李超军. 基于迁移学习的小样本交警关键手势识别算法研究[J]. 智能计算机与应用,2026,16(7):77-85. DOI:10.20169/j.issn.2095-2163.25121803

基于迁移学习的小样本交警关键手势识别算法研究

李超军
(山西能源学院 计算机与信息工程系, 山西 晋中 030600)

摘 要: 随着智能交通系统的发展,交警手势识别成为实现人机协同指挥的重要研究方向。然而,由于手势数据获取困难、样本数量有限,传统深度学习方法在小数据集上易出现过拟合问题。本文提出一种基于 VGG16 迁移学习的交警关键手势识别算法,通过对自采集的小样本关键手势数据进行数据增强,并结合 fine-tune 策略优化模型性能。实验结果表明,该方法在识别准确率、收敛速度和泛化能力方面均优于传统的从头训练和仅训练顶层分类器的方法,分类准确率达到 97.73%,具有良好的实际应用前景。

关键词: 交警关键手势; 迁移学习; VGG16; 数据增强; fine-tune
中图分类号: TP391.41 **文献标志码**: A **文章编号**: 2095-2163(2026)07-0077-09

Research on a small-sample traffic police key gesture recognition algorithm based on transfer learning

LI Chaojun
(Department of Computer and Information Engineering, Shanxi Institute of Energy, Jinzhong 030600, Shanxi, China)

Abstract: With the development of intelligent transportation systems, traffic police gesture recognition has become an important research direction for achieving human-machine collaborative command. However, due to the difficulty in obtaining gesture data and limited sample size, traditional deep learning methods are prone to overfitting on small datasets. This paper proposes a traffic police key gesture recognition algorithm based on VGG16 transfer learning, by enhancing the self-collected small-sample key gesture data and combining the fine-tune strategy to optimize model performance. Experimental results show that this method outperforms traditional methods that train from scratch and only train the top-level classifier in terms of recognition accuracy, convergence speed, and generalization ability. The classification accuracy reaches 97.73%, indicating good practical application prospects.

Key words: key gestures of traffic police; transfer learning; VGG16; data augmentation; fine-tune

0 引 言

交警手势识别是智能交通系统中的关键任务之一,广泛应用于自动驾驶、交通监控与人机协同指挥等场景[1]。由于交警手势具有高度的动态性、视角多样性和环境依赖性,传统基于手工特征的图像分类方法难以满足实际应用对准确性与鲁棒性的要求[2]。近年来,卷积神经网络(CNN)在图像识别领域取得了突破性进展,其表现要显著优于传统机器学习算法[3]。然而,CNN 模型的性能高度依赖于大规模高质量的训练数据,而交警手势数据在实际采集中面临样本稀缺、类别不平衡和场景多样性不足等问题,直接从头训练模型易导致过拟合,且计算成本高昂[4-6]。

针对上述问题,本文聚焦于 8 种交警关键手势的分类任务,提出一种基于迁移学习的小样本学习方案。具体而言,本文在自采集的交警手势小数据集基础上,采用多种图像增强策略(如亮度、对比度、饱和度调整、裁剪、加噪及标准化等)对原始数据进行扩充,有效提升数据多样性与模型泛化能力。进一步地,引入在 ImageNet[7] 上预训练的 VGG16 模型[8],结合 fine-tune 策略对网络顶层进行微调,充分利用源域模型的通用特征提取能力,同时适应目标域手势数据的特异性特征,从而在小数据集上实现高精度、低成本的分类模型训练。

基金项目: 山西能源学院校级教学改革创新项目(NJ202433)。

作者简介: 李超军(1993—),男,硕士,助教。主要研究方向:图像处理,机器学习,深度学习等。Email:lichaojun1130@qq.com。

收稿日期: 2025-12-18

1 交警关键手势数据准备

1.1 数据采集

为使采集的数据更加丰富,本文在不同路口、光照环境下采集交警关键手势图片,使图片在光照、背景上存在差异。采集时,用摄像机对 4 名实验者的 8 种交警关键手势拍照,共获 2 000 张,每种约 250 张,分辨率 800×480,图 1 为部分采集图片。

