

DOI: 10.13382/j.jemi.B2003005

# 基于卷积神经网络的 LED 灯类字体数字识别\*

王立刚<sup>1</sup> 张志佳<sup>1</sup> 李 晋<sup>2</sup> 范莹莹<sup>1</sup> 刘立强<sup>2</sup>

(1. 沈阳工业大学 信息科学与工程学院 沈阳 110870; 2. 辽宁航天凌河汽车有限公司 朝阳 122599)

**摘要:**针对自然场景下由 LED 灯组合形成的数字具有易受光照、背景和成像扭曲等因素影响识别困难的特点,提出了一种 LED-LeNet 卷积网络识别算法。对自采集 LED 灯类字体数据集按数字进行分类,将图像 ROI 操作、分辨率调整至 32×32 和数据增强等预处理后,在 LeNet-5 网络架构上通过卷积核重构、使用 Swish 激活函数并引入 Dropout 正则化等方法改进网络。采用自然场景下采集的交通信号灯倒计时数字图像数据库 TST 对算法进行了验证,算法识别正确率可达 99.52%,识别速度为 1 ms。实验结果表明在调整网络结构与卷积核参数并通过改变训练策略后算法识别 LED 灯类字体具有明显优势。

**关键词:** LED-LeNet; 自然场景; 卷积神经网络; 数字识别

**中图分类号:** TP391.41 **文献标识码:** A **国家标准学科分类代码:** 510.40

## Digital recognition of LED lights based on convolutional neural networks

Wang Ligang<sup>1</sup> Zhang Zhijia<sup>1</sup> Li Jin<sup>2</sup> Fan Yingying<sup>1</sup> Liu Liqiang<sup>2</sup>

(1. School of Information Science and Engineering, Shenyang University of Technology, Shenyang 110870, China;

2. Liaoning Aerospace Linghe Automobile Co., Ltd., Chaoyang 122599, China)

**Abstract:** In order to solve the LED recognition problem that the number formed by the factors such as illumination, background, and image distortion in natural scene, a recognition algorithm of LED-LeNet convolutional network is proposed. Firstly, the self collected LED light font data set was classified according to the number. Image data preprocessing includes image ROI operation, resolution adjustment to 32 × 32 and data enhancement. The network was reconstructed by convolution kernel, swish activation function and dropout regularization which referred to LeNet-5 network. The algorithm was verified by TST digital image database of traffic signal countdown collected in natural scene. The recognition accuracy of the algorithm can reach 99.52%, and the recognition speed was 1 ms. The experimental results show that the algorithm has obvious advantages in recognizing LED light fonts after adjusting the network structure and convolution kernel parameters and changing the training strategy.

**Keywords:** LED-LeNet; natural scene; convolutional neural network; digital recognition

## 0 引言

自然场景下文本识别已在计算机视觉领域引起越来越多的关注<sup>[1]</sup>。由于 LED 灯珠排列组合成数字形状,易受光照不均、背景掺杂、细节模糊、成像扭曲等因素影响,这使得 LED 灯组成的数字识别并不像室内理想环境下相关仪表字符识别,因此在自然场景下 LED 灯类字体类

数字识别极具挑战。

LED 灯类字体数字与印刷体数字相似。关于清晰字体识别,例如室内环境下仪表数显数字、LED 数码管字体识别已取得较好的识别结果;关于模糊、带有复杂背景的数字识别,如街景门牌号(street view house numbers, SVHN)、自然场景下的数字识别等仍存在问题。

