



# 智能系统学报

CAAI TRANSACTIONS ON INTELLIGENT SYSTEMS

## 一种改进的深度学习的路面交通标识识别算法

何锐波, 狄岚, 梁久祯

引用本文:

何锐波, 狄岚, 梁久祯. 一种改进的深度学习的路面交通标识识别算法[J]. 智能系统学报, 2020, 15(6): 1121–1130.

HE Ruibo, DI Lan, LIANG Jiuzhen. An improved deep learning algorithm for road traffic identification[J]. *CAAI Transactions on Intelligent Systems*, 2020, 15(6): 1121–1130.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.201811009>

## 您可能感兴趣的其他文章

### 基于生成式对抗网络的道路交通模糊图像增强

Enhancement of blurred road-traffic images based on generative adversarial network

智能系统学报. 2020, 15(3): 491–498 <https://dx.doi.org/10.11992/tis.201903041>

### 深度学习的双人交互行为识别与预测算法研究

Human interaction recognition and prediction algorithm based on deep learning

智能系统学报. 2020, 15(3): 484–490 <https://dx.doi.org/10.11992/tis.201812029>

### 基于小样本学习的LCD产品缺陷自动检测方法

An automatic small sample learning-based detection method for LCD product defects

智能系统学报. 2020, 15(3): 560–567 <https://dx.doi.org/10.11992/tis.201904020>

### 基于改进卷积神经网络的多标记分类算法

A multi-label classification algorithm based on an improved convolutional neural network

智能系统学报. 2019, 14(3): 566–574 <https://dx.doi.org/10.11992/tis.201804056>

### 卷积神经网络的贴片电阻识别应用

Chip resistance recognition based on convolution neural network

智能系统学报. 2019, 14(2): 263–272 <https://dx.doi.org/10.11992/tis.201710005>

### 基于改进型BP神经网络的手部动作识别

Hand-motion recognition based on improved BP neural network

智能系统学报. 2018, 13(5): 848–854 <https://dx.doi.org/10.11992/tis.201703018>

微信公众平台



关注微信公众号, 获取更多资讯信息

DOI: 10.11992/tis.201811009

# 一种改进的深度学习道路交通标识识别算法

何锐波<sup>1,2</sup>, 狄岚<sup>1</sup>, 梁久祯<sup>3</sup>

(1. 江南大学 人工智能与计算机学院, 江苏 无锡 214122; 2. 中国电子科技集团公司第二十八研究所, 江苏 南京 210007; 3. 常州大学 信息科学与工程学院, 江苏 常州 213164)

**摘要:** 针对复杂的环境, 结合图像预处理与深度学习神经网络, 提出了一种道路交通标识识别算法。该方法不仅利用图像分割技术, 而且利用卷积神经网络模型对道路交通标识进行了更准确的识别。首先, 通过调节光照影响、去除复杂背景、数据增强和归一化等批量预处理操作, 形成一个完整的数据集; 然后, 结合 squeeze-and-excitation 思想和残差网络结构, 充分训练出自己的卷积神经网络模型; 最后, 将优化的网络模型用于道路交通标识的识别。实验结果表明, 该方法使训练时间缩短了 12% 左右, 识别精度可达 99.26%。

**关键词:** 道路交通标识识别; 图像分割; 卷积神经网络; 去除复杂背景; 数据增强; 归一化; 压缩和激励网络; 残差连接

中图分类号: TP391.4 文献标志码: A 文章编号: 1673-4785(2020)06-1121-10

中文引用格式: 何锐波, 狄岚, 梁久祯. 一种改进的深度学习道路交通标识识别算法 [J]. 智能系统学报, 2020, 15(6): 1121-1130.

英文引用格式: HE Ruibo, DI Lan, LIANG Jiuzhen. An improved deep learning algorithm for road traffic identification[J]. CAAI transactions on intelligent systems, 2020, 15(6): 1121-1130.

## An improved deep learning algorithm for road traffic identification

HE Ruibo<sup>1,2</sup>, DI Lan<sup>1</sup>, LIANG Jiuzhen<sup>3</sup>

(1. School of Artificial Intelligence and Computer Science, Jiangnan University, Wuxi 214122, China; 2. The 28th Research Institute of China Electronics Technology Group Corporation, Nanjing 210007, China; 3. School of Information Science and Engineering, Changzhou University, Changzhou 213164, China)

**Abstract:** This study proposes a road traffic identification algorithm based on image preprocessing and deep-learning neural networks for complex environments. The proposed method uses not only the image segmentation technology but also the convolutional neural network model to more accurately identify the road traffic signs. First, a complete dataset is obtained via batch preprocessing operations, including illumination effect adjustment, complex background elimination, data enhancement, and normalization. Next, the convolutional neural network model is sufficiently trained based on the combination of the squeeze-and-excitation network and residual network structure concepts. Finally, the optimized network model is used to identify the road traffic signs. The experimental result shows that the proposed method reduces the training time by approximately 12% and that the recognition accuracy can reach 99.26%.

**Keywords:** road traffic identification; image segmentation; convolutional neural network; complex background elimination; data enhancement; normalization; squeeze-and-excitation network; residual connection

道路交通标识的探测和检验在快速发展的智能交通领域受到了广泛的关注, 其含有的限速、限行、限高提示、车道方向的指示等提示信息, 为

道路交通的通行效率和通行安全提供了保障, 对其内容的研究具有重要意义。

现今有许多针对路标的检测方法, 如: 采用不同色彩空间<sup>[1-3]</sup>进行匹配检测的方法, 对形状及轮廓<sup>[4-6]</sup>进行分割提取的方法, 或融合形状与色彩特征<sup>[7]</sup>的检测方法、或采用深度学习、神经网络有

收稿日期: 2018-11-11.

基金项目: 江苏省研究生科研与实践创新计划项目 (KYCX18\_1872).

通信作者: 狄岚. E-mail: dilan@jiangnan.edu.cn.

