

基于多路卷积神经网络的多光谱场景识别方法

江泽涛^{1,2} 秦嘉奇¹ 胡 硕³

(桂林电子科技大学广西图像图形智能处理高校重点实验室 广西 桂林 541004)¹

(桂林电子科技大学广西可信软件重点实验室 广西 桂林 541004)²

(南昌航空大学 南昌 330063)³

摘 要 现有的基于卷积神经网络的场景识别算法无法处理目标场景图形是多光谱图像的情况,在数据量较小的情况下,该算法的识别率不高。针对以上问题,提出一种基于多路卷积神经网络的多光谱场景识别方法。多路卷积神经网络接受三通道可见光彩色图像(RGB 图像)以及单通道的近红外图像(NIR 图像)共四通道输入。所提方法能够有效提取可见光图像特征、红外光图像特征以及可见光和红外光图像之间的关联特征,并将特征在全连接层进行融合,合理利用了各个光谱图像之间的相关信息,并通过结合预训练的方法来提高识别精度。在 NIR_RGB 数据集上的实验表明,与 AlexNet、InceptionNet、ResNet 以及人工设计特征描述子方法相比,该网络的平均识别率较高。并且,对此网络稍加改动,就能将其推广到其他多光谱图像分类任务中。

关键词 多光谱,卷积神经网络,图像分类,场景识别

中图法分类号 TP391.4 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2019.09.040

Multi-spectral Scene Recognition Method Based on Multi-way Convolution Neural Network

JIANG Ze-tao^{1,2} QIN Jia-qi¹ HU Shuo³

(The Key Laboratory of Image and Graphic Intelligent Processing of Higher Education in Guangxi, Guilin University of Electronic Technology, Guilin, Guangxi 541004, China)¹

(The Key Laboratory of Dependable Software of Guangxi, Guilin University of Electronic Technology, Guilin, Guangxi 541004, China)²
(Nanchang Hangkong University, Nanchang 330063, China)³

Abstract The existing scene recognition algorithm based on convolution neural network can't deal with the multi spectral image of the target scene and can't implement ideal accuracy in the case of insufficient data. In view of the above problems, this paper proposed a multi-spectral convolution neural network based multispectral scene recognition method. The multi-way convolution neural network accepts three channels of visible light color image (RGB image) and a single channel near infrared image (NIR image) with a total of four channels. The proposed method can effectively extract the features of visible light image, infrared image and the correlation between visible and infrared images, and combine the features in the full connection layer, utilizing the correlation information among spectral images reasonably. The pre-training method is combined to improve the accuracy. Experiment results on the NIR_RGB dataset show that the average accuracy of the network is higher than that of AlexNet, InceptionNet, ResNet and artificial design feature descriptors. Moreover, this network can be extended to other multi-spectral image classification tasks with slight modification.

Keywords Multi-spectral, Convolution neural network, Image classification, Scene recognition

1 引言

场景识别是计算机视觉领域的重要研究课题,对于智能监控、汽车自动驾驶等应用具有重要意义。场景识别问题实

际上是对不同场景图像的分类问题。对于图像分类问题的算法主要分为两类:基于人工设计的人工描述子分类算法以及基于深度学习的图像分类算法。相较于传统人工设计的特征描述子算法,卷积神经网络提取图像特征有很强的鲁棒性,因

收稿日期:2018-08-03 返修日期:2018-11-12 本文受国家自然科学基金(61572147),广西科技计划项目(AC16380108),广西图像图形智能处理重点实验室项目(GIIP201701),广西研究生教育创新计划资助项目(2018YJCX46),江西省自然科学基金资助项目(20171BAB212015)资助。

江泽涛(1961—),男,博士,教授,CCF 会员,主要研究方向为计算机视觉;秦嘉奇(1993—),男,硕士生,主要研究方向为计算机视觉,E-mail:1445155606@qq.com(通信作者);胡 硕(1983—),博士生,讲师,主要研究方向为智能计算。

