

# 基于方向梯度直方图特征的车脸识别方法研究

曲爱妍<sup>1,2</sup>, 张正<sup>1</sup>, 梁颖红<sup>1</sup>, 吴秋玲<sup>1</sup>, 黄晓婷<sup>3</sup>

(1.金陵科技学院网络安全学院, 江苏 南京 211169;  
2.陆军工程大学指挥控制工程学院, 江苏 南京 210001;  
3.南京航空航天大学电子工程学院, 江苏 南京 210016)  
✉quaiyan@jit.edu.cn; zhangzheng@jit.edu.cn; liangyh@jit.edu.cn;  
Wuqiuling@jit.edu.cn; 934246903@qq.com



**摘要:** 随着道路监控系统的数字化和智能化发展, 车辆类型识别成为智能交通系统的研究重点之一。本文基于道路监控系统中的视频图像, 对车前脸进行粗定位, 并在粗定位车脸的基础上对车脸进行精确定位。对精确定位的车脸进行水平梯度水平投影和水平梯度垂直投影, 提取车标的方向梯度直方图——HOG特征, 由此提出了基于HOG特征描述的车脸特征点提取方法, 并采用支持向量机(SVM)对车标特征向量进行了分类识别。实验结果表明, 只有选择适当的HOG特征参数, 才能提高车型识别率。

**关键词:** 车脸识别; 视频识别; 方向梯度直方图; 支持向量机

**中图分类号:** TP311.5 **文献标识码:** A

## Research on Vehicle Face Recognition Method based on Histogram of Oriented Gradients Features

QU Aiyan<sup>1,2</sup>, ZHANG Zheng<sup>1</sup>, LIANG Yinghong<sup>1</sup>, WU Qiuling<sup>1</sup>, HUANG Xiaoting<sup>3</sup>

(1.College of Network Security, Jinling Institute of Technology, Nanjing 211169, China;  
2.College of Command and Control Engineering, Army Engineering University, Nanjing 210001, China;  
3.College of Electronic Information Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China)  
✉quaiyan@jit.edu.cn; zhangzheng@jit.edu.cn; liangyh@jit.edu.cn;  
Wuqiuling@jit.edu.cn; 934246903@qq.com

**Abstract:** With the digital and intelligent development of road monitoring system, vehicle type recognition has become one of the research focuses of intelligent traffic systems. Based on video images of the road monitoring system, this paper proposes to roughly locate the front face of the car, and then accurately locate the face of the car on the basis of the rough positioning. Based on the horizontal gradient horizontal projection and the horizontal gradient vertical projection of the accurately located front face, Histogram of Oriented Gradients (HOG) feature of the vehicle logo is extracted. Thus, a car face feature point extraction method based on the HOG feature description is proposed, and the vehicle logo feature vector is classified and recognized by Support Vector Machine (SVM). Experimental results show that only by selecting proper HOG feature parameters can the vehicle recognition rate be improved.

**Keywords:** car face recognition; video recognition; HOG; SVM

## 1 引言(Introduction)

随着生活水平的提高, 汽车越来越成为我们生活中不

可缺少的交通工具, 汽车数量的剧增促使智能化交通系统(Intelligent Traffic System, ITS)<sup>[1]</sup>也迅速发展。智能交通系

统中重要的部分有车型、车牌、车标识别等。车型识别方法主要分为两种，一种是基于车辆本身固有参数进行识别的方法，主要有环形感应线圈法、射频识别法、激光识别法<sup>[2]</sup>等；另外一种是基于视频技术的CCD(Charge Coupled Device)视频识别法<sup>[3]</sup>，在智能交通系统中越来越常见，主要基于实时拍摄的目标公路信息视频录像进行识别。ITS探索的重点是基于图像分析的车型识别技术，但目前在技术环节上相应地比较单薄。现今在国内外基于图像分析的车型识别技术研究方向主要分为模板匹配法<sup>[4]</sup>、神经网络识别法<sup>[5]</sup>以及支持向量机(Support Vector Machine, SVM)识别法<sup>[6-7]</sup>。本文主要研究SVM的车型识别方法。

