叠分块, 设每块大小为 $k \times L$, 第 ij 块的矩阵形式 $G_{k,L}^A(i, j)$, 则其构造如下:

$$G_{k,L}^A(i, j) = (a_{st})_{k \times L} \quad (4)$$

$$(i-1) \times k + 1 \leq s \leq i \times k, (j-1) \times L + 1 \leq t \leq j \times L$$

$G_{k,L}^A(i, j)$ 下采样的定义如下:

$$\text{down}(G_{k,L}^A(i, j)) = \frac{1}{k \times L} \sum_{s=(i-1) \times k + 1}^{i \times k} \sum_{t=(j-1) \times L + 1}^{j \times L} a_{st} \quad (5)$$

用大小为 $k \times L$ 不重叠块对矩阵 A 的下采样定义为:

$$\text{down}_{k \times L}(A) = \text{down}(G_{k,L}^A(i, j)) \quad (6)$$

对矩阵 A 进行倍数为 $k \times L$ 的不重叠上采样定义为

$$\text{up}_{k \times L}(A) = A \odot 1_{k \times L} \quad (7)$$

其中, $1_{k \times L}$ 表示为元素全 1 的矩阵, \odot 表示克罗内克积.

1.2 树状卷积神经网络

传统的卷积神经网络^[7-9]在图像分类上取得了很好的效果, 但影响卷积神经网络正确识别率和鲁棒性的关键参数是卷积核大小和下采样层的采样间距. 文献^[10]证明了卷积核的大小对识别率有很大的影响, 文献^[11]中证明了采样间距较小时, 网络输出的激活值仍能重构出与原始输入相似的图像. 而下采样间隔过小会导致不变性丧失, 过大丢失大量有用信息. 由于传统的网络结构是在单通路内完成的, 只能单一的设置卷积核大小和池化层的采样间距, 这样就限制了参数设置的灵活性, 不能有效的提取局部和全局的特征. 并且传统的卷积网络只是使用顶层提取的特征进行训练分类器, 并没有充分的利用底层提取的特征. 为了提高识别率和鲁棒性, 提出了 T-CNN 网络结构, 用于建立车标分类模型.

T-CNN 网络采用多通路树状结构, 即在传统卷积网络单一卷积核的卷积层上, 使用多种类型的卷积核进行卷积操作, 形成树状网络结构.

网络结构如图 2 所示. 该模型包含 1 个输入层 x , 3 个卷积层 (h_1, h_3, h_5) 和 (h_2, h_4, h_6) 3 个池化层, 1 个全连接层 (h_7) 和 1 个输出层 o . 特别, 该网络首先输入一副图像, 经过第一个包含不同大小卷积核的卷积层 (h_1) 时, 分为了两个通道. 第二个卷积层 (h_3), 分别对经过池化层 (h_2) 后的特征图用不同大小卷积核进行卷积操作, 得到四个不同的通道. 第三个卷积层 (h_5), 分别对经过池化层 (h_4) 的特征图, 对每个特征图使用两种不同大小的卷积核进行卷积操作, 得到八个不同的通道. 最后把池化层 (h_6) 提取的特征直接作为全连接层的输入进行分类. 输出层有 20 个节点. 代表输入图像的分类. 表 4 给出了该模型的具体描述.

1.2.1 树状卷积网络的学习过程

对于第 a 个样本 x^a , 树状卷积网络的前向传播过程

$$h_{1,j}^a = \sigma(u_{1,j}^a) = (x^a \otimes w^{1,j} + b^{1,j}), 1 \leq j \leq 12$$

$$h_{2,j}^a = g(u_{2,j}^a) = g(\text{down}(h_{1,j}^a)), 1 \leq j \leq 12$$

$$h_{3,j}^a = \sigma(u_{3,j}^a) = \sigma(\sum_{i=1}^{12} h_{2,i}^a \otimes w^{3,j} + b^{3,j}), 1 \leq j \leq 48$$

$$h_{4,j}^a = g(u_{4,j}^a) = g(\text{down}(h_{3,j}^a)), 1 \leq j \leq 48$$

$$h_{5,j}^a = \sigma(u_{5,j}^a) = \sigma(\sum_{i=1}^{24} h_{4,i}^a \otimes w^{5,j} + b^{5,j}), 1 \leq j \leq 192$$

