

引用格式: Zheng Zongsheng, Hu Chenyu, Huang Dongmei, *et al.* Research on Transfer Learning Methods for Classification of Typhoon Cloud Image[J]. Remote Sensing Technology and Application, 2020, 35(1): 202-210. [郑宗生, 胡晨雨, 黄冬梅, 等. 基于迁移学习及气象卫星云图的台风等级分类研究[J]. 遥感技术与应用, 2020, 35(1): 202-210.]
doi:10.11873/j.issn.1004-0323.2020.1.0202

基于迁移学习及气象卫星云图的台风等级分类研究

郑宗生, 胡晨雨, 黄冬梅, 邹国良, 刘兆荣, 宋巍

(上海海洋大学 信息学院, 上海 201306)

摘要: 针对传统卫星云图特征提取方法复杂且深度卷积神经网络(Deep Convolutional Neural Network, DCNN)模型开发困难的问题, 提出一种基于参数迁移的台风等级分类方法。利用日本气象厅发布的近40 a 10 000多景台风云图数据, 构建了适用于迁移学习的台风云图训练集和测试集。在大规模ImageNet源数据集上训练出3种源模型VGG16, InceptionV3和ResNet50, 依据台风云图低层特征与高层语义特征的差异, 适配网络最佳迁移层数并冻结低层权重, 高层权重采用自适应微调策略, 构建出了适用于台风小样本数据集的迁移预报模型T-typCNNs。实验结果表明: T-typCNNs模型在自建台风数据集上的训练精度为95.081%, 验证精度可达91.134%, 比利用浅层卷积神经网络训练出的精度高18.571%, 相比于直接用源模型训练最多提高9.819%。

关键词: 台风等级; 迁移学习; 深度卷积神经网络; 迁移层数; 自适应微调

中图分类号: TP79 **文献标志码:** A **文章编号:** 1004-0323(2020)01-0202-09

1 引言

台风是一种破坏力极强的灾害性天气系统, 对沿海地区造成了巨大的人员伤亡和经济损失。目前, 国内外对台风的研究主要集中在路径和强度的预测上^[1], 相对于台风路径而言, 强度预测更加困难^[2-3]。传统的台风强度预测大多基于数值预报模型, 由于对台风强度变化的物理过程了解不够、模式分辨率不高^[4], 严重影响了模型性能。随着观测技术与探测手段的不断发展, 特别是卫星遥感技术通过大面积、同步及非接触方式直接获取台风数据, 克服了传统数据获取的不足^[5]。遥感卫星云图中蕴含丰富的气象信息, 通过分析各类台风云系图像, 能够有效地进行云类识别^[6]、台风监测和预报^[7]。台风大多是由热带辐合区(ITCZ)云带中的

云团扰动引起, 在卫星云图上表现为不同特征的台风(眼)螺旋云系, 其与台风强度密切相关^[8], 因此对不同的螺旋云系进行特征提取, 能够实现台风强度的等级分类。Dvorak分析法是目前较成熟的卫星云图分析法。通过人工直接分析可见光及红外线云图, 获得云系中心和气旋形态相关参数, 进行全天候热带气旋强度评估^[9]。但是计算过程复杂且需要大量专业知识, 主观性较强, 增加了分析误差。

随着机器学习和深度学习在图像处理领域的广泛应用, 研究者开始将此类算法运用到遥感图像上。杨蒙蒙等^[10]提出了一种基于双树复小波分解的BP神经网络的遥感影像分类方法, 解决了高分辨率多光谱遥感影像分类的模糊性和不确定性, 与基于纹理、光谱等特征的神经网络算法对比精度达到93.10%。针对云种特征的模糊性, 赵亮亮等^[11]建

收稿日期: 2018-09-25; 修订日期: 2019-12-21

基金项目: 国家自然科学基金项目(41671431), 上海市科委地方院校能力建设项目(17050501900)和国家海洋局数字海洋科学技术重点实验室开放基金项目共同资助。

作者简介: 郑宗生(1979—), 男, 河北唐山人, 博士, 副教授, 主要从事海洋信息化及深度学习、迁移学习应用方面的研究。

E-mail: zszheng@shou.edu.cn。

通讯作者: 胡晨雨(1995—), 女, 江苏盐城人, 硕士研究生, 主要从事迁移学习应用研究。E-mail: 1105814265@qq.com。

立了具有大卷积核的浅层神经网络进行云分类任务,分类精度比传统的 K 近邻机器学习算法高 15.6%。为了进一步提高卫星遥感图像场景分类的准确度,崔先亮等^[12]提出一种指导学习卷积神经网络 GL-CNN,将高低频子带样本融合后,用 GL-CNN 网络自动提取图像的高低频子带包含的高层特征来实现场景分类,分类准确率高达 94.52%。但目前利用深度学习方法对气象云图的分类存在如下问题:①对于复杂的气象云图,浅层卷积网络特征提取不充分,分类效果不佳;②深层卷积网络开发难度大,在台风云图数据样本小的情况下,过拟合现象严重。