## 2 车辆的车脸定位技术 (Vehicle face positioning technology)

### 2.1 车脸的粗定位技术

粗定位车脸<sup>[8]</sup>主要使用开运算和闭运算。开运算主要用于消除图像中物体的外部噪声，闭运算主要用于增强图像中物体之间的连接。实验主要提取运动车辆的车头部分，其中车脸定义为车灯左右边沿和上边沿，以及车头下边沿的矩形框部分。采用开运算能够将车辆的引擎盖跟车窗部分分离。车脸的开运算提取结果如图1所示。

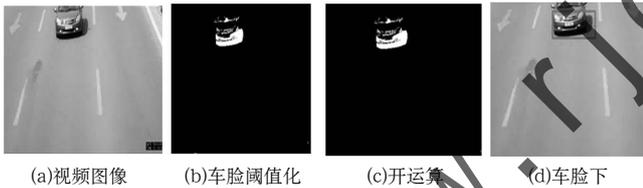


图1 车脸的开运算提取结果

Fig.1 Extraction results of vehicle face operation

由于背景的颜色、光照对车辆前景图像的影响，经过阈值化处理后的像素分布有很大的差异，例如，车辆在光照下的阴影会被当成车辆的一部分被检测出来。因此，开运算提取结果只能粗定位车脸的大致区域。

### 2.2 车脸的精确定位技术

车脸的精确定位<sup>[9]</sup>是在粗定位车脸的基础上，提取图像的水平梯度值，并将其在水平方向上和竖直方向上进行投影。图2第一行图像为原始图像  $I(i, j)$ ，图2第二行图像为水平梯度图像  $G_x$ ，图2第三行图像为  $G_x$  经过式(1)的阈值化处理得到二值结果  $G_{x'}$ 。

$$G_{x'}(i, j) = \begin{cases} 255 & G_x(i, j) \geq th \\ 0 & G_x(i, j) < th \end{cases} \quad (1)$$

其中， $th$  为阈值。

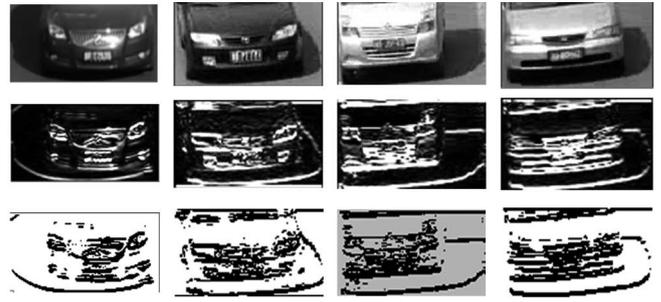


图2 不同车脸在不同背景下的水平梯度及其二值结果图

Fig.2 Horizontal gradient of different vehicle faces in different backgrounds and its binary results