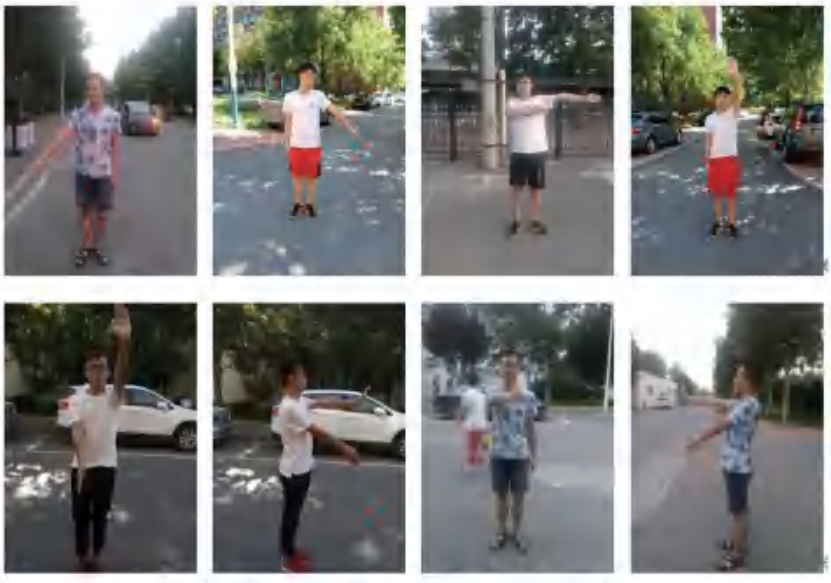


图 1 交警关键手势数据

Fig. 1 Key gesture data of traffic police

1.2 数据增强

数据增强技术是通过对数据进行特殊处理从而增加数据量的一种技术[9]。在深度学习模型训练中,充足且丰富的数据可以防止模型对数据的过拟合,所以要确保数据量足够大和数据集质量足够好[10]。

本文采集的 2 000 张交警手势图片,对训练神经网络模型通常不够。为此,本文采用 TensorFlow 中的图片预处理技术进行数据增强,方式包括裁剪、亮度与对比度调整、加噪声、标准化处理等,因本文训练的模型网络对肢体位置及角度敏感,故不采用翻转、仿射变换。具体数据增强方案如下:

(1)裁剪。在交警关键手势图片中,交警的双臂属于感兴趣区域。在裁剪过程中,必须保留感兴趣区域,而感兴趣区域一般位于图片的中央位置,所以找到图片中央位置向四周扩展 L 和 H,L 和 H 表示裁剪后的图片长和宽的 1/2。裁剪前后的图片如图 2 所示。

(2)亮度、对比度、饱和度调整。在很多图像识别应用中对图像进行亮度、对比度和饱和度的调整都不会影响到识别结果[11]。所以通过调整训练图片的这些属性,可以使得最终模型对含有这些因素的图像更具有泛化能力[12]。对图像进行亮度、对比度和饱和度调整后的结果,如图 3~图 5 所示。



(a) 裁剪前 (b) 裁剪后

图 2 裁剪操作

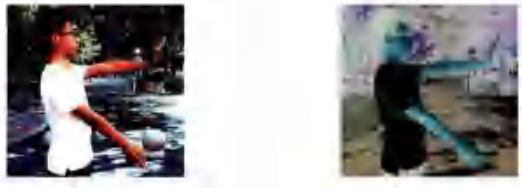
Fig. 2 Cropping operation



(a) 亮度增加 (b) 亮度减小

图 3 亮度调整

Fig. 3 Brightness adjustment



(a) 对比度增加 (b) 对比度减小

图 4 对比度调整

Fig. 4 Contrast adjustment



（a）饱和度增加　　（b）饱和度减小

图 5　饱和度调整

Fig. 5　Saturation adjustment

（3）加噪（高斯噪声和椒盐噪声）。图像具有冗余信息量大的特征，利用该特征可以对图像进行加噪处理，如果一个训练好的模型能够将加噪后的图片进行准确的识别，那么该模型必然具有很强的鲁棒性[13]。本文分别利用高斯噪声和椒盐噪声对图像进行处理。处理结果如图 6 所示。



（a）加入高斯噪声　　（b）加入椒盐噪声

图 6　加噪处理

Fig. 6　Noise processing

（4）数据标准化处理。在图像中，相邻像素之间具有很强的相关性，当这些图像像素输入模型中进行训练时，输入是冗余的，经过数据标准化处理可以去除这些冗余，使得模型学习到的特征之间的相关性降低[14]。数据的标准化处理能够取消量纲，使得训练模型时梯度始终朝着最小值的方向前进，加速收敛[15]。本文通过将图像的亮度均值变为 0、方差变为 1 的方式对图像进行标准化处理。计算公式为：

$$\boldsymbol{x}_{\text{normalization}} = \frac{\boldsymbol{x} - \boldsymbol{\mu}}{\sigma} \tag{1}$$

其中，$\boldsymbol{x}$ 表示图像像素矩阵；$\boldsymbol{\mu}$ 表示图像的均值；σ 表示图像的方差。标准化处理结果如图 7 所示。



图 7　标准化处理

Fig. 7　Standardization processing

通过以上数据增强的方式处理数据集后，数据集得到了成倍的增加，由 2 000 张扩充到 20 000 张。数据集增强前后统计见表 1。

表 1　数据集增强前后统计

Table 1　Statistics before and after data set enhancement　张

数据集	扩充前数量	扩充后数量
直行手势	259	2 590
停止手势	266	2 660
左转弯手势	238	2 380
左转弯待转手势	219	2 190
右转弯手势	254	2 540
变道手势	273	2 730
靠边停车手势	255	2 550
减速慢行手势	236	2 360
总计	2 000	20 000