迁移学习提供一种将源环境中学到的知识运用到相关环境,以辅助完成新任务的方法。迁移学习放宽了传统机器学习中必须拥有足够的训练样本,且训练和测试样本必须服从同概率分布的前提条件^[13]。随着深度卷积神经网络在自动分析和图像特征识别方面取得的卓越成果,迁移学习被广泛应用于卷积神经网络,通过共享源环境中的模型结构或先验参数以实现目标任务。如王文朋等^[14]迁移 VGG16 源模型的整体网络参数用于提高烟雾场景识别的精度。黄洁等^[15]将预迁移源模型 AlexNet 与支持向量机(SVM)相结合,用于舰船遥感图像的目标检测,一定程度上避免了过拟合现象。孙超等^[16]改进了基于 CNN 的自然图像超分辨率方法以提高图像的分辨率,并将改进的网络结构及参数迁移到少数高质量红外图像上,改善了红外图像的超分辨率性能。考虑到源域和目标域的差异,Long 等^[17]将 ImageNet 数据集上已训练好的 AlexNet 网络结构和对应权重参数迁移到目标任务上,并在高层网络间进行多层适配,以减小源域和目标域之间的差异。为了数值化预迁移源模型的迁移能力,Muhammad 等^[18]将源域与目标域之间样本的互信息作为衡量源模型迁移能力的指标,并将迁移率最高的源模型网络结构及参数迁移到目标域上进行分类任务。迁移学习能够很好地解决现实中多数领域中标注数据有限、特征异构甚至异类等问题。但迁移学习在各方面应用及迁移效果的研究中多数考虑网络参数的全部迁移,实际上由于源域及目标域的差异,必然存在最佳迁移层数。如何适配源模型的最佳迁移层数,在充分考虑目标域样本特征的同时提高迁移效果,也是迁移学习研究的关键性问题。

鉴于 VGG16、InceptionV3、ResNet50 3 种源模型在大样本 ImageNet 自然数据集上具有良好的特征提取能力,本文通过这 3 种具有代表性的源模型在小样本自建台风数据集上迁移效果的对比实验,确立最佳迁移源模型及迁移层数,构建了适应于台风小样本数据集的迁移预报模型 T-tycNNs,有效减缓深度网络在小样本数据集上出现的过拟合问题。实验结果表明,在迁移源模型基础上通过最佳迁移层数适配与自适应微调相结合的方法,能够在遥感数据不足的情况下,实现台风等级快速、高精度分类,为台风预警预报提供了一种有效的辅助决策方法。

2 数据获取及研究方法

2.1 数据集构建与预处理

本文所用的自建台风数据由日本国立情报学研究所(National Institute of Informatics,简称 NII)提供,气象云图数据取自“Himawari-8 号”气象卫星,时间分辨率为 10 min。“H-8”搭载的高级成像仪(AHI)有 16 个可见光和红外通道,可见光和红外图像的空间分辨率分别可达到 0.5~1 km 和 1~2 km。本文选取近 40 a 来台风云图作为数据样本。由于源域 ImageNet 数据集的自然图像超过百万张,为避免域样本数量及内容上差异引发的过拟合问题,首先对图像进行增强和归一化处理,通过随机旋转、缩放、偏移和翻转等数据扩增技术提高样本数量。由于源模型输入为三通道彩色图像,将气象云图的单通道灰度图像进行通道扩增,形成 RGB 三通道图像,最后将图片重采样为 224×224 和 299×299 像素。依照国际台风分类标准,根据台风中心风速将数据集分为 5 类:热带低压、热带风暴、强热带风暴、台风和强台风,如表 1 所示。

表 1 台风等级标准

Table1 Typhoon grading standards

台风等级分类	最大风速(10 min 平均值)		
	kt	m/s	km/h
热带低压(TD)	<33	<17	<62
热带风暴(TS)	≥33~<48	≥17~<25	≥62~<89
强热带风暴(STS)	≥48~<64	≥25~<33	≥89~<118
台风(TY)	≥64~<85	≥33~<42	≥118~<150
强台风(STY)	≥85	≥42	≥150

台风云图数据集采用 5 类标签标记,每类图像 2 500 张,其中 2 000 张训练集,500 张验证集,数据集总样本达 12 500 张,部分云图样本如图 1 所示。

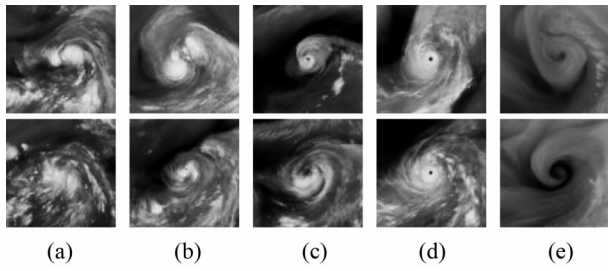


图1 台风云图样本

注:(a)为一级台风部分数据样本,(b)为二级台风部分数据样本,(c)为三级台风部分数据样本,(d)为四级台风部分数据样本,(e)为五级台风部分数据样本

Fig.1 Typhoon samples

2.2 研究方法与模型构建

2.2.1 研究方法

深度卷积神经网络通过卷积层与采样层交替堆叠,实现对原始图像的低层抽象特征提取和高层语义表达。

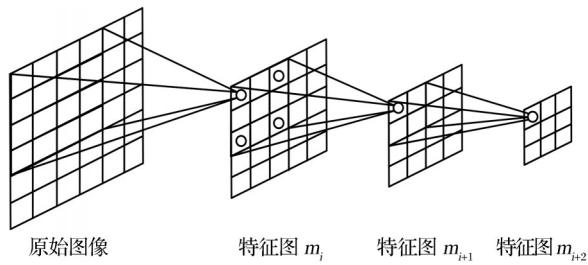


图2 深度卷积神经网络

Fig.2 Deep convolutional neural network

如图2所示,卷积核前向映射到前一层特征图 m_i 的感受野(视觉感受区域大小) r_i 的计算过程具体如下:

$$stride_l = stride_1 \times stride_2 \times \dots \times stride_{l-1} \quad (1)$$

$$r_l = (r_{l-1} - 1) \times stride_l + c_l \quad (2)$$

从式(1)和(2)可以看出,每一层感受野大小 r_l 与它之后的所有层的卷积核尺寸 c_l 和卷积步长 $stride_l$ 有关。随着网络层数 l 的增加,映射到原始图像的感受野不断变大,低层特征提取得更加全面,高层语义表达得更加具体,输出结果越能表达原始图像含义。