#### (1)水平梯度水平投影

基于二值化水平梯度图像  $G_{x'}(i, j)$ ，灰度值沿水平方向上将进行叠加，得到水平方向上的投影  $Y(j)$ ：

$$Y(j) = \sum_{i=1}^m G_{x'}(i, j) \quad (2)$$

水平投影的均值为  $persum\_y$ ：

$$persum\_y = \frac{1}{n} \sum_{i=0}^n Y(j) \quad (3)$$

为了均衡误差，连续取三次水平投影的均值为  $mean\_y$ ：

$$mean\_y = \frac{1}{3}(Y(j) + Y(j+1) + Y(j+2)) \quad (4)$$

车脸的上边界定义为： $mean\_y > persum\_y$ ，其中  $Y(j)$  从图像的第一行至图像2/3高度处的值进行计算。

车脸的下边界定义为： $mean\_x \leq persum\_x$ ，其中  $Y(j)$  从图像的1/2高度处至最后一行进行计算。

#### (2)水平梯度垂直投影

在二值化水平梯度图像  $G_{x'}(i, j)$  的基础上，沿垂直方向上将灰度值进行叠加，得到垂直方向上的投影  $X(i)$ ：

$$X(i) = \sum_{j=1}^n G_{x'}(i, j) \quad (5)$$

垂直投影的均值为  $persum\_x$ ：

$$persum\_x = \frac{1}{m} \sum_{i=1}^m X(i) \quad (6)$$

由于在垂直方向上进行车脸的精确定位时，需要对车脸左、右这两个边界条件进行区分，连续取三次垂直投影的均值为  $mean\_x$ ：

$$mean\_x = \frac{1}{3}(X(i) + X(i+1) + X(i+2)) \quad (7)$$

从图像  $X(i)$  第一列开始，扫描到图像的1/4宽度处，当  $mean\_x > persum\_x$  时，定义其为车脸左边界；从图像1/2宽度处开始，扫描至最后一列，当满足  $mean\_x \leq persum_x$  时，定义其为车脸右边界。

车脸定位结果如图3所示。不同车脸在不同背景下的水平和垂直投影如图4所示。图4中第二行图像是车脸的垂直投影

图，第三行图像是车脸的水平投影图。第二行图像和第三行图像中的白色横线和白色竖线分别代表水平投影的均值和垂直投影的均值。

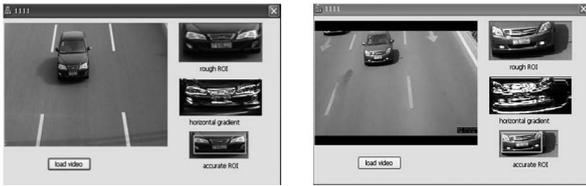


图3 车脸的精确定位图

Fig.3 Precise location map of vehicle face



图4 不同车脸在不同背景下的车脸精确定位图

Fig.4 Precise location map of different vehicle faces in different backgrounds

### 3 基于HOG+SVM的车脸识别(Vehicle face recognition based on HOG+SVM)

#### 3.1 提取车脸的HOG特征

方向梯度直方图(Histogram of Oriented Gradient, HOG)特征<sup>[10]</sup>是一种描述图像处理中物体检测的数学特征。它通过计算和统计图像局部区域的方向梯度直方图来构成特征。它的主要原理是基于统计图像中局部目标的表象,以及形状的梯度或边缘的方向密度来提取特征的。本文图像默认为64×64像素, HOG特征提取的方法流程如图5所示。

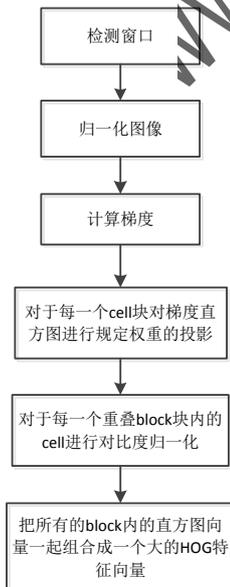


图5 HOG特征提取的方法流程

Fig.5 Flow chat of HOG feature extraction method

(1)将选取的车脸图像灰度化成为一个(x,y)的二维图像。

(2)对输入图像利用gamma校正法进行颜色空间的归一化处理。归一化可以对图像的对比度进行调整,也降低了阴影和光照对图像造成的影响,同时噪音的干扰也可以被抑制。归一化公式为:

$$I(x,y) = I(x,y)^{gamma} \tag{8}$$

式(8)中gamma取值1/2。

(3)计算图像中每个像素点的梯度(包括水平方向和垂直方向),关键是轮廓的信息,与此同时弱化光照的干扰。计算图像中像素点(x,y)的水平梯度值与垂直梯度值如下所示。

