深度学习在TSDR中的应用报告

作者姓名：李欣芮

作者单位：电子科技大学

# 摘要

随着社会的发展和进步，良好的交通同时促进着社会的快速发展，但现代交通的繁忙、堵塞和复杂的路况，在我们不断用更多的交通标识去规范出行时，也便增加了人们出行的麻烦，加大了交通事故的发生频率。在此背景下,各国开展了智能交通系统(ITS: Intelligent Transportation System)的研究。ITS被认为是解决以上问题的基本手段,其核心技术涉及数字信号处理、通信技术、图像处理、模式识别、人工智能和系统工程技术等,是一门综合性技术。道路交通标志识别(TSR)系统是智能车辆的重要组成部分,它在车辆行驶中对道路出现的交通标志进行信息采集和识别,及时地向驾驶员做出指示或警告,甚至接替驾驶员直接对车辆控制,以确保交通顺畅和防止事故的发生。

交通标志被广泛应用于道路交通中，由于受到许多不可控制因素如天气光照变化、物理遮挡、运动模糊等的影响，交通标志的准确检测和快速识别对研究者来说是一个重大的挑战。然而传统的目标检测方法在图像识别领域存在较高表达力的特征难以提取，特征分类的准确度和速度都较低，深度学习中卷积神经网络的发展，使得其在图像识别领域普遍存在两种应用方式，分别为基于分类的卷积神经网络目标检测，基于回归的卷积神经网络的目标检测，本文将就目前深度学习在交通标志目标识别领域的应用做出概述型的报告，以便实验室了解当前该领域的研究现状并有助于做进一步深入的研究。

**关键字: 深度学习, 交通标志, 目标识别**

# Abstract

With the development and progress of the society, good transportation also promotes the rapid development of society. However, the busy, blocked and complicated road conditions of modern transportation have increased people as we continue to use more traffic signs to regulate travel. The trouble of travel has increased the frequency of traffic accidents. In this context, countries have conducted research on the Intelligent Transportation System (ITS). ITS is considered to be the basic means to solve the above problems. Its core technology involves digital signal processing, communication technology, image processing, pattern recognition, artificial intelligence and system engineering technology. It is a comprehensive technology. The Road Traffic Sign Recognition (TSR) system is an important part of intelligent vehicles. It collects and identifies the traffic signs appearing on the road while the vehicle is running, promptly gives instructions or warnings to the driver, and even takes over the driver directly, for ensuring smooth traffic and prevent accidents.

Traffic signs are widely used in road traffic. Due to many uncontrollable factors such as weather illumination, physical occlusion, motion blur, etc., accurate detection and rapid identification of traffic signs is a major challenge for researchers. However, the traditional target detection method is difficult to extract in the field of image recognition, and the accuracy and speed of feature classification are low. The development of convolutional neural network in deep learning makes it popular in image recognition. The application methods are classification-based convolutional neural network target detection and regression-based convolutional neural network target detection. This paper will give an overview report on the application of deep learning in the field of traffic sign target recognition. The room understands the current state of research in this area and will help to further research.

**Keywords: Deep Learning, Traffic Signs, Target Recognition**

# 绪论

## 研究背景及意义

随着改革开放的深入进行，我国社会经济进入高速发展阶段，人民的生活水平日益提升，汽车的需求量近些年来保持着持续增长，汽车己经成为人们出行的主要交通工具。据交管部门统计，截至2017年4月底，我国汽车的总量己经超过两亿辆，汽车驾驶人数高达3.2亿。如今，无论是在发达的城市或者偏远的郊区乡村，交通道路上都挤满了汽车。汽车给人们生活带来便利的同时，也给交通管理带来了很大的挑战，同时也不可避免地造成了很多交通安全的问题。根据世界卫生组织机构公布的《全球道路安全状况报告》显示，我国每年因交通事故死亡的人数已经超过2.5万，这些频频发生的交通事故不仅给当今社会造成了重大损失，更是给亲人带去了无尽的痛苦。根据交通心理学者研究发现，人们对于突发性的事件的反应具有一定的滞后性，从发现目标到采取相应的措施有一定的时间延迟，如果没有足够的时间留给驾驶员去反应其面临的交通路况，就很可能导致悲剧的发生。正是在这样的背景下，高级驾驶员辅助系统(Advanced Driver Assistance System, ADAS)应运而生，而自动交通标志牌检测和识别系统(Automatic traffic sign detection and recognition, TSDR)是高级ADAS的重要组成部分。

