

基于迁移学习的交通标志识别系统设计

刘玉民^a, 张雨虹^b

(唐山学院 a. 机电工程学院; b. 计算中心, 河北 唐山 063000)

摘要: 针对自动驾驶领域对交通标志识别的需求, 设计了一种基于迁移学习的交通标志识别系统。该系统采用了预训练的 MobileNetV3(去掉输出层)作为特征提取网络, 然后添加两个自定义的全连接层以实现信号的分类和输出。由于采用迁移学习方法, 深度学习网络中需训练的模型参数大幅减少, 训练所需时间大为缩短。该系统使用经典的中国交通标志数据库(CTSDB)中的数据作为交通标志的训练数据和测试数据, 训练结果表明, 损耗低至 0.024 3, 准确率高达 99.88%; 测试结果表明, 可以对 58 类交通标志进行识别, 准确率为 55.3%。

关键词: 迁移学习; 交通标志; 识别系统; 自动驾驶

中图分类号: TP391.41; TP183 **文献标志码:** A **文章编号:** 1672-349X(2023)06-0001-04

DOI: 10.16160/j.cnki.tsxyxb.2023.06.001

Design of Traffic Sign Recognition System Based on Transfer Learning

LIU Yu-min^a, ZHANG Yu-hong^b

(a. School of Mechanical and Electrical Engineering; b. Computation Center, Tangshan University, Tangshan 063000, China)

Abstract: To meet the demand for traffic sign recognition in the field of autonomous driving, a traffic sign recognition system based on transfer learning has been designed, which uses pre-trained MobileNetV3 (excluding the output layer) as the feature extraction network, and then adds two self-defined fully connected layers for signal classification and output. Due to the adoption of transfer learning, the number of model parameters to be trained in the deep learning network are significantly reduced, resulting in a considerable reduction in training time. The system uses data from the classic Chinese Traffic Sign Database (CTS-DB) as training and testing data for traffic signs. The training results indicate a low loss of 0.0243 and a high accuracy of 99.88%, while the testing results show that this traffic sign recognition system can recognize 58 different classes of traffic signs, with an accuracy rate of 55.3%.

Key Words: transfer learning; traffic sign; recognition system; autonomous driving

0 引言

道路交通标志作为重要的道路交通安全附属设施, 可向驾驶人或自动驾驶系统提供各种

引导和约束信息, 以保障行车安全。因此, 基于计算机视觉的交通标志识别在自动驾驶领域极为重要。现有的交通标志识别方法大体可分为

作者简介: 刘玉民(1975—), 男, 河北唐山人, 教授, 博士, 研究方向为机器学习;

张雨虹(1978—), 女, 河北唐山人, 实验师, 硕士, 研究方向为计算机软件。

三类^[1]:第一类为采用传统图像处理技术的识别方法,此方法主要利用颜色、形状等特征,采用模板匹配的方法进行识别;第二类为采用机器学习技术的识别方法,此方法通过提取方向梯度直方图(HOG)、哈尔(Harr)等特征,采用支持向量机(SVM)、BP神经网络、决策树等分类器完成识别;第三类为采用深度学习技术的识别方法(如 RCNN, Fast RCNN^[2], Faster RCNN^[3], SSD, YOLOv3, YOLOv4 等),此方法采用深度卷积神经网络的架构,利用海量样本数据进行训练,从中提取最相关的、最有用的特征,并自动挖掘其他适合预测的特征,然后利用这些特征完成识别。此方法因通过从数据中直接学习特征,替代了复杂的人工提取特征的过程,因此得到了较好的应用。但是采用深度学习技术的识别方法,如果可供训练的数据集中数据较少,则无法从零开始完成对深度卷积神经网络的训练任务。对于此缺点,可以使用迁移学习方法解决,即利用预训练的深度网络模型(如 VGG, ResNet, DenseNet, InceptionV3, MobileNetV3 等)作为特征提取网络,在保留其基本结构和参数不变的基础上增加一些新的、自定义的、可训练的网络层。本设计正是应用了迁移学习方法,设计了一种交通标志识别系统。该系统采用预训练的 MobileNetV3(去掉输出层)作为特征提取网络,然后添加两个全连接层以实现信号的分类和输出。这样的学习方法能够减少训练样本数据的需求量,简化网络架构设计,而且还能够缩短网络参数的训练时间。

1 网络架构

本设计主要面向无人驾驶系统的交通标志识别,因此对目标检测的实时性要求较高,且不能占用过多的软硬件资源。基于上述考虑,选用了轻量级的目标检测网络 MobileNetV3^[4],它是在 MobileNetV1, MobileNetV2 基础上经过系列改进而成的,主要面向移动应用场合的视觉识别与检测,其参数已在 ImageNet 数据集上完成了预训练,默认情况下可对 1 000 类目标进行分类。

