# VehicleLogoRecognitionandAttributesPredictionbyMulti-taskLearningwithCNN

用CNN进行多任务学习的车辆标识识别和属性预测

# 摘要：

车辆标识识别（VLR）一直是智能交通系统（ITS）的重要研究领域。本文提出通过结合卷积神经网络（CNN）和多任务学习（MTL）来识别车辆标志和预测标志属性。为了加速多任务模型的收敛,采用自适应的权重训练策略。为了验证算法,将厦门大学车辆标识识别数据集扩展到一个较大的车辆标识数据集中,其中包括15个品牌,6个视觉属性和3个非视觉属性。实验结果表明,提出的多任务CNN模型在标识分类和属性预测方面表现良好,总体准确率为98.14％。

关键词：车辆标识识别（VLR）,卷积神经网络（CNN）,多任务学习（MTL）

# 简介

智能交通系统包括一个重要的应用,车辆制造商识别对政治或商业机构的汽车所有权统计有很大的好处。VMR系统通常利用标识进行识别,因为车辆标志是车辆上的明显标志。

已经发布了许多关于车辆标识识别的论文。大多数以前的提议方法将一些手工特征HOG与可训练分类器SVM结合[2]。这些方法有许多限制,例如1）不足以同时满足各种成像条件的手工特征（例如,HOG）,例如半灯,旋转,视点等等;2）精确的标志检测边界框不可或缺的要求,因为不精确的边界框会引入噪声,从而导致识别精度显着降低。

近年来,机器学习的一个新领域,即深度学习吸引了全世界的关注,其目的在于代表性学习。一种特殊的深度学习模型,即卷积神经网络（CNNs）,已经在各种计算机视觉任务中表现出优异表现[3]。最近的研究还表明,CNN可以通过输入和输出之间的多个卷积层非线性映射从原始像素中提取稳健和通用的特征[4]。由于其分层结构,CNN对光照变化,车辆姿态,污点等具有很强的鲁棒性[5],[6]。基于这些特点,CNN在识别车辆标识方面实现了高性能

用语义属性或属性来描述对象是一项在计算机视觉问题中受到许多研究者关注的技术[7]。属性不仅从多个方面描述对象,还充当了低级特征与高级语义的桥梁。各种多媒体任务可以从属性中受益,如知识转移和不同属性间的知识共享[7]。在我们的工作中,我们调查了CNNs学习框架中的6个视觉属性和3个非视觉属性。

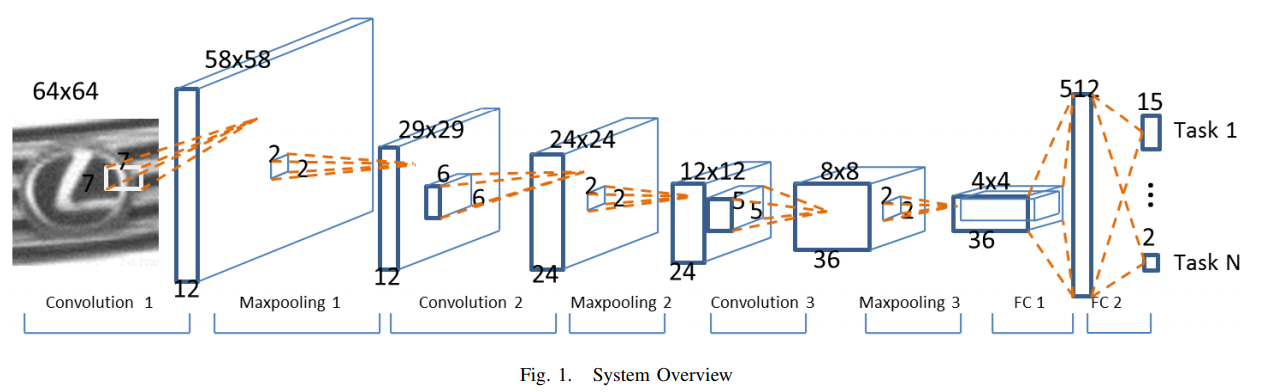
近来,MTL已被用于许多视觉识别研究,特别是当任务具有一些共同性并且通常略微缺乏各种数据时[8],[9]。这项技术试图迫使多个相关任务共享知识。MTL旨在通过共享相关和不相关的信息来扩大类间距离和缩短类内部距离,从而提高每项任务的效果[9]。现有的CNNMTL可以简单地分为两种。在第一个[10]中,一个任务被优化并且其他任务被固定,然后选择另一个任务来优化其他任务保持不变,这样的过程随着训练的进行而迭代地重复。然而,许多研究人员采用第二种类型[11],它同时优化所有任务。我们遵循第二种方式,并提出了一种由[11]启发的新的自适应加权MTL。