迁移学习旨在利用源领域中训练出的良好学习能力解决新领域中的学习任务。具体可以公式化定义为: D_s 表示源数据样本构成的源域, T_s 表示在源域上训练的源任务, D_t 表示目标数据样本构成的目标域, T_t 表示在目标域上训练的目标任务。 D_s 和 D_t 可以定义为一个二元对 $\{\chi, P(X)\}$ 。其中 χ 表示特征空间, $P(X)$ 是 X 的边缘概率分布, X 是数据样

本集合, $X = \{x_1, x_2, \dots, x_n\} \in \chi$ 。 T_s 和 T_t 也是一个二元对 $\{\gamma, f(\cdot)\}$ 。其中 γ 是域中每一个样本对应的真实标签空间, $f(\cdot)$ 是目标函数映射,即在数据集上训练出的模型 $f(\cdot): x_i \rightarrow y_i$, 其中 $x_i \in X, y_i \in \gamma$ 。迁移学习本质就是将源域 D_s 上表现优异的模型算法 $f_s(\cdot)$ 迁移到目标域 D_t 上,通过微调来辅助完成目标任务 T_t 的方法,迁移并微调后的模型算法用 $f'_s(\cdot)$ 表示。

深度卷积神经网络中特定的结构及参数代表网络良好的特征提取能力,本文提出一种基于参数迁移的网络自适应微调学习方法,将深度网络算法 $f_s(\cdot)$ 训练出的特征提取能力迁移到小样本自建台风数据集上。

鉴于目标域中台风云图特有的台风眼、云墙以及螺旋云带等特征难以通过源模型算法 $f_t(\cdot)$ 直接提取,且深度网络提取到的低层和高层特征存在差异,因此目标模型算法的分类性能取决于迁移源模型的层数。本文在保留低层特征提取能力的同时,将高层卷积各参数设置为可调节状态,并在台风数据集上做自适应微调,微调后的模型算法 $f'_s(\cdot)$ 可进一步提取到台风云图中的细微特征,提高目标模型分类精度。其中高层可调节参数并非随机初始化,而是源模型的对应层参数,便于通过小幅度自适应调整使模型快速收敛,达到最佳分类精度。实现在高资源利用率和低开发成本的条件下,训练出性能良好的台风等级分类模型,完成跨领域快速建模方法。

如图3所示,在预处理后的云图数据上对比两种源模型迁移方式的分类精度:一种是迁移源模型的整体网络结构及参数;另一种是迁移整体网络结构,依据适配的最佳迁移层数冻结相应层参数,并自适应微调剩余网络参数。对比分析两种目标模型的性能,最终构建出适用于台风等级分类的 T-typCNNs 模型。

2.2.2 T-typCNNs 模型构建

由于台风卫星云图中不同强度等级的螺旋云系特征关联相似性很高^[19],为了能充分区分台风图像的高层特征及不同螺旋云系间低层特征的细微差别,本文选取了三种在 ImageNet 图像识别上具有良好特征学习能力的深度卷积网络模型进行实验对比。

VGG 网络^[20]是一个多层卷积神经网络,特征提取层均采用尺寸为 3×3 卷积滤波器。VGG 网络采

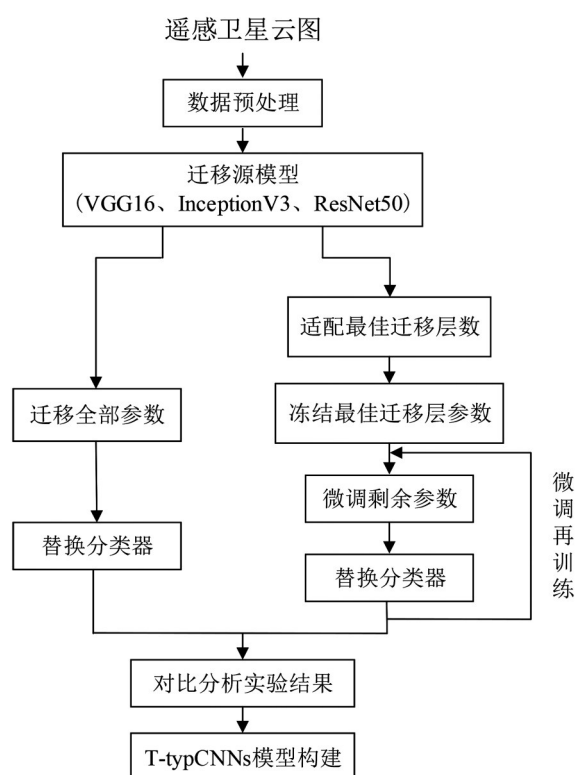


图3 研究方法流程图

Fig.3 Flow chat of research method

用原始的网络堆积方法,网络参数过多会导致训练耗时长、易发生过拟合。Inception结构^[21]有效缓解了增加网络的深度和每层神经元的数目而造成的网络参数冗余、模型过拟合、梯度消失等问题,提高

了非线性扩展模型的表达能力。ResNet(Residual Neural Network)网络^[22]引入了残差结构块,允许原始输入信息直接传输到中间的网络层,有效缓解了由于层数不断加深而造成网络对权值的更新不灵敏问题。

针对上述3种源模型各自的特点,本文选取了网络深度相对较浅的VGG16、网络层数较深的InceptionV3及ResNet50,用于对比网络深度对迁移效果的影响。为了对比网络结构对迁移效果的影响,选取结构不同的深度网络InceptionV3和ResNet50。各源模型的具体参数详见表2。

