

交通标志识别特征提取研究综述

薛 搏, 李 威, 宋海玉, 方安琪, 彭京涛, 王鹏杰, 郭宏烨

(大连民族大学计算机科学与工程学院, 辽宁 大连 116605)

摘 要: 交通标志识别(TSR)是智能交通系统(ITS)的一个重要研究方向, 而特征提取是交通标志识别研究中的重点。聚焦交通标志识别的特征提取, 综述了常见的人工特征(颜色直方图、尺度不变特征变换特征、局部二值模式特征、方向梯度直方图特征、Haar-like 特征、Gabor 小波特征、Canny 特征等)和深度特征(提取自 AlexNet, VGG16, Inception 等), 并在同一数据集(GTSRB)上提取多种特征, 采用相同分类器, 通过相同评价指标体系进行定量比较与分析, 并以图表方式, 针对不同特征和不同交通标志类别, 进行直观的性能比较研究, 以期对交通标志识别时特征向量的选择和深入研究提供参考。

关 键 词: TSR; ITS; 特征提取; 人工特征; 深度特征

中图分类号: TP 391

DOI: 10.11996/JGJ.2095-302X.2019061024

文献标识码: A

文 章 编 号: 2095-302X(2019)06-1024-08

Review on Feature Extraction of Traffic Sign Recognition

XUE Bo, LI Wei, SONG Hai-yu, FANG An-qi, PENG Jing-tao,
WANG Peng-jie, GUO Hong-ye

(College of Computer Science and Engineering, Dalian Minzu University, Dalian Liaoning 116605, China)

Abstract: Traffic sign recognition (TSR) is an important research direction of intelligent transportation system (ITS). Feature extraction is the key point of traffic sign recognition research. This paper focuses on feature extraction of traffic sign recognition and summarizes common manual features and depth features. Manual features include color histogram, scale invariant feature transformation feature, local binary pattern feature, directional gradient histogram feature, Haar-like feature, Gabor wavelet feature, Canny feature, etc. Depth features are extracted from AlexNet, VGG16, Inception, etc. Various features are extracted from the same data set (GTSRB). Various features are compared and analyzed quantitatively by using the same classifier and the same evaluation index system. This paper makes an intuitive comparative research of performance for different features and different types of traffic signs by means of charts and graphs, aiming at providing a reference for the selection of feature vectors and for the further research of traffic sign recognition.

Keywords: TSR; ITS; feature extraction; manual features; depth features

收稿日期: 2019-07-31; 定稿日期: 2019-09-14

基金项目: 国家自然科学基金项目(6130089); 辽宁省自然科学基金项目(201602199); 辽宁省高等学校创新人才支持计划(LR2016071); 国家级大学生创新训练项目(201912026022)

第一作者: 薛 搏(1999-), 男, 辽宁大连人, 本科生。主要研究方向为图像理解、机器学习。E-mail: 307305049@qq.com

通信作者: 宋海玉(1971-), 男, 河南安阳人, 副教授, 博士。主要研究方向为图像理解、计算机视觉、机器学习。E-mail: shy@dlmu.edu.cn

交通标志识别(traffic sign recognition, TSR)是无人驾驶汽车和智能汽车等领域的基础性技术,也是自动辅助驾驶系统的关键技术^[1]。交通标志的检测与识别最早开始于20世纪70年代,但发展比较缓慢,到20世纪80年代时,日本开始了道路交通标志检测和识别的研究,并在1987年研发出了针对限速TSR的系统^[2],而到2010年左右,一些汽车已经安装了交通标志检测和识别系统。近年来,受智能汽车和无人驾驶汽车技术的推动,TSR不仅成为交通学科和计算机学科的研究热点^[3],也是汽车产业界的研究应用热点。TSR系统主要分为2个阶段:①交通标志检测阶段,主要利用交通标志的颜色和形状等信息的特殊性,从原始图像中提取出感兴趣区域(region of interest, ROI);②交通标志识别阶段,TSR本质上是图像分类问题,包括特征提取和分类2部分,其中,图像特征提取是重点和难点。

很多学者围绕TSR的特征提取开展了大量工作,所完成工作的主要贡献是特征提取,但由于采用不同数据集(大部分是研究者自己采集的数据集)、不同的分类方法和不同评价体系,所采用特征向量的贡献无法直接量化对比,不利于该领域后续的进一步研究。本文聚焦于TSR领域的特征提取方法,首先介绍该领域常用方法,然后,在同一个数据集上选择有代表性的特征进行提取,采用相同的评价指标进行对比,以便为从事该领域研究的相关学者提供一个清晰的对比,为该领域特征向量的选择提供依据。