2　迁移学习

2.1　迁移学习理论

近年来，CNN 在图片分类识别领域取得重要成果。通常，训练优秀的 CNN 模型需海量数据支撑，只有从中提取足够丰富特征，才能拥有良好鲁棒性。若缺乏充足训练样本，要训练出高识别率模型，无疑是一项难以达成的任务[16-17]。事实正是如此，在小数据集上训练得到的 CNN 模型通常是不理想的[18]。

为解决小数据集上训练的深度学习模型识别率低的问题，专家学者随即陆续开启了任务间模型迁移方法研究。研究表明，已训练好的深度学习分类模型可用于新任务，不同领域的 CNN 模型经网络微调也能用于新数据集，此方式被称作迁移学习[19-22]。图 8 展示了迁移学习的过程，迁移学习就是将源任务中学习到的知识迁移到目标任务中，从而有效解决问题，达成任务目标。

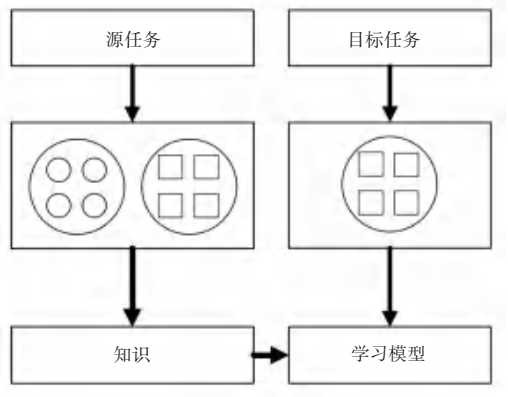


图 8　迁移学习过程

Fig. 8　Process of transfer learning

2.2 迁移学习方法

迁移学习通常有以下 3 种模型迁移方式：

(1)方式 1：只训练顶层分类器。将源任务中的模型用作特征提取工具，保留预训练模型中除顶层分类器之外的参数，将输出层换成目标任务适用的输出层。

(2)方式 2：重新训练。只将预训练好的模型的网络结构迁移过来，重新初始化参数，并在目标任务数据集上进行训练。

(3)方式 3：fine-tune。对预训练好的模型的部分网络层进行微调，将底层网络参数进行冻结，微调顶部网络层。将底层网络参数冻结可以达到防止过拟合的效果，而微调顶部网络层可以学习到数据更抽象的特征。

在实际应用中如何选取迁移学习方法与数据集大小和数据集相似度(源任务数据集与目标任务数据集的相似度)有关。

(1)对于小数据集、数据集相似度高的情况。由于源任务与目标任务数据集高度相似(特征高度相似)，所以只需要将源任务中的模型作为特征提取器即可提取到目标数据集的特征，此后只训练网络的顶层分类器即可。

(2)对于小数据集、数据集相似度不高的情况。在小数据集从头训练的方式容易造成过拟合且数据集相似度不高，不宜采用只训练顶层分类器的方式。同时，考虑到神经网络的底层学习到的特征具有通用性，而顶层学习到的特征更能代表特定数据集的特征，所以可以采用 fine-tune 的方式，将预训练好的模型的前 k 层网络参数进行冻结，只训练后 $n-k$ 层网络的参数。

(3)对于大数据集、数据集相似度不高的情况。由于数据集比较大，模型的训练过程效率会下降，并且数据集的相似度不高，预训练模型中保留的参数不能有效地提取目标数据集的特征，因此对预训练模型的参数重新初始化并从头开始训练模型将是最好的方式。

迁移学习因能解决小数据集训练模型识别率低的问题，得到了许多专家学者的广泛关注。使用迁移学习，只要将在某一个数据集上表现优异的模型应用到另一个问题上，就可以取得不错的性能[23-25]。从头训练一个卷积神经网络会遇到过拟合或者损失函数收敛慢等问题，而迁移学习由于其参数是预训练好的，只需要对参数进行微调即可，这样一来就加快了损失函数的收敛速度[26]。迁移学习通过微调预训练好的模型不仅加快了训练速度，也提高了分类准确率，避免了数据集不足带来的短板问题[27-28]。预训练模型一般是在一些图像分类应用中已经取得了良好成果，比如 AlexNet[29]、VGGNet[8]、GoogleNet[30] 等网络在 ImageNet 数据集上表现出卓越性能，这些模型经常被用来作为迁移学习的预训练模型。