关知识进行检测的方法<sup>[8-10]</sup>。以色彩空间为模板的检测方法最早由 Yang 等<sup>[11]</sup>提出,通过提取 R 层信息,构建待选模板,对路标图像进行粗略分割;因 YCbCr 色彩空间<sup>[12]</sup>具有受亮度变化影响较小的特点,采用依据图像 Cb、Cr 分量进行图像分割;提取形状特征模板的检测识别方法最早由 Piccioli 等<sup>[13]</sup>提出,后逐渐改进为 SNCC 算法<sup>[14]</sup>。现今采用多尺度的算法来识别检测,如马尔科夫随机场和基于贝叶斯的<sup>[15]</sup>、基于特征点采样和 OSPA 距离的<sup>[16]</sup>识别检测路标的算法,这些算法虽然在速度和准确率上有着特有的优势,但却有着严苛的限制条件,缺乏普适性。深度学习模型可以看作由多个人工神经网络层组成<sup>[17]</sup>,通过构建含有多个隐层的神经网络模型,将低层特征通过逐层非线性特征组合变换,形成更加抽象的高层特征表达。加深模型网络是常用于提高检测识别准确率的方法之一,但随之而来会出现“退化”问题。He 等<sup>[18]</sup>就提出了 ResNet (residual neural network),缓解了“退化”现象,但并没有解决梯度爆炸问题。WMW 团队提出了 SENet (squeeze-and-excitation networks) 模型<sup>[19]</sup>,使用特征通道融合的方式,通过建模展现特征通道间依赖关系,从而强化有用特征并抑制无用特征。本文结合 SENet 和 ResNet 模型的原理,取各自优点,使用较小的网络层数来达到较高的识别水平。

因此针对上述存在的问题,本文所做的工作是:

1) 提出了将 HU 不变量与色彩空间多尺度相结合进行处理的方法,减少光照影响,并除去复杂背景,降低无用特征被学习到的几率;

2) 通过旋转数据增强及归一化操作,对小型数据集进行增强,满足网络训练要求,提高识别准确率;

3) 结合 SE 网络和残差网络结构,构建了自己的网络模型。同时,通过调整相应的参数以适应路标图像的特点,采用浅层网络模型使训练时间得到大幅减少,保证具有较高识别准确率。

## 1 图像预处理

### 1.1 数据集的光照处理

环境因素会给交通标识的图像采集工作带来很大的负面影响,从而造成识别困难。本文从 RGB 和 YCbCr 两个不同的颜色空间对图像进行阈值划分,从而实现对图像的预处理,降低环境因素干扰。RGB 色彩空间模型通过构建红、绿、蓝 3 个通道的笛卡尔坐标系,来表示色彩图像,通

道的取值范围为  $[0,1]$ 。YCbCr 色彩空间较 RGB 色彩空间能显示图像整体的亮度信息。将 RGB 图像转换成 YCbCr 空间表示的具体转换公式<sup>[20]</sup>为

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} + \begin{bmatrix} 0.257 & 0.504 & 0.098 \\ -0.148 & -0.291 & 0.439 \\ 0.439 & -0.368 & -0.071 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

经过空间换算,经过大量实验得到图像亮度阈值在  $Y \in [85, 200]$  内较为正常,通过图像增强等方式对阈值外图像进行处理。通过 RGB 色彩空间的叠加或删减来对图像进行增强,具体效果如图 1 所示,对曝光不足图像的处理公式为

$$R = R + k(1 - R) \cdot R$$

式中  $k$  为一个可调节参数。

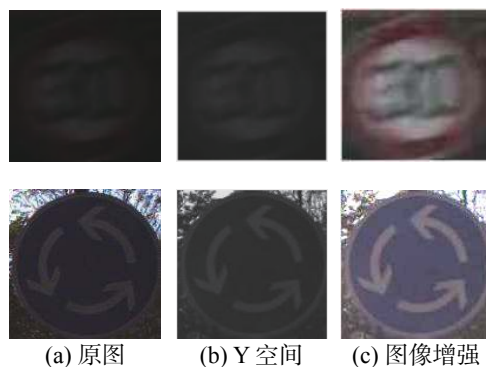


图 1 曝光不足图像增强

Fig. 1 Insufficient exposure image enhancement

对于过曝图片的增强公式  $R = R - k(1 - R) \cdot R$ ,  $k$  是一个可调节参数,实验效果如图 2。



图 2 过曝图像增强

Fig. 2 Overexposure image enhancement

### 1.2 特征匹配

数字图像处理中,通过提取轮廓边界及形状特征进行分割是图像识别领域的一种常用的方法。本文整合交通标识图像的共有形状特征进行特征提取,降低神经网络在数据集训练时的无用特征学习。图 3 所示为交通标识形状特征图。





图3 交通标识形状特征图

Fig. 3 Traffic-sign shape feature chart

通过形态学的开闭操作降低图像中的噪声, 后用最大类间方差法寻找图像阈值, 以阈值为界灰度图像转换为二值图像; 最后使用孔洞填充提取图像特征形状。图4为操作示意图。



(a) 原图 (b) 开操作 (c) 闭操作 (d) 最大类间方差 (e) 孔洞填充

图4 图像形状特征提取

Fig. 4 Shape feature extraction

在取得图像的形状特征后, 可通过对比轮廓矩<sup>[21]</sup>的方式判断形状相似度。轮廓的  $(p, q)$  矩定义为

$$m_{p,q} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} I(x,y) x^p y^q f(x,y) dx dy$$

式中:  $p = 1, 2, \dots; q = 1, 2, \dots$ 。对于离散图像  $f(x,y)$ , 其阶标准矩<sup>[22]</sup>定义公式为

$$m_{p,q} = \sum_{y=1}^N \sum_{x=1}^M x^p y^q f(x,y), \quad p = 0, 1, 2, \dots; q = 1, 2, \dots$$