ADAS是一种结合了包括汽车操控、数字通信、信息处理、图像处理等多个领域先进技术的高级驾驶辅助系统。它主要是通过安装在车上的众多传感器在最快的时间内收集道路周边的环境状况信息，然后分析采集到的信息并进行侦测和物体辨别，从而提示驾驶者当前路况存在的危险并给出正确规范的提示。为了提高驾驶辅助系统的可靠性和稳定性，国内外研究者己经围绕ADAS的各个方面展开研究，包括对周围环境路况的信息检测、驾驶员驾驶状态的侦测以及道路上车辆、行人、路标线、交通标志的识别等。TSDR是ADAS的一个重要组成部分，TSDR通过对道路上交通标志牌的准确识别能够有效提示和指导驾驶员规范驾驶，它对现代交通智能管理有着重要意义，也越来越受到人们的关注和重视。

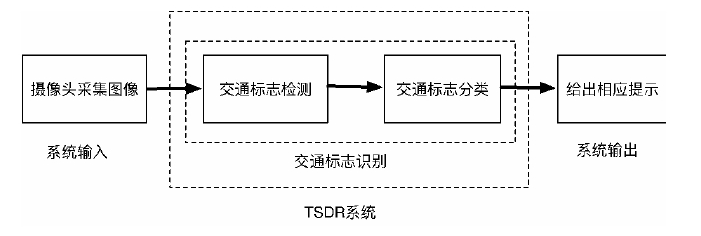


图 1-1 系统基本框架

如图1-1所示，TSDR系统工作的基本流程为:首先通过配置在汽车内的摄像头拍摄道路前方路况图像，这些图像是TSDR系统的输入;接着利用机器学习相关算法和图像处理技术对输入图像进行交通标志的检测，检测指的是提取出包含交通标志的区域;然后对检测出来的那些包含交通标志的区域进行具体类别的判断，最后TSDR将根据交通标志分类出的结果做出不同的应对提示反馈给驾驶员，比如，道路前方200米处有障碍物、前方是学校区域应禁止鸣喇叭并限速等。TSDR系统能够有效帮助驾驶员尽早地发现路况可能存在的风险，让驾驶员有充裕的时间去执行正确规范的操作，这将极大地帮助减轻驾驶员的驾驶压力，降低交通事故发生的可能性。

除了可以帮助驾驶员缓解驾驶压力之外，TSDR系统还有助于提高交通管理部门的工作效率。目前，交通管理相关的工作主要由人工完成，每天需要大量的交通管理工作人员站岗指示行人和车辆的行驶、定期检查道路上交通标志牌的损毁情况等，这些机械、琐碎但又对保证交通安全起着重要作用的工作消耗了大量的人力物力，而且容易出现人为操作失误。随着TSDR系统的不断发展，这些工作都能交给TSDR系统智能地完成，并且比人工操作更加准确高效，这将能解放大量的人力物力。

在无人驾驶技术中，TSDR系统同样扮演着不可或缺的作用，它是指导汽车掌握周边路况、做出规范操作的核心技术。在2017年7月\_5日举行的百度 AI开发者大会上，互联网巨头公司百度现场展示了百度创始人Robin乘坐百度无人车前往会场的画面，其中重点介绍了TSDR系统在无人驾驶技术中起着支撑性作用。

综上所述，TSDR系统对于指导驾驶员安全行驶、提高交通部门的管理效率、促进无人车技术的发展等方面起着重大作用，对交通标志识别系统的算法研究，具有重要的理论意义和商业价值。

## 国内外研究现状

针对于TSDR系统的研究主要倾向于交通标志检测和交通标志分类这两方面的研究，在复杂的外部环境中实时准确的检测到交通标志牌，才能为交通标志的分类作铺垫，精准的将检测到的交通标志进行分类最终反馈才能使得减少交通安全事故的发生，目前，国内外都相继在这两个领域做了相应的研究。