更具体地说,我们使用自适应加权MTL训练CNN模型以识别车辆标志并同时预测其属性。并从厦门大学车辆标识识别数据集[6]中扩展出一个包含15个品牌,6个视觉属性和3个非视觉属性的大型车辆标识数据库来评估算法。根据实验结果,得到满意的结果

本文的其余部分安排如下。第二部分解释了整个系统并阐述了自适应加权MTL。第三部分分析了实验,接下来是第四部分的结论和未来的工作。

# 系统概述

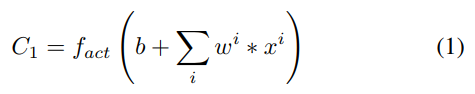
整个系统如图1所示。原始图像被转换为灰度,然后归一化为统一大小,64\*64。然后CNN从图像中提取特征。最后,应用几个多层感知器（MLP）来识别车辆标志并预测一些二进制属性。下面将在II-A中概述CNN,并将在II-B中解释自适应加权MLT。



## 使用CNN进行车辆标志识别

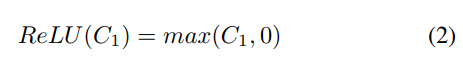
可以使用CNN从图像中分层提取特征,如图1所示：

C1：在CNN前馈过程中,卷积运算使用一组局部感受域（核或滤波器）来重复滑动整个视场从左上角到右下角以提取物体的显着图。为了减少CNN的参数,卷积运算将相同的局部感受野（包括权重和偏置）与输入一起应用,这被称为参数共享。这种技术不仅使得培训在有限的数据中成为可能,而且增加了CNN的泛化性能。在训练过程中,卷积层的核心将逐渐对边缘,颜色或车辆标识的特定模式检测器敏感。卷积操作被形式化为：



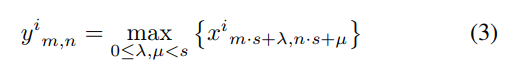
其中Xi和C1分别是第i个输入和第i个输出特征映射。对于第一个卷积操作,输入是原始图像,输出是第一个图层特征。对于其他卷积运算,输入是来自最后一个池化层或激活层的输出。Wi是内核的权重向量。\*表示点乘操作。B代表偏置和fact（·）代表非线性映射函数。滤波器的权值随机初始化,然后用熟知的反向传播算法进行训练。

R1：通常,非线性激活层在卷积层之后,因为卷积操作产生线性映射。然而,实际问题通常需要非线性复杂模型来解决非线性问题。为了将非线性映射引入到CNN中,非线性激活层是必不可少的。该层将来自卷积层的线性输入转换为非线性表示。过去,sigmoid函数扮演着重要角色,而在CNN中为了消除梯度问题而放弃了它。近年来,Glorot等人通过提出一个新的激活函数,称为整流线性单元（ReLU）,在方程2显示



随着CNN模型的深入,考虑到这一优势,最近基于CNN的方法[13]采用了ReLU作为卷积层和全连接层之后的非线性激活层,以缩短训练时间,大致如[13]中所述。

S1：如果CNN仅通过卷积层提取特征,将导致模型变得太深或特征的维度变得太高。为了解决这个问题,[13]利用池操作,也可以保持空间方差。广泛的池化功能是最大池化,即最大池化层,如等式3所示