表2 3种源模型参数配置

Table2 Parameter configuration of each source model				
模型	原始图像	参数	网络深度	大小
VGG16	224×224×3	138 357 544	23	528 MB
InceptionV3	299×299×3	23 851 784	159	92 MB
ResNet50	224×224×3	25 636 712	168	99 MB

比较上述3种源模型在迁移网络整体结构以及全部层参数到台风数据集上的分类精度,选择出性能最佳的源模型,并进一步适配其最大迁移层数。如图4所示,将最后的分类器层中神经元个数设置为5,表示台风的5个等级。通过冻结此部分层参数权重并微调再训练剩余网络参数,构建出适用于台风小样本集的迁移预报模型T-typCNNs。

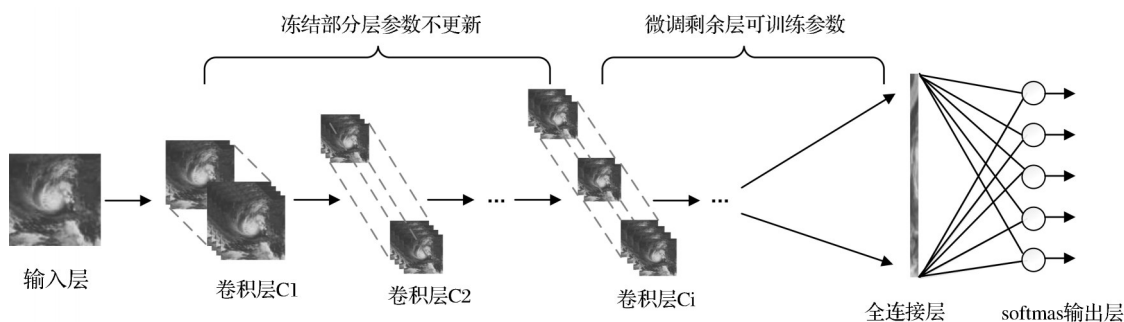


图4 T-typCNNs迁移模型结构

Fig.4 Transferred model structure of the T-typCNNs

3 实验结果与分析

本文实验环境为 windows10 操作系统, CPUintelXeonX5650@2.67 GHz, 内存为 16 G, 软件采用 python3.6 版本, 实验基于 Tensorflow 的 Keras 框架。

实验主要分3个部分:①评估自建的CNN模型与3种源模型深度卷积网络深度对台风云图等级分类精度的影响;②迁移源模型深度网络结构和参数

辅助台风等级分类任务,进一步寻找最佳迁移层数并在台风数据集上对剩余网络参数进行自适应微调;③针对源模型的特殊网络结构,探究不同数量级的台风样本对迁移并微调后的源模型分类性能的影响。

3.1 基于网络层深度的对比与分析

实验自建了8层卷积网络CNN_8,包含4个卷积层、2个池化层和和2个全连接层。卷积核尺寸参

考VGG16模型,均选用 3×3 ,步长为2的滤波器对台风云图中的螺旋云带进行特征提取。深度卷积网络结构借用VGG16、InceptionV3、ResNet50三个源模型。迁移源模型整体网络结构,参数均随机初始化,并根据台风数据集进行自适应调整。训练的基本设置为迭代100次,批数量64张。采用随机梯度下降法,模型超参数设置:浅层模型学习率为 $1e-2$,深度模型学习率为 $1e-4$,学习动量均为0.9。实验1中4个对比实验的详细参数以及在台风数据集上的训练和测试精度详见表3。

表3 不同深度模型仅迁移网络结构的分类性能

Table 3 Classification performance of models with different depth when only transfer network structures

网络结构	代码 层数	网络参数	训练精 度/%	测试精 度/%	时间 /epoch
CNN_8	6	35 474	75.314	72.563	71 s
VGG16	20	138 357 544	81.647	74.031	153 s
InceptionV3	311	23 851 784	85.033	78.472	304 s
ResNet50	175	25 636 712	86.951	81.315	212 s

从表3看出,在12 500张自建台风云图上训练出的4种网络由于层数和结构不同,呈现出不同的训练效果。自建的CNN_8网络由于卷积层数少于VGG16,平均测试精度低于3种迁移源模型5.67%。随着网络层数加深,模型的分类精度明显提高,证明深度网络可以提取到原始台风图像中不同螺旋云系之间特征的细微差别。但VGG16和InceptionV3深度网络均出现过拟合现象。ResNet50凭借特殊的残差结构和较深网络层数,将测试精度提升了2.84%,但随着网络层数的加深,训练时间大幅增加,过拟合程度随着迭代次数的增加越来越严重。

3.2 网络参数与结构全部迁移对比与分析

针对实验1中深度卷积网络存在的问题,实验2迁移3种源模型结构和全部卷积层权重参数到自建台风数据集上。只替换模型的Softmax分类层,替换分类器后的3个源模型在自建台风数据集上的分类效果详见表4。

从表4可以看出,3种模型的训练时长均有明显缩短。测试精度比实验1平均提高了6.87%,并有效缓解了过拟合现象,将训练和测试精度差控制在3.02%左右,ResNet50的最高分类精度可达88.611%。

表4 深度模型迁移整体网络结构和参数的分类性能

Table 4 Classification performance of deep models when transfer all the network structures and parameters

网络结构	训练精度/%	测试精度/%	训练时间/s
VGG16	83.131	81.092	137s/epoch
InceptionV3	87.772	84.726	253s/epoch
ResNet50	92.649	88.611	176s/epoch