本设计使用的迁移学习网络架构由如下三部分串联组成。

(1) MobileNetV3 层(去掉输出层):已剪去 MobileNetV3 完整网络上层的全连接层,输出信号为 960 个,共有 2 996 352 个参数,全部设置为不可训练参数。

(2) dense 层:新增自定义的由 98 个神经元组成的全连接层,激活函数为修正线性单元(relu),输出信号为 128 个,共有 $(960 + 1) \times 128 = 123\ 008$ 个参数,全部为可训练参数。

(3) dense_1 层:新增自定义的由 58 个神经元组成的全连接层,激活函数为 softmax,输出信号为 58 个,对应 58 个不同类别,共有 $(128 + 1) \times 58 = 7\ 482$ 个参数,也全部为可训练参数。

由此可见,该迁移学习网络共有 3 126 842 个参数,其中有 130 490 个参数是需要训练的,需训练参数占参数总量的 4.17%,说明仅需使用较少的训练数据和训练时间即可完成训练任务,这是迁移学习方法最大的优势。

2 数据集

为了对构造好的迁移学习网络中的参数进行训练,需要选用或自己构建所需的交通标志数据集。目前可用的国内外交通标志数据集主要有国内的 CTSDB, CCTSDB, ApolloScape, TsinghuaTT100K;德国的 GTSDB, BSTLD;美国的 LISA 等。本设计使用经典的中国交通标志数据库(CTSDB)^[5]中的数据作为交通标志的训练数据和测试数据。CTSDB 包含由 4 170 个图像组成的训练数据库和由 1 994 个图像组成的测试数据库两个子交通标志数据库,共计 6 164 个 58 类交通标志图像。由于 CTSDB 中包含 58 类交通标志图像,所以所设计的交通标志识别系统输出信号的数量确定为 58 个,据此确定 dense_1 层中神经元的数量为 58 个。CTSDB 数据库中部分示例图像如图 1 所示。

从图 1 可以看出,CTSDB(包含两个子库)中的图像名称是有意标记过的,即文件名的前三个字符(000—057)对应图像所属的类别(0—57)。此外,数据库中的图像均为便于网络传输的 png 格式,但它无法直接提供彩色图像的



图1 CTSDB 数据库中部分示例图像

RGB 分量,而 MobileNetV3 网络输入层要求输入包含 RGB 三个分量的彩色图像,因此需对 CTSDB 中的图像进行格式转换,即将 png 格式转换为 jpg 格式文件。实现此类格式转换的方法比较多,可以借助第三方图像处理软件完成,也可以利用 python 编写程序调用 PIL 库或者 opencv 库中相关函数来实现。本设计是通过调用 PIL 库中 `image.open()` 和 `image.save()` 两个函数来实现格式转换的。

3 模型训练和保存

在构建网络模型、下载 CTSDB 中数据集、转换图像格式等环节之后即可对网络模型进行训练和保存^[6]。

模型训练过程可使用 `tf.keras` 提供的 `model.fit()` 函数完成,在这之前需要确定训练所需的各种参数,如 `epochs`, `steps_per_epoch`, `batch_size` 等。图 2 给出了模型训练过程中 epoch1—10 的用时、损耗(loss)、准确率(accuracy)的数值。由图 2 可见,在训练结束时损耗已低至 0.024 3,准确率高达 99.88%,收到了较好的训练效果。

训练结束后,可使用 `model.save()` 函数来保存训练好的用于交通标志识别的深度学习模型(`mobile_net_TrafficSign`),以供调用。保存的网络模型中包含 `assets`, `variables` 两个

```
Epoch 1/10
105/105 [====] - 13s 86ms/step - loss: 1.9313 - accuracy: 0.5468
Epoch 2/10
105/105 [====] - 9s 83ms/step - loss: 0.6595 - accuracy: 0.8411
Epoch 3/10
105/105 [====] - 9s 83ms/step - loss: 0.3365 - accuracy: 0.9263
Epoch 4/10
105/105 [====] - 9s 89ms/step - loss: 0.2023 - accuracy: 0.9592
Epoch 5/10
105/105 [====] - 10s 91ms/step - loss: 0.1344 - accuracy: 0.9766
Epoch 6/10
105/105 [====] - 10s 92ms/step - loss: 0.0847 - accuracy: 0.9922
Epoch 7/10
105/105 [====] - 9s 88ms/step - loss: 0.0591 - accuracy: 0.9949
Epoch 8/10
105/105 [====] - 9s 88ms/step - loss: 0.0430 - accuracy: 0.9970
Epoch 9/10
105/105 [====] - 9s 83ms/step - loss: 0.0308 - accuracy: 0.9985
Epoch 10/10
105/105 [====] - 9s 85ms/step - loss: 0.0243 - accuracy: 0.9988
```