$p+q$  阶中心矩定义为

$$\mu_{p,q} = \sum_{y=1}^N \sum_{x=1}^M (x - \bar{x})^p (y - \bar{y})^q f(x,y)$$

$$p = 0, 1, 2, \dots; q = 1, 2, \dots$$

式中  $\bar{x}$  和  $\bar{y}$  代表图像的重心, 归一化的中心矩定义为  $\eta_{pq} = \mu_{pq} / \mu_{00}^r$ , 其中,  $r = (p+q)/2 + 1$ 。

通过二阶和三阶归一化中心矩可以构造出 7 个不变矩  $M_1 \sim M_7$ :

$$M_1 = \eta_{20} + \eta_{02}$$

$$M_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2$$

$$M_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2$$

$$M_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2$$

$$M_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{21} + \eta_{03})^2 - (\eta_{21} + \eta_{03})^2]$$

$$M_6 = (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] +$$

$$4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03})$$

$$M_7 = (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] +$$

$$4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03})$$

这 7 个不变矩构成了一组特征向量, 通过将交通标识图像与标准轮廓图像形状进行特征向量的匹配, 形成 mask 掩膜, 进行路标特征的提取。图5表示 mask 掩膜的匹配生成及数据集图片的特征提取去背景。

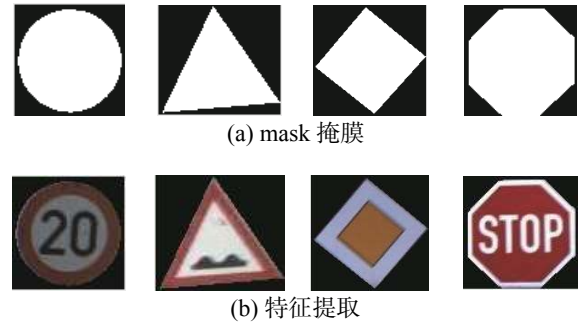


图5 掩膜与特征提取

Fig. 5 Mask and decontextualization

### 1.3 归一化与数据增强

本文采用三次样条插值的方法<sup>[23]</sup>将深度学习模型所需的训练集归一化处理到  $128 \times 128$  的尺寸大小。对小于尺寸的图片, 根据深度学习模型的特性, 可以通过图像的几何变换 (如平移、旋转、尺度拉伸、对比度调整以及颜色变换) 进行数据增强; 对于大尺寸的, 则采用均值缩减降低图像尺寸。本文对数据量较少的进行旋转变换来扩充数据集。

## 2 模型与算法

### 2.1 卷积神经网络结构

随着人们对神经网络的深入探索研究, 衍生出了名为深度学习的思维理念, 其中较为常用的叫卷积神经网络模型。其核心思想是通过反复的迭代, 逐层反向调节网络中的权重参数来最小化损失函数同时提高识别率。为解决参数多、复杂度高问题采用权值共享策略, 而其稀疏连接的特性提高了并行处理能力及误差的容忍程度。卷积层 (convolution layer)、池化层 (pooling layer) 和全连接层 (Full-connection Layer) 是卷积神经网络的主要构成部分; 卷积神经网络结构示意图如图6。

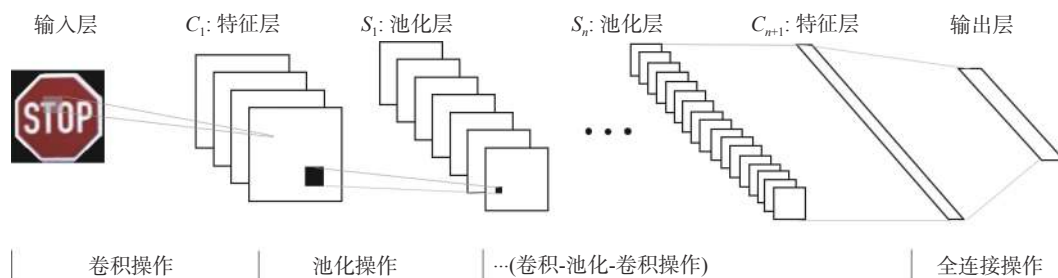


图6 卷积神经网络结构

Fig. 6 Structure of the convolutional neural network

## 2.2 局部响应归一化与残差连接

文献[24]针对梯度爆炸问题,提出了一种局部响应归一化的思想,为提高网络模型的泛化能力减少过拟合程度,采用竞争机制,增大局部特征值的差距分化;文献[25]提出一种恒等映射拟合的方法优化深层网络的训练学习。局部响应归一化(local response normalization)如图7所示:借

鉴别抑制思想,利用附近几层对特征中间层进行平滑处理,计算公式为

$$b_{x,y}^i = \frac{a_{x,y}^i}{\left( k + \alpha \sum_{j=\max(0, i-\frac{n}{2})}^{\min(N-1, i+\frac{n}{2})} (a_{x,y}^j)^2 \right)^\beta}$$

式中:  $k$ 、 $n$ 、 $\alpha$ 、 $\beta$  为超参数;  $a_{x,y}^i$  表示第  $i$  个核在  $(x,y)$  位置的输出;  $N$  表示该层中核的总数。

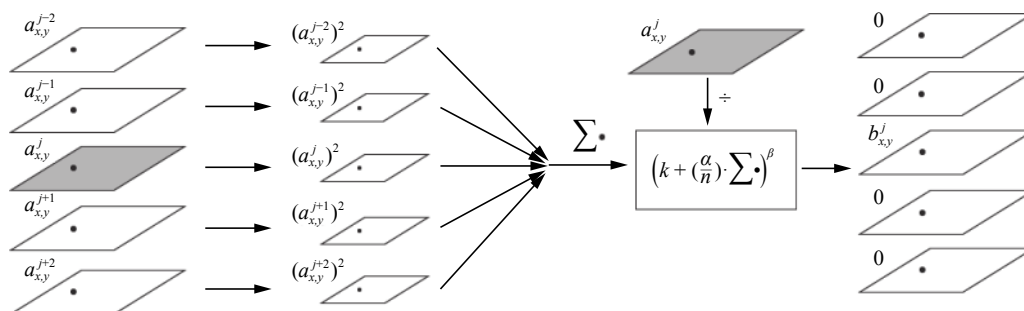


图7 局部响应归一化示意图

Fig. 7 Local response normalization diagram

图8为残差连接示意图,为降低梯度爆炸问题,简化训练学习过程,利用数学运算把式  $H(x) = F(x) + x$  转化成  $F(x) = H(x) - x$  的思想,其中当  $F(x) = 0$  时  $H(x) = x$ ; 若  $F(x) \approx 0$ , 则只考虑  $H(x)$  与  $x$  的相差部分,残差连接用此思想,保留了学习到的深层特征,又简化了层级梯度。