3.3 网络参数部分迁移对比与分析

3.3.1 最佳迁移层数

以迁移源模型VGG16为例,VGG16的特征提取层被划分为5个模块,分别从block1到block5,每个模块包含多个卷积层。以模块为单位,将每一个模块中的卷积层及其层参数迁移到目标数据集上,并记录实验数据。据此方法,源模型InceptionV3和ResNet50在台风数据集上的最大迁移层数均可以获得,如图6和7所示。

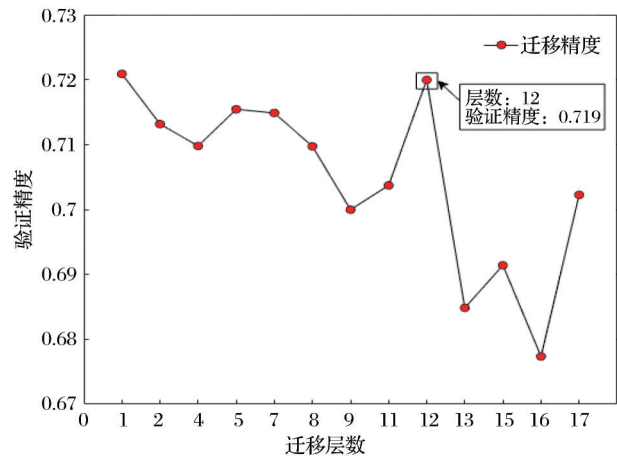


图5 迁移VGG16模型各层的分类精度

Fig.5 Classification accuracy of each layer of the transferred VGG16 model

VGG16有13个卷积层,分别迁移每一层结构及参数到自建台风数据集,并比较13个子源模型训练出的验证精度。横坐标表示用Keras实现的代码层数。如图5所示,迁移前1~12层时,验证精度的波动幅度不大,振幅在3.0%内。从迁移13层开始,精度下降到67.7%。模型迁移能力明显下降,精度上下浮动较震荡,证明源模型的前12个卷积层在ImageNet数据集上训练出的特征提取能力在自建台风数据集上具有普适性,能够提取到台风云图底层像素级别的公共特征信息。13~17层的卷积核在ImageNet上提取特征的能力并不能适用于台风云图数据,即提取的特征因数据集的不同开始特殊化,特征表示越来越抽象化。所以即使网络层数增

加了,模型性能反而下降。此现象验证了Yosinski等^[23]通过逐层分析 AlexNet 卷积神经网络的迁移能力所得出的结论。

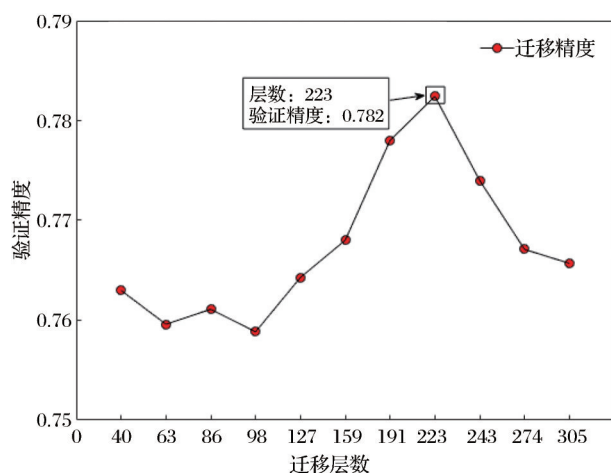


图6 迁移 InceptionV3 模型各层的分类精度

Fig.6 Classification accuracy of each layer of the transferred InceptionV3 model

从图6看出,当迁移 InceptionV3 的 1~223 层,特别是 98~223 层时,子源模型在台风数据集上的验证精度从 75.8% 提高到 78.2%,当迁移到 224~305 层时精度一直下降,证明提取到的台风特征越来越具体特殊化。因此,InceptionV3 模型在台风数据集上的最佳迁移层数为 223。

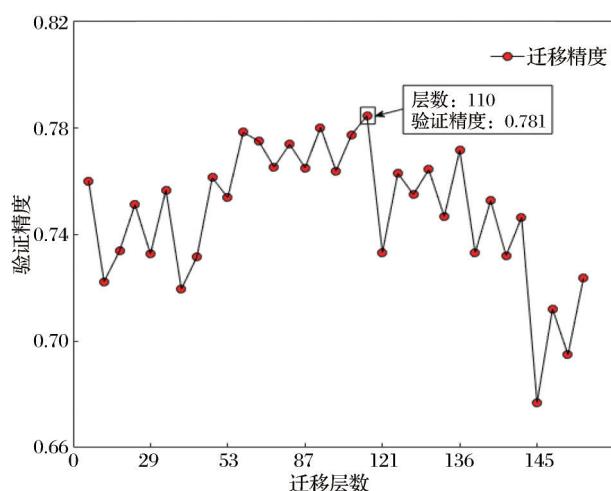


图7 迁移 ResNet50 模型各层的分类精度

Fig.7 Classification accuracy of each layer of the transferred ResNet50 model

图7展示了迁移 ResNet50 各卷积层到台风数据集上的分类性能。在迁移 1~110 层时,子源模型的验证精度高达 78.1%,从 111~156 层迁移精度下降到低至 67.6%。因此,ResNet50 模型在台风数据

集上的最佳迁移层数为 110。

通过比较各子源模型在目标域上的分类精度,确定能提取到源域和目标域公共特征的最大网络层数。因此,源模型的迁移层参数应该选取合理值。

3.3.2 自适应微调方法

通过上述方法可以得出各源模型的最佳冻结层数,即分别冻结源模型 VGG16、InceptionV3 和 ResNet50 的 12 层、223 层和 110 层权重参数,并设置剩余层参数为可更新状态,在目标数据集上进行自适应微调。表 5 列出了当 3 种源模型分别冻结最佳层数权重,剩余层参数在台风数据集上进行微调再训练后的模型精度。实验参数设置与实验 1 相同,训练迭代 100 次,批数量为 64,学习率均设置为 $1e-4$,学习动量均为 0.9。