此使用卷积神经网络提取到的图像特征进行分类能够得到很好的效果。自2006年Hinton等提出深度置信网络^[1]以来,深度学习得到了快速发展,卷积神经网络逐渐得到各个学科领域的关注,在自然语言处理、语音处理、视频图像处理中都得到了广泛应用。近年来,在图像分类、目标识别等多项任务中,卷积神经网络都取得了优异成果,越来越受到研究者的关注,成为目前的研究热点。

随着新的卷积神经网络不断被提出,研究者们通过在结构上改进卷积神经网络^[2-7]、分类器优化等方法^[8-9],使卷积神经网络在计算机视觉任务上有了更广泛的普适性^[10-11],在ImageNet上的识别率突破了人眼的极限。但是,一般卷积神经网络仅针对单通道灰度图像或者三通道的彩色图像^[12],无法处理灰度红外图像和三通道彩色图像这样一组多光谱图像对。目前,已经有多种用于多光谱图像的特征描述子被提出^[13-16],这些算法在红外图像的分类、识别方面取得了较好的效果,但仍然存在对场景中光照变化的鲁棒性不强、计算量大的问题。为了解决该问题,Zhang等^[17]建立了红外光和可见光数据集,并使用VGG-16网络来进行图像识别,当某光谱图像无法识别时,采用切换到另外一个光谱下进行识别的策略;刘峰等^[18-19]直接使用多个并列的卷积神经网络,将各个光谱段的图像单独送入网络进行特征提取,并在网络末端进行特征融合,从而实现了多光谱图像特征融合的舰船识别;Hajhassan等^[20]设计了一种能接受多光谱直肠活检图像的卷积神经网络,用于在早期诊断直肠癌。然而,上述方法都是对图像的各个通道单独考虑,待特征提取后再将最终特征进行

融合,或者在可见光彩色场景图像和红外光灰度场景图像之间做出取舍,并没有考虑各个光谱图像通道之间的关联,因而损失了部分特征信息;并且上述方法没有考虑到多源图像训练数据太少而导致网络难以训练的问题。

针对以上问题,本文提出一种多光谱卷积神经网络,该网络不受卷积神经网络只接受单一光谱图像作为输入的限制,使得卷积神经网络能够同时接受红外光图像和可见光图像作为输入。对于同样的场景,彩色图像中所含有的信息量比灰度图像中含有的信息量多,这是因为彩色图像中有更多的颜色通道。可以想象,如果在红绿蓝三色通道的基础上再加上波长比红色光更长的红外光通道,场景图像中含有的信息量将更加丰富,有助于提高场景识别的精度。同时,为了解决红外光图像数据较少使得网络训练困难的问题,将预训练技术^[21]加入到多光谱卷积神经网络中。基于多光谱图像的多路卷积神经网络能够分别提取到红外光图像和可见彩色光图像的特征,并且通过通道合并,能提取到红外光图像与可见光图像之间的关联特征。最后,本文通过实验证明了所提方法对多光谱场景有较好的识别率。

2 卷积神经网络

卷积神经网络(Convolutional Neural Network, CNN)由特征提取网络和分类器网络两大部分组成。典型的CNN是一种多层神经网络,主要由输入层、卷积层、降采样层、全连接层和输出层组成。用于图像分类的卷积神经网络的基本结构如图1所示。