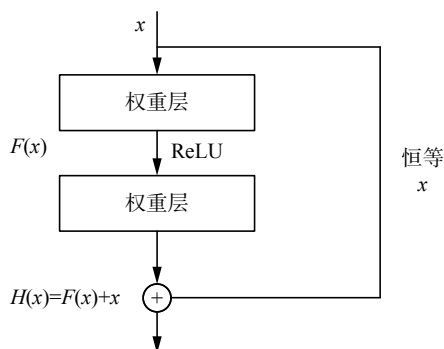


图8 残差连接

Fig. 8 Residual connection

## 2.3 Squeeze-and-Excitation 模块

深度学习模型的改进方法中,许多都从空间维度进行考量;而 SENet(squeeze-and-excitation networks)通过使用特征重标定方法,不引入新的空间维度并考虑特征通道间的关系来进行特征通道融合。Squeeze-and-Excitation 可以被任何形式如:  $F_{tr}: X \rightarrow U, X \in R^{H' \times W' \times C'}, U \in R^{H \times W \times C}$  的运算定义,本质是个计算模块,把  $F_{tr}$  视为一种卷积运算,让  $V = [v_1 v_2 \cdots v_C]$  表示滤波核的集合,其中  $v_c$  为第  $c$  层滤波的参数。输出结果可以被表示为  $U = [u_1 u_2 \cdots u_C]$ , 其中

$$u_c = v_c * X = \sum_{s=1}^{C'} v_c^s * x^s$$

式中:  $v_c = [v_c^1 v_c^2 \cdots v_c^{C'}]$  和  $X = [x^1 x^2 \cdots x^{C'}]$  (为了简化符号,省略了偏差项),其中  $v_c^s$  是一个二维空间内核,  $v_c$  中包含了通道与通道之间的相互依赖关系。

图9是 Squeeze-and-Excitation 模块思想的直观图。

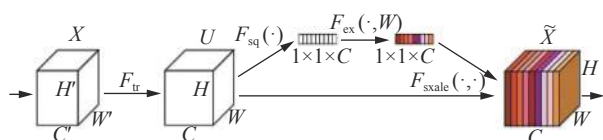


图9 Squeeze-and-Excitation 模块  
Fig. 9 A Squeeze-and-Excitation block

## 2.4 基于深度学习改进的路标识别算法

通过整合上述理论与公式,优化网络结构、调节模型参数、结合 Squeeze-and-Excitation 与残

差连接思想,形成本文的算法,具体流程如图10:

- 1) 对数据集进行特征提取及亮度值批量预处理操作;
- 2) 对数据集进行归一化及数据增强操作;
- 3) 结合 Squeeze-and-Excitation 模块及残差连接思想,调节参数优化网络模型结构,构建自己的网络模型,并对训练集进行充分训练学习;
- 4) 利用训练好的模型结构对任意交通标识图像进行识别检测,验证模型算法。

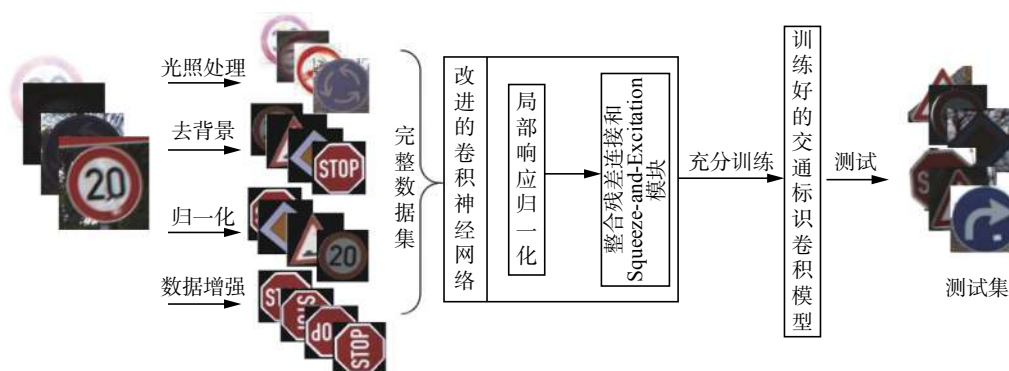


图10 基于深度学习改进的路标识别算法流程

Fig. 10 Improved road sign recognition algorithm flow using deep learning

## 3 实验结果与分析

实验采用德国交通标识标准数据库 (GTSRB), 其包含 39209 张训练集路标图像及 12630 张测试集路标图像, 共有 43 类; 以及比利时交通标识数据库 (BelgiumTSC), 其包含 4590 张训练集图像及 2520 张测试集图像, 共 62 类。数据集中的图像大小从  $15 \times 15$  到  $250 \times 250$  不等, 通过对训练集进行归一化和数据增强, 得到增强后的 BelgiumTSC 训练集含图像 6272 张, GTSRB 的训练集含图像 51989 张, 图像大小统一为  $(128 \times 128)$ 。实验所用电脑型号为 Dell Precision T7910, 中央处理器为 E5-2600v4 八核, 内存大小为 32 GB, 图形处理器为 NVIDIA QUADRO M2000, 具有 4 GB 显存, 使用软件为 MATLAB2016b 及 VisualStudio 2010, 并采用 Caffe 框架<sup>[26]</sup>, 为减少对比实验误差, 将非对比项参数固定, 如初始学习率设置为 0.001, 训练集图像一次处理量设为 100, 测试集图像一次处理量为 10, 优化算法为随机梯度下降法 (stochastic gradient descent, SGD), 学习率下降函数为 step 函数, momentum 值为 0.9。模型的好坏通过损失函数来衡量, 当损失函数值趋于平稳则代表模型已充分学习。