水平梯度值:

$$G_x(x,y) = H(x+1,y) - H(x-1,y) \tag{9}$$

垂直梯度值:

$$G_y(x,y) = H(x,y+1) - H(x,y-1) \tag{10}$$

式(9)与式(10)中,  $H(x,y)$ 表示输入图像中像素点(x,y)的像素值。像素点(x,y)的梯度幅值和梯度方向分别如式(11)、式(12)所示:

$$G(x,y) = \sqrt{G_x(x,y)^2 + G_y(x,y)^2} \tag{11}$$

$$\alpha(x,y) = \tan^{-1} \left( \frac{G_y(x,y)}{G_x(x,y)} \right) \tag{12}$$

(4)将block分成多个单元格(cell),每个cell占8×8像素大小,用每个block块中每个cell中的每个像素构建方向梯度直方图。

(5)针对每个block中的cell进行方向梯度直方图的统计如下:分为9个方向角,初始方向角为0°至180°。每个cell经过9个方向的计算后得到9个特征值。假设block为16×16像素大小,则每个block可以分为4个cell,将block内4个cell的9个特征值合并成一个特征向量,就得到该block的36个特征值。

(6)将整个车脸图像用block扫描,  $g$ 表示一个block的特征向量维数,  $h$ 表示车脸图像的block个数,则车脸图像的HOG特征用  $g \times h$  维特征向量表示。

#### 3.2 支持向量机(SVM)

SVM主要用于小样本、非线性、高维以及提高泛化性能的模式识别方法,它是一种基于“结构风险最小化(Structural Risk Minimization, SRM)”原则的分类方法<sup>[11]</sup>。支持向量机最初用于解决两类训练样本的分类问题,二分类线性最优分

剖面示意图如图6所示。

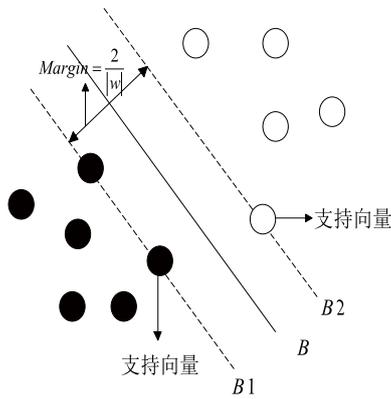


图6 二分类最优分割面示意图

Fig.6 Schematic diagram of optimal partition surface for two classification

图6中，两类样本的最优线性分割线为B，样本中离B距离最近，且与B平行的两类样本分别用B1、B2表示。B1和B2之间的距离被称为分类间隔。推广于高维空间时，最优分割线变成最优分割面，则最优分割面为：

$$g(x) = w \cdot x + b = 0 \quad (13)$$

式(13)中，w为权重向量，x为输入样本向量，b为偏移量。在 $|g(x)| \geq 1$ 条件约束下，式(13)可变为分类判别函数，则分类判别的规则为：

$$\begin{aligned} w \cdot x_i + b &\geq 1, \quad y_i = 1 \\ w \cdot x_i + b &< 1, \quad y_i = -1 \end{aligned} \quad (14)$$

式(14)中， $x_i$ 表示的是输入样本向量中的一个样本， $y_i$ 表示的是输入样本的类别。在 $|g(x)| \geq 1$ 条件约束下，式(13)、式(14)可以表示为：

$$y_i [(w \cdot x_i) + b] - 1 \geq 0, \quad i = 1, 2, \dots, n \quad (15)$$

分类间隔B1、B2之间的距离为 $2/\|w\|$ ，其与 $\|w\|$ 的大小成反比。因此，在 $\|w\|^2$ 最小的情况下，式(15)所得到的最小值就是最优分类面。

$$\min \frac{1}{2} \|w\|^2 \quad \text{s.t.} \quad y_i [(w \cdot x_i) + b] \geq 1, \quad i = 1, 2, \dots, n \quad (16)$$

针对式(16)，构造拉格朗日函数(Lagrange)式(17)求解最小值：

$$L(w, b, \alpha) = \frac{1}{2} (w \cdot w^T) - \sum_{i=1}^n \alpha_i \{y_i [(w \cdot x_i) + b] - 1\} \quad (17)$$