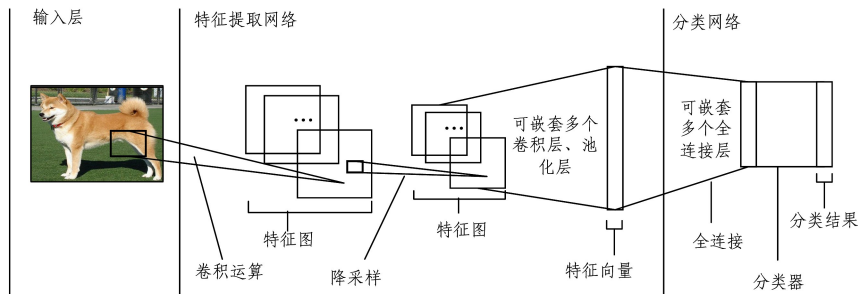


图1 用于分类的卷积神经网络的基本结构

Fig. 1 Basic structure of convolution neural network for classification

特征提取网络是CNN网络的核心,主要由卷积层和降采样层组成。在多数情况下,卷积层包括卷积(Convolution)、激活(Activation)、池化(Pooling)3个操作。其中,卷积操作用于提取图像的特征,卷积核的数量决定了该层输出特征图的特征数量。经过卷积得到的特征图被送入激活层进行激活操作,然后对特征数据进行非线性映射,最后为了降低特征图的大小,将激活后的特征图送入池化层,以减少整个输出特征数量。卷积神经网络会嵌套多个这样的卷积层,最终得到一个特征向量。也就是说,卷积神经网络中的若干层完成了从图像到特征之间的转换。根据不同的任务,将特征向量送入不同的网络中得到最后结果。对于图像分类任务,则将特征向量送入分类网络,得到最后的分类概率。

随着卷积神经网络的发展,各层的具体操作也有一些变

化。例如对于InceptionNet,使用 1×3 和 3×1 大小的卷积来代替以往的 3×3 卷积核以减少参数数量;激活函数由最初的sigmoid衍变为RELU、maxout等;池化函数也从平均池化变成现在常用的最大池化。但是,上述变化都没有改变卷积神经网络本质,只是使卷积特征表述变得更加高效。同时,为了减少过拟合现象,dropout, batch normalization等手段也加入到卷积神经网络中。为了使卷积神经网络中的权重和偏执符合图像特征提取的要求,我们需要使用反向传播算法(BP算法)对网络进行训练。

卷积神经网络使用多个卷积核进行多次卷积与降采样来提取特征。单个卷积核只能提取一种特征,但多个卷积核可以提取到图像的多个特征。同时,多个卷积核的堆叠能够提取到上层特征图的特征,也就是图像的高维特征,这就使得提

取到的特征具有很强的尺度不变性和旋转不变性。与 SIFT^[22], BRISK^[23] 等人工设计的特征描述子相比,卷积神经网络能够更好地提取到图像的特征信息。

3 多路、多光谱卷积神经网络

经典的卷积神经网络,诸如 AlexNet, VGG-Net, InceptionNet, ResNet 以及 DenseNet 等,都接受三通道的彩色图像输入。以 50 层的 ResNet-50 为例,ResNet-50 网络模型接受 224×224 大小的三通道图像输入,输入图像经过一个简单的卷积层和池化层,得到 $56 \times 56 \times 64$ 的特征图,这 64 个特征图再经过 48 个残差学习单元,最终将图像转化为 1000 维的特

征向量。由于经典卷积神经网络都是接受三通道的图像输入,无法适用于可见光图像(三通道)和近红外图像(单通道)共 4 个通道的输入,因此为使可见光图像+近红外图像也能够使用卷积神经网络,本文提出一种基于多路卷积神经网络的多光谱场景图像识别方法。

3.1 多路、多光谱卷积神经网络结构

为了使卷积神经网络适用于多于三通道的图像输入,本文提出一种适用于多光谱图像的卷积神经网络结构,其能够接受 RGB+NIR 一共 4 个通道的输入,并充分利用各个光谱段图像的信息提取特征。多路、多光谱卷积神经网络结构如图 2 所示。