### 交通标志检测方法研究现状

交通标志检测问题的研究，最早可以追溯到二十世纪八十年代。随后美国首先提出了智能交通系统的概念，此后，智能交通的研究领域开始迅速发展。在庞大的道路交通系统中，交通标志的自动识别是智能交通系统的重要部分，而交通标志自动识别的首要任务就是在获取的道路场景图像上分割出交通标志，也就是在交通道路场景中检测交通标志，所以交通标志的检测无疑是识别系统中关键的一步。直至目前为止，随着科学技术的快速发展，智能交通系统逐步成长起来，很多研究机构和研究学者已经对交通标志检测这一领域做了深入研究，提出了很多交通标志检测的方法，并取得了大量的学术研究成果。

交通标志的作用是指导或警告驾驶员和行人，提示并帮助驾驶员更加安全地行驶。为了使交通标志更为醒目，各个国家都有自己独特的设计方式。我国的交通标志基本可以概括分为五类:禁止标志、警告标志、指示标志、指路标志和道路施工标志。禁止标志大多为红色边缘，白色填充，黑色内容，包括限速标志，禁止通行标志等;警告标志一般为黄色填充，黑色边缘，黑色内容;指示标志和指路标志均为蓝色填充，白色内容;道路施工标志同警告标志类似。交通标志有其固定的几何形状和规定的较为醒目的颜色，因此交通标志的检测方法可以大致归为以下几类:基于颜色的方法、基于形状(模板匹配)的方法、基于颜色和形状的结合的方法、基于特征的方法。这些也是目前使用最为普遍的交通标志检测方法，本文将对这四类方法做详细介绍。

(1)基于颜色的交通标志检测方法

颜色信息是交通标志最明显的特征之一，根据特定的颜色信息可以很快去除掉道路场景中的非感兴趣区域，而且利用颜色信息运算过程较为简单，所以研究学者们最早开始了基于颜色的交通标志检测方法研究。

行车记录仪的所获取的视频中，最常用的颜色空间为RGB模式。早在2003年，HsiuMing等人利用交通标志中的红色信息，通过红色像素点对图像进行分割，成功获取了含有红色像素点的候选区域。在这之后Cyganek在红色像素点分割的基础上，针对含红色信息的交通标志，设计了基于支持向量机(SupportVector Machine，简称SVM)的分类器对含红色区域的交通标志进行分割。Asakura等还提出一种基于颜色信息的交通标志检测方法，该方法通过使用一个简单向量滤波器((Simple Vector Filter，简称SVF)快速提取指定颜色，并能够去除无关像素。RGB彩色空间图像在算法上的优点表现为计算速度快，实时性好，但它的缺点是受光照影响较大，不能够精准地模拟人眼的感知。为了克服RGB颜色空间图像在交通标志检测时对颜色变化比较敏感的缺点，Kehtarn等人指出将RGB空间图像转换到HSI颜色空间，利用H分量受光照影响小的特点检测禁止类型的交通标志，但是当像素信息为灰度信息(R=G=B)时，这种方法的检测效果较差。Cyganek在HSI颜色空间下，采用直方图分析的方法，准确估计了在HSI彩色空间颜色分割所适应的阂值。在此基础上，通过模糊推理的算法对交通标志进行颜色分割。沙莎等人提出了一种基于多通道颜色空间的交通标志检测方法，该方法首先将图像转换到HSI颜色空间，再对其颜色和饱和度进行闭值分割处理，然后结合RGB颜色空间，对其颜色特征进行聚类，最终完成交通标志的分割。尽管基于HSI彩色空间的多种检测算法有效解决了RGB颜色空间对光照的敏感性，但是HSI彩色空间在使用时也存在缺点:彩色图在处理时虽然信息量较多，但颜色信息并不稳定，白天和夜晚的图像的像素值变化较大。除此之外，在光照和天气变化时，RGB信号会有少量变化，转换HSI彩色空间的色调信号就会产生较大的变化，这也在一定程度上影响了HSI颜色空间上的分割效果。

(2)基于形状的交通标志检测方法

形状是交通标志的重要特征，基于形状的交通标志检测方法很早就受到研究者的关注并投入研究。利用交通标志的形状特征可以有效地将交通标志与其背景分离。基于形状特征的交通标志检测方法避免了颜色信息对天气光照变化敏感的问题，在一定程度上增强了检测算法的鲁棒性。