王晓东等<sup>[2]</sup>提出一种基于支持向量机的仪表数字识别方法。对原始图像预处理后使用支持向量机对单个

LED 数码管数字进行识别。该方法对变电站数显仪表数字识别准确率在 90% 以上。崔家瑞等<sup>[3]</sup>提出一种识别数字式动作次数数字字符的识别方法。由于字符边界无粘连且单个字符清晰,所以该方法在图像增强后使用垂直投影法将数字区域分割成单个字符,使用 BP 神经网络识别单个字符。该方法对避雷针动作次数数字字符进行识别,结构特征明显的数字识别准确率接近 100%。对于结构特征相似的数字如 0、6、9 或字符粘连数据识别效果不佳。刘进博等<sup>[4]</sup>提出一种使用 ROI 将车牌单个字符分割出,使用 BP 神经网络识别的方法。通过对自然场景下 203 帧图像中的 236 个车牌进行识别,准确率高达 95.3%。白睿等<sup>[5]</sup>提出一种基于卷积神经网络(convolutional neural network, CNN)的道路场景字符识别方法。该方法使用最新改进版 LeNet-5 神经网络识别道路场景字符,此算法相较于 K 近邻法(k-nearest neighbors, KNN)、方向梯度直方图+支持向量机(HOG+SVM)准确率有所提高。Sermanet 等<sup>[6]</sup>提出一种使用二阶卷积神经网络在自然场景下识别门牌号的方法,通过构建二阶卷积神经网络以及使用 LP 池化使算法在 SVHN 数据集上准确率可达 94.85%。这些方法利用特征提取等方法将字符串分割成单个字符,然后利用 SVM、BP 神经网络等对单个字符进行识别。在实际应用中,数据种类繁多且数据背景复杂,这些方法受光线等干扰因素影响在进行字符分割与识别方便有一定的局限性。

在室内等外界干扰因素小的环境下成像清晰的印刷体数字可采用传统方法识别,例如穿线法、投影法和模板匹配等方法;成像环境复杂成像不清晰的印刷体数字可以使用 SVM<sup>[7]</sup>、BP 神经网络和卷积神经网络为代表的人工神经网络<sup>[8]</sup>,这几种方法相对于传统算法在识别不清晰印刷体字符有着较高的准确率。近年来,随着计算机算力的提升,卷积神经网络得到飞速发展。卷积神经网络在识别自然场景下的印刷体数字准确率相较于支持向量机与 BP 神经网络的传统机器学习算法有着较大提升。因此,本文针对 LED 灯类字体数字特点设计一种卷积神经网络识别方法。该方法通过对原始图片进行 ROI 操作提取数字区域,调整分辨率后直接使用卷积神经网络对数字区域进行识别。试验对比证明,该网络模型在兼顾识别速度快的同时,具有较好的泛化能力和较高的识别率。

## 1 预处理

图像预处理由对原始图像数据进行 ROI 操作提取数字区域、调整分辨率、数据增强去均值等 4 部分组成,如图 1 所示。

本文对 LED 灯类字体数字进行识别。使用 ROI 操

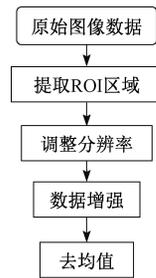


图 1 预处理流程

Fig. 1 Preprocessing flowchart

作对数字区域进行提取,可以简化识别网络复杂度和提高识别准确度。

神经网络训练与识别时,为提取更多有效价值特征信息通常使用尺寸相同的数据样本。ROI 操作提取的数字区域图片大小尺寸不固定,因此需要对数字区域进行调整分辨率操作。为保持数字原有形状信息,减少畸变对识别准确率的影响。本文对 ROI 区域短边区域进行填充白色信息,调整分辨率后的尺寸为 32×32,该尺寸根据经验值得出。

训练卷积神经网络时,由于样本数量较少,为提高训练后网络模型的泛化能力与鲁棒性,一般对数据集进行数据增强操作。通过旋转、缩放、任意裁切、平移和添加噪声等操作对数据集进行样本增加。

数据各维度减去各维度均值使得各维度数据都中心化为 0,这有助于缓解过拟合的发生。

## 2 LeNet-5 网络

LeNet 作为识别手写体数字最经典的神经网络,是 Yann LeCun 在 1998 年设计并提出应用于手写体识别,在 MNIST 数据集上的识别率可达 99.1%,该网络在被提出后被广泛用于银行及邮政的手写体数字识别系统。LeNet 卷积神经网络可以看成是一个多隐层的人工神经网络,其基本结构主要包括输入层、卷积层、池化层、全连接层和输出层,其中卷积层和池化层交替出现,构成了隐层<sup>[9]</sup>。网络模型结构如图 2 所示。