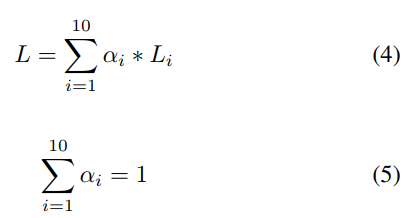


最大池化操作与卷积操作类似,可以将输入特征映射Xi上的矩形SxS从左上角平滑到右下角以选择局部最大值yi。在本文中,使用了一个不重叠的范式。

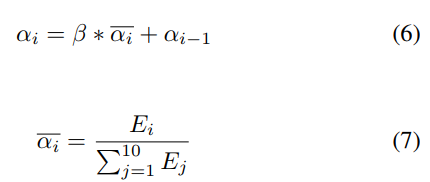
对于VLR,可以使用原始像素获取CNN特征提取器,以自动学习低级和中级特征,从而减轻对手工特征的需求,从而实现良好的识别性能。CNN结构将三种建设性思想结合起来,以承受某种程度的规模,移位和失真不变性,即卷积操作,激活操作和池化操作。因此,VLR提高了对各种复杂成像条件的高分类性能和鲁棒性。

## 具有CNN的自适应加权MTL

受到[10]的启发,我们应用MTL对车辆标志品牌进行分类并同时预测9个属性。如图1所示,第六层是一个完全连接的层,具有512个神经元,并被分成10个分支。我们提出一种自适应加权MLT来增加类别间的距离。当MTL运行时,我们将每项任务的损失的线性组合最小化：



在这个等式中,Li是第i个任务的损失,αi是第i个任务在总损失中的权重,L许多研究人员[10,14]在训练CNN时修正α。另一方面,Wu[11]用α初始化α,并根据训练中的验证性能进行修改。我们引入冲量惯性概念来平滑训练。自适应加权MLT可以在下面进行公式化。



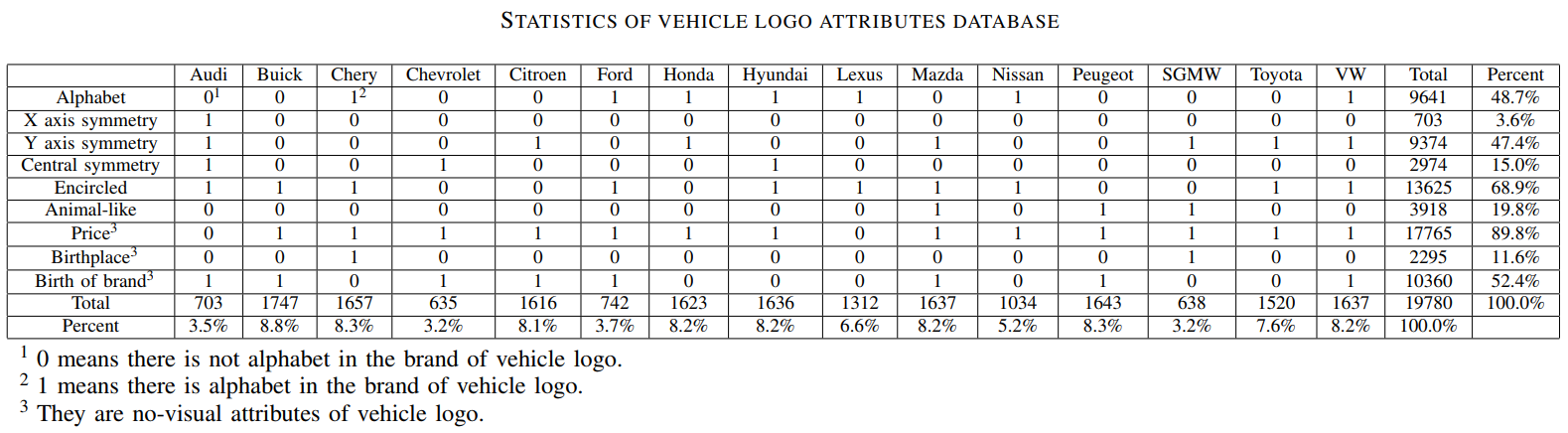
其中β\*αi是每个时期任务权重的变化,β是控制这个时间误差影响下一次任务权重的量的实验参数。在我们的实验中,β被设置为0.1。在等式7,Ei是每个批次第i个任务的错误。在初始状态下,所有的Ei初始值都是0.1。实验证明自适应加权MTL是有效的