1 国内外现状

1.1 数据集

长期以来,制约TSR研究的一个问题是缺少领域内广泛接受的通用数据集,很多成果所完成的实验是在研究者自己搜集的数据集上验证完成。仅有的几个数据集包括德国的GTSRB数据集^[4-5]、瑞典的STS数据集^[6]、比利时的KUL数据集^[7]和法国的Stereopolis数据集^[8]。其中STS数据集包括7种类别交通标志、4 000幅图像样本,KUL包括100多种交通标志、9 000幅图像样本,Stereopolis只有10个类别、847幅图像样本。以上3种数据集规模较小、交通标志种类少,且图像场景的多样性不够。GTSRB是2011年在国际会议IJCNN2011上由德国学术界和产业界联合举办的交通标志检测与识别大赛所构建的TSR数据集^[9],该数据集由5万多幅

图像样本组成,包括了43种交通标志,包含大量低分辨率、不同光照强度、部分遮挡、倾斜、运动模糊等各种不利条件下的图像,能够反映交通真实场景,且识别难度较大^[10]。

Tsinghua-Tencent 100 K数据集是清华大学和腾讯公司于2015年联合发布的一个交通标志检测数据集,也是国内仅有的公开交通标志数据集。该数据集总共包含10万幅图像,其中大约有1万幅图像包含有交通标志,图像分辨率高^[11]。Tsinghua-Tencent 100 K数据集出现较晚,目前采用该数据集的实验结果偏少,该数据集清晰,实现难度要低于GTSRB数据集^[12]。

图1显示的GTSRB数据集中的图像来源于真实场景,经检测并保存为独立的交通标志图像。该数据集中图像的质量区别很大,而且大部分图像清晰度很差,以至于很多图像即便交通专家也无法识别,其是TSR技术应用的真实场景,准确识别难度较大,本文实验部分会给出采用不同特征的识别效果。



图1 GTSRB数据集中的交通标志样例

尽管2011年之后,GTSRB已经成为国内外TSR领域应用最广泛的TSR数据集,但由于多种原因,该领域已发表成果所使用实验数据集依然呈现多样性,或采用较小规模数据集,或研究者自己搜集数据集,因此,不同方法之间很难直接比较性能。基于GTSRB数据种类多、样本规模大、包含各种复杂场景、识别难度大等特点,本文所完成的实验均采用该数据集完成。

1.2 评价指标

现有文献中,TSR评价指标有3种评价指标体系:①采用查准率(Precision)、查全率(Recall)、F1-measure和 N^{+} ^[13];②采用均值平均精度(mean average precision, MAP);③采用精度(Accuracy)。因为TSR本质上是图像分类问题,而以上3种评价指标均为分类评价指标,所以,理论上这3种评价指标均可采用。但第①种评价指标更适合于多标签图像分类、图像标注和图像检索等领域;

第②种评价指标较多应用于信息检索领域, 定量分析了同一个对象多个标签的相关性的次序; 第③种评价指标计算的是测试样本的正确分类率(correct classification rate, CCR), 适用于单标签对象分类, 如深度学习网络模型默认的功能是单标签对象分类问题, 因此, 目前的深度学习模型所提供的性能评价指标往往都是精度(accuracy)。

TSR 在检测之后, 每个识别对象都是单标签对象, TSR 问题属于单标签多类分类问题, 因此, 尽管前 2 种也可以分析 TSR 性能, 但更能准确刻画 TSR 性能的指标是精度(accuracy)。IJCNN2011 所举办的交通标志数据集识别竞赛中所采用的评价指标为样本的 CCR, 即分类精度(accuracy)。但由于目前文献更多选用第一种评价指标体系, 本文所完成的实验同时给出这 2 种评价指标结果。

1.3 分类模型

影响 TSR 效果的除了特征向量外, 就是分类/识别模型, 但不同方法往往采用不同的分类技术, 甚至有的方法采用多种分类技术分阶段完成。如支持向量机(support vector machine, SVM)^[14]、AdaBoost^[15]、随机森林(random forest, RF)^[16]等。TSR 性能是对特征向量和分类模型的综合性能评价, 分类方法不同, 则难以量化特征向量的贡献。为定量分析特征向量对 TSR 的贡献, 对于任何特征向量, 本文均采用完全相同的分类模型。

本文统一采用模型复杂度较低的 K 近邻(K nearest neighbor, KNN)分类方法, 旨在突出特征向量对最终性能的影响, 但图像相似度计算可以选用 L1 距离(街区距离)、L2 距离(欧式距离)或余弦距离, 对于每种特征向量均选择效果最优的距离公式。

2 特征提取综述

TSR 是在交通标志检测后得到的交通标志感兴趣区域 ROI 的分布和定位的基础上, 通过识别算法来进一步确定标志的具体类型。识别过程一般包括 2 部分: ①特征提取, 即提取交通标志感兴趣区域的特征信息; ②交通标志分类, 即将交通标志感兴趣区域的类别信息进行判断。在自然场景中, 光照、天气、背景干扰、运动模糊、遮挡和标志不全等问题, 给 TSR 系统的研究带来了非常大的挑战。