输入层将手写体数字图像以二维向量形式传递到卷积层。卷积层将输入图像数据经过  $n$  个卷积模板得到  $n$  个特征图谱。通过这些图谱可获得图像的边缘及轮廓信息<sup>[10]</sup>。池化层通常被设置在卷积层之后,通过对特征矩阵的局部区域进行池化操作,降低卷积层输出特征矩阵的维度<sup>[11]</sup>。利用池化对卷积后的图像进行下采样,池化可以减少网络参数从而减少计算量,并且池化能在一定程度上可控制过拟合。全连接层通过将不同的特征信息压缩成  $N \times 1$  向量形式利用权值分类输出。

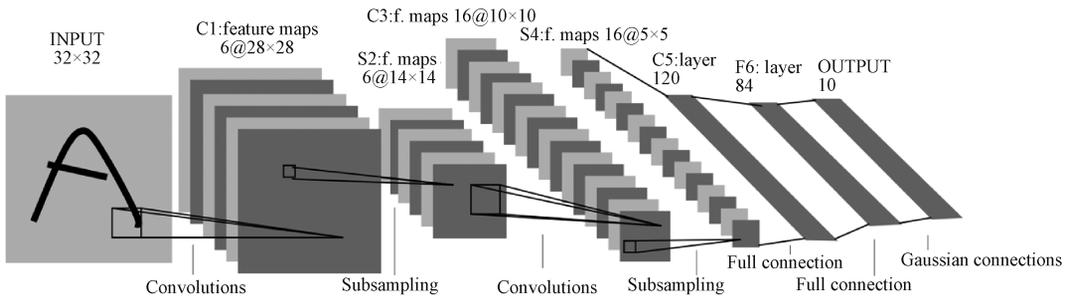


图 2 LeNet-5 网络模型

Fig. 2 LeNet-5 network model

近年来研究人员对 LeNet-5 的改进主要有两个方面。首先是在每层卷积层后增加 ReLU 激活函数,减少计算量使网络收敛更快<sup>[10]</sup>。其次 LeNet-5 输出层由原径向基函数输出改由 softmax 输出<sup>[12]</sup>。经改进后的网络在 MNIST 数据集上的识别率可达 99.6%。

### 3 LED-LeNet 设计

相比于 LeNet-5 识别的二值图像样本,本文识别的自然场景下 LED 灯类字体数字具有图像亮度不均、笔画不连续和字体模糊等特点,因此可增加网络深度提取更多特征信息。网络深度增加会一定程度导致网络对数据样本的过拟合,使用正则化减缓过拟合的发生。

#### 3.1 卷积核重构

通过卷积核对输入对象进行特征提取,是 CNN 的核心<sup>[13]</sup>。卷积核在浅层次网络提取纹理、细节等特征。网络层次越深,卷积层输出特征图越抽象,提取特征维度越高。LeNet-5 网络模型中使用 5×5 大小的卷积核。大尺寸的卷积核能够提取更多信息,但参数计算量也随之增大。Karen Simonyan 等在 VGG 网络中使用 2 个 3×3 卷积核代替 1 个 5×5 卷积核<sup>[14]</sup>。假设特征图大小为  $C$ ,由式(1)可知 2 个 3×3 卷积层的参数为  $18C^2$ 。由式(2)可知 1 个 5×5 的卷积层的参数为  $25C^2$ 。

$$2 \times (3 \times 3 \times C \times C) = 18C^2 \quad (1)$$

$$1 \times (5 \times 5 \times C \times C) = 25C^2 \quad (2)$$

通过计算可知使用 2 个 3×3 卷积核进行卷积操作比使用单个 5×5 卷积核拥有更少的参数。此外,多个小卷积核组合可以拥有更多非线性,使得网络的鲁棒性更好。5×5 的有效区域使用 5×5 的卷积核输出大小为 1×1 的图像;相同大小的有效区域使用 2 个 3×3 的卷积核输出图像大小为 1×1,如图 3 所示。因此,使用 2 个 3×3 卷积核与 1 个 5×5 的卷积核视野大小相同。