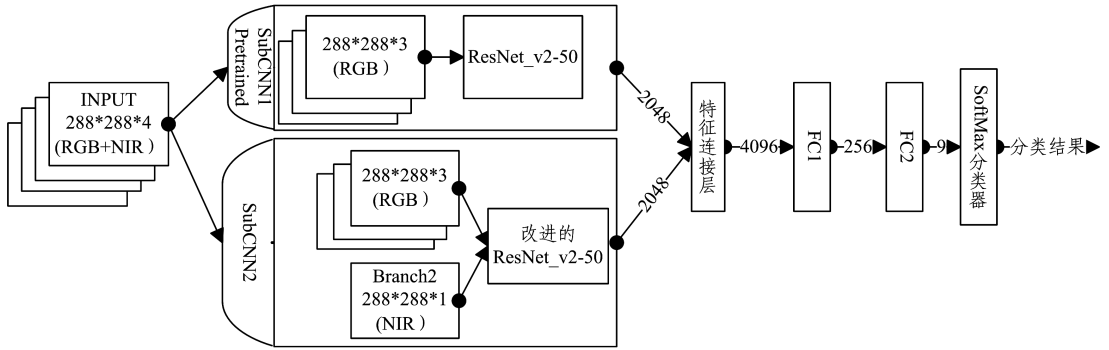


图 2 多路多光谱卷积神经网络结构

Fig. 2 Structure of multichannel and multispectral convolution neural network

多路、多光谱卷积神经网络将可见光图像和 NIR 图像组成的四通道图像送入网络后,分成两组进行处理。三通道可见光图像送入网络的 SubCNN1, RGB+NIR 四通道图像送入网络的 SubCNN2。为了解决训练数据太少的问题,SubCNN1 需要使用预训练方法。在本文中,子网络 SubCNN1 使用经典的 ResNet-50 网络,该子网络使用在 ImageNet 数据集上训练得到的权重作为初始化权重,最后输出 2048 维的可见光图像特征向量。SubCNN2 子网络是经过改进的 Resnet-50 网络。改进的 ResNet-50 网络将输入层改为四通道图像输入,网络之后的各层在权重和输出上均按照 ResNet-50 的结构做出相应改变。SubCNN2 最终也输出 2048 维的特征向量。SubCNN1 和 SubCNN2 两个子网络中都不包含全连接层,仅有卷积层、池化层以及激活层。由 SubCNN1 和 SubCNN2 得出的特征组合,得到 4096 维的特征向量,将其送入两个全连接层 FC1 和 FC2,最终将特征输出为 9 分类结果表示。

要将原本三通道输入的 ResNet-50 改成四通道的输入,主要是对其第一个卷积层进行改进。一个卷积层的卷积运算可以由卷积核的大小、卷积核的数量、卷积步长(stride)和卷积填充(padding)来表示。原 ResNet-50 的第一个卷积层由 64 个 $7 \times 7 \times 3$ 的卷积核组成,卷积步长(stride)为 2。为了将 ResNet-50 网络用于四通道的输入,改进的 ResNet-50 网络将第一个卷积层的卷积核大小变成 $7 \times 7 \times 4$,卷积核数量、stride、padding 都不变。如此处理后,第一个卷积层卷积运算的第 3 个维度的长度变为 4,但是该层的输出仍然为 64 个特征图,因此,之后的残差学习单元都不需要进行修改。

由于输入时将可见光图像和 NIR 图像一同作为四通道图像输入,SubCNN2 提取的特征不仅包含了 NIR 图像的特征信息,还包括了可见光图像与 NIR 图像之间的关联特征,这些关联特征能够辅助神经网络进行识别分类,提高分类的正确率。两个子网络均得到 2048 维的特征向量,经过特征连接层后,得到 4096 维的最终特征向量。该特征向量经过两个全连接层之后,送入分类网络,最终得到 9 分类结果。