近年来研究者提出了很多 TSR 方法, 传统的方法主要是采用人工特征和机器学习相结合的方法。

采用的人工特征有各种颜色特征^[17], 尺度不变特征变换特征(scale-invariant feature transform, SIFT)^[18]、方向梯度直方图特征(histogram of oriented gradient, HOG)^[19-20]、Haar-like 特征^[21]、Gabor 小波特征^[22]、不变矩^[23-24]等。

尽管 TSR 领域已经尝试了各种人工特征, 但大多数成果是在改进特征的同时也改进了分类模型^[18,20,24-25], 或多种特征融合^[18-19,26], 或多种特征级联^[27-28]。

以上特征都是人工提取的底层特征直接表示图像。2003 年后, 在计算机视觉领域, 在底层特征基础之上视觉词袋(bags of words, BOW)稀疏编码方法给计算机视觉领域各种应用带来了不同程度的性能提升, 此后近 10 年里, 以 BOW 为代表的稀疏编码表示方法表现出了明显的优越性, 直到 2012 年深度学习的出现之前基于 SIFT 的 BOW 模型一直被视为最优秀的特征表示方法^[29]。也有很多学者尝试把 BOW 模型用于 TSR^[30]。还有一些学者把塔式词袋直方图(pyramid histogram of visual words, PHOW)、塔式边缘方向梯度直方图(pyramid histogram of oriented gradient, PHOG)等与其他底层特征相结合用于交通标志识别^[31]。

TSR 领域是较早使用深度学习技术的领域之一, 虽然 2012 年提出的通用图像分类模型 AlexNet 被公认是最早取得成功的深度学习模型, 但事实上, 早在一年前的 IJCNN2011 的 TSR 竞赛中, 很多排名靠前的方法均采用了卷积神经网络方法, 其中就包括因深度学习而获得 2018 年图灵奖的 LeCun 及其团队的成果, 其提出了多尺度卷积神经网络用于识别交通标志并取得了很大成功^[32], 但目前该领域所提出或采用的大多数深度学习模型往往都是其他特征向量或复杂的分类模型相结合^[32-33]。

3 本文特征提取实验与分析

3.1 基于颜色特征的实验

区域选择方法包括全局特征、区域特征、局部特征, 提取内容包括颜色信息、纹理信息和形状信息, 区域和局部特征所提取的原始特征往往还要通过稀疏编码方法表示。本文淡化各种特征向量的原理和提取方法, 而聚焦于在相同数据集上不同特征向量在 TSR 效果上的对比。

不论国内还是国外的交通标志, 一般都具有特定的醒目的颜色, 因此基于颜色的特征是交通

标志的重要特征, 也被很多研究者所使用。彩色图像中颜色信息最丰富、提取方法最简单, 而且, 由于颜色直方图等特征具有缩放、旋转和平移不变性的特点, 因此, 成为计算机视觉领域最常用的特征之一。颜色信息可以在不同的颜色空间中表示, 如 RGB 空间、HSV 空间和 Lab 空间, 其中, RGB 更加适合于硬件设备表示场合, 而 HSV 和 Lab 更适合于人的视觉感知描述场合^[34]。由于颜色直方图对观察视角和缩放变换有很强的健壮性^[35], 已成为计算机视觉领域最常用的视觉特征向量之一。在 RGB、HSV 和 Lab 3 种颜色空间上提取颜色直方图, 每个通道提取 16 维直方图, 每种颜色空间形成 48 维的颜色直方图。此外, 提取 3 种颜色空间每个通道的均值和方差, 每种颜色空间形成 6 维的特征向量。

SWAIN 和 BALLARD^[35]于 1991 年首次把颜色直方图用于图像检索中, 认为颜色直方图对观察视角变化和缩放也有很强的健壮性。表 1 的实验结果表明, 尽管颜色直方图简单且有很好的健壮性, 但其特征往往受到了成像环境的影响, 如光照变化等, 在 TSR 中颜色直方图效果并不理想。而全局颜色统计量由于统计信息过于简单, 其效果明显低于颜色特征向量。

从表 1 的实验结果看出, 在对全局颜色特征向量进行提取的情况下, 采用 3 种颜色空间中的任意一种, 得到的各项评价指标的值的差别很小, 为了进一步验证颜色空间对识别结果的影响, 本文将交通标志中的 43 个子类别在 3 种颜色空间下分别进行颜色直方图特征的提取, 得到其 F1 值对应的柱状图如图 2 所示。