针对部分 LED 灯类字体数字轮廓非连续为提取完整轮廓信息第 1 层卷积层使用与 LeNet-5 相同的 5×5 卷积核,第 2 层卷积层更改为两层卷积核为 3×3 的卷积层,

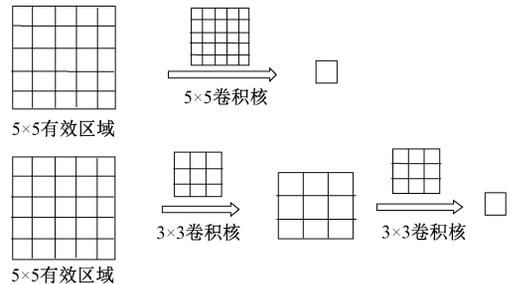


图 3 卷积核感受野

Fig. 3 Convolution nuclear receptive field

增加网络深度提取更多的图像信息的同时减少网络参数且保持相同的感受野。样本特征与手写体数字相比存在数字轮廓模糊、粘连现象,因此采用更少的卷积核可在充分学习样本特征的同时减少网络的复杂度。

#### 3.2 激活函数

使用非线性映射的激活函数可提升整个神经网络的表达能力。常用激活函数有 Sigmoid、Relu 以及 Softmax 等。

Swish 激活函数是由谷歌研究人员在 2017 年提出<sup>[15]</sup>,该激活函数会在损失一定训练时间的同时增加网络模型准确率。激活函数对比如图 4 所示。

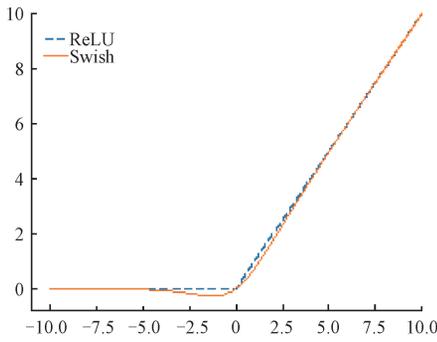
#### 3.3 正则化

使用数据样本较少和样本特征相似时会导致网络过拟合,而正则化有利于缓解过拟合问题,提高泛化能力。正则化包括数据增强、L1 正则化、L2 正则化、Dropout、DropConnect 和早停(early stopping)等方法。

Dropout 是由 Hinton 等<sup>[16]</sup>在 2012 年提出,当一个复杂的前馈神经网络训练在小的数据集时易造成过拟合。Dropout 通过关闭一部分神经元减少对部分不太依赖的局部特征使用,模型泛化能力增强。本文网络结构引入 Dropout 正则化,在防止过拟合的同时增加网络的拟合能力<sup>[17]</sup>。