表5 自适应微调后的各模型性能

Table5 Performance of each model after fine-tuning

网络结构	最佳冻结 层数	训练精度 /%	测试精度 /%	训练时间/s
VGG16	12	89.323	82.712	146s/epoch
InceptionV3	223	92.744	87.398	294s/epoch
ResNet50	110	95.081	91.134	182s/epoch

从表5可以看出,采用参数迁移并微调再训练的方法可以对模型性能起明显提升作用,尤其对 ResNet50 模型在适当增加训练时间的情况下,测试精度从 88.611% 提高到 91.134%,优于 VGG16 模型 8.422%,比 InceptionV3 分类性能高出 3.736%。ResNet50 模型凭借其特殊的残差结构,获得较高分类精度的同时,模型深度和结构上优于 VGG16,占用内存资源和训练时间上优于 InceptionV3。

3.4 T-typCNNs 模型性能分析

将分类精度与训练时间作为综合评价指标,迁移 ResNet50 模型整体网络结构,冻结前 110 层网络参数并将剩余层参数设置为可更新状态,构建出台风等级预测模型 T-typCNNs。图 8 和 9 展示了 T-typCNNs 在自建台风数据集上的模型分类性能。

由上图看出,T-typCNNs 模型的训练和验证精度分别在迭代 16 次和 34 次之后趋于收敛,达到 95.081% 和 91.175%,没有发生过拟合现象。将迭代次数增加到 200 次精度并没有明显改善,进一步证明模型已经训练到最优性能,模型损失值在迭代 42 次后下降到 0.106 左右并趋于平稳。

T-typCNNs 模型中残差结构内部的并行传输

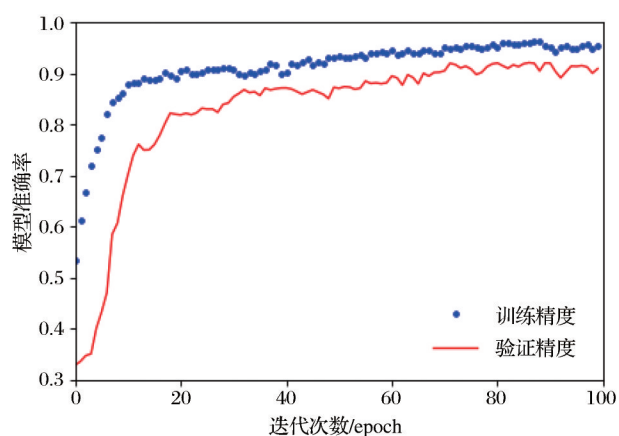


图8 T-typCNNs模型准确率曲线

Fig.8 Accuracy curve of T-typCNNs model

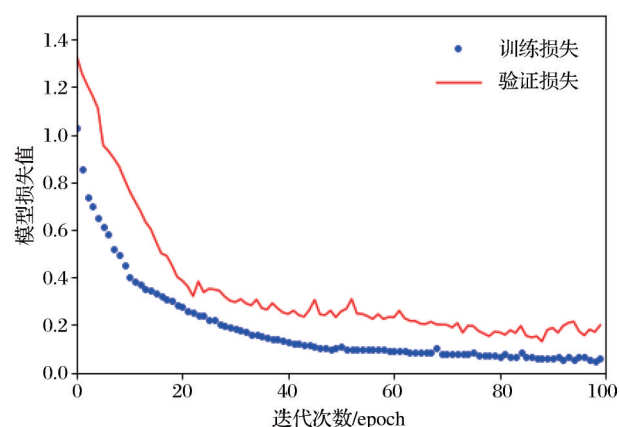


图9 T-typCNNs模型损失值曲线

Fig.9 Loss curve of T-typCNNs model

方式,通过舍弃局部对图像的特征提取能力来换取比 InceptionV3 更深的网络层数,理论上适当增加目标域中有标签训练样本,模型的性能会有所提升。

为证实上述理论,实验在不同数量级的台风云图样本上比较 T-typCNNs 模型的性能。如图 10 所示,进行 10 组对比实验,每组台风样本的训练集和验证集按 4:1 的比例进行分配,将每类样本的训练集数量初始设置为 1 500 张,每组实验增加 500 张。

实验表明,每类台风云图样本从 1 500 张增加到 4 000 张时,对 T-typCNNs 模型的性能均有所改善。尤其当每类有 3 000 张云图样本时,测试精度高达 93.26%。但随着训练样本数不断增加,模型的测试精度有所下降并出现过拟合的趋势。可以得出结论,在一定范围内适当增加台风数据集的样本数量,可以提高 T-typCNNs 模型对台风云图的分类性能。

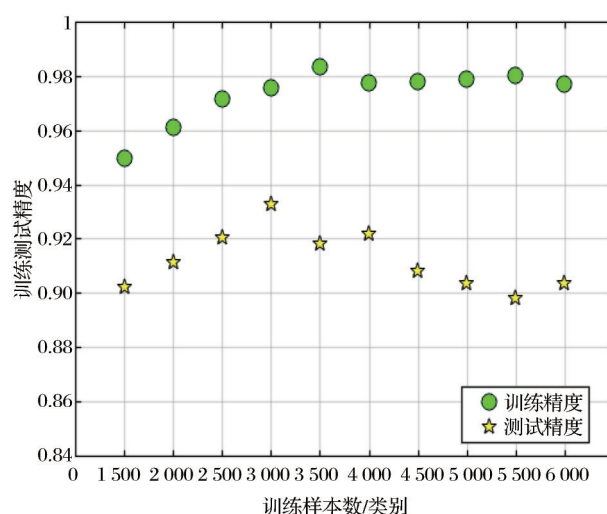


图10 不同样本数下的模型性能

Fig.10 Model performance of different sample numbers

4 结 语

本文提出了一种基于参数迁移的网络自适应微调学习方法,将深度卷积神经网络参数迁移到台风云图数据集上,辅助完成台风强度等级分类任务。解决了气象领域小样本台风数据等级分类问题,为台风预测预报提供了一种辅助决策方法。