表 1 全局颜色特征向量实验结果

特征向量	向量维数	相似度计算	查准率(%)	查全率(%)	F1 值(%)	N ⁺	Accuracy (%)
RGB 直方图	48	街区距离	22.42	20.29	21.30	39	26.95
HSV 直方图	48	街区距离	24.08	21.27	22.59	41	27.03
Lab 直方图	48	街区距离	21.72	19.68	20.65	42	25.72
RGB 统计量	6	欧式距离	12.64	12.34	12.49	41	17.51
HSV 统计量	6	欧式距离	12.89	12.41	12.64	41	16.86
Lab 统计量	6	欧式距离	13.08	12.52	12.79	40	17.28

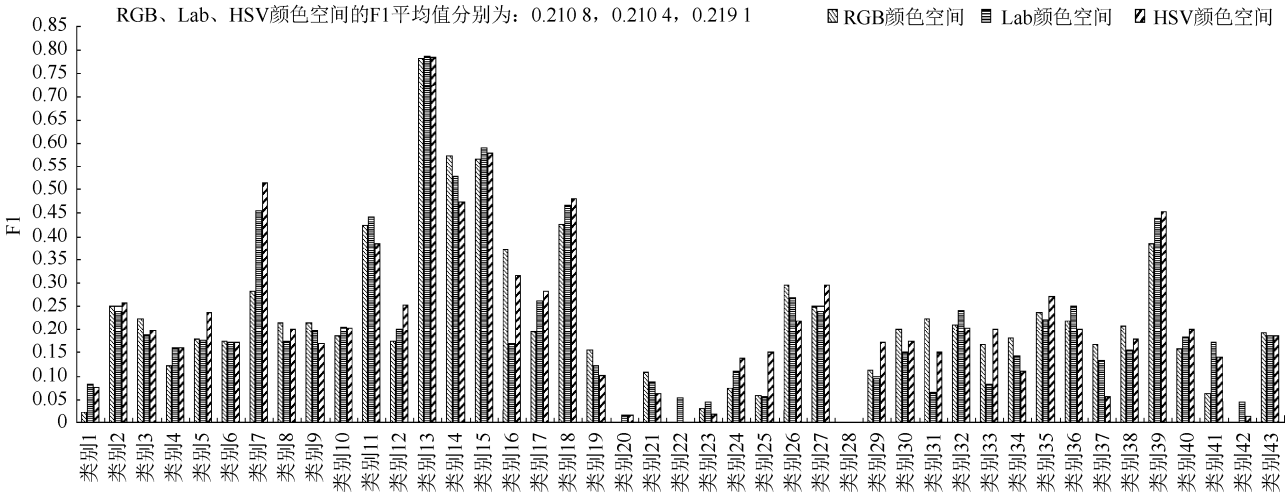


图 2 3 种不同颜色空间下颜色直方图特征向量的性能比较

图 2 表明, 在交通标志领域, 颜色直方图向量特征提取过程中, 不同颜色空间特征向量的总体性能差异很小, 在不同子类型上也没有明显的规律。因此, 颜色特征空间的选择不应该成为特征提取阶段重点考虑因素。

基于全局特征存在的不足, 在 TSR 中更多学者采用分块方法, 本文把图像等分为 16 块, 提取各

个图块特征向量, 然后拼接成整幅图像的特征向量, 所完成的 TSR 效果见表 2。表 2 中各特征提取方法与表 1 中对应方法相同, 区别在于表 2 为分块处理。无论直方图方法还是颜色统计量方法, 分块提取效果均优于全局特征提取方法, 特别是分块颜色统计量方法明显优于全局颜色统计量特征提取方法。

表 2 分块颜色特征向量实验结果

特征向量	向量维数	相似度计算	查准率(%)	查全率(%)	F1 值(%)	N ⁺	Accuracy (%)
RGB 直方图	768	街区距离	48.90	44.46	46.22	43	50.76
HSV 直方图	768	街区距离	42.66	40.59	41.60	43	46.83
Lab 直方图	768	街区距离	44.54	41.30	42.86	42	49.22
RGB 统计量	96	欧式距离	46.36	43.15	44.70	41	49.18
HSV 统计量	96	欧式距离	35.42	34.35	34.87	41	40.40
Lab 统计量	96	欧式距离	40.95	38.27	39.56	40	45.44

最早期的计算机视觉领域研究中,往往采用全局描述子作为图像特征表示,此后,针对存在的不足,相继提出了区域描述和局部描述等^[35]。当通过交通标志检测后,TSR 往往是单对象分类问题,因此,理论上采用全局描述子作为特征向量进行交通标志分类是可行的。但通过对比表 2 和表 1 不难发现,基于分块的区域特征描述方法效果远优于全局特征,主要是由于 GTSRB 在多种不同天气、速度、光照等实际场景中成像,成像效果受清晰度、遮挡等影响,因此,全局描述子的实际效果往往不如区域描述子效果。