整合 Squeeze-and-Excitation Networks 思想和残差连接到经典 AlexNet 网络模型中进行对比实

验。表1为经典网络及本文模型训练时间表。图11为模型层级及参数设定, 图12和图13为不同深度模型在 BelgiumTSC 和 GTSRB 两数据集上的实验结果, 迭代次数设置为 30000, 是多次实验抽取的结果展示, 通过实验可得准确率与网络层级并非线性正相关关系, 因为梯度爆炸问题, 导致实验层级增加到 10 层之后模型开始无法收敛。

表1 训练时间 (比利时数据库)  
Table 1 Training time (BelgiumTSC)

| 模型名称      | 训练时间( $\pm 10$ min) |
|-----------|---------------------|
| 5层LeNet模型 | 1 h 25 min          |
| AlexNet   | 1 h 30 min          |
| 8层网络结构模型  | 2 h                 |
| 残差网络      | 6 h                 |
| SENet     | 4 h 40 min          |
| 本文网络模型    | 2 h 20 min          |

本文的网络模型结构如图14所示, 分别为与经典的网络进行原始数据、数据增强及预处理操作不同训练集的对比分析。表2和表3给出了多次实验下不同网络模型的准确率及损失值。由表格实验结果横向对比可知, 对数据集进行归一化处理、数据增强操作以及预处理操作能防止过拟合现象, 增加模型的识别准确率; 同时通过实验结果纵向对比发现残差连接及 SENet 的思想均能

提高网络模型的识别率。综合上述实验结果可得,神经网络模型学习需要大量数据支持,对数据集进行数据增强能有效避免过拟合现象,而对数据的预处理能够减少无效特征的学习从而增加识别准确率; Squeeze-and-Excitation block 思想是一个特征重标定的计算模块,为提高识别准确率

增加了模型的计算复杂度,延长了训练时间,而残差连接思想简化了网络层级间的运算,降低了训练时间,同时保障了识别准确率。结合两种算法思想形成的本文网络模型结构,在 GTSRB 数据集上的实验结果表明其识别准确率可以达到 99.26%, 并且有着较优的训练时长。

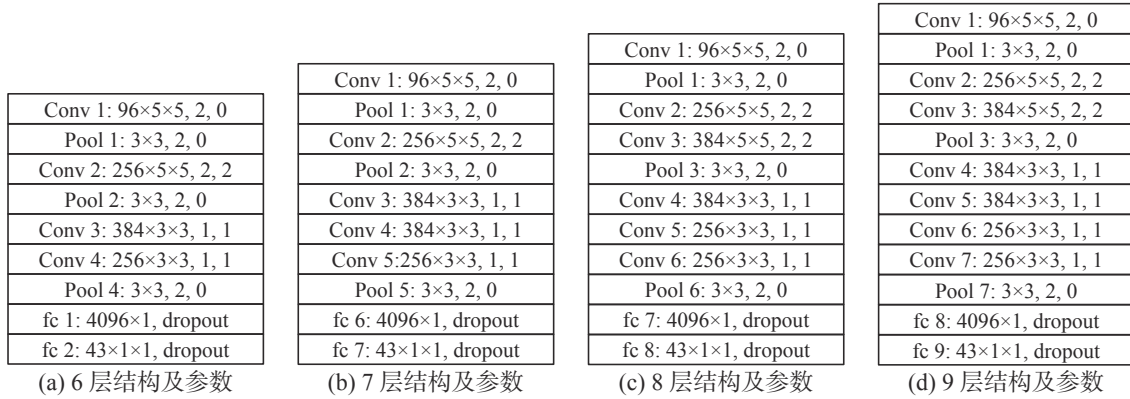


图 11 网络结构模型及参数设置图

Fig. 11 Network structure model and parameter setting

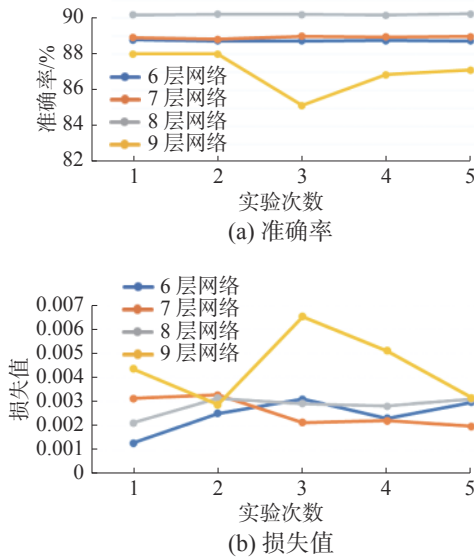


图 12 不同网络深度下的准确率以及损失值 (比利时数据库)

Fig. 12 Accuracy and loss values of different network depths (BelgiumTSC)

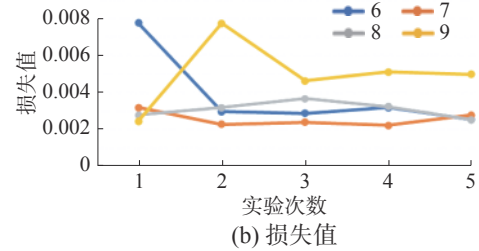
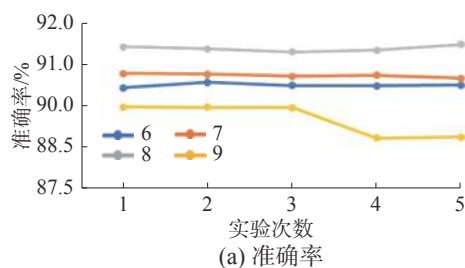