3 基于 VGG16 的迁移学习模型

3.1 VGG16 模型

VGGNet 作为卷积神经网络模型[31]通过反复使用 3×3 和 2×2 的卷积核，构筑了 16~19 层深的卷积神经网络。VGGNet 由于其强大的拓展和泛化能力，经常被用来作为图像特征提取工具[32-33]。基于此，本文将选择在 ImageNet 数据集上表现优异的 VGG16 模型作为迁移学习对象，下面对 VGG16 模型进行介绍。

VGG16 模型如图 9 所示，VGG16 网络模型由 13 个卷积层与 3 个全连接层构成，13 个卷积层可以分为 5 个卷积块，每个卷积块有 2~3 个卷积层与 1 个最大池化层构成。在 5 个卷积块的头部接有一个 224×224×3 的输入层，在尾部有 3 个全连接层和输出层。

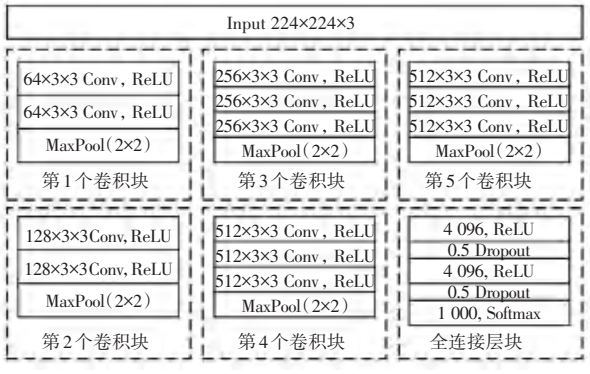


图 9 VGG16 模型

Fig. 9 VGG16 model

3.2 迁移学习网络结构设计

本文采集的交警关键手势数据量少，而由第 2 节可以得出，迁移学习可以有效解决小数据集上模型识别率低的问题，所以本文采用了迁移学习方式对交警关键手势进行分类识别。本文选用在 ImageNet 数据集上表现优异的 VGG16 模型作为迁移对象。

ImageNet 数据集与交警关键手势数据集相比，相似度不大，且关键手势数据量小。根据 2.2 节对

迁移学习方法的分析，本文将采用 fine-tune 网络的方式对交警关键手势数据集进行训练。同时，为了能够验证 fine-tune 训练方式在小数据集上进行迁移学习的优势，本文还将另外 2 种迁移学习方式应用到了交警关键手势数据集上进行训练，从而形成对比。

(1) fine-tune 网络结构。在卷积神经网络中，底部的卷积层提取的是数据的通用性特征，顶部的卷积层提取的是数据的抽象特征，本文在对 VGG16 进行 fine-tune 的时候冻结底部 4 个卷积块，保留其在 ImageNet 数据集上训练得到的权重参数，从第 5 个卷积块起进行重新训练。对于输出层，因 VGG16 原用于 ImageNet 分类，最后一层有 1 000 个节点，本文对 8 种交警手势进行分类，故替换为 8 个节点。搭建好的网络结构如图 10 所示。

(2) 只训练顶层分类器。将在 ImageNet 数据集上预训练好的 VGG16 网络模型只训练顶层的分类器，并保留前面所有卷积层的权重参数，不进行训练，只训练顶层分类器，将输出层的 1 000 个结点改为 8 个结点，网络结构如图 11 所示。