图2 模型训练过程

文件夹以及 `fingerprint.pb`, `keras_metadata.pb`, `saved_model.pb` 三个文件。

4 模型测试

交通标志识别系统模型 `mobile_net_TrafficSign` 在正式使用前还需进行模型测试,以检验或评价其识别性能的优劣。此处应用 `mobile_net_TrafficSign` 模型对 CTSDB 数据库测试集中的 1 994 个图像进行识别。测试结果表明,单个交通标志图像的识别平均耗时仅约 34 ms,较好地满足了系统实时性的要求。在定量描述系统识别性能时,使用 `sklearn.metrics` 中的 `accuracy_score`, `confusion_matrix` 和 `classification_report` 工具来统计识别的准确率、混淆矩阵和分类报告。本系统的准确率和混淆矩阵的统计结果如图 3 所示。由图 3 可见,系统在对 58 类共计 1 994 个的测试图像进行分类时,识别的准确率为 55.3%,可以对交通标志进行识别。受篇幅所限,图 3 中仅显示了第 0,1,2,55,56,57 六类的混淆矩阵数值。在混淆矩阵中,第 1 行第 1 列的元素 10,表明图像真实类别为 0 类,且被正确识别为 0 类的图像数为 10 个;

第 1 行第 3 列的元素 2,表明图像真实类别为 0 类,却被错误识别为 2 类的图像数为 2 个。

```

Accuracy:0.553
Confusion matrix:
[[10  0  2 ...  0  0  0]
 [ 0  6  0 ...  0  0  0]
 [ 2  4 20 ...  0  0  0]
 ...
 [ 0  0  0 ... 54  0  0]
 [ 0  0  0 ...  0 28  0]
 [ 0  0  0 ...  0  0  4]]

```

图 3 准确率和混淆矩阵

分类报告如表 1 所示。表中仅给出了 0—6 类图像对应的 precision(查准率)、recall(召回率)、f1-score(查准率和召回率的调和平均数值)以及 support 数值(属于此类的图像的个数)。表 1 中,0—5 类的识别分值较高,但 6 类的识别分值均为 0,这说明系统在对该类交通标志图像进行识别时出现了较大的问题。因此,系统的识别性能还有待进一步提升。

表 1 分类报告

class	precision	recall	f1-score	support
0	0.38	0.71	0.50	14
1	0.12	0.17	0.14	12
2	0.20	0.17	0.18	60
3	0.32	0.62	0.42	84
4	0.33	0.59	0.42	58
5	0.27	0.12	0.17	50
6	0.00	0.00	0.00	30
...	...	...	...	...

5 结论

所设计的交通标志识别系统,使用迁移学习方法,保留特征提取网络 MobileNetV3(去掉输出层)的基本结构和参数不变,仅对新增两个全连接层进行了设计和训练,使得训练参数的数量大幅减少、训练时间大为缩短。系统采用经典的中国交通标志数据库(CTSDb)中的数据作为交通标志的训练数据和测试数据,训练结果表明,损耗低至 0.024 3,准确率高达 99.88%;测试结果表明,可以对 58 类交通标志进行识别,准确率为 55.3%。

客观地讲,交通标志识别系统识别的准确率还有待进一步提升,今后会从以下几个方面进行改进:(1)增加对数据集中图像的预处理,如增加直方图均衡、滤波处理等,最大限度增强图像有关信息的可检测性,提升深度神经网络提取特征的可靠性;(2)提升 CTSDb 中某些类别图像的数量和质量,如第 19 类即“解除限速标志 50”类中图像清晰度不够且仅有 4 张图像,造成系统对其识别性较差;(3)在训练深度卷积神经网络时有可能出现了过拟合现象,即训练集性能远远好于测试集性能,应通过交叉验证避免;(4)可在此基础上增加微调(fine-tuning)环节,对全网络(包括 MobileNetV3)的参数进行微调,以进一步提升系统的识别性能。

参考文献:

- [1] 王强. 基于深度学习的交通标志检测和识别[D]. 南京:南京邮电大学,2021.
- [2] GIRSHICK R. Fast R-CNN[C]. IEEE International Conference on Computer Vision(ICCV),2015:1440-1448.
- [3] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks [C]. Neural Information Processing Systems(NIPS),2015:1-14.
- [4] HOWARD A, SANDLER M, CHU G, et al. Searching for mobilenetv3[C]. Proceedings of the IEEE International Conference on Computer Vision,2019:1314-1324.
- [5] 中国科学院自动化研究所模式识别国家重点实验室. 中国交通标志数据库[EB/OL]. (2015-03-26). <http://www.nlpr.ia.ac.cn/pal/trafficdata/recognition.html>.
- [6] 平震宇,匡亮. TensorFlow 深度学习实例教程[M]. 北京:机械工业出版社,2022:125-130.

(责任编辑:李秀荣)