(1)将 VGG16、InceptionV3 和 ResNet50 3 种源模型迁移到小样本台风数据集上。通过较短的训练时间、较深的网络层数,训练出性能较好的分类模型。实验证明,迁移源模型整体网络结构和参数到台风数据集上能将精度提高 7%。

(2)适配最佳迁移层数,将源模型按模块划分并迁移对应层参数到目标数据集上,对比子模块网络精度,找到能提取公共特征的最大卷积层数并冻结权重参数,剩余层参数在台风数据集上微调再训练。实验表明,从深度和速度上综合考虑,ResNet50 模型具有最好的分类性能。

(3)通过冻结 ResNet50 部分层参数,并将剩余网络参数作自适应微调,本文构建了适用于台风小样本数据集的迁移预报模型 T-typCNNs。该模型将台风数据集上训练出的精度从 88.611% 提高到 91.134%,有效缓解过拟合现象。另外,本文发现在一定数量级范围内增加台风训练样本,有利于提高 T-typCNNs 迁移模型的分类性能。但随着样本数量的增加,模型分类精度有下降趋势。

(4)提出了适用于自建台风数据集的迁移学习方法,由于只针对台风数据集,当源域和目标域发生改变时,模型的鲁棒性可能有所降低,分析原因

主要是由于源域和目标域的共同特征减少,导致模型的迁移性能下降,这也是今后重点研究的方向。但本文提出的迁移策略可对特定领域小样本数据集的迁移提供指导。在后期的研究中,将重点探究在迁移卷积神经网络时,如何通过域自适应方法来缓解域差异对分类精度的影响。

参考文献(References):