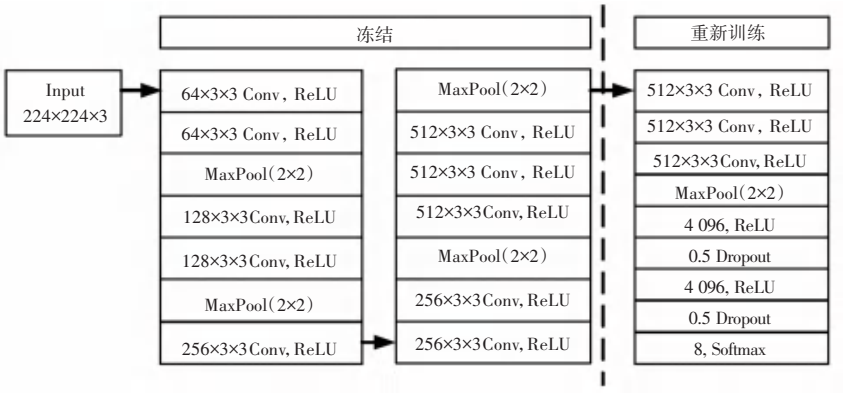


图 10　fine-tune 迁移学习

Fig. 10　fine-tune transfer learning

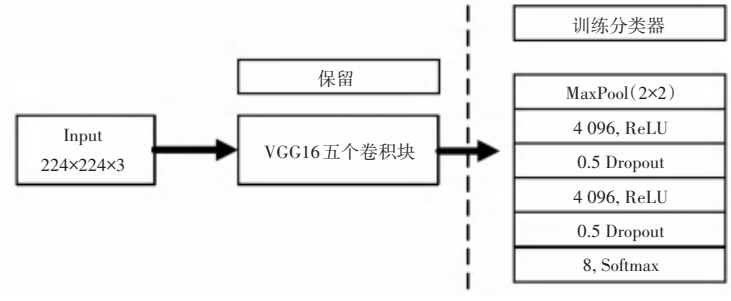


图 11　只训练顶层分类器

Fig. 11　Train only the top level classifier

(3) 从头训练。对 VGG16 网络从头进行训练，对网络中的参数进行随机初始化，将输出层的 1 000 个结点改为 8 个结点。网络结构如图 12 所示。

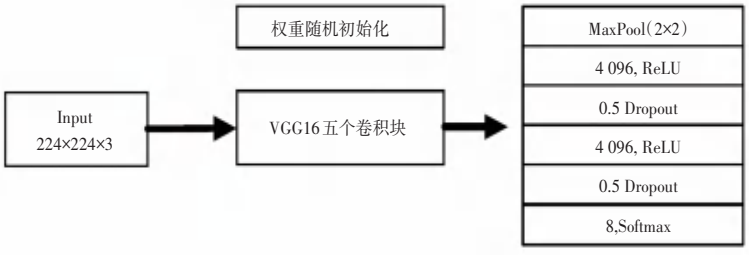


图 12　重新训练

Fig. 12　Retraining

3.3 迁移学习的对比实验

对 3.2 节中的 3 种网络结构进行实验,在做实验之前需要进行数据集的准备、参数的设置等工作。

(1)数据集准备。为了适应 VGG16 网络的输入格式,本文将数据的大小进行了调整,调整过程中使用了 TensorFlow 框架中提供的图像大小调整函数 tf. image. resize_images( ),将图片调整为 224×224 的 RGB 图片。数据集中共有 20 000 张图片,本文将数据集平均分成 5 份存放,每份对应一个文件夹,每个文件夹中又有 8 个文件夹,分别存放 8 种交警关键手势图片,如图 13 所示。数据集平均分成 5 份是为了后续做交叉验证使用。



图 13 数据集存放格式

Fig. 13 Dataset storage format

(2)训练参数设置。训练过程中,3 种网络的训练参数设置见表 2。

表 2 参数设置

Table 2 Parameter settings

参数名称	参数值
学习率(Learning_rate)	0.000 1
Batch	50
epoch	400
分类函数	SoftMax
损失函数	交叉熵损失函数(cross_entropy)

(3)K 折交叉验证。K 折交叉验证是将数据集平均分成 K 份,每次训练时,将其中 $K-1$ 份作为训练集,1 份作为验证集,最后将在验证集上得出的结果求平均得到最后的结果[34]。数学定义公式为:

$$E = \frac{1}{K}\sum_{i=1}^{K} e_i \tag{2}$$

其中,e_i 表示在第 i 次交叉验证中,验证集上得到的结果。

K 折交叉验证可以避免过拟合和欠拟合的发生,得到的结果具有说服力。本文将数据集中的 20 000 张图片,平均分成 5 份,做 5 次交叉验证实验,如图 14 所示。

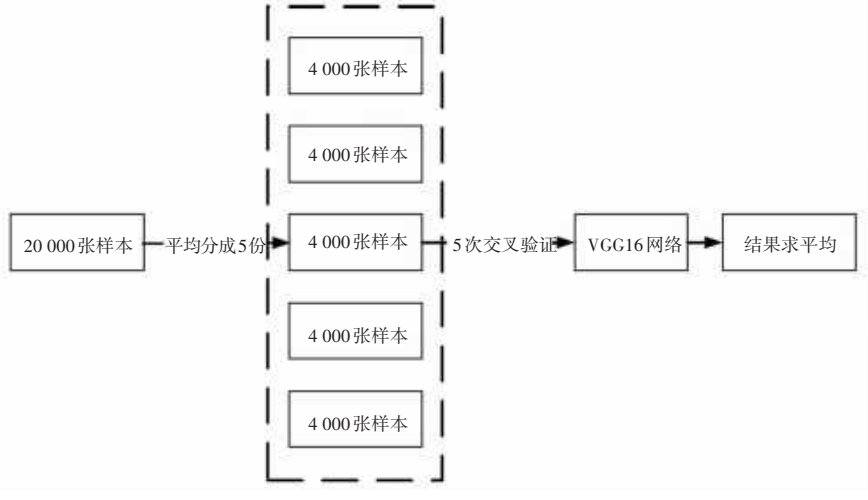


图 14 交叉验证

Fig. 14 Cross validation

4　实验结果及分析

4.1　3 种迁移学习网络实验结果对比

本文在 Pycharm 环境下搭建了 TensorFlow 框架以及 Keras 框架，对 3 种迁移学习方式进行实验，每种方式分别做 5 次交叉验证，在验证集上得到的结果见表 3。