$$h_{6,j}^a = g(u_{6,j}^a) = g(\text{down}(h_{5,j}^a)), 1 \leq j \leq 192$$

$$H_7^a = (h_{6,1}^a, h_{6,2}^a, \dots, h_{6,48}^a)$$

$$o_8^a = \sigma(u_8^a) = \sigma(w^8 H_7^a + b^8) \quad (8)$$

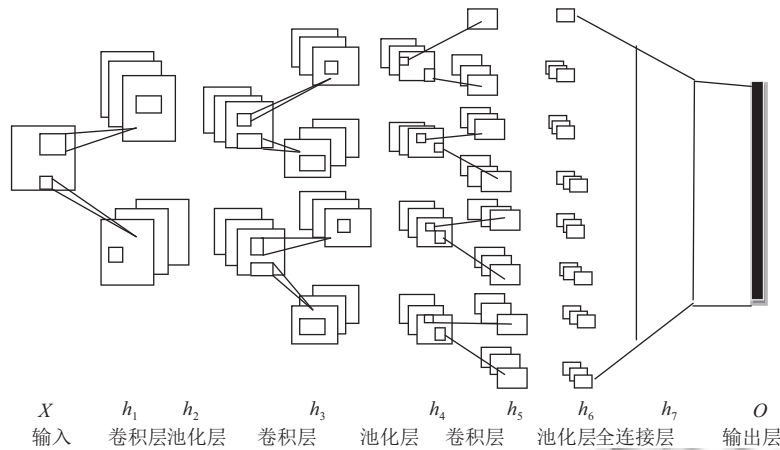


图2 T-CNN网络结构

函数 $\sigma(x) = 1/(1 + e^{-x})$, $g(x) = x$. $w^{i,j}$ ($i = 1, 3, 5$)表示第 i 个隐层的第 j 个卷积核, $b^{k,j}$ ($k = 1, 3, 5$)表示第 k 个隐层的第 j 个卷积核的偏置. $h_{r,j}^a$ ($r = 1, 3, 5$)标识第 r 个隐层的第 j 个特征图. $h_{t,j}^a$ ($t = 2, 4, 6$)表示第 t 个隐层的第 j 个池化面. H_7^a 表示第七个隐层的输出值. o_8^a 表示输出层的输出值.

假设有 N 个训练样本 (x^a, y^a) , $1 \leq a \leq N$, 则输入为 $x^a = (x_1^a, x_2^a, \dots, x_m^a)$, 预期输出 $y^a = (y_1^a, y_2^a, \dots, y_c^a)$, 实际输出, 均方误差为 $o^a = (o_1^a, o_2^a, \dots, o_c^a)$, $L_N = \sum_{a=1}^N \sum_{n=1}^c (o_n^a - y_n^a)^2$.

反向传播过程, 首先计算网络中的每一层反馈传递误差:

$$\begin{aligned} \delta_k & (1 \leq k \leq 8) \\ \delta_8^a & = (o^a - y^a) \circ \sigma'(u_8^a) \\ \delta_7^a & = [(w^8)^T \delta_8^a] \circ H_7^a \circ (1 - H_7^a) \\ \delta_7^a & = (\delta_{6,1}^a, \delta_{6,2}^a, \dots, \delta_{6,192}^a) \\ \delta_{5,j}^a & = \frac{1}{k \times L} (\sigma'(u_{5,j}^a) \circ up_{k \times L}(\delta_{6,j}^a)) \\ \delta_{4,j}^a & = [\delta_{5,j}^a \oplus w^{5,j}] \circ g'(u_{4,j}^a), 1 \leq j \leq 48 \\ \delta_{3,j}^a & = \frac{1}{k \times L} (\sigma'(u_{3,j}^a) \circ up_{k \times L}(\delta_{4,j}^a)) 1 \leq j \leq 48 \\ \delta_{2,j}^a & = [\delta_{3,j}^a \oplus w^{3,j}] \circ g'(u_{2,j}^a), 1 \leq j \leq 12 \\ \delta_{1,j}^a & = \frac{1}{k \times L} (\sigma'(u_{1,j}^a) \circ up_{k \times L}(\delta_{2,j}^a)) 1 \leq j \leq 12 \end{aligned} \quad (9)$$