图 13 不同网络深度下的准确率以及损失值 (德国标准数据库)

Fig. 13 Accuracy and loss values of different network depths (GTSRB)

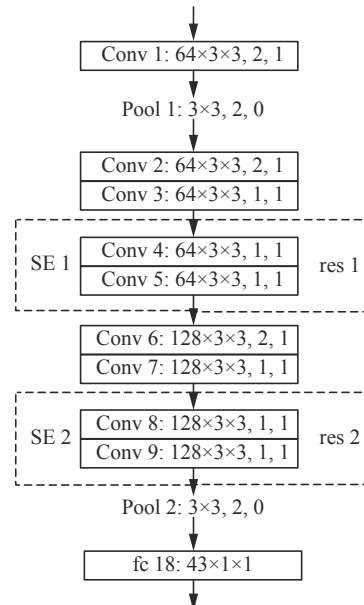


图 14 本文网络结构模型及参数设置

Fig. 14 Network structure model and parameter setting

表 2 图像预处理与残差连接对网络模型影响 (比利时数据库)  
Table 2 Influence of image preprocessing and residual connection on a network model (BelgiumTSC)

| 模型名称      | 原始数据      |          | 归一化与数据增强  |          | 预处理数据     |          |
|-----------|-----------|----------|-----------|----------|-----------|----------|
|           | 损失值       | 准确率/%    | 损失值       | 准确率/%    | 损失值       | 准确率/%    |
| LeNet-5   | 0.012 691 | 80.976 7 | 0.012 205 | 82.321 5 | 0.008 716 | 85.410 7 |
|           | 0.011 784 | 80.895 7 | 0.010 177 | 83.001 2 | 0.009 872 | 85.125 8 |
|           | 0.010 335 | 81.345 8 | 0.017 484 | 81.982 4 | 0.011 178 | 84.842 4 |
|           | 0.016 348 | 80.881 6 | 0.014 532 | 82.195 8 | 0.009 867 | 85.019 1 |
|           | 0.009 871 | 81.218 3 | 0.012 031 | 82.650 7 | 0.009 915 | 84.943 8 |
| ResNet-50 | 0.003 014 | 93.218 4 | 0.003 717 | 93.428 1 | 0.003 928 | 95.993 8 |
|           | 0.002 958 | 93.104 8 | 0.002 458 | 93.521 4 | 0.003 215 | 96.174 1 |
|           | 0.003 234 | 92.943 7 | 0.002 714 | 94.169 7 | 0.003 001 | 96.348 1 |
|           | 0.002 917 | 93.441 4 | 0.002 848 | 94.154 8 | 0.003 128 | 96.174 9 |
|           | 0.003 418 | 92.218 4 | 0.003 171 | 93.628 1 | 0.002 714 | 97.128 4 |
| SENet     | 0.003 478 | 92.648 1 | 0.002 917 | 94.248 3 | 0.003 288 | 96.204 2 |
|           | 0.002 951 | 93.100 8 | 0.003 058 | 94.123 1 | 0.002 114 | 97.142 5 |
|           | 0.003 104 | 92.986 1 | 0.003 014 | 94.192 7 | 0.003 018 | 96.765 7 |
|           | 0.002 719 | 93.426 6 | 0.004 048 | 93.633 5 | 0.003 141 | 96.528 3 |
|           | 0.002 814 | 93.152 8 | 0.004 175 | 92.628 1 | 0.002 101 | 97.217 7 |
| AlexNet   | 0.004 827 | 88.082 7 | 0.003 527 | 90.218 0 | 0.004 282 | 91.847 2 |
|           | 0.004 897 | 87.341 3 | 0.003 208 | 90.138 2 | 0.003 019 | 92.458 2 |
|           | 0.005 316 | 88.863 2 | 0.004 083 | 89.647 2 | 0.002 318 | 92.006 9 |
|           | 0.003 037 | 90.001 2 | 0.002 544 | 91.040 1 | 0.001 827 | 93.310 1 |
|           | 0.004 096 | 89.082 5 | 0.004 218 | 89.543 8 | 0.004 148 | 91.981 1 |
| 8层结构      | 0.003 028 | 90.105 8 | 0.005 487 | 90.328 4 | 0.004 281 | 91.528 7 |
|           | 0.004 017 | 90.952 1 | 0.005 284 | 90.624 8 | 0.005 414 | 90.185 7 |
|           | 0.004 857 | 88.763 2 | 0.004 018 | 91.518 7 | 0.001 217 | 94.001 0 |
|           | 0.004 247 | 89.821 1 | 0.003 528 | 92.528 4 | 0.002 108 | 93.828 4 |
|           | 0.003 848 | 89.384 4 | 0.004 281 | 91.258 4 | 0.003 031 | 93.322 6 |
| 本文模型      | 0.002 561 | 94.101 1 | 0.002 017 | 95.080 7 | 0.001 727 | 97.961 5 |
|           | 0.003 151 | 93.821 7 | 0.003 093 | 93.846 8 | 0.001 958 | 97.766 8 |
|           | 0.003 039 | 93.012 4 | 0.003 442 | 93.764 9 | 0.002 018 | 97.582 1 |
|           | 0.003 096 | 92.719 2 | 0.003 031 | 93.524 6 | 0.003 001 | 96.991 8 |
|           | 0.002 189 | 94.215 8 | 0.002 515 | 94.121 8 | 0.002 364 | 98.083 7 |



表3 图像预处理与残差连接对网络模型影响(德国数据库)