3.2 多光谱图像分类算法流程

多光谱图像分类卷积神经网络需要接受多个图像的输入。本文算法的输入是一组图像对。在图像处理中,对于任何一张图像,其数据规格可以表示为 $W \times H \times C$,其中 W 和 H 表示图像的宽度和高度, C 表示图像的通道数。对于可见光图像,一般其通道数 $C=3$,有 RGB 3 个通道,而近红外图像(NIR 图像)只有一个通道。在多路、多光谱卷积神经网络中,可以认为输入图像是由可见光图像的 RGB 通道再加上 NIR 图像一共 4 个通道组成的。由于数据集提供的图像大小不一定相同,如果将图像缩小后输入,则会影响分类结果,而图像不经降采样又会导致网络参数太大,因此,将输入图像统一变换为 288×288 像素。每一组图像对由同一取景角度的可见光图像和近红外(NIR)图像组成,可见光图像是三通道的 RGB 图像,NIR 图像是单通道的灰度图像。本文算法流程主要包含 3 个步骤:1)利用多路卷积神经网络提取图像对特征;2)对提取到的特征进行特征连接,得到最终特征;3)将最终特征送入分类网络,得到最终分类结果。网络在训练之前,需要将预训练的权重参数加载到 CNN Branch1 中。算法流程示意图如图 3 所示。

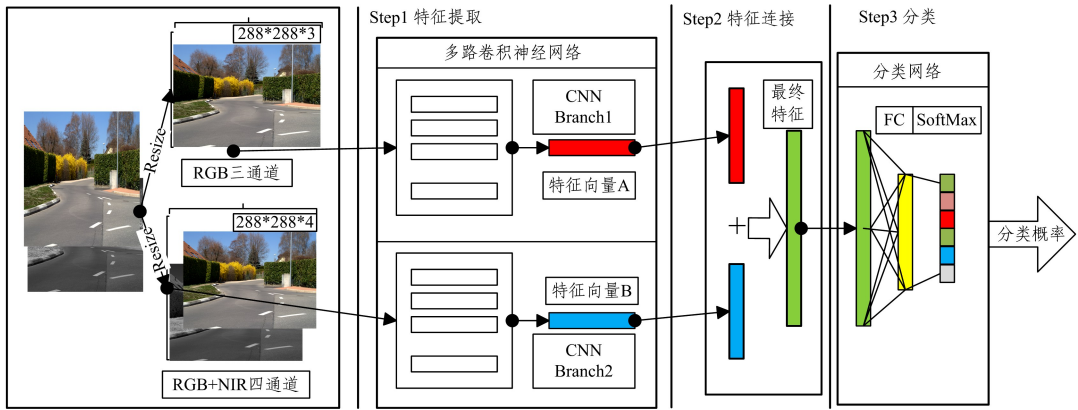


图3 基于多路卷积神经网络的多光谱图像分类算法流程

Fig. 3 Multi-spectral image classification algorithm based on multi-convolution neural network

本文算法实现了多光谱图像端到端的训练,使得卷积神经网络的使用不受输入图像的通道数量的限制,并且有效地利用了图像中各个光谱段的特征信息以及光谱段之间的关联信息。将本文算法与预训练神经网络结合使用,解决了训练数据较少的问题,即使在仅有较少训练数据的情况下,也能够得到较好的分类结果。

4 实验结果与分析

4.1 实验研究

本文使用公共 NIR-RGB 数据集作为实验数据集。该数据集包含了 477 对 NIR-RGB 图像,这些图像被分为 9 类场景。每一对图像包含同一场景、同一角度拍摄的 RGB 和 NIR 图像,图像的大小都在 1024×600 左右。数据集的详细分类情况以及样图如图 4 所示。