表 3　交叉验证的实验结果

Table 3　Experimental results of cross validation　%

网络类型	第 1 次	第 2 次	第 3 次	第 4 次	第 5 次	平均值
fine-tune 网络	98.67	96.54	97.72	98.10	97.62	97.73
只训练顶层分类器	83.47	82.63	84.65	82.18	83.67	83.32
从头训练	79.78	80.45	81.74	80.65	80.58	80.64

从表 3 中可以看出 fine-tune 网络的实验结果更优，达到了 97.73%。而另外 2 种网络的实验结果较差。对于 fine-tune 方式的迁移学习网络，通过冻结 VGG16 模型的前 4 个卷积块学习到了目标数据集的一般性特征，通过重新训练最后一个卷积块学习到了目标数据集更抽象的特征，最终使得模型获得了比较优异的识别率。对于只训练顶层分类器的方式，由于 ImageNet 数据集与交警关键手势数据集相似度较小，且没有对原有网络参数进行重新训练，使得网络只能学习到目标数据集一般性特征，学习不到比较优秀的抽象特征，而顶层分类器几乎没有特征学习的能力，使得最终的识别率低于 fine-tune 网络。对于从头训练的网络结构，由于数据集有限，导致网络训练得到的参数无法很好地提取目标数据集的一般性和特殊性特征，最终识别率也不高。

表 3 只是以数字化的形式比较了 3 种网络的实验结果，下面从性能方面对 3 种网络进行比较。图 15 和图 16 分别展示了 3 种网络模型在训练过程中的准确率曲线图和损失曲线图。从图 15—16 中可以看出 fine-tune 网络的收敛速度最快，在 50 轮时，曲线开始趋于平缓，稍后曲线在 97%上下浮动。其次，是只训练顶层分类器的网络，在 80 轮之后，曲线开始趋于平缓，并且在 83%上下浮动。最后，是从头训练的网络，在 200 轮之后，曲线开始趋于平缓，并且在 80%上下浮动。fine-tune 网络和只训练顶层分类器网络由于将 VGG16 模型参数直接迁移过来，可以让网络模型在新数据集上很快收敛；而从头训练网络需要经过多次迭代才能学习到目标数据集的特征，收敛速度较慢。

通过上面的分析发现，fine-tune 方式的迁移学习在性能上以及识别率上都要优于其他 2 种网络。说明 fine-tune 方式的网络通过重新训练最后一个卷积块的方式，使得收敛速度更快，识别率更高。