目标形状轮廓分析法作为交通标志形状检测的传统方法，其核心思想是将交通标志分为若干个区域，再对每个区域进行轮廓检测。Besserer等人就提出利用构建金字塔，在金字塔结构的图像中进行边缘检测，通过分析边缘轮廓来区分交通标志为三角形、圆形还是其他。朱双东等人针对红色倒三角形标志提出了一种智能检测方法，首先他将图像转换到红色HIS颜色空间，边缘检测后利用标准LOG模板判别顶点位置，对边缘点使用分类器分类，判断顶点是否满足三角形的几何特点，进而标记三角形的3种顶点，最终通过三角形模板匹配准则提取三角形交通标志候选区。该方法的具有较好的鲁棒性、检测速度快、检测准确率高，但是检测目标的种类比较局限。针对圆形交通标志，鲍朝前提出一种加强型椭圆最小二乘拟合法，在随机椭圆最小二乘拟合算法的基础上，增加了图像细化算法，该检测算法对形变和部分遮挡的圆形交通标志提高了检测率。对于圆形交通标志，张静等人提出了一种多特征融合的圆形交通标志分割方法，首先利用非彩色分解技术，将图像中的非彩色区域去除，过滤掉大部分的非感兴趣区域，再通过综合利用圆形目标物的展弦比和圆形度等特定特征精确提取出圆形区域，从而将圆形交通标志成功分割出来，获得了满意的研究成果。

(3)基于特征提取的交通标志检测方法

针对传统基于模板匹配和颜色特征算法在实际应用中存在的问题，虎亚玲等人结合Labor滤波器和傅里叶变换，提出了一种基于多尺度视觉显著性的交通标志检测算法。该方法首先利用高斯金字塔对输入图像进行多尺度分解，提取每个子级的特征之后，使用Labor滤波器提取相应的颜色特征和纹理特征用于构造多尺度四元图像，再通过计算其四元数傅里叶变换的相位谱，得到最终的交通标志显著图。该方法虽然有效提高了检测率，但其计算复杂度较高，不利于达到工程实时性的要求。Jesmin等人提出了一种基于层聚类并结合相关变换的道路交通标志自动检测系统，首先通过经验模态分解(EMD)检测到兴趣点，再通过Labor滤波器提取亮度和颜色特征，计算相关性，进而使用层聚类算法聚类这些兴趣点从而成功提取候选物。该算法对于轻度的局部遮挡和旋转有较强的适应性，但当交通标志与背景对比度不足或者有较大程度的闭塞和旋转时无法准确检测。

基于HOG特征对交通标志进行检测的方法也受到研究者的青睐用提取CIELAB和YCbCr颜色空间，统计颜色信息的直方图，计算特征向量时合并相同颜色的信息，还可以通过改变颜色标准降低计算负荷，将24位RGB图像转换为8位图像，只有感兴趣的颜色分量被提取用于过滤候选区域，然后用HOG描述符来检测候选区域的特征，最后将这些描述符作为输入，用支持向量机((SVM)进行分类。该算法与经典的颜色空间变换相比，可以防御强烈的光照，阴影，阴雨天等情况带来的噪声影响，有效增强了算法的鲁棒性。