式中, \circ 表示矩阵 (或者向量) 对应元素相乘, δ_8^a 和 δ_7^a 分别表示输出层和全连接层的反馈传递误差, $\delta_{i,j}^a$ ($i = 6, 4, 2$) 表示第 i 个隐层的第 j 个池化面的反馈传递误差. $\delta_{k,j}^a$ ($k = 5, 3, 1$)表示第 k 个隐层的第 j 个特征图的反馈传递误差, 在得到每层的反馈传递误差后, 关于权值和

偏置的偏导数:

$$\begin{aligned} \frac{\partial L_N}{\partial w^8} & = \sum_{n=1}^N \delta_8^n (H_7^n)^T, \frac{\partial L_N}{\partial b^8} = \sum_{n=1}^N \delta_8^n \\ \frac{\partial L_N}{\partial w^{5,j}} & = \sum_{n=1}^N h_{4,j}^n \otimes \delta_{5,j}^n, \frac{\partial L_N}{\partial b^{5,j}} = \sum_{n=1}^N \delta_{5,j}^n, 1 \leq j \leq 192 \\ \frac{\partial L_N}{\partial w^{3,j}} & = \sum_{n=1}^N h_{2,j}^n \otimes \delta_{3,j}^n, \frac{\partial L_N}{\partial b^{3,j}} = \sum_{n=1}^N \delta_{3,j}^n, 1 \leq j \leq 48 \\ \frac{\partial L_N}{\partial w^{1,j}} & = \sum_{n=1}^N x^n \otimes \delta_{1,j}^n, \frac{\partial L_N}{\partial b^{1,j}} = \sum_{n=1}^N \delta_{1,j}^n, 1 \leq j \leq 12 \end{aligned} \quad (10)$$

式中, $\frac{\partial L_N}{\partial w^i}$ ($i = 8, 5, 3, 1$), $\frac{\partial L_N}{\partial b^j}$ ($j = 8, 5, 3, 1$), 表示损失函数对权值 w^i 和偏置 b^j 的偏导数.

训练步骤:

输入: 训练集 $S = \{(x^a, y^a), 1 \leq a \leq N\}$, 网络结构, 最大迭代次数 10000, 学习率 0.01

输出: 网络的权值和偏置

- Step 1. 随机初始化所有权重和偏置;
- Step 2. 前向传播, 计算实际输出, 公式 (8);
- Step 3. 反向传播, 计算每层的反馈传递误差, 公式 (9).
- Step 4. 计算网络的权值和偏置的偏导数, 公式 (10);
- Step 5. 梯度下降算法更新权值和偏导数.

2 实验结果与分析

2.1 数据

本文使用的数据集是从 LPR 数据库和网络上收集得到的图片. 对数据进行增强, 如图像的平移, 旋转等图像变换, 最终获得了 20 类共 40000 张图片. 对于这 40000 张图片, 30000 张用于训练, 10000 张用于测

试, 车标图像示例如图3所示. 所有图片都被归一化到64×64大小. 分别为阿尔法罗密欧, 奥迪, 宝马, 奔驰, 比亚迪, 东风标致, 别克, 金杯, 大众, 菲亚特, 丰田, 福特, 现代, 雷诺, 铃木, 马自达, 尼桑, 欧宝, 西亚特, 雪铁龙.

2.2 实验结果

2.2.1 车标测试集测试

测试集中有20类车标类型, 每类有500张图片, 表1给出了T-CNN网络模型对测试集测试的结果.

同时, 在传统的卷积神经网络进行测试, 在其深度上与T-CNN保持一致. 3个卷积层, 卷积核大小分别为12×5×5, 48×3×3, 192×3×3, Stride为1. 3个下采样层, 采样间距2×2, stride为2. 测试结果如表2所示.