为了验证 fine-tune 网络模型的泛化能力以及是否有过拟合的现象，本文对 fine-tune 网络在训练集和验证集上的准确率进行对比，如图 17 所示。

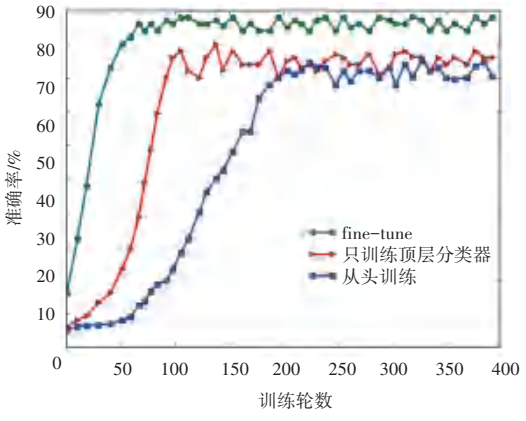


图 15　3 种网络的准确率曲线

Fig. 15　The accuracy curves of the three networks

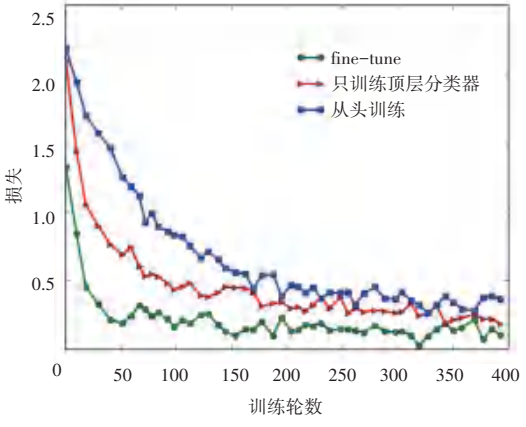


图 16　3 种网络的损失曲线

Fig. 16　Loss curves of the three networks

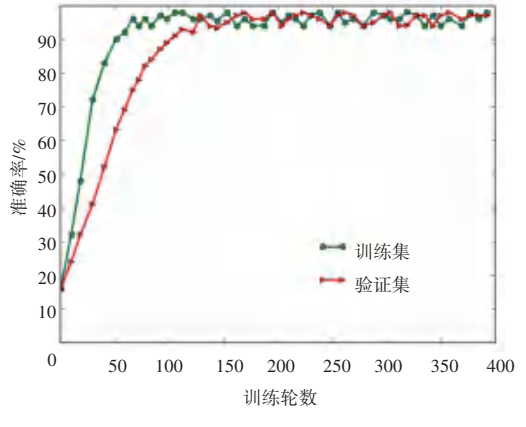


图 17　fine-tune 网络训练集与验证集准确率对比

Fig. 17　Comparison of the accuracy of training set and verification set

从图 17 可以看出，验证集上的准确率从 100 轮

开始与训练集上的准确率基本吻合。说明 fine-tune 模型的泛化能力很好,没有出现过拟合的现象,从而证明了迁移学习可以避免过拟合现象的发生。

4.2 数据增强前后实验结果对比

在 4.1 节得出了 fine-tune 网络具有很好的泛化能力,识别率高,性能优异。本小节将在该网络的基础上,对数据增强前后网络的识别率进行对比,验证数据增强的有效性。

将数据增强前的 2 000 张交警关键手势图片输入 fine-tune 网络中进行训练,得出的结果与数据增强后的训练结果进行对比。数据增强前后准确率对比如图 18 所示。

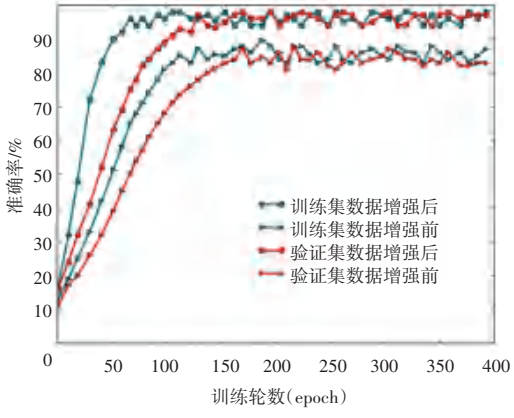


图 18 数据增强前后准确率对比

Fig. 18 Comparison of accuracy before and after data enhancement

从图 18 中可以看出,数据增强前、无论是训练集上还是验证集上的准确率要低于数据增强后的。另外,数据增强使得模型的收敛速度加快。数据增强前,由于数据集少,网络学习不到更多丰富的抽象特征,导致了识别率低;数据增强后,由于通过数据增强方式,在原有数据集上加入了更多的与实际相符的因素,模型训练之后,学习到了更多的数据集上的特殊性特征,使得网络更具有泛化能力。

4.3 与其他文献实验结果对比

进一步与文献[1]、文献[15]中的识别算法进行对比分析,见表 4。

表 4 识别算法对比

Table 4 Comparison of recognition algorithms

实验对比	识别算法	准确率/%
文献[1]	基于时间注意模块(TAM)的 GCN 改进方法	90.73
文献[15]	GCPMAGRU 模型	93.70
本文	迁移学习	97.73

文献[1]提出了基于时间注意模块(TAM)的 GCN 改进方法对交警手势进行识别,识别率为 90.73%;文献[15]提出 GCPMAGRU 模型识别交警手势,通过改进 CPM,并提出基于注意力机制的 GRU 获取时间信息,结合时空特征进行交警手势识别,识别准确度达 93.7%。本文采用迁移学习方法进行交警手势识别,识别率可达 97.73%,与文献[1]和文献[15]相比,本文算法识别效果更好。

5 结束语

本文采用基于 VGG16 的 fine-tune 迁移学习网络对小样本交警关键手势进行分类识别,通过实验证明了将预训练好的 VGG16 模型的部分网络进行冻结,可以使得迁移学习网络自带特征提取功能,能更好地提取到优异的特征。对另一部分网络进行重新训练,使网络可以学习到数据集上特殊性的特征,最终使网络的泛化能力得以提升,识别率达到 97.93%。另外,通过数据增强的策略对小样本数据集进行处理,将迁移学习应用到小数据集上,避免了小数据集导致的过拟合问题,并且收敛速度快,性能比从头训练网络更优。