3.2 基于纹理和形状特征的实验

纹理是在对象的表面表示、区域区分和对象识别等应用中最重要属性,因此,在计算机视觉领域,纹理是仅次于颜色的最常用的特征表示,最常用的纹理特征表示包括 Gabor 小波和 SIFT (scale-invariant feature transform)。Gabor 小波特征提取方法为,首先使用 12 个 Gabor 滤波器(滤波器因子在 3 个尺度、4 个不同方向变化)完成整幅图像的滤波,然后,把图像等分为 16 个图块,计算每个图块滤波结果的均值和方差,从而形成 24 维的 Gabor 小波特征向量。SIFT 特征提取方法为,首先把原始图像栅格化等分为 16 个图块,提取每个图块的中心点作为种子点,统计每个中心点 8 个方向的向量信息,从而每个图块提取出一个 128 维度特征向量,整幅图像的特征向量为 2 048 维。基于 BOW 方法在计算机视觉领域的巨大成功,在提取 SIFT 基础之上,采用 SIFT 向量聚类等方法构造视觉词典,从而把每幅图像表示成 BOW 所表示的稀疏特征向量,其效果见 BOW (SIFT)效果。

形状是描述对象最重要的特征,高质量的形状特征依赖于高质量的图像分割,受制于当前图像分割质量制约,无论基于边界的形状信息描述质量往往不高,因此,在计算机视觉领域往往采用基于图像分割后的区域特征的形状信息,本文比较了交通标志识别领域应用较多的 Canny 形状特征、HOG

特征、Hu 矩。Canny 形状特征,首先使用 Canny 算子进行边缘检测,然后,把图像 16 等分,计算每个图块内容边缘方向信息(角度步长为 45°,共 8 个角度),统计边缘方向每个生成一个 8 维的直方图,整幅图像形成 128 维的特征向量。HOG 特征采用 GTSRB 数据集附带的特征向量,维数为 1 568 维^[36]。Harr-like 特征采用 GTSRB 数据集附带的特征向量,维数为 11 584^[36]。不变矩(invariant moments)是一种高度浓缩的图像特征,具有平移、灰度、尺度、旋转不变性,Hu 矩是不变矩的代表性方法。其提取方法,首先把原始图像栅格化等分为 16 个图块,然后,计算每块的 7 阶矩,从而形成 112 维的特征向量。

表 3 中,SIFT 特征并未采用传统的高斯微分方法(difference of Gaussian, DoG)^[34]等方法检测显著点(saliency),而是选取各个图块中心点作为显著点,归功于 SIFT 特征具有缩放、平移和旋转不变性,而且具有光照和拍照视角的部分不变性等^[37-40],SIFT 描述子的效果明显优于其他方法。但与传统的计算机视觉领域应用不同的是,采用 BOW 编码的表示方法并未在 SIFT 基础上提升性能反而明显降低,究其原因在于每幅图像栅格化后图块个数过少,主要由于 BOW 方法表示的特征向量过于稀疏造成的。

基于 SIFT 特征的 TSR 的总体性能明显优于其他手工特征,图 3 进一步比较了 SIFT 和另外 2 种性能较好的手工特征 HOG 和 Haar-like 在各个子类别上的性能。如图 3 所示,在 GTSRB 交通标志数据集中,除第 22 类外,在其他 42 个类别中,SIFT 特征明显优于其他 2 种特征,说明 SIFT 特征对各种视觉特征均具有较强的描述能力和区分能力。另外,SIFT 特征的缩放、平移和旋转不变性,以及具有光照和视角的部分不变性,使其能够适应于各种复杂实际场景。因此,在诸多手工特征中,基于 SIFT 的特征应该是重点研究和应用对象。

表 3 纹理和形状特征向量实验结果

特征向量	向量维数	相似度计算	查准率(%)	查全率(%)	F1 值(%)	N ⁺	Accuracy (%)
SIFT	2048	欧式距离	94.35	92.36	93.35	43	94.16
BOW(SIFT)	500	欧式距离	70.80	65.43	68.01	43	70.55
Gabor	384	欧式距离	62.62	59.09	60.80	42	65.04
Canny	128	街区距离	67.51	63.92	65.67	43	68.80
HOG	1568	欧式距离	74.14	69.06	71.51	43	73.36
Harr-like	11584	余弦距离	75.45	71.22	73.27	43	74.38
Hu 矩	112	欧式距离	52.24	48.53	50.32	43	55.38