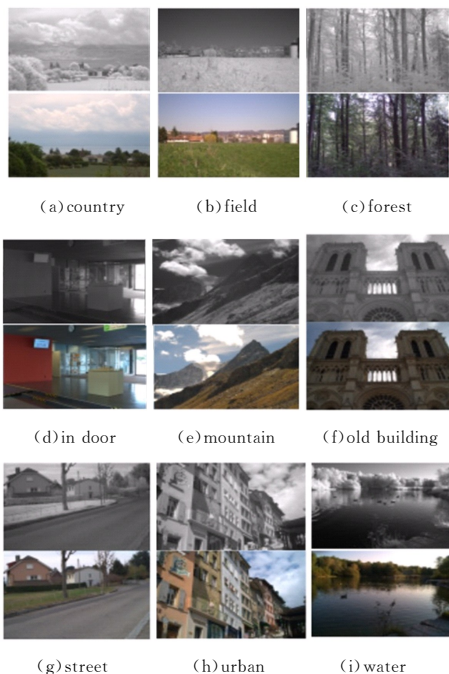


图4 NIR-RGB数据集

Fig. 4 NIR-RGB dataset

实验使用 tensorflow 搭建实验环境,并使用两块 GTX1080TiGPU 进行训练。网络结构如图 2 所示,在实验

中,子网络 SubCNN1 使用 ResNet_v2_50,并且将该网络在 ImageNet 数据集上得到的权重作为其初始化权重。数据集图像全部规范化到 288×288 像素,采样方法使用两次立方,并且将 NIR 图像拼接到可见光图像的 RGB 通道之后,组成 RGB_NIR 规格的四通道图像。在训练时,将提取到的特征组合后进行 0.2 的 dropout 后送入全连接层。实验中使用多种方法与本文方法进行对比,包括预训练的 VGG-16、Inception-Net、ResNet-50 以及四通道的 ResNet-50,另外还对比了基于人工设计的特征描述子算法的效果。由于使用普通卷积神经网络只能输入三通道或者单通道图像,因此使用普通卷积神经网络的实验均仅使用了可见光图像。实验数据集中,477 组数据随机打乱,取其中 80% 作为训练集,20% 作为测试集。每次实验结果为网络收敛后若干次测试集测试精度的平均值,最终平均精度为 5 次实验的平均值。除了 InceptionNet 网络在 300 代之后收敛外,其余网络均在 30 代左右收敛,取 30 代到 70 代之间的测试集平均精度为一个实验结果。实验结果如表 1 所列。

表1 本文方法与其他方法的实验结果

Table 1 Experimental results of proposed method and other methods

方法	预训练	初始学习率	使用数据	数据增强	平均精度/%
MSIFT with PCA ^[9]	—	—	RGB+NIR	—	73.1
FV+SVM with PCA ^[12]	—	—	Only RGB	—	87.9
mCENTRIST with PCA ^[10]	—	—	RGB+NIR	—	84.5
VGG16	是	0.05	Only RGB	否	82.15
InceptionNet	是	0.02	Only RGB	否	84.50
ResNet-50	是	0.012	Only RGB	否	83.34
四通道 ResNet-50	否	0.008	RGB+NIR	否	52.10
多路、多光谱 ResNet-50	部分	0.012	RGB+NIR	否	88.94
多路、多光谱 ResNet-50	部分	0.012	RGB+NIR	是	93.35

扩展的多路、多光谱卷积神经网络模型使用了同一场景的 NIR 光谱图像和可见光图像,数据存在一定冗余(主要集中在 NIR 光谱和可见光的 R 通道),且面对的数据集数据量很少,全部数据仅有 477 组,极有可能出现过拟合现象。为验证扩展的网络模型是否出现了过拟合现象,通过实验得出 ResNet-50 网络以及多路、多光谱 ResNet-50 网络在没有数据增强的情况下训练集和测试集的收敛曲线,如图 5 所示。

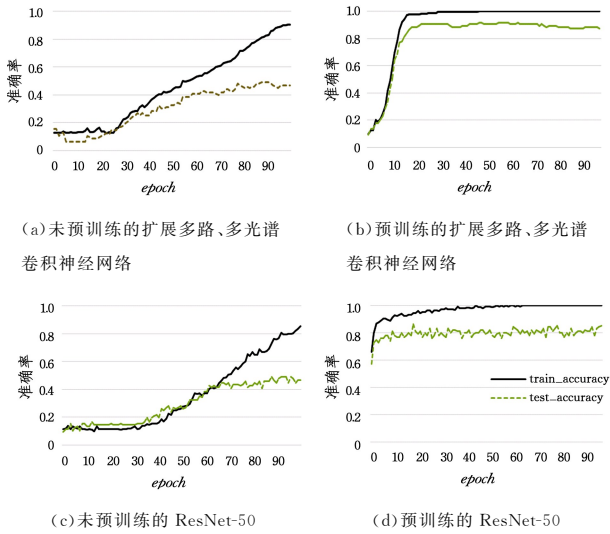


图 5 扩展多路、多光谱卷积神经网络和 ResNet-50 的准确率收敛曲线

Fig. 5 Accuracy rate convergence curves of extended multichannel, multispectral convolution neural networks and ResNet-50