### 交通标志识别方法研究现状

目前对于交通标志识别的方法，基于神经网络的方法表现出其优势。基于双层神经网络的交通标志识别方法将识别问题简化，采用神经网络群分别组成两级智能分类器，第一级分类器用以实现交通标志的粗分类，即按照大类别分类;第二级分类器用以细分类，即判定候选物为哪种具体标志。通过模糊特征训练集训练BP神经网络，用小规模的训练集也可以实现交通标志的粗分类且实验结果证明该分类器具有较好的鲁棒性。相比于神经网络算法，基于卷积神经网络的方法能够自主学习特征，从大量数据中学习到最有利于实现分类的特征，能够克服交通标志对环境变化敏感的缺点。近年来，深度卷积神经网络算法在交通标志分类任务中取得突破性进展，国际神经网络会议关于交通标识牌识别的竞赛吸引了许多学者和研究人员，竞赛第一名的Ciresan提出多纵深度神经网络算法，识别率高达99.46% 。MCDNN算法的实现依赖于快速实现的图像处理器GPU，网络结构复杂，识别过程非常耗时。Yujun等人提出极端学习结合深度卷积神经网络算法对交通标志进行识别，同时指出该算法的高识别率依赖于复杂的网络结构对待识别物体的特征提取和所选用的分类器。以上交通标志的识别在深度学习中的运用主要是分类算法，其识别率高但不具备实时性，而且过程复杂。2016年12月29日，Wei Liu, Dragomir Anguelov等人提出了一种使用单个深度网络来检测图片中对象的方法为SSD，并且在VOC2007的300X300的数据集上实现了74.3%mAP的准确率，在512X512的数据集上实现了76.9%mAP。本文则是根据这两类深度学习的方法在交通识别上的应用的研究分析。

# 基于分类的CNN交通目标识别应用

## 交通标志检测

交通标志的检测是交通标志识别的第一步，检测的任务是从采集的图像中分割出候选区域，并从候选区域中筛选出真正包含交通标志的区域。由于在分类阶段，会把检测阶段输出的位置区域赋予具体的类别标签，一旦检测阶段出现错检，会直接导致分类阶段的错误分类。因此，检测阶段效果的好坏直接影响整个交通标志识别系统的准确率，检测的速度快慢也是整个识别系统能否满足实时性要求的关键。交通标志检测的基本流程如图1-2所示。

图 2 交通标志检测基本流程

目前在交通标志，普遍采用基于HOG的特征提取方法，对图像感兴趣的局部特征进行提取，并训练分类器进行分类判别交通标志区域的正样本子候选区，而分类器现常采用基于SVM的二分类分类器，级联分类器，和基于CNN的布尔判定分类器BCNN。

在交通标志检测阶段，CNN对候选区的负样本的筛选过滤有着很好的效果，

目前所研发的HOG-BCNN的算法可以很好的过滤级联分类器的误检样本。

BCNN的样本规模设计合理的网络结构。由于BCNN模型的目的是区分交通标志与非交通标志，相比于较多种类的识别任务相对简单。其改变网络参数从简单到复杂，通过实验并对比结果选择合理的网络结构作为BCNN的网络模型。

首先第一种网络结构(Network)选择使用两个卷积层，每层分别使用4个和8个单尺度的5\*5卷积核;第二种网络结构((Network2)再使用较多的、且不同尺度的卷积核，第一层采用16个5\*5的卷积核，第二层采用32个3\*3的卷积核。在第一层卷积时当输入图像较大时选取较大尺度的卷积核，不仅可以提取局部特征，还能够快速降低特征维数以加快网络的运算速度;而第二层卷积使用较小的卷积核是由于输入的特征图维数已经比较小，且较小尺度的卷积核能够提取较为细节的特征。

为了能够尽可能多的过滤掉误检窗口，BCNN需要达到一个很高的识别率。Network 1的最高识别率为95.2%，而Network2的识别率己经高达99.5%，该网络结构已经符合本文过滤误检的要求，无需再增加网络的复杂度，所以本文选择Network2作为BCNN的网络模型。两种网络结构的识别性能如图3所示。

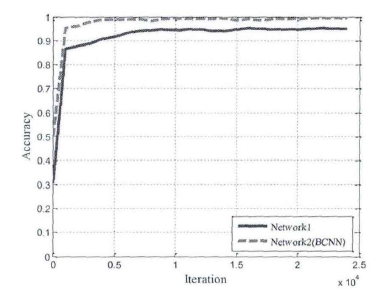


图 3 不同网络结构的识别率

在检出率的比较中，图3是不同的方法在检测行车记录仪获取的视频时的precision-recall(精确率一召回率)曲线。精确率也叫查准率，是指系统检测到的正确目标数占系统检测到的所有物体数的比例;召回率也叫查全率，是指系统检测到的正确目标数占系统所有目标数的比例;相比于基于HOG特征训练级联分类器检测交通标志，利用径向对称变换或BCNN过滤误检都能够有效排除误检提高检测率。

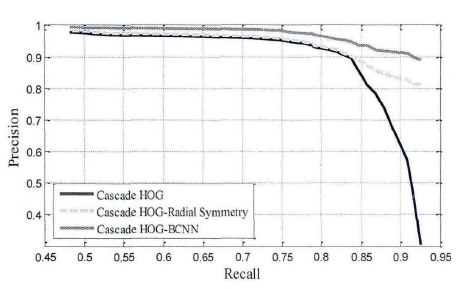


图 4 不同方法检测交通标志的精确率一召回率曲线

由图4可知，当训练级联分类器的召回率较高时，基于HOG的检测存在较多误检而使得检测率较低，而径向对称变换和BCNN则因为能够有效过滤部分误检对检测率有所改善，并且BCNN的过滤效果要明显优于径向对称算法。这是因为HOG特征虽然能够利用图像局部区域的梯度信息很好地描述物体轮廓，但场景中存在较多干扰，树叶树梢、建筑物等物体也有明显且丰富的边缘信息。而径向对称变换能够根据其特征性质过滤掉部分误检区域，但由于特征提取时噪声的干扰，仍存在一些误检难以排除。而BCNN通过自主学习最有助于区分交通标志与非交通标志的特征，能够过滤掉大部分误检，显著提高了交通标志的检测率。

BCNN没有直接采用基于CNN的算法进行交通标志的检测，这是因为基于CNN的交通标志检测方法往往需要构建候选窗口提取的网络，且网络结构比较复杂。与之相比，BCNN网络结构简单，网络的计算复杂度大大降低且运算速度快。除此之外，CNN在图像分类任务中己经取得突出的成绩，用CNN来过滤交通标志检测中的误检区域体现了它的优势。

## 交通标志分类

### 数据集

2011年国际神经网络联合会议提供的数据集德国交通标志识别标准GTSRB是由摄像机实时场景获取，由相机在每一个场景中连续拍摄1S，每一类均由多个场景中的连续帧作为样本，由于相机由远及近接近目标物，分辨率逐渐由低到高，既有清晰可辨也有运动模糊人眼难以识别的。数据集GTSRB集中图片的分辨率在15X 15到250 X 250的范围内变化。GTSRB中含训练集样本39209张，测试集12630张，共包含43类交通标志。

数据集Tsinghua-Tencent 100K是由清华大学博士生朱哲等人与腾讯公司合作制作的数据集。该数据集是将腾讯街景中的交通道路图像保存，标注交通标志的位置及其类别。数据集Tsinghua-Tencent LOOK所含样本的分辨率为2048 X2048。本文根据标注文件将含有交通标志的图像中的待识别目标完整抠出作为验证本文方法的数据集之一。数据集Tsinghua-Tencent 100K中含训练样本共16524张，测试样本8181张，共包含交通标志类别超过100类。

两个数据集中交通标志的具体类别如图5所示。



图 5 数据集样本类别

### CNN网络

基于分类的卷积神经网络应用于交通目标识别的有FTSR-CNN，结合卷积神经网络算法原理，根据已有数据的规模与特点设计基于卷积神经网络的快速交通标志识别(Fast Traffic Sign Recognition based onConvolution Neural Network, FTSR-CNN)的网络结构。Taohong Lu等人指出，CNN网络设计的深度和宽度并非越复杂越好，当模型参数的规模远远大于数据集规模时，会出现维数灾难。维数灾难就是分类器的性能随着维度漪征个数)的增加逐渐增强，但在到达某点后，其性能便会逐渐下降。维数灾难的现象也可以用数学中的方程组求解来解释，网络模型的参数个数即代表了特征的维度。如果特征的数量过多，相当于方程组没有唯一最优解，会出现过拟合现象，使网络只记住了训练样本，没有学习到训练样本的共性，而且会增加耗时;与之相反，当模型参数的规模小于数据集时，相当于对无解方程组进行求解，容易出现欠拟合。因此只有设计合理的网络结构，平衡好网络深度宽度和样本数据量的关系才能取得较好的识别结果。

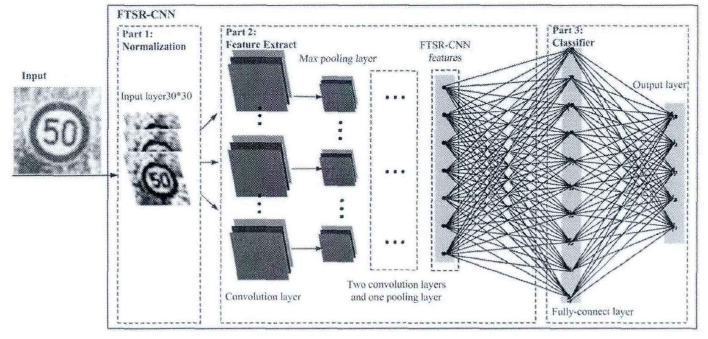


图 6 FTSR-CNN网络结构

FTSR-CNN所包含的3个卷积层(convl-conv3)，其每层使用的卷积核大小(kernel size)分别为5\*5, 3\*3, 3\*3个像素。多层结构可以对输入对象逐层抽象，获得更高层次的分布式特征表达。在convl和conv2层的特征图后，均连接了最大值采样层，以获取图像在一定范围内畸变时的特征不变性，并大大降低特征维数，提高计算效率。FTSR-CNN模型的最后两层为全连接层(fully6, fully7), fully6层为包含300个神经元的隐藏层，而fully7层为包含43类的Softmax输出层。交通标志识别效果如图7。



图 7 交通标志识别效果图

# 基于回归的CNN交通目标识别应用

基于回归的CNN交通标志识别目前主要有：RCNN,借鉴了滑动窗口的思想，采用了区域识别的方案，特征提取和分类分离的过程；FastCNN在RCNN的基础上将SVM的回归方法加入了ROI Pooling算法将最终提取到的特征加入到CNN网络中进行分类，FasterRCNN在FastRCNN的基础上采用了RPN算法筛选提取建议的候选框，大大的减少了提取的候选框，同时也减少了需要进行回归运算的候选框，加快了运算时间；YOLO是一种实时性的图像识别的深度学习网络，它使用了仅仅一个CNN网络，使用一组固定比率的框来进行学习，并加入了NMS算法来筛选框使得训练方便快捷，虽然识别率不如前面所提到的两阶段算法，但它是一种全新的便捷的深度学习网络，拓宽了图像识别的应用；SSD在YOLO的基础上增加了不同比率的组框，并且使用多特征图的回归运算，大幅提高了其识别率。以上基于回归的CNN的深度学习方法中RCNN，FastRCNN，FasterRCNN是两阶段的学习过程，不是实时性的，而且运算效率低，但识别率高，YOLO，SSD是单阶段的图像识别算法，具备实时性，便捷性。

# SSD交通标志分类简单实验

本次实验采用了SSD-Tensorflow的框架搭建了实验平台，对该计算框架进行了相应的参数调整，并使用了CSUST Chinese Traffic Sign Detection Benchmark 中国交通数据集由长沙理工大学综合交通运输大数据智能处理湖南省重点实验室张建明老师团队制作完成的数据集，到目前为止，已经上传图像15734张，全部的groundtruth也已经上传。目前的标注数据只有三大类：指示标志、禁止标志、警告标志。由于该实验目前正在搭建过程中，未得到实验结果，则主要介绍实验的整体实验过程。

## 实验框架介绍

SSD-TF框架是Paul Balanca团队在tensorflow的框架平台上对SSD算法的一个实现，并对原始的caffeSSD算法做了一个对比，其整体性能基本一致。目前，它仅实现基于VGG的SSD网络（具有300和512输入），但该项目的架构是模块化的，并且应该易于实现和培训其他SSD变体（例如，基于ResNet或Inception）。目前的TF检查点已直接从SSD Caffe模型转换。整个框架的灵感来自于tensorflow-slim模型包，主要包括以下三部分：

数据集：流行数据集的接口（Pascal VOC，COCO，...）和将前者转换为TF-Records的脚本

网络：SSD网络的定义，以及常见的编码和解码方法（我们参考关于这个精确主题的论文）

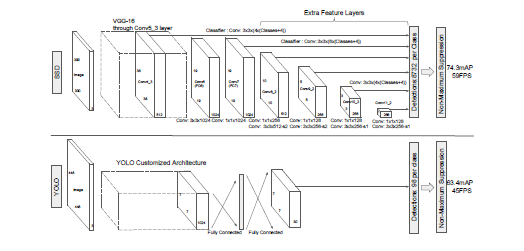
预处理：预处理和数据增强例程，受原始VGG和Inception实现的启发。

## 数据集的介绍

TF-SSD数据集使用的是VOC数据集，tenorflow在训练的中使用的是tfrecord格式的数据集，因此该框架也提供了将VOC数据集转化成tfrecord的方法文件，该框架采用了工厂模式的设计模式，可以灵活的提供不同的类型的数据集进行实验和可拓展性。本次交通标志实验采用中国交通数据集由长沙理工大学综合交通运输大数据智能处理湖南省重点实验室张建明老师团队制作的CCTSDB作为数据集，并且在此基础上谢了相应的格式脚本文件，将该数据集转变为VOC2007的格式。

## SSD模型介绍

SSD 是基于一个前向传播 CNN 网络，产生一系列 固定大小（fixed-size） 的 bounding boxes，以及每一个 box 中包含物体实例的可能性，即 score。之后，进行一个 [非极大值抑制（Non-maximum suppression）](http://www.pyimagesearch.com/2014/11/17/non-maximum-suppression-object-detection-python/) 得到最终的 predictions。



SSD 模型的最开始部分，本文称作 base network，是用于图像分类的标准[架构](http://lib.csdn.net/base/architecture)。在 base network 之后，本文添加了额外辅助的网络结构：

* Multi-scale feature maps for detection   
  在基础网络结构后，添加了额外的卷积层，这些卷积层的大小是逐层递减的，可以在多尺度下进行 predictions。
* Convolutional predictors for detection   
  每一个添加的特征层（或者在基础网络结构中的特征层），可以使用一系列 convolutional filters，去产生一系列固定大小的 predictions，具体见 Fig.2。对于一个大小为 ，具有  通道的特征层，使用的 convolutional filters 就是  的 kernels。产生的 predictions，那么就是归属类别的一个得分，要么就是相对于 default box coordinate 的 shape offsets。   
  在每一个  的特征图位置上，使用上面的  的 kernel，会产生一个输出值。bounding box offset 值是输出的 default box 与此时 feature map location 之间的相对距离（YOLO 架构则是用一个全连接层来代替这里的卷积层）。
* Default boxes and aspect ratios   
  每一个box相对于与其对应的 feature map cell 的位置是固定的。 在每一个 feature map cell 中，我们要 predict 得到的 box 与 default box 之间的 offsets，以及每一个 box 中包含物体的 score（每一个类别概率都要计算出）。 因此，对于一个位置上的boxes 中的每一个 box，我们需要计算出20个类，每一个类的score，还有这个 box 相对于 它的默认 box 的4个偏移值（offsets）。于是，在 feature map 中的每一个 feature map cell 上，就需要有80个 filters。对于一张大小的8X8feature map，即会产生 5120个输出结果。

这里的 default box 很类似于 [Faster R-CNN](https://arxiv.org/abs/1506.01497) 中的 [Anchor boxes](https://github.com/rbgirshick/py-faster-rcnn/issues/112)，关于这里的 Anchor boxes，详细的参见原论文。但是又不同于 Faster R-CNN 中的，本文中的 Anchor boxes 用在了不同分辨率的 feature maps 上。

## 实验过程

实验过程分为平台的搭建、数据集的制作、分类的调整、训练参数的调整、脚本的编写。在windows的平台下，搭建anaconda+tensorflow+matplot+opencv环境，制作VOC2007数据集。（后面的实验还未开展）

# 总结与展望

TSDR领域的研究对与人们未来的交通密切相关，使用深度学习的方法解决TSDR问题，已经成为当下热门而且比较成功的方法。尽管使用深度学习的方法在特征提取和图像识别的准确率上都得到了大幅的提升，但对于真正用于生活中这种技术还需要在这些方面继续提高，并且尽管各种深度学习的图像识别框架都得到了相应的改善，但实时性对小目标的精准识别还是很大一难题，并且对于TSDR这种交通方面的识别，其准确率的要求也是非常高的，甚至不能允许它出错，这也是在交通的自动驾驶方面的硬性要求。