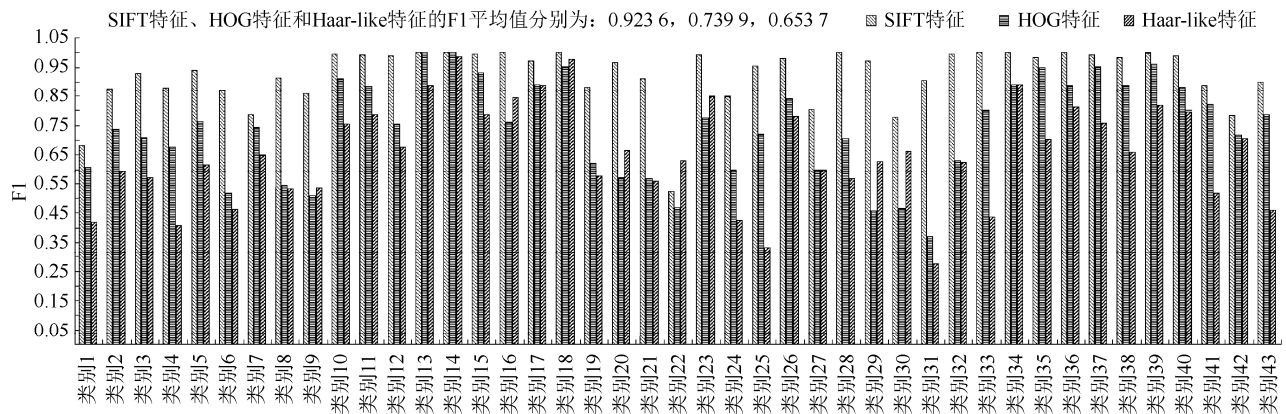


图 3 基于 SIFT、HOG 和 Haar-like 特征提取方法的性能比较

3.3 基于深度特征的实验

2012 年后, 基于深度学习的特征在计算机视觉领域的多种应用中取得突破性进展, 本文使用深度学习领域最有代表性的 AlexNet、VGG16 和 Inception 3 个模型完成交通标志的特征提取及识别。AlexNet 是第一个在通用数据集中取得显著效果的深度学习模型, VGG16 是目前为止图像分类等领域使用最频繁的模型, 尽管其模型复杂度并不高, 但实际效果往往好于很多复杂度高的深度学习模型, Inception 网络是一种更复杂的网络模型, 可以在同一层中学习不同大小的卷积特征。对于每种深度学习模型, 分别采用种不同方式完成特征提取和分类: ①端到端的方式, 通过训练好的深度学习模型直接得到图像的分类结果(网络模型给出 accuracy 性能指标), 即端到端分类; ②提取训练好的深度学习模型的全连接特征, 而不是直接得到分类结果, 并采用 KNN 方法完成交通标志候选区域的识别。

本文所采用的 3 个深度学习网络模型都在 ImageNet2012 上进行过预训练, 然后, 采用 GTSRB 的训练数据集完成了模型微调, 表 4 中任何一个模型均提取 2 种特征, 如 AlexNet 为采用 AlexNet 提取特征, 并完成端到端的分类(方法为 Softmax), 而 AlexNet-FC7 表示采用 AlexNet 模型的全连接层(FC7)的生成图像的特征向量, 然后再使用 KNN 完

成测试图像的分类。Inception-FC 表示 Inception 网络的最后一个全连接层(FC)作为特征向量。

表 4 深度特征向量实验结果

网络模型/特征	分类方法	查准率 (%)	查全率 (%)	F1 值 (%)	N ⁺	Accuracy (%)
AlexNet	Softmax	-	-	-	-	94.69
AlexNet-FC7	KNN	90.96	88.92	89.92	43	92.42
VGG16	Softmax	-	-	-	-	97.18
VGG16-FC7	KNN	94.96	95.32	95.14	43	96.73
Inception	Softmax	-	-	-	-	95.59
Inception-FC	KNN	94.15	92.27	93.20	43	95.61

如表 4 所示, 得益于高质量的深度特征, 即是采用最简单的 KNN 分类模型, 也可以取得较好的识别效果。若进一步设计针对性强的复杂的分类模型等, 基于通用深度学习模型的 TSR 效果有望取得较理想效果。

从表 5 的第 1~7 行可以看出, 即使在光线强度和成像角度变化较大, 以及噪声显著的场景中, VGG16 的总体性能依然很高, 表现出较强的鲁棒性, 其识别性能已经超过了人类专家。但在个别类型的图像中, 如最后一行, SIFT 特征依然能优于 VGG16 等深度特征, 说明尽管深度特征总体上明显优于手工特征, 但高质量手工特征可以有较强的针对性。因此, 如何设计统一框架使深度特征和手工特征相结合, 使 2 类特征优势互补, 以便应用于复杂应用场景, 将成为今后 TSR 领域的研究重点。

表 5 针对不同特征的检测结果比较

原始 图像	标签 类别	Haar-like	HOG	SIFT	VGG16
	3	1	6	6	3
	22	32	12	32	22
	2	6	8	5	2
	22	1	12	32	22
	5	1	6	6	5
	19	1	27	12	19
	25	31	19	28	25
	23	1	26	23	26