针对式(17)，利用相应的极值定理，分别将w、b求偏导数并使其等于0，如式(18)、式(19)所示：

$$\frac{\partial L(w, b, \alpha)}{\partial w} = w - \sum_{i=1}^n \alpha_i y_i x_i = 0 \quad (18)$$

$$\frac{\partial L(w, b, \alpha)}{\partial b} = \sum_{i=1}^n \alpha_i y_i = 0 \quad (19)$$

将计算结果代入拉格朗日函数 $L(w, b, \alpha)$ ，式(16)可以变为约束条件式(20)情况下求解式(21) $Q(\alpha)$ 的极大值问题。

$$\sum_{i=1}^n y_i \alpha_i = 0, \quad \alpha_i \geq 0, \quad i = 1, 2, \dots, n \quad (20)$$

$$Q(\alpha) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j (x_i \cdot x_j) \quad (21)$$

由于绝大多数样本都满足 $|g(x)| \geq 1$ ，因此 $\alpha_i$ 必须为零。若使 $\alpha_i > 0$ ，只有满足 $[y_i (w \cdot x_i) + b] - 1 = 0$ 的样本，这样的样本满足支持向量，此时 $\alpha_i$ 标记为 $\alpha_i^*$ ，那么

$$w^* = \sum_{i=1}^n \alpha_i^* y_i x_i \quad (22)$$

则最优分类函数最终可表达为式(23)：

$$f(x) = \text{sgn} \left\{ (w^* \cdot x) + b^* \right\} = \text{sgn} \left\{ \sum_{i=1}^n \alpha_i^* y_i (x_i \cdot x) + b^* \right\} \quad (23)$$

式(23)中，支持向量所对应的 $\alpha_i > 0$ ，则 $f(x)$ 的求和过程仅与支持向量相关。将满足支持向量的任意值代入可求出 $b^*$ 。

上述是支持向量机线性可分的场景，但实际应用中很多都是非线性可分的。这样，需要将原样本中的特征数据非线性映射到一个特征空间z中，则特征空间z中存在一个广义的最优分界面，用 $z = \varphi(x)$ 来表示。

支持向量机的非线性多分类问题如图7所示。

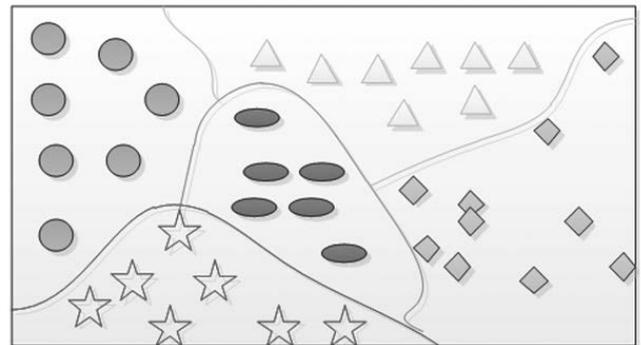


图7 支持向量机的非线性多分类问题

Fig.7 Nonlinear multi-classification problem of Support Vector Machine

在式(15)中加入松弛变量 $\xi_i$ ，且 $\xi_i \geq 0 (i = 1, 2, \dots, n)$ ，则式(15)变为：

$$y_i [(w \cdot x_i) + b] \geq 1 - \xi_i \quad (24)$$

其中， $\xi_i$ 表示的是 $x_i$ 允许的偏移量， $\xi_i$ 越小越好。因此，在求解广义最优分类面时，希望 $\xi_i$ 总和最小。所以，求解最优分类面的问题就转化成求解式(25)最小值的问题：

$$\phi(w, \xi) = \frac{1}{2}(w \cdot w) + C \left( \sum_{i=1}^n \xi_i \right) \quad (25)$$