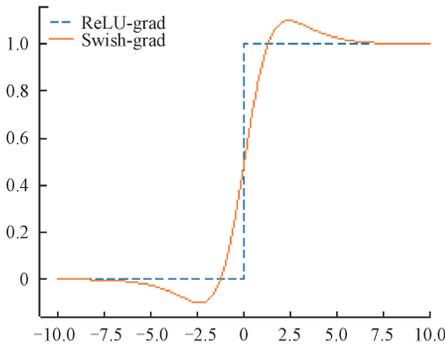
#### 3.4 LED-LeNet 网络模型

重构卷积核在减少了网络的参数量的同时增加网络



(a) 激活函数

(a) Activation function



(b) 激活函数梯度

(b) Activation function gradient

图 4 激活函数对比示意图

Fig. 4 Comparison of activation functions

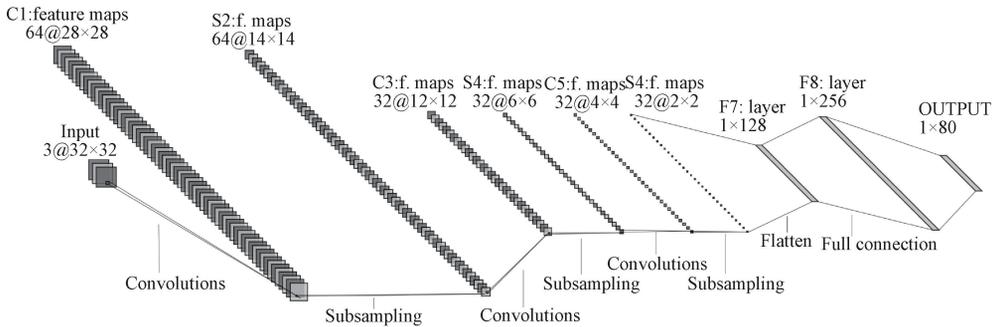


图 5 LED-LeNet 网络模型

Fig. 5 LED-LeNet network model diagram



图 6 LED 灯类字体数字样本

Fig. 6 LED light font digital sample

深度。为解决增加网络深度导致的过拟合问题,在全连接层之间使用 Dropout 正则化,按照固定概率关闭神经元从而减少保存无关信息的神经元的使用,减缓过拟合。基于以上改进,本文设计 LED- LeNet 网络模型如图 5 所示,其中每个卷积层后对应一个激活层,激活层不改变图像关键特征信息,原始图像经过激活层后非线性与模型泛化能力会得到增强<sup>[18]</sup>。

### 4 实验

#### 4.1 数据集

训练样本的大小将直接影响卷积神经网络的训练结果的好坏<sup>[19-20]</sup>,为验证本文设计的 LED-LeNet 网络模型的有效性,使用自然场景下同为 LED 灯类字体的交通灯倒计时数字进行验证,如图 6 所示。

本文采用道路测试采集的交通信号灯倒计时数字建立的数据集(traffic sign timing, TST)进行训练与测试。数据集包含数字 01~09、数字 1~69 共 80 类数字。数据集包括训练集 9 843 张,测试集 2 516 张,图片大小为 32×32。数据集样张如图 7 所示。

#### 4.2 实验环境

本文实验硬件平台为 Intel® Core™ i5-8400 CPU, 2.80 GHz,6 核处理器,NVIDIA GeForce GT 710。软件配

框架为 Keras,其后端为 Tensorflow 深度学习架构。Python 计算环境发行版及 Opencv3.4 计算视觉库,集成开发环境为 Jupyter notebook。

#### 4.3 实验过程

训练时,采用由 Geoff Hintont 提出的 RNSpr-op 自适应学习方法。该方法相比 Adagrad 会累加之前多有的平方,而 RMSpropj 仅计算对应的平均值,因此可缓解 Adagrad 算法学习率下降较快的问题。训练集 bath-size 设置为 64,验证集 Bathsize 设置为 32,数据集迭代 50 次可学习所有图像的特征信息。



图 7 数据集部分样本

Fig. 7 Dataset partial sample

通过对比试验,如表 1 所示,在第 1 层卷积层卷积核数量设置为 64,第 2 层与第 3 层卷积层卷积核数量各设置 32 个时,模型识别准确率最高。Dro-pout 关闭率经试验对比,在 0.5 时网络识别率最高,因此设置为 0.5。网络最后使用 Softmax 分类输出识别结果。

表 1 调节参数准确率

Table 1 Adjust parameter accuracy

C1	C3	Dropout	正确率/%
5×5×64	3×3×32	0.5	99.523
5×5×64	3×3×32	1.0	98.524
5×5×64	3×3×32	0.7	99.244
5×5×64	3×3×32	0.3	98.205
5×5×32	3×3×32	0.5	98.927
3×3×64	3×3×32	0.5	99.126
3×3×32	3×3×32	0.5	99.244
5×5×64	-	0.5	99.403

网络训练时,卷积核在卷积层提取样本的特征权重信息,通过卷积层输出融合特征图。特征图不仅可以观察每层卷积核提取的特征信息,还可判断网络结构设计是否合理。数据样本在卷积网络各卷积层输出特征图如图 8 所示。在训练过程中以某一概率随机失活一些神经元,可以减小 CNN 发生过拟合可能,并提高神经网络分类的准确率<sup>[19]</sup>。