实验过程中,基于卷积神经网络的方法均使用随机梯度下降法(Stochastic Gradient Descent,SGD)。由于网络结构不同,在实验中,不同的卷积神经网络使用的学习率初始值不一定相同。所有卷积神经网络的学习率递减设置为每 10 步降低 5%,训练时,网络每次训练 32 个数据,即 BatchSize 为 32。最后,在实验中对原始数据使用了数据增强技术,将每一个训练数据做垂直翻转,使得一个训练数据扩充为两个。

4.2 实验结果分析

对比上述实验结果可知,基于卷积神经网络的方法得到的效果已经接近于特征描述子方法,使用多路、多光谱的 ResNet-50 网络能够得到较好的效果。结果表明,使用了数据增强后,识别精度得到再一次提高。

从图 5(a)、图 5(c)可以看出,无论是 ResNet-50 还是扩展的多路、多光谱卷积神经网络,在没有经过大数据集预训练的情况下,都出现了严重的过拟合现象,原因在于用于训练的数据量太少。相对地,图 5(b)、图 5(d)是经过 ImageNet 数据集预训练的网络准确率收敛曲线,可见经过预训练的网络即使在数据量很少的情况下也仅仅是在 70 次迭代之后出现了轻微的过拟合。由于实验网络并未使用除预训练以外的抗过拟合技术,在如此小的数据集上出现轻微的过拟合现象是可以接受的。

在没有经过大数据集训练的情况下,由于数据量较少导致出现过拟合现象,即使加入了 NIR 通道的四通道 ResNet-50 也不能取得较好的效果。尽管扩展的网络模型规模变大,理论上其训练难度增加,更加容易出现过拟合现象,但从图 5(b)、图 5(d)的对比可知,相对于未经扩展的 ResNet-50,扩展的多路、多光谱卷积神经网络在同等条件下并未表现出更加严重的过拟合现象。本文方法出现过拟合现象是由于数据量不足,而非扩展网络模型自身所导致。

对比预训练的 VGG16、InceptionNet 和 ResNet-50 网络

可知,三者的准确率是相近的,说明在三通道网络中图像数据的特征已经被有效提取,限制精度的是网络通道数量而非网络结构。与之对比,使用了预训练的多路、多光谱 ResNet-50 网络能够取得更好的效果,原因在于多路、多光谱网络在结合了预训练模型特征的情况下,还提取到了 NIR 图像与可见光图像这两种光谱图像之间的相关特征,图像数据的利用率更高。尽管 NIR 通道与可见光的 R 通道波段相近,但是通过加入 NIR 通道,整体精度提高了将近 4.5%。这说明在同一组图像中,NIR 图像与 RGB 3 个通道值之间尽管有冗余信息,但是卷积运算以及特征全连接运算将这些冗余信息去除,并将这 4 个通道之间的关联信息及其特征提取、保留下来。人们通常更加关注图像中像素邻域之间的关系,也就是图像像素的空间相关信息;从另外一个角度看,图像通道之间同样存在着关系,要找出图像通道之间的关系,就需要使用三维卷积操作,即对图像长度、宽度和深度 3 个方向进行卷积,本文将 NIR 图像作为彩色图像的通道一同进行卷积,正是出于这样的想法。