其中, C 为常量。

由Hilbert Schmidt原理可知, 希望找到核函数  $K(x, x')$ , 该核函数  $K(x_i, x_j) = (z_i^* z_j)$ , 则式(26)成为目标函数。

$$Q(\alpha) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j K(x_i, x_j) \quad (26)$$

其中,  $z_i^* z_j$ 是内积函数, 表示为:

$$z_i^* z_j = \sum_{k=1}^m \phi_k(x_i) \phi_k(x_j) \quad (27)$$

则最优分类面判别函数为:

$$f(x) = \text{sgn} \left\{ \sum_{i=1}^n \alpha_i^* y_i K(x_i, x) + b^* \right\} \quad (28)$$

从以上分析过程可以看出, 核函数实现了SVM从非线性可分空间到新的线性可分特征空间的非线性变换。根据不同的算法, 核函数可分为线性、多项式、前馈神经网络以及径向基等。

### 3.3 车脸识别系统设计

本文车脸识别系统前端是基于DM8168开发板开发的, 主要实现交通监控视频中运动目标的检测<sup>[12]</sup>。系统的框架图如图8所示。对运动车辆进行检测, 将检测到的图像进行灰化处理, 然后针对车辆的前面板和车牌位置进行定位。在车牌定位的基础上, 针对车辆的前脸进行车型识别; 在形态学与模板匹配进行车标定位的基础上, 对车标进行SVM分类。最后, 对比车脸识别与车标识别的结果, 确定输出车型。

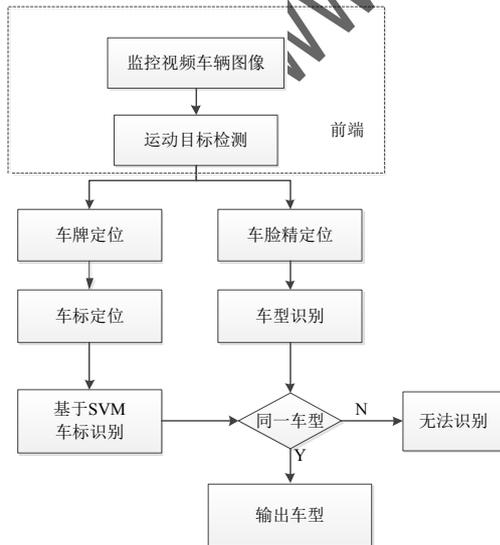


图8 车脸识别系统框架图

Fig.8 Frame diagram of vehicle face recognition system

车辆视频中车脸的识别结果如图9所示。



图9 车辆视频的识别结果

Fig.9 Recognition results of vehicle video

由车辆视频的识别结果可以看出, 通过视频图像可实现车脸定位、车标定位, 最终实现车型的识别和车牌的识别。

## 4 实验与结果分析(Experiment and result analysis)

### 4.1 车前脸库选取

本文选取从监控视频中定位的车脸图像, 对此进行归一化大小为128×64像素, 如图10所示。



图10 车脸图像库

Fig.10 Library of vehicle face images

### 4.2 HOG特征参数分析

HOG特征主要由block和cell的水平梯度和垂直梯度向量组成, 所以图像分割不同的网格所对应的block和cell的大小直接影响车脸识别的结果。本次实验选取6种车型: 别克、本田、福特、上海大众、雪佛兰、尼桑, 36种车系, 提取225张车前脸的图像。不同网格大小车型识别的误差率如表1所示。

表1 不同网格大小车型识别的误差率

Tab.1 Error rate of vehicle recognition with different grid sizes

block \ cell	cell			
	16×16	8×8	4×4	2×2
2×1	12%	10%	9%	10.5%
4×2	9%	8%	6%	11%
8×4	10%	9%	8%	8.5%
4×4	7%	6%	5%	6.5%
2×2	10%	9%	7.5%	9.5%

从实验结果可以看出：当block为 $4 \times 4$ 、cell为 $4 \times 4$ 大小的像素时，识别车型的误差率最小。当cell变大时，车型识别的误差率也随之变大；当cell变小时，车型的误差率变大，且性能变差。因此，需要分割合适的block和cell大小，过大或过小的分割都会使性能变差。