网络模型在前两层卷积网络提取数据样本边缘、形状等信息,卷积核权重特征图与卷积层融合图平滑。最后一层网络卷积核提取更高维度信息表现为色块按序排列色块,以上信息表明网络参数设置合理,网络拟合较好。同时第 3 层大约 50%卷积核权重特征图没有权重信息或权重信息较少,进一步印证 Dropout 关闭率设置为 0.5 合理。

综上,本文网络模型参数如表 2 所示。

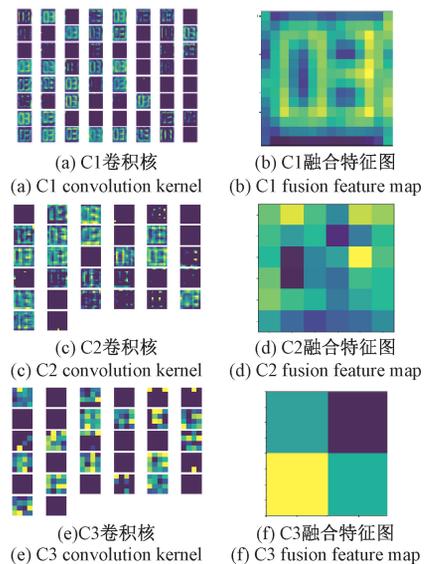


图 8 卷积特征图

Fig. 8 Convolution feature map

表 2 网络模型各层参数

Table 2 Network model layer parameters

层结构	输入数据	输出数据
卷积层 1	32×32×3	28×28×64
最大池化层	28×28×64	14×14×64
卷积层 2	14×14×32	12×12×32
最大池化层	12×12×32	6×6×32
卷积层 3	6×6×32	4×4×32
最大池化层	4×4×32	2×2×32
扁平层	2×2×32	128
全连接层	128	256

#### 4.4 实验结果与分析

将本文提出的 LED-LeNet 网络结构与在 MNIST 上测试的手写体数字的 LeNet-5 识别算法和在 SVHN 数据库上测试的二阶 CNN 卷积神经网络算法对比,各算法在对应数据库上的准确率如表 3 所示(以上算法均在相同环境测得)。

表 3 数据库算法对比

Table 3 Database algorithm comparison

数据集	数据库	准确率/%
LeNet-5	MNIST	98.42
LED-LeNet	MNIST	97.16
二阶 CNN	SVHN	94.85
LED-LeNet	SVHN	96.10

在 MNIST 数据库上,本文算法相较于 LeNet-5 识别准确率较低,由于本文算法针对于背景复杂,干扰因素多的 LED 灯类字体设计,因此神经元及参数量多于 LeNet-

5. 在图像信息较少情况下,过多的神经元数量会导致网络对过度拟合,因此新输入图像识别准确率降低。而在 SVHN 数据库上,LED-LeNet 针对背景复杂、干扰因素多的 LED 灯类字体设计结构发挥作用,识别率高于二阶卷积网络。

在本文建立的 TST 数据库上,对文献[4-6](文献[4,6]未提供代码带来可重复问题,本文根据其算法模型进行自编码复现)和 LED-LeNet 进行了识别实验,结果如表 4 所示。

表 4 不同算法测试集结果

Table 4 Different algorithm test set results

算法	损失率	准确率/%	平均用时/ms
BP 算法 <sup>[4]</sup>	0.402	90.30	0.61
二阶 CNN <sup>[6]</sup>	0.172	96.17	1.60
LeNet-5 <sup>[5]</sup>	0.152	97.09	2.00
LED-LeNet	0.036	99.52	1.00

通过对比 BP 算法、二阶卷积神经网络算法、最新版 LeNet-5 网络模型和本文算法在测试集准确率,本文提出的 LED-LeNet 卷积神经网络在识别 LED 字体类数字方面优于现有算法。