4 总 结

TSR 是单对象识别,理论上全局特征描述是可行的,但实际交通环境成像存在光照、形变遮挡等情况,从而导致全局特征健壮性较差。因此,人工特征的设计应该更倾向于区域特征。在通用图像分类等领域取得较大成功的 BOW 模型,并未在 TSR 中表现出明显优势,反而低于底层的 SIFT 特征。人工特征中, SIFT 特征明显优于其他底层特征,与 SIFT 的一些通用优点有关,也与交通标志数据集特点有关, SIFT 更善于刻画和区分交通标志的关键特征。针对交通标志数据集的特点,聚焦局部显著点,设计具有更强的表现能力和区分能力的显著点特征表示方法,可以成为今后人工特征研究的重点。

在 GTSRB 数据集上的实验表明,很多学者在基于深度学习的 TSR 研究方面做了有益探索,或提出了新的模型,或改进了原有通用模型,或人工特征和深度特征相结合,或在深度特征基础之上采用复杂分类方法,但实际效果并未取得明显优于 VGG16 等通用模型的结果。因此,在研究通用模型基础之上,融合多种现有经典网络模型,或基于通用模型的深度特征设计合适的分类模型等,都可以成为该领域的研究内容。

本文采用相同数据集、相同分类模型、相同评价指标完成了主要特征向量的实验对比,为今后该领域研究者的特征选择提供了依据,也希望能够对该领域相关学者的后续深入研究提供参考。在后续的相关工作中,将对多个手工特征的组合进行实验,并试图提取更高层次的深度特征,以期得到综合的分析比较数据,再结合时间因素进行总体分析。

参 考 文 献

- [1] SAADNA Y, BEHLOUL A. An overview of traffic sign detection and classification methods [J]. International Journal of Multimedia Information Retrieval, 2017, 6(3): 193-210.
- [2] YUAN Y, XIONG Z T, WANG Q. An incremental framework for video-based traffic sign detection, tracking, and recognition [J]. IEEE Transactions on Intelligent Transportation Systems, 2017, 18(7): 1918-1929.
- [3] GUDIGAR A, CHOKKADI S, RAGHAVENDRA U. A review on automatic detection and recognition of traffic sign [J]. Multimedia Tools and Applications, 2016, 75(1): 333-364.
- [4] STALLKAMP J, SCHLIPSING M, SALMEN J, et al. the german traffic sign recognition benchmark: A multi-class classification competition [C]//The 2011 International Joint Conference on Neural Networks. New York: IEEE Press, 2011: 1453-1460.
- [5] 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.
- [6] LARSSON F, FELSBERG M. Using Fourier descriptors and spatial models for traffic sign recognition [M]//Image Analysis. Heidelberg: Springer, 2011: 238-249.
- [7] TIMOFTE R, ZIMMERMANN K, GOOL L V. Multi-view traffic sign detection, recognition, and 3D localization [EB/OL]. [2019-06-02]. <https://ieeexplore.ieee.org/document/5403121>.
- [8] FOUCHER P, TAREL J P, SOHEILIAN B, et al. Road sign detection in images: A case study [EB/OL]. [2019-05-10]. https://www.researchgate.net/publication/48418194_Road_Sign_Detection_in_Images_A_Case_Study.
- [9] 高歌, 谭兵, 陈心睿. 道路交通标志检测研究综述[J]. 现代商贸工业, 2017, 29(12): 176-178.
- [10] 梁敏健, 崔啸宇, 宋青松, 等. 基于 HOG-Gabor 特征融合与 Softmax 分类器的交通标志识别方法[J]. 交通运输工程学报, 2017, 17(3): 151-158.
- [11] 张淑芳, 朱彤. 基于残差单发多框检测器模型的交通标志检测与识别[J]. 浙江大学学报: 工学版, 2019, 53(5): 940-949.
- [12] XIAO D G, LIU L. Super-resolution-based traffic prohibitory sign recognition [EB/OL]. [2019-04-04]. <http://www.paper.edu.cn/releasepaper/content/201904-59>.
- [13] 樊星, 沈超, 徐江, 等. 基于级联端对端深度架构的交通标志识别方法[J]. 计算机测量与控制, 2019, 27(4): 143-148.
- [14] ZHANG T. An introduction to support vector machines and other Kernel-based learning methods [J]. AI Magazine, 2001, 22(2): 103-104.
- [15] RUTA A, LI Y M, LIU X H. Robust class similarity measure for traffic sign recognition [J]. IEEE Transactions on Intelligent Transportation Systems,