结束语 本文通过改进经典的卷积神经网络,提出了一种多路、多光谱卷积神经网络。该网络能够适用于多个不同成像技术得到的图像对数据;通过在网络中加入预训练网络分支的方法,解决了多源图像对数据不足导致过拟合的问题。通过多路、多光谱卷积神经网络的数据不仅能够提取到多个光谱中每个光谱的特征,还能够提取到多个光谱之间的关联特征。通过实验证明,本文提出的方法相较于其他方法在准确度上具有一定优越性。本文提出的网络结构也能够很容易地应用到其他多光谱图像分类、识别任务上。然而,目前多光谱图像太少,能用于训练的数量较多的多光谱图像数据集更是少之又少。因此,下一步工作计划建立一个数量庞大的 NIR_RGB 图像对数据集;充分利用该网络的可迁移性来改进本文网络结构,力图将多路、多光谱卷积神经网络用于单一光谱拍摄图像不清晰的情况。另外,在可见光图像成像的基础上,还可以将热红外图像、紫外线强度传感器数据甚至声波成像等信息数据加入到场景识别中,以提高场景识别的精度和速度。同时,还需进一步对多路、多光谱卷积神经网络结构进行研究,以更加合理地组织、利用多光谱图像信息。

参考文献

- [1] HINTON G E, OSINDERO S, TEH Y W. A Fast Learning Algorithm For Deep Belief Nets[J]. *Neural Computation*, 2006, 18(7):1527-1554.
- [2] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[C]// *International Conference on Neural Information Processing Systems*. Curran Associates Inc., 2012:1097-1105.
- [3] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[C]// *Proceedings of the International Conference on Learning Representations*. Oxford, USA, 2015.

- [4] SZEGEDY C, LIU W, JIA Y, et al. Going deeper with convolutions[C]//IEEE Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, 2015:1-9.
- [5] HE K M, ZHANG X Y, REN S Q, et al. Deep Residual Learning for Image Recognition[C]//IEEE Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, 2015: 770-778.
- [6] GAO H, ZHUANG L, MAATEN L V D, et al. Densely Connected Convolutional Networks[C]//IEEE Conference on Computer Vision & Pattern Recognition. IEEE Computer Society, 2017.
- [7] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[J/OL]. <https://arxiv.org/abs/1709.01507>.
- [8] ROUTRAY S, RAY A K, MISHRA C, et al. Efficient hybrid image denoising scheme based on SVM classification[J]. *Optik*, 2018, 157:503-511.
- [9] LEE S J, CHEN T, YU L, et al. Image Classification Based on the Boost Convolutional Neural Network[J]. *IEEE Access*, 2018, 6(99):12755-12768.
- [10] SHI L, WANG Y M, CAO Y J, et al. Vehicle Type Recognition Based on Deep Convolution Neural Network[J]. *Computer Science*, 2018, 45(5):280-284.
- [11] GONG A, FEI F, ZHENG J. Multi-person Behavior Recognition Method Based on Convolutional Neural Networks[J]. *Computer Science*, 2018, 45(2):306-311, 321. (in Chinese)
龚安, 费凡, 郑君. 基于卷积神经网络的多人行为识别方法[J]. *计算机科学*, 2018, 45(2):306-311, 321.
- [12] ZHOU F Y, JIN L P, DONG J. Review of convolutional neural network[J]. *Chinese Journal of Computers*, 2017, 40(6):1229-1251. (in Chinese)
周飞燕, 金林鹏, 董军. 卷积神经网络研究综述[J]. *计算机学报*, 2017, 40(6):1229-1251.
- [13] BROWN M, SÜSSTRUNK S. Multi-spectral SIFT for scene category recognition[C]//IEEE Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, 2011: 177-184.
- [14] XIAO Y, WU J, YUAN J. mCENTRIST: A Multi-Channel Feature Generation Mechanism for Scene Categorization[J]. *IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society*, 2014, 23(2):823-836.
- [15] VAPNIK V N. An overview of statistical learning theory[J]. *IEEE Transactions on Neural Networks*, 1999, 10(5):988-999.
- [16] SALAMATI N, LARLUS D, CSURKA G. Combining Visible and Near-Infrared Cues for Image Categorisation[C]//British Machine Vision Conference. BMVA Press, 2011.