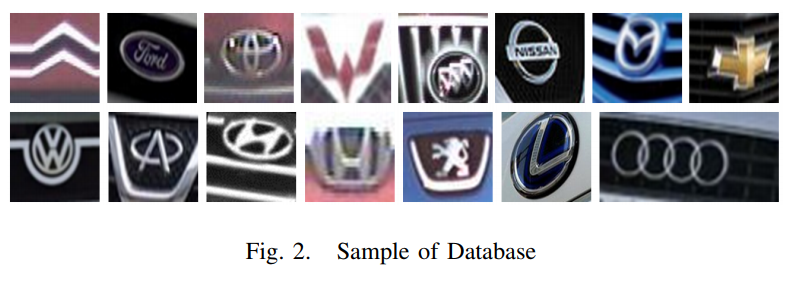
# 实验

我们评估了车辆标识分类任务中的系统性能,III-B和标识属性预测III-C。我们对CNN的训练遵循[10],即随机梯度下降,批量为128例,学习率为0.01。在我们的内部实施中,每个批次的培训都需要大约20分钟的时间在一台英特尔至强E5CPU上,并且网络在大约100个批次时收敛。

## 车牌数据集

我们将厦门大学汽车标志识别数据库[6]扩展到一个由15个品牌组成的更大的数据集,其中包含6个视觉属性和3个无视觉属性。从交通监控中捕获数千辆户外车辆正面图像。然后,我们以前的工作[15]大致检测到了标志。最后,生成了来自十五个制造商的19780幅图像的车辆标志数据库。图2显示了数据库中15家制造商的一些样本图像。此外,我们将9种分类为视觉属性和非视觉属性的分类属性标为[16]。所有分割的图像均匀归一化为64\*64像素。车辆标识识别数据库的统计列于表一。





## 车牌识别

表1中描述的数据库用于评提出议的系统。80％,10％,其余图像用作训练样本,验证和测试样本。在实验中,我们比较了四种不同的方法来证明自适应加权MTLCNN在VLR任务中的好处：

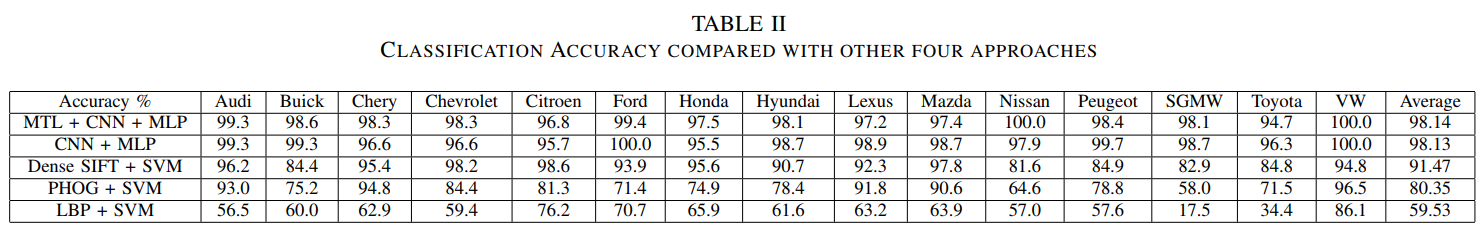
•单任务CNN

•密集SIFT和SVM

•PHOG和SVM

•LBP和SVM

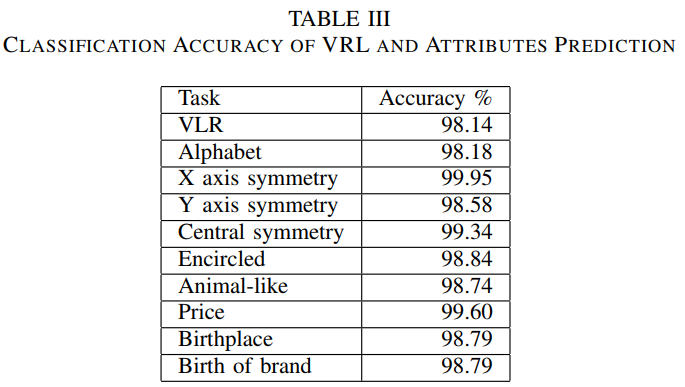
结果如表II所示。我们将训练数据在训练过程中的批次次数限制在120次。可以看出,CNN总体上优于传统特征提取器,LBP,PHOG和密集SIFT。比较我们的MLTCNN和单一任务CNN,尽管它们都获得了高性能,但我们的系统可以同时预测另外九个二进制属性。