顾伟等<sup>[13]</sup>的实验采用了block重叠形式，由于重叠block会增加特征向量的维度和计算量，因此本实验采用的是block不重叠的方式进行车型识别。图11显示了不重叠方式对车型的识别率。

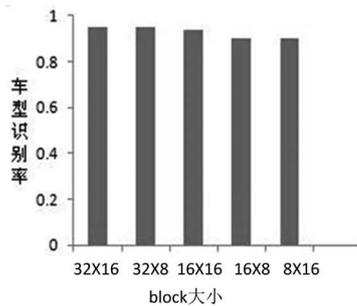


图11 不同block大小不重叠方式车型的识别率

Fig.11 Vehicle model recognition rate of non-overlapping blocks with different sizes

## 5 结论(Conclusion)

本文针对道路监控系统中的汽车视频图像，运用HOG特征对车前脸区域的图像进行特征提取，并在SVM中进行训练识别。本文设计实验分析对比了不同模式下车型正确识别率和不同重叠方式对车型识别的影响，结果表明选择合适的分割block非常重要，适当的block识别率才最高，过大或是过小识别率都会变差。以block不重叠计算HOG特征向量，可以减少特征向量的维度和计算量，能够有效应用于车型的识别过程中。

## 参考文献(References)

[1] 胡亮.用于车辆辅助驾驶系统中的前车追踪算法研究[D].北京:北京工业大学,2017.

[2] DECKERT V, DECKERT-GAUDING T, CIALLAMAY D, et al. Laser spectroscopic technique for direct identification of a single virus I: FASTER CARS[J]. Proceedings of the National Academy of Sciences of the United States of America, 2020, 117(45):27820-27824.

[3] 李卫江.基于线阵CCD成像交通信息采集和检测技术的研究[J].长安大学学报,2008,10(3):22-24.

[4] 唐红强.基于GRM模板匹配算法的车型和车系识别[D].成都:电子科技大学,2014.

[5] 晏世武,罗金良,严庆.基于卷积神经网络车型分类的研究[J].智能计算机与应用,2020,10(01):67-70.

[6] 黄晓婷.基于DM8168双核DSP的车型识别系统[D].南京:南京航空航天大学,2015.

[7] 黄灿.基于局部特征的汽车识别[J].微型电脑应用,2010,26(8):51-56.

[8] 欧先锋,晏鹏程,王汉谱,等.基于深度帧差卷积神经网络的运动目标检测研究方法[J].电子学报,2020,48(12):2384-2393.

[9] 张文鹏.复杂条件下多通道SAR运动目标检测与参数估计方法研究[D].长沙:国防科技大学,2018.

[10] ERAZO-AUX J, LOAIZA-CORREA H, RESTREPO-GIRON A D. Histograms of oriented gradients for automatic detection of defective regions in thermograms[J]. Applied Optics, 2019, 58(13):3620-3629.

[11] 刘渭滨,邹智元,邢薇薇.模式分类中的特征融合方法[J].北京邮电大学学报,2017,40(04):1-8.

[12] 张红兵.基于TMS320DM8127的车型识别系统的研究与实现[D].南京:南京航空航天大学,2015.

[13] 顾伟,李菲菲,陈虬.基于多特征融合的行人检测方法[J].电子科技,2021,34(05):29-34.

## 作者简介:

曲爱妍(1976-),女,硕士,高级工程师.研究领域:图像处理,人工智能,数据安全,数据分析.

张正(1973-),男,本科,研究员.研究领域:专用网络安全,加解密技术,无线通信.本文通讯作者.

梁颖红(1970-),女,博士,教授.研究领域:自然语言处理,智能对话系统.

吴秋玲(1976-),女,博士,副教授.研究领域:信息隐藏技术,无线网络与信息安全.

黄晓婷(1990-),女,硕士,工程师.研究领域:图像处理,DSP.