- 2010, 11(4): 846-855.
- [16] ZAKLOUTA F, STANCIULESCU B, HAMDOUN O. Traffic sign classification using K-d trees and random forests [C]//The 2011 International Joint Conference on Neural Networks. New York: IEEE Press, 2011: 2151-2155.
- [17] 韩习习, 魏民, 徐西义, 等. 基于多特征融合的交通标志识别算法[J]. 计算机工程与应用, 2019, 55(18): 195-200.
- [18] REN F X, HUANG J S, JIANG R Y, et al. General traffic sign recognition by feature matching [EB/OL]. [2019-05-27]. <https://ieeexplore.ieee.org/document/5378370>.
- [19] GREENHALGH J, MIRMEHDI M. Real-time detection and recognition of road traffic signs [J]. IEEE Transactions on Intelligent Transportation Systems, 2012, 13(4): 1498-1506.
- [20] HUANG Z Y, YU Y L, GU J, et al. An efficient method for traffic sign recognition based on extreme learning machine [J]. IEEE Transactions on Cybernetics, 2017, 47(4): 920-933.
- [21] PRISACARIU V A, TIMOFTE R, ZIMMERMANN K, et al. Integrating object detection with 3D tracking towards a better driver assistance system [C]//Proceedings of 20th International Conference on Pattern Recognition, (ICPR) 2010. Washington, D C: IEEE Computer Society, 2010: 3344-3347.
- [22] DOUVILLE P. Real-time classification of traffic signs [J]. Real-Time Imaging, 2000, 6(3): 185-193.
- [23] HOSSAIN M S, HASAN M M, AMEER ALI M, et al. Automatic detection and recognition of traffic signs [C]//2010 IEEE Conference on Robotics, Automation and Mechatronics. New York: IEEE Press, 2010: 286-291.
- [24] KHAN J F, BHUIYAN S M A, ADHAMI R R. Image segmentation and shape analysis for road-sign detection [J]. IEEE Transactions on Intelligent Transportation Systems, 2011, 12(1): 83-96.
- [25] YIN S Y, OUYANG P, LIU L B, et al. Fast traffic sign recognition with a rotation invariant binary pattern based feature [J]. Sensors, 2015, 15(1): 2161-2180.
- [26] 李阳, 丁辉, 王云飞, 等. 多特征融合的交通标志检测与分类研究[J]. 光学技术, 2012, 38(5): 564-568.
- [27] TANG S S, HUANG L L. Traffic sign recognition using complementary features [C]//2013 2nd IAPR Asian Conference on Pattern Recognition. New York: IEEE Press, 2013: 210-214.
- [28] 王刚毅. 交通标志检测与分类算法研究[D]. 哈尔滨: 哈尔滨工业大学, 2013.
- [29] ZHENG L, YANG Y, TIANQ. SIFT meets CNN: A decade survey of instance retrieval [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 40(5): 1224-1244.
- [30] VIRUPAKSHAPPA K, HAN Y, ORUKLU E. Traffic sign recognition based on prevailing bag of visual words representation on feature descriptors [C]//2015 IEEE International Conference on Electro/Information Technology (EIT). New York: IEEE Press, 2015: 489-493.
- [31] 刘亚辰, 陈跃鹏, 张赛硕, 等. 融合式空间塔式算子和 HIK-SVM 的交通标志识别研究[J]. 交通运输系统工程与信息, 2017, 17(1): 220-226.
- [32] 张建明, 王伟, 陆朝铨, 等. 基于压缩卷积神经网络的交通标志分类算法[J]. 华中科技大学学报: 自然科学版, 2019, 47(1): 103-108.
- [33] ZENG Y J, XU X, SHEN D Y, et al. Traffic sign recognition using extreme learning classifier with deep convolutional features [J]. IEEE Transactions on Intelligent Transportation Systems, 2017, 18(6): 1647-1653.
- [34] 宋海玉. 自动图像标注及标注改善算法的研究[D]. 长春: 吉林大学, 2012.
- [35] SWAIN M J, BALLARD D H. Color indexing [J]. International Journal of Computer Vision, 1991, 7(1): 11-32.
- [36] STALLKAMP J, SCHLIPSING M, SALMEN J, et al. The German traffic sign recognition benchmark [EB/OL]. [2012-03-16]. <https://ieeexplore.ieee.org/document/6033395?arnumber=6033395>.
- [37] MIKOLAJCZYK, SCHMID C. A performance evaluation of local descriptors [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005, 27(10): 1615-1630.
- [38] LOWE D G. Object recognition from local scale-invariant features [C]//Proceedings of the Seventh IEEE International Conference on Computer Vision. New York: IEEE Press, 1999: 1150-1157.
- [39] LOWE D G. Distinctive image features from scale-invariant keypoints [J]. International Journal of Computer Vision, 2004, 60(2): 91-110.
- [40] KE Y, SUKTHANKAR R. PCA-SIFT: A more distinctive representation for local image descriptors [C]//Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. New York: IEEE Press, 2004: 506-513.