虽然本文提出算法拥有较高的准确率,但在图像高度模糊、字体笔画缺失时仍误识别,误识别图像如图 9 所示。



图 9 测试集识别错误图像

Fig. 9 Test set recognition error image

## 5 结论

本文结合自然场景下识别 LED 灯类字体数字需求提出了一种 LED-Net 卷积神经网络。该网络模型以 LeNet-5 为基础,借鉴 VGG 网络使用 2 个  $3 \times 3$  卷积核代替  $5 \times 5$  卷积核思想,进一步提高网络识别准确率。作为验证,该网络在识别自然场景下 LED 灯类字体的交通灯倒计时数字时明显优于现有识别算法。下一步将着重优化网络识别速度,通过量化减少网络模型体积,增加训练集以增加网络泛化能力。

## 参考文献

[1] HE P, HUANG W L, HE T, et al. Single shot text detector with regional attention[C]. IEEE International Conference on Computer Vision, 2017:3066-3074.

[2] 王晓东,魏成保,冯海荣,等.基于 SVM 的变电站保护室数显仪表数字识别方法[J].电子测量技术, 2019,42(2):92-95.  
WANG X D, WEI CH B, FENG H R, et al. SVM-based digital identification method for digital display instrument of substation protection room [J]. Electronic Measurement Technology, 2019, 42(2): 92-95.

[3] 崔家瑞,纪建才,胡广大,等.基于图像的换流站用避雷器仪表识别[J].电工技术学报,2015,30(S1): 377-382.  
CUI J R, JI J C, HU G Y, et al. A recognition method of lightning arrester instrument in converter station based on image [J]. Transactions of China Electrotechnical Society, 2015, 30(S1): 377-382.

[4] 刘进博,朱新新,伍越,等.基于神经网络和支持向量机的多机动车车牌在线检测方法[J/OL]. https://doi.org/10.16383/j.aas.c180753, 2019-07-24.  
LIU J B, ZHU X X, WU Y, et al. An on-line method for multi-license plates recognition based on neural network and support vector machine [J/OL]. https://doi.org/10.16383/j.aas.c180753, 2019-07-24.

[5] 白睿,徐友春,李永乐,等.智能车道路场景数字字符识别技术[J].激光与光电子学进展,2019,56(19): 211-219.  
BAI R, XU Y CH, LI Y L, et al. Digital character recognition technique for intelligent vehicles in road scenes [J]. Laser & Optoelectronics Progress, 2019, 56(19): 211-219.

[6] SERMANET P, CHINTALA S, LECUN Y. Convolutional neural networks applied to house numbers digit classification [C]. Proceedings of the 21st International Conference on Pattern Recognition, IEEE, 2012: 3288-3291.

[7] KUMAR T P, SAIMURUGAN M. A multi-sensor information fusion for fault diagnosis of a gearbox utilizing discrete wavelet features [J]. Measurement Science and Technology, 2019 30(8): 085101.

[8] GUNERKAR R S, JALAN A K. Fault diagnosis of rolling element bearing based on artificial neural network [J]. Journal of Mechanical Science and Technology, 2019, 33(2): 505-511.

[9] 刘笑楠,白雨辰,尹思璐,等.基于类卷积神经网络的可见光虹膜识别方法[J].仪器仪表学报,2017, 38(11): 2651-2658.  
LIU X N, BAI Y CH, YIN S L, et al. Iris recognition of visible light based on analogous convolutional neural network [J]. Chinese Journal of Scientific Instrument, 2017, 38(11): 2651-2658.