- [1] Liao Fei, Li Wenting, Zhang Ziran, *et al.* Statistical Characteristics of Rapid Changes in the Intensity and Track of Tropical Cyclones in the South China Sea from 1949 to 2017[J]. *Journal of Ocean*, 2019, 41(9): 126-135. [廖菲,李文婷,张子然,等. 1949~2017年南海海域热带气旋强度和路径快速变化统计特征[J]. *海洋学报*, 2019, 41(9): 126-135.]
- [2] Yu Jinhua, Sheng Siwei. Comparison of Tropical Cyclone Intensity Data in the Northwest Pacific[J]. *Meteorological Science*, 2010, 30(6): 778-784. [余锦华,盛思伟. 西北太平洋热带气旋强度资料的对比[J]. *气象科学*, 2010, 30(6): 778-784.]
- [3] Yin Chengtuan, Zhang Jinshan, Xiong Mengjie, *et al.* Analysis on the Disaster trend of typhoon and tide in south China sea [J]. *Journal of Tropical Oceanography*, 2019, 38(1): 38-45. [殷成团,张金善,熊梦婕,等. 我国南海沿海台风及暴雨灾害趋势分析[J]. *气象科学*, 2019, 38(1): 38-45.]
- [4] Tang Tingting, Li Qingxi, Li Guangxin, *et al.* Research on Typhoon Intensity Prediction Statistical Model based on Meteorological Big Data[J]. *Integration Technology*, 2016, 5(2): 73-84. [汤婷婷,李晴岚,李广鑫,等. 基于气象大数据的台风强度预测统计模型研究[J]. *集成技术*, 2016, 5(2): 73-84.]
- [5] Zhao Xiaoli. Research on Cloud Image Cloud Detection and Classification of Meteorological Satellites [D]. Guangzhou: South China University of Technology, 2014. [赵晓利. 气象卫星云图云检测及分类的研究[D]. 广州:华南理工大学, 2014.]
- [6] Tian Wenzhe, Fu Randi, Jin Wei, *et al.* Adaptive Fuzzy Support Vector Machine for Classification of Clouds in Satellite Imagery[J]. *Journal of Wuhan University*, 2017, 42(4): 488-495. [陈田文哲,符冉迪,金炜,等. 面向卫星云图云分类的自适应模糊支持向量机[J]. *武汉大学学报*, 2017, 42(4): 488-495.]
- [7] Cao Minjie, Liu Zenghong, Wang Zhenfeng, *et al.* A Review of Real Time Marine Monitoring and Its Application in Typhoon Area [J]. *Advances in Meteorological Technology*, 2017(4): 47-52. [曹敏杰,刘增宏,王振峰,等. 台风海域实时海洋监测及其应用研究综述[J]. *气象科技进展*, 2017(4): 47-52.]
- [8] Shi Lanhong, Cui Linli, Zhao Bingke, *et al.* Study on the Vertical Distribution of Cloud Properties in the Eye Sall and Surrounding Spiral Cloud Belt of Typhoon[J]. *Journal of Tropical Meteorological*, 2015(1): 53-64. [史兰红,崔林丽,赵兵科,等. 台风眼壁及周围螺旋云带云属性垂直分布研究[J]. *热带气象学报*, 2015(1): 53-64.]
- [9] Dvorak V F. Tropical Cyclone Intensity Analysis Using satellite Data[R]. NOAA Technical Report, 1984, NESDUS, 11.
- [10] Yang Wei, Ouyang Sida, Fan Kuikui, *et al.* Classification of Remote Sensing Image based on BP Neural Network based on Double Tree Complex Wavelet Decomposition [J]. *Remote Sensing Technology and Application*, 2018, 33(2): 313-320. [杨朦朦,欧阳斯达,范奎奎,等. 基于双树复小波分解的BP神经网络遥感影像分类[J]. *遥感技术与应用*, 2018, 33(2): 313-320.]
- [11] Zhao Liangliang. Research and Application of Cloud Image Recognition and Ultra-short-term Direct Solar Radiation Prediction based on Neural Network [D]. Nanjing: Southeast University, 2017. [赵亮亮. 基于神经网络的云图识别与超短期直接太阳光辐射预测研究及应用[D]. 南京:东南大学, 2017.]
- [12] Cui Xianliang, Chen Lifu, Xing Xuemin, *et al.* Remote Sensing Image Scene Classification based on Frequency Band Feature Fusion and GL-CNN. *Remote Sensing Technology and Application*, 2019, 34(4): 712-719. [崔先亮,陈立福,邢学敏,等. 基于频带特征融合的GL-CNN遥感图像场景分类[J]. *遥感技术与应用*, 2019, 34(4): 712-719.]
- [13] Zhuang Fuzhen, Luo Ping, He Qing, *et al.* Progress in Migration Learning Research [J]. *Journal of Software*, 2015, 26(1): 26-39. [庄福振,罗平,何清,等. 迁移学习研究进展[J]. *软件学报*, 2015, 26(1): 26-39.]
- [14] Wang Wenpeng, Mao Wentao, He Jianliang, *et al.* Smoke Recognition Method based on Deep Migration Learning [J]. *Journal of Computer Applications*, 2017, 37(11): 3176-3181. [王文朋,毛文涛,何建樑,等. 基于深度迁移学习的烟雾识别方法[J]. *计算机应用*, 2017, 37(11): 3176-3181.]
- [15] Huang Jie, Jiang Zhiguo, Zhang Haopeng, *et al.* Ship Target Detection based on Convolutional Neural Network for Remote Sensing Image [J]. *Journal of Beijing University of Aeronautics*, 2017, 43(9): 1841-1848. [黄洁,姜志国,张浩鹏,等. 基于卷积神经网络的遥感图像舰船目标检测[J]. *北京航空航天大学学报*, 2017, 43(9): 1841-1848.]
- [16] Sun Chao, Lu Junwei, Liu Feng, *et al.* Research on Infrared Image Super-resolution Method based on Migration Learning [J]. *Laser & Infrared*, 2017, 47(12): 1559-1564. [孙超,吕俊伟,刘峰,等. 基于迁移学习的红外图像超分辨率方法研究[J]. *激光与红外*, 2017, 47(12): 1559-1564.]
- [17] Long M S, Cao Y, Wang J M, *et al.* Learning Transferable Features with Deep Adaptation Networks[C]// In International Conference on Machine Learning, 2015: 97-105.
- [18] Afridi M J, Ross A, Shapiro E M. On Automated Source Selection for Transfer Learning in Convolutional Neural Networks[J]. *Pattern Recognition*, 2017, 73: 65-75.
- [19] Gao Song, Zhao Peng, Pan Bin, *et al.* The Typhoon Path Nowcasting Model based on LSTM Neural Network [J]. *Acta Oceanologica Sinica*, 2016, 37(5): 8-12. [高松,赵鹏,潘斌,等. 基于长短期记忆神经网络的台风路径临近预报模型[J].

- 海洋学报, 2016, 37(5):8-12.]
- [20] Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-scale Image Recognition[J]. Computer Science, 2014, arxiv:1409.1556.
- [21] Szegedy C, Vanhoucke V, Ioffe S, *et al.* Rethinking the Inception Architecture for Computer Vision[C]// In International Conference on Learning Representation, Computer Science, 2015: 2818-2826.
- [22] He K, Zhang X, Ren S, *et al.* Deep Residual Learning for Image Recognition[C]// Computer Vision and Pattern Recognition, 2015: 770-778.
- [23] Yosinski J, Clune J, Bengio Y, *et al.* How Transferable are Features in Deep Neural Networks?[C]// Advances in Neural Information Processing Systems, 2014:3320-3328.

Research on Transfer Learning Methods for Classification of Typhoon Cloud Image

Zheng Zongsheng, Hu Chenyu, Huang Dongmei, Zou Guoliang,
Liu Zhaorong, Song Wei

(School of Information, Shanghai Ocean University, Shanghai 201306, China)

Abstract: Aiming at the complexity of traditional methods for feature extraction about satellite cloud images, and the difficulty of developing deep convolutional neural network from scratch, a parameter-based transfer learning method for classifying typhoon intensity is proposed. Take typhoon satellite cloud images published by Japan Meteorological Agency, which includes 10 000 scenes among nearly 40 years to construct training and test typhoon datasets. Three deep convolutional neural networks, VGG16, InceptionV3 and ResNet50 are trained as source models on the large-scale ImageNet datasets. Considering the discrepancy between low-level features and high-level semantic features of typhoon cloud images, adapt the optimal number of transferable layers in neural networks and freeze weights of low-level network. Meanwhile, fine-tune surplus weights on typhoon dataset adaptively. Finally, a transferred prediction model which is suitable for small sample typhoon datasets, called T-tycNNs is proposed. Experimental results show that the T-tycNNs can achieve training accuracy of 95.081% and testing accuracy of 91.134%, 18.571% higher than using shallow convolutional neural network, 9.819% higher than training with source models from scratch.

Key words: Typhoon grade; Transfer learning; Deep convolutional neural network; Number of transferable layers; Adaptive fine-tuning