图3 20类车标图片实例

2.2.2 与其他方法的比较

跟其它车标识别方法的比较中, 采样相同的训练集和测试集. 20类车标30000张图片进行训练, 10000张测试集上进行测试. 与文献[2]所提到的HOG+SVM方法, 文献[4]中使用传统的CNN+SVM方法, 及文献[5]提出的联合卷积网络Multi-CNN特征的识别方法进行比较. 比较的结果如表3所示.

2.2.3 实验结果分析

从表1和表2的结果分析得到, T-CNN能够在自然场景下进行图像的识别, 在光照不均, 图像在一定程

度的形变, 扭曲和缺失的情况下能够提取较好的特征进行分类任务. 比较传统的卷积网络, 分类的性能有一定的提升. 从表3分析可知, 传统的特征提取受到外界因素的影响较大, 不利于分类特征的提取. 而Multi-CNN利用多层采样层的特征进行分类, 但单通路下不能提取多种尺度范围的特征.

表1 T-CNN车标测试识别率

车型	正确数	识别率(%)	车型	正确数	识别率(%)
阿尔法	489	97.80	丰田	493	98.60
奥迪	490	98.00	福特	487	97.40
宝马	492	98.40	吉利	491	98.20
奔驰	489	97.80	雷诺	496	99.20
比亚迪	498	99.60	铃木	492	98.40
标致	493	98.60	马自达	492	98.40
别克	490	98.00	尼桑	498	99.60
宾利	490	98.00	欧宝	490	98.00
大众	496	99.20	西亚	485	97.00
菲亚特	500	100.80	雪铁	492	98.40
总计	10000	正确数	9843	识别率	98.43%

表2 CNN车标测试识别率

车型	正确数	识别率(%)	车型	正确数	识别率(%)
阿尔法	486	97.20	丰田	490	98.00
奥迪	484	96.80	福特	484	96.80
宝马	487	97.40	吉利	483	96.60
奔驰	480	96.00	雷诺	495	99.00
比亚迪	491	98.20	铃木	492	98.40
标致	487	97.40	马自	490	98.00
别克	490	98.00	尼桑	486	97.20
宾利	480	96.00	欧宝	488	97.60
大众	489	97.80	西亚	473	94.60
菲亚特	494	98.80	雪铁	483	96.60
总计	10000	正确数	9732	识别率	97.32%

表3 不同方法车标测试结果

方法	HOG+SVM	CNN+SVM	Multi-CNN	T-CNN
识别率(%)	95.80	97.56	98.35	98.43

2.2.4 T-CNN性能分析

T-CNN网络结构在测试集上的性能分析, 分类准确率和损失函数与迭代次数之间的关系, 如图4.

3 结语

本文提出了一种树状结构的卷积神经网络T-CNN

用于车标识别,与传统的卷积神经网络相比较,能够提取有效的识别特征,能够自主的进行特征的提取,避免了人工提取特征的繁琐.但在实验中使用的车标类型较少,对于工程应用有一定的局限性.而且与一般的卷

积神经网络相比,树状结构卷积网络层数过多会引起参数过大,不利于计算.接下来的工作中,扩大样本数量和类型的数量.扩大网络的深度,减少各个通路的卷积核的数量进行 T-CNN 的验证.

表4 T-CNN 具体描述

Layer	type	Patchsize/stride/padding	Output size	Layer	type	Patchsize/stride/padding	Outputsize
h_1	C	3×3/1/1	64×64×6	h_2	mean pool	2×2/2/0	32×32×6
		5×5/1/0	60×60×6			2×2/2/0	30×30×6
h_3	C	5×5/1/0	28×28×12	h_4	mean pool	2×2/2/0	14×14×12
		3×3/1/1	32×32×12			2×2/2/0	16×16×12
		3×3/1/0	28×28×12			2×2/2/0	14×14×12
		1×1/1/1	32×32×12			2×2/2/0	16×16×12
h_5	C	1×1/1/0	14×14×24	h_6	mean pool	2×2/2/1	8×8×24
		3×3/1/1	14×14×24			2×2/2/1	8×8×24
		1×1/1/0	16×16×24			2×2/2/0	8×8×24
		3×3/1/1	16×16×24			2×2/2/0	8×8×24
		1×1/1/0	14×14×24			2×2/2/1	8×8×24
		3×3/1/1	14×14×24			2×2/2/1	8×8×24
h_7	FC	1×1/1/0	16×16×24	o	O	2×2/2/0	8×8×24
		3×3/1/1	16×16×24			2×2/2/0	8×8×24
		3×3/1/1	16×16×24			2×2/2/0	8×8×24
			12288				20