Table 3 Influence of image preprocessing and residual connection on a network model (GTSRB)

| 模型名称      | 原始数据      |          | 归一化与数据增强  |          | 预处理数据     |          |
|-----------|-----------|----------|-----------|----------|-----------|----------|
|           | 损失值       | 准确率/%    | 损失值       | 准确率/%    | 损失值       | 准确率/%    |
| LeNet-5   | 0.092 814 | 83.938 2 | 0.008 746 | 85.927 1 | 0.009 894 | 86.281 4 |
|           | 0.008 988 | 85.562 0 | 0.011 276 | 83.524 7 | 0.008 892 | 87.018 5 |
|           | 0.010 205 | 84.582 1 | 0.009 687 | 86.369 1 | 0.010 583 | 86.001 9 |
|           | 0.010 018 | 84.641 1 | 0.009 825 | 86.182 7 | 0.009 571 | 86.514 7 |
|           | 0.012 530 | 82.158 2 | 0.009 825 | 85.242 8 | 0.011 759 | 85.486 2 |
| ResNet-50 | 0.002 727 | 94.012 1 | 0.001 978 | 96.214 2 | 0.001 872 | 97.950 1 |
|           | 0.002 715 | 94.728 3 | 0.002 214 | 95.582 7 | 0.001 378 | 98.260 5 |
|           | 0.002 428 | 94.715 1 | 0.003 014 | 94.851 7 | 0.001 457 | 97.756 8 |
|           | 0.002 823 | 94.218 4 | 0.002 013 | 96.085 1 | 0.001 311 | 98.664 7 |
|           | 0.003 410 | 93.928 1 | 0.002 401 | 95.472 4 | 0.001 554 | 98.025 8 |
| SENet     | 0.002 979 | 94.071 3 | 0.003 478 | 94.921 2 | 0.001 002 | 99.001 3 |
|           | 0.002 807 | 94.841 1 | 0.002 814 | 95.528 7 | 0.001 498 | 98.260 5 |
|           | 0.003 101 | 93.715 1 | 0.002 717 | 95.754 7 | 0.002 052 | 97.756 8 |
|           | 0.002 914 | 94.263 4 | 0.003 071 | 95.589 8 | 0.002 118 | 97.664 7 |
|           | 0.003 075 | 93.982 6 | 0.002 401 | 96.480 1 | 0.001 559 | 98.025 8 |
| AlexNet   | 0.003 248 | 89.825 5 | 0.003 421 | 90.078 5 | 0.002 952 | 91.828 7 |
|           | 0.003 180 | 90.240 2 | 0.003 718 | 90.828 7 | 0.004 859 | 91.258 4 |
|           | 0.002 861 | 90.304 2 | 0.004 182 | 89.938 7 | 0.002 148 | 93.001 9 |
|           | 0.003 225 | 89.982 1 | 0.003 655 | 90.007 2 | 0.002 281 | 92.180   |
|           | 0.002 962 | 90.390 2 | 0.004 557 | 89.871 8 | 0.002 881 | 91.885 1 |
| 8层结构      | 0.003 058 | 91.282 7 | 0.004 272 | 91.855 2 | 0.003 048 | 94.528 4 |
|           | 0.003 180 | 91.354 2 | 0.003 229 | 92.463 2 | 0.002 917 | 93.481 2 |
|           | 0.003 055 | 91.210 2 | 0.003 138 | 92.043 9 | 0.003 017 | 94.002 1 |
|           | 0.004 085 | 90.908 7 | 0.002 927 | 93.384 7 | 0.002 837 | 95.052 8 |
|           | 0.003 017 | 91.292 1 | 0.004 317 | 91.928 2 | 0.002 827 | 93.922 3 |
| 本文模型      | 0.002 176 | 96.002 0 | 0.002 488 | 96.221 2 | 0.000 972 | 99.050 0 |
|           | 0.002 707 | 95.812 3 | 0.001 914 | 96.528 7 | 0.000 198 | 99.260 5 |
|           | 0.002 801 | 95.715 1 | 0.002 418 | 95.855 7 | 0.001 458 | 98.756 8 |
|           | 0.002 018 | 96.218 4 | 0.002 041 | 96.585 3 | 0.001 510 | 98.664 7 |
|           | 0.002 675 | 95.991 6 | 0.001 901 | 96.482 7 | 0.001 154 | 99.025 8 |

## 4 结束语

本文提出用基于深度学习方法来实现道路交通标志识别。该方法分为两部分:图像预处理和卷积神经网络模型。在第一部分中,利用不变矩运算、多色彩空间融合和形状特征提取等方法对数据集图像进行归一化、数据增强及去背景特征提取。在第二部分中,结合残差连接思想及 Squeeze-and-Excitation block 思想形成本文网络模型。由实验可得,本文建立的网络模型识别准确率可达 99.26%,并且训练学习时间 SENet 减少了 12%。进一步工作,通过结合主成分分析法思想及自适应阈值调整方法的思想,提取图像特征并进行增强,从而达到尽量减少人工干预的方式,减少神经网络参数的调整和输入,并保证识别的准确率。