- [10] 李勇, 林小竹, 蒋梦莹. 基于跨连接 LeNet-5 网络的面部表情识别 [J]. 自动化学报, 2018, 44(1): 176-182.  
LI Y, LIN X ZH, JIANG M Y. Facial expression recognition with cross-connect LeNet-5 network [J]. Acta Automatica Sinica, 2018, 44(1): 176-182.
- [11] 余琼芳, 胡亚倩, 杨艺. 基于小波特征及深度学习的故障电弧检测 [J]. 电子测量与仪器学报, 2020, 34(3): 100-108.  
YU Q F, HU Y Q, YANG Y. Arc fault detection based on wavelet feature and deep learning [J]. Journal of Electronic Measurement and Instrumentation, 2020, 34(3): 100-108.
- [12] 曲景影, 孙显, 高鑫. 基于 CNN 模型的高分辨率遥感图像目标识别 [J]. 国外电子测量技术, 2016, 35(8): 45-50.  
QU J Y, SUN X, GAO X. Remote sensing image target recognition based on CNN [J]. Foreign Electronic Measurement Technology, 2016, 35(8): 45-50.
- [13] 常亮, 邓小明, 周明全, 等. 图像理解中的卷积神经网络 [J]. 自动化学报, 2016, 42(9): 1300-1312.  
CHANG L, DENG X M, ZHOU M Q, et al. Convolutional neural networks in image understanding [J]. Acta Automatica Sinica, 2016, 42(9): 1300-1312.
- [14] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [J]. Computer Vision and Pattern Recognition, 2014, arXiv: 1409.1556.
- [15] RAMACHANDRAN P, ZOPH B, LE Q V. Swish: A self-gated activation function [J]. Computer Science, 2017, arXiv: 1710.05941.
- [16] HINTON G E, SRIVASTAVA N, KRIZHEVSKY A, et al. Improving neural networks by preventing coadaptation of feature detectors [J]. Computer Science, 2012, 3(4): 212-223.
- [17] 卜令正, 王洪栋, 朱美强, 等. 基于改进卷积神经网络的多源数字识别算法 [J]. 计算机应用, 2018, 38(12): 3403-3408.  
BU L Z, WANG H D, ZHU M Q, et al. Multi-source digital recognition based on improved convolutional neural network [J]. Journal of Computer Applications, 2018, 38(12): 3403-3408.
- [18] 周波, 李俊峰. 结合目标检测的人体行为识别 [J/OL]. 自动化学报, 2019, <https://doi.org/10.16383/j.aas.c180848>.  
ZHOU B, LI J F. Human action recognition combined

with object detection [J/OL]. Acta Automatica Sinica, 2019, <https://doi.org/10.16383/j.aas.c180848>.

- [19] 杨钊, 陶大鹏, 张树业, 等. 大数据下的基于深度神经网络的相似汉字识别 [J]. 通信学报, 2014, 35(9): 184-189.  
YANG Z, TAO D P, ZHANG SH Y, et al. Similar handwritten Chinese character recognition based on deep neural networks with big data [J]. Journal on Communications, 2014, 35(9): 184-189.
- [20] 王金甲, 陈浩, 刘青玉. 大数据下的深度学习研究 [J]. 高技术通讯, 2017, 27(1): 27-37.  
WANG J J, CHEN H, LIU Q Y. The study of deep learning under big data [J]. Chinese High Technology Letters, 2017, 27(1): 27-37.
- [21] 陈庆, 闫斌, 叶润, 等. 航拍绝缘子卷积神经网络检测及自爆识别研究 [J]. 电子测量与仪器学报, 2017, 31(6): 942-953.  
CHEN Q, YAN B, YE R, et al. Insulator detection and recognition of explosion fault based on convolutional neural networks [J]. Journal of Electronic Measurement and Instrumentation, 2017, 31(6): 942-953.

## 作者简介



王立刚, 2017 年于沈阳工业大学获得学士学位, 现为沈阳工业大学硕士研究生, 主要研究方向为深度学习及自然场景文本检测。  
E-mail: ganggang233@hotmail.com

**Wang Ligang** received his B. Sc. degree from Shenyang University of Technology in 2017. Now he is a M. Sc. candidate at Shenyang University of Technology. His main research interests include deep learning and natural scene text detection.



张志佳, 分别在 1996 年和 2002 年在东北大学获得学士学位和硕士学位, 2006 年于中国科学院沈阳自动化研究所获得博士学位, 现为沈阳工业大学教授、博士生导师, 主要研究方向为机器视觉检测技术、图像处理与模式识别。

E-mail: zzjsut@126.com

**Zhang Zhijia** received B. Sc. and M. Sc. both from Northeastern University in 1996 and 2002, Ph. D. from Shenyang Institute of Automation of Chinese Academy of Sciences in 2006, respectively. Now he is a professor and Ph. D. supervisor at Shenyang University of Technology. His main research interests include machine vision, image processing and pattern recognition.