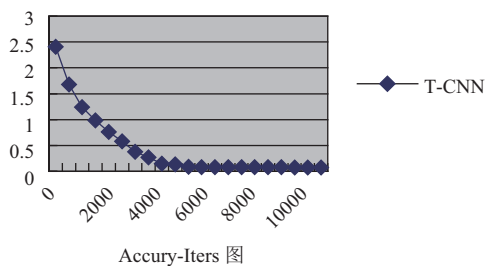
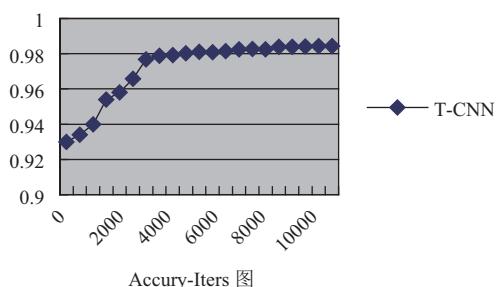


图4 T-CNN 性能图

参考文献

1 王枚, 王国宏, 房培玉, 等. 基于 PCA 与不变矩的车标定位与识别. 武汉大学学报·信息科学版, 2008, 33(1): 36-40.
 2 Dai SJ, Huang H, Gao ZY, et al. Vehicle-logo recognition method based on Tchebichef moment invariants and SVM. Proc. of the 2009 WRI World Congress on Software Engine-

ring. Xiamen, China. 2009. 18-21.
 3 Llorca DF, Arroyo R, Sotelo MA. Vehicle logo recognition in traffic images using HOG features and SVM. Proc. of the 16th International IEEE Conference on Intelligent Transportation System. The Hague, Netherlands. 2013. 2229-2234.
 4 彭博, 藏笛. 基于深度学习的车标识别方法研究. 计算机科学, 2015, 42(4): 268-273. [doi: 10.11896/j.issn.1002-137X.2015.04.055]
 5 Zhang L, Zhang DM, Zheng H. Vehicle logo recognition using convolutional neural network combined with multiple layer feature. Journal of Computer Application, 2016, 36(2): 444-448.
 6 Sun Q, Lu XB, Chen L, et al. An improved vehicle logo recognition method for road surveillance images. Proc. of the 7th International Symposium on Computational Intelligence and Design (ISCID). Hangzhou, China. 2014. 373-376.
 7 LeCun Y, Bottou L, Bengio Y. LeNet-5, Convolutional neural networks. IEEE Communication, 1989: 41-46.
 8 Szegedy C, Liu W, Jia YQ, et al. Going deeper with convolutions. Proc. of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Boston, MA, USA. 2015. 1-9.
 9 Krizhevsky A, Sutskever I, Hinton GE. ImageNet classifi-

- cation with deep convolutional neural networks. Proc. of the 25th International Conference on Neural Information Processing Systems. Lake Tahoe, Nevada, USA. 2012. 1097–1105.
- 10 Coates A, Lee H, Ng AY. An analysis of single-layer networks in Unsupervised feature learning. Proc. of the 14th International Conference on Artificial Intelligence and Statistics. Ft. Lauderdale, FL, USA. 2011. 215–223.
- 11 Zeiler MD, Fergus R. Visualizing and understanding convolutional networks. Proc. of the 13th European Conference on Computer Vision. Zurich, Switzerland. 2014. 818–833.
- 12 张婷, 李玉鑑, 胡海鹤, 等. 基于跨连卷积神经网络的性别分类模型. 自动化学报, 2016, 42(6): 858–865. [doi: [10.16383/j.aas.2016.c150658](https://doi.org/10.16383/j.aas.2016.c150658)]

www.c-s-a.org.cn

www.c-s-a.org.cn