## 参考文献:

- [1] KHAN J F, ADHAMI R R, et al. Traffic sign recognition based on pictogram contours[J]. IEEE ninth international workshop on image analysis for multimedia interactive services, 2011, 1(1): 83–96.
- [2] YANG H M, LIU Chaolin, LIU Kunhao, et al. Traffic sign recognition in disturbing environments[C]//Proceedings of the 14th International Symposium on Methodologies for Intelligent Systems. Maebashi City, Japan, 2003: 252–261.
- [3] BARTNECK N, RITTER W. Colour segmentation with polynomial classification[C]//Proceedings of the 11th IAPR International Conference on Pattern Recognition. Vol. II. Conference B: Pattern Recognition Methodology and Systems. The Hague, Netherlands, 1992: 635–638.
- [4] SANDOVAL H, HATTORI T, KITAGAWA S, et al. Angle-dependent edge detection for traffic signs recognition[C]//Proceedings of 2000 IEEE Intelligent Vehicles Symposium. Dearborn, USA, 2000: 308–313.
- [5] BARNES N, ZELINSKY A, FLETCHER L S. Real-time speed sign detection using the radial symmetry detector[J]. IEEE transactions on intelligent transportation systems, 2008, 9(2): 322–332.
- [6] PAULO C F, CORREIA P L. Traffic sign recognition based on pictogram contours[C]//Proceedings of 2008 IEEE Ninth International Workshop on Image Analysis for Multimedia Interactive Services. Klagenfurt, Austria, 2008: 67–70.
- [7] GAO X W, PODLADCHIKOVA L, SHAPOSHNIKOV D, et al. Recognition of traffic signs based on their colour and shape features extracted using human vision models[J]. Journal of visual communication and image representation, 2006, 17(4): 675–685.
- [8] STALLKAMP J, SCHLIPSING M, SALMEN J, et al. Man vs. computer: benchmarking machine learning algorithms for traffic sign recognition[J]. Neural networks, 2012, 32: 323–332.
- [9] SEEBÖCK P. Deep learning in medical image analysis[D]. Vienna, Vienna: Vienna University of Technology, 2015: 333–338.
- [10] 丁颖洋, 刘格, 王贤哲, 等. 基于深度学习的游行示威视频图像检测方法 [J]. 指挥信息系统与技术, 2018, 9(6): 75–79.  
DING Weiyang, LIU Ge, WANG Xianzhe, et al. Demonstration video image detection method based on deep learning[J]. Command information system and technology, 2018, 9(6): 75–79.
- [11] YANG H M, LIU C L, LIU K H, et al. Traffic sign recognition in disturbing environments[C]//Foundations of Intelligent Systems, 14th International Symposium, ISMIS 2003. Maebashi, Japan, 2003: 28–31.
- [12] ISLAM K T, RAJ R G. Real-time (vision-based) road sign recognition using an artificial neural network[J]. Sensors, 2017, 14(4): 853.
- [13] PICCIOLI G, DE MICHELI E, PARODI P, et al. Robust method for road sign detection and recognition[J]. Image and vision computing, 1996, 14(3): 209–223.
- [14] 房泽平, 段建民, 郑榜贵. 基于特征颜色和 SNCC 的交通标志识别与跟踪 [J]. 交通运输系统工程与信息, 2014, 14(1): 47–52.  
FANG Zeping, DUAN Jianmin, ZHENG Banggui. Traffic signs recognition and tracking based on feature color and SNCC algorithm[J]. Journal of transportation systems engineering and information technology, 2014, 14(1): 47–52.
- [15] ZHOU Jun, BAO Xu, LI Dawei, et al. Traffic video image segmentation model based on Bayesian and spatio-temporal Markov random field[J]. Journal of physics: conference series, 2017, 910(1): 012041.
- [16] 陈利, 刘伟峰, 杨爱兰. 基于 OSPA 距离和特征点采样的路标识别算法 [J]. 哈尔滨师范大学自然科学学报, 2017, 33(2): 55–57.  
CHEN Li, LIU Weifeng, YANG Ailan. A recognition algorithm for road signs based on the OSPA metric and characteristic point sampling[J]. Natural Science Journal of Harbin Normal University, 2017, 33(2): 55–57.
- [17] SHEN Dinggang, WU Guorong, SUK H I. Deep learning in medical image analysis[J]. Annual review of biomedical engineering, 2017, 19: 221–248.
- [18] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, et al. Deep residual learning for image recognition[C]//Proceed-

- ings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA, 2016: 770–778.
- [19] HU Jie, SHEN Li, SUN Gang. Squeeze-and-excitation networks[C]//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA, 2018: 7132–7141.
- [20] 张懿, 刘旭, 李海峰. 数字 RGB 与 YCbCr 颜色空间转换的精度 [J]. 江南大学学报(自然科学版), 2007, 6(2): 200–202.
- ZHANG Yi, LIU Xu, LI Haifeng. The precision of RGB color space convert to YCbCr color space[J]. Journal of JiangNan University (Natural Science Edition), 2007, 6(2): 200–202.
- [21] HU M K. Visual pattern recognition by moment invariants[J]. *IRE transactions on information theory*, 1962, 8(2): 179–187.
- [22] 孙即祥. 模式识别中的特征提取与计算机视觉不变量 [M]. 北京: 国防工业出版社, 2001.
- [23] 温佩芝, 陈晓, 吴晓军, 等. 基于三次样条插值的 GrabCut 自动目标分割算法 [J]. *计算机应用研究*, 2014, 31(7): 2187–2190.
- WEN Peizhi, CHEN Xiao, WU Xiaojun, et al. Automatic target segmentation algorithm of GrabCut based on cubic spline interpolation[J]. *Application research of computers*, 2014, 31(7): 2187–2190.
- [24] ZEILER M D, FERGUS R. Visualizing and understanding convolutional networks[C]//Proceedings of the 13th European Conference on Computer Vision. Zurich, Switzerland, 2014: 818–833.
- [25] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural net-

works[C]//Proceedings of the 25th International Conference on Neural Information Processing Systems. Lake Tahoe, USA, 2012: 1097–1105.

- [26] JIA Yangqing, SHELHAMER E, DONAHUE J, et al. Caffe: convolutional architecture for fast feature embedding[C]//Proceedings of the 22nd ACM International Conference on Multimedia. Orlando, USA, 2014: 675–678.

#### 作者简介:



何锐波, 硕士研究生, 主要研究方向为人工智能和数字图像处理。



狄岚, 副教授, 中国人工智能学会粒计算与知识发现专业委员会委员, 主要研究方向为数字图像处理和计算机仿真。近年主持及参与国家级、省部级科研项目 7 项, 主持校级科研项目 4 项、企业合作项目近 20 项, 获得省级自然科学学术奖 1 次, 行业联合会技术奖 3 次。发表学术论文 50 余篇。



梁久祯, 教授, 博士, 中国计算机学会多媒体专业委员会委员, 江苏省人工智能学会理事, 主要研究方向为计算机视觉和数字图像处理。主持项目 10 余项, 曾获得浙江省青年英才奖。取得专利成果 57 项, 软件著作权 7 项。发表学术论文 160 余篇, 出版教材及专著 4 部。