参考文献

[1] 张庆, 梁涛年, 时培成. 基于改进图卷积网络的交警指挥手势识别研究[J]. 安徽工程大学学报, 2025, 40(2): 32-38.
[2] 方吴逸, 陈章进, 唐英杰. 基于改进 YOLOX-tiny 算法的交警手势识别[J]. 电子测量技术, 2024, 47(8): 100-109.
[3] 王芳, 李强. 基于 CNN 的小样本手势识别方法[J]. 模式识别与人工智能, 2023, 36(5): 456-464.
[4] 张丞, 侯义斌, 何坚. 高度分层分区的图卷积交警手势识别技术[J]. 计算机辅助设计与图形学学报, 2022, 34(7): 1037-1046.
[5] 黄涛, 王磊. 面向自动驾驶的手势识别技术综述[J]. 交通运输工程学报, 2023, 23(2): 123-134.
[6] 张磊, 王丽. 深度学习在交通手势识别中的实际应用案例[J]. 人工智能与实际应用, 2023 (2): 34-42.
[7] Deng J, Dong W, Socher R, et al. ImageNet: A large-scale hierarchical image database[C]//Proceedings of IEEE Conference on Computer Vision & Pattern Recognition. Piscataway, NJ: IEEE, 2009: 248-255.
[8] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[J]. arXiv preprint arXiv, 1409.1556, 2014.
[9] 冯晓硕, 沈樾, 王冬琦. 基于图像的数据增强方法发展现状综述[J]. 计算机科学与应用, 2021, 11(2): 370-382.
[10] 杨光, 赵晨. 交通手势识别中的数据增强技术研究[J]. 计算机应用研究, 2024, 41(1): 256-261.
[11] 刘超, 王伟. 小样本条件下的手势识别特征提取方法[J]. 电子学报, 2023, 51(8): 1921-1927.
[12] 刘艺峰, 罗亮. 基于 CNN 与迁移学习实现肺部影像分类识别[J]. 现代信息科技, 2024, 8(7): 86-90.
[13] 刘洋, 张伟. 基于深度学习的交通手势识别数据集构建[J]. 数据采集与处理, 2023, 38(2): 405-412.

[14]刘洋，张伟. 基于深度学习的交通手势识别数据集标注方法[J]. 数据采集与处理，2024，39(1)：123-130.
[15]杜兵，赵骥. 基于时空特征融合的交通警察手势识别[J]. 计算机工程与应用，2024，60(8)：250-257.
[16]Tajbakhsh N，Shin J Y，Gurudu S R，et al. Convolutional neural networks for medical image analysis：Full training or fine tuning[J]. IEEE Transactions on Medical Imaging，2016，35(5)：1299-1312.
[17]Sargno A B，Wang X，Angelov P，et al. Human action recognition using transfer learning with deep representations[C]// Proceedings of International Joint Conference on Neural Networks. Piscataway，NJ：IEEE，2017：463-469.
[18]Zeiler M D，Fergus R. Visualizing and understanding convolutional networks[M]//Fleet D，Pajdla T，Schiele B，et al. Computer Vision. Lecture Notes in Computer Science. Cham：Springer，2013，8689：818-833.
[19]李华，刘洋. 基于迁移学习的多模态手势识别[J]. 多媒体工具与应用，2024，83(5)：12345-12357.
[20]孙伟，李静. 基于迁移学习的跨域手势识别[J]. 人工智能与机器人研究，2024，11(2)：123-132.
[21]陈晨，马超. 迁移学习在图像识别领域的应用进展[J]. 计算机研究与发展，2024，61(1)：150-162.
[22]李明，王伟. 基于深度迁移学习的交通手势识别研究[J]. 计算机工程与应用，2023，59(12)：123-129.
[23]吴波，刘静. 基于迁移学习的实时手势识别系统[J]. 计算机工程与设计，2023，44(6)：1789-1795.
[24]赵敏，李明. 基于深度迁移学习的手势识别性能评估[J]. 软件学报，2024，35(2)：345-357.
[25]周杰，吴华. 交通场景中手势识别的迁移学习策略[J]. 交通运输系统工程与信息，2023，23(3)：145-152.
[26]王强，李娜. 基于迁移学习的手势识别在智能交通中的应用前景[J]. 交通运输系统工程与信息，2023，23(5)：45-52.
[27]王强，李娜. 基于 VGG19 迁移学习的手势识别研究[J]. 计算机科学与应用，2024，14(3)：567-575.
[28]李华，刘洋. 基于迁移学习的动态手势识别[J]. 多媒体工具与应用，2024，83(7)：23456-23467.
[29]Krizhevsky A，Sutskever I，Hinton G. ImageNet classification with deep convolutional neural networks[J]. Advances in Neural Information Processing Systems，2012，25(2)：1097-1105.
[30]Szegedy C，Vanhoucke V，Ioffe S，et al. Rethinking the inception architecture for computer vision[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence，2016，39(5)：879-894.
[31]沈展，蔡露，杨刚，等. 基于 VggNet 16 和特种光纤散斑的曲率测量方法[J]. 光学学报，2025，45(15)：56-66.
[32]赵雷，杨静. 基于 VGG16 迁移学习的手势识别优化[J]. 电子与信息学报，2022，44(10)：3021-3027.
[33]张亚军，王聪聪，贾广成，等. 改进 VGG-16 的轻量级手势识别研究[J]. 现代电子技术，2025，48(23)：137-143.
[34]胡局新，张功杰. 基于 K 折交叉验证的选择性集成分类算法[J]. 科技通报，2013，29(12)：115-117.