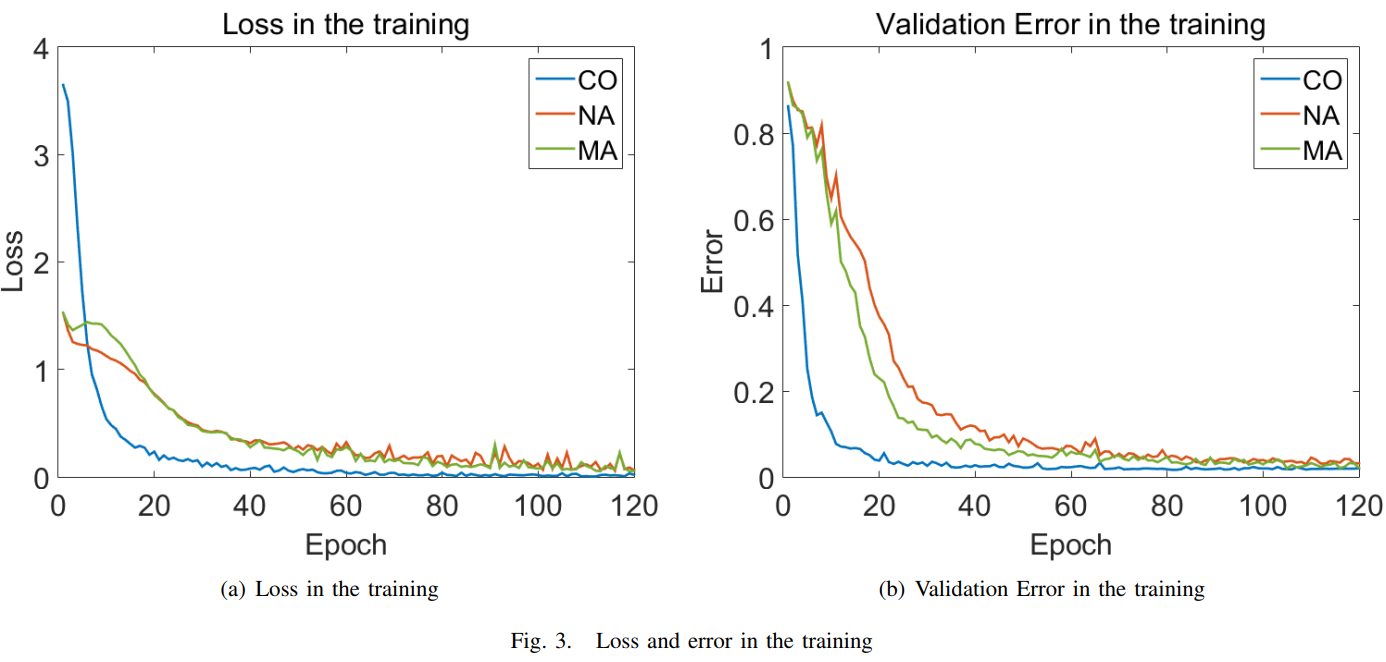


## 用自适应加权MTL预测属性

我们还通过使用VLR预测九个二进制属性来评估我们数据集上的自适应加权MTL CNN。 属性包括六个视觉属性和三个非视觉属性。 表3列出了结果，表明我们的算法对于视觉属性和非视觉属性都获得了高预测性能。

我们还将我们的算法与固定加权训练策略[10]与单任务CNN作为基线进行了比较。 从训练过程可以看出，图3，我们的算法具有更低和更稳定的损失和验证错误。 与单任务CNNs相比，我们的算法性能相似。





# 总结

本文提出了一种基于CNN的车辆标志属性预测的VMR方法。 我们的系统通过MTL同时识别车辆标识并预测标识属性。 我们提出了一种自适应加权MTL来增加类间距离。 通过实验，该算法加速并稳定了多任务CNN的训练过程。 最终结果表明车辆标识识别和标识属性预测取得了令人满意的性能。

进一步的工作有两个方向。 一个是完成车辆标识的检测和识别。 另一种是通过组合交替方式和同时方式来改善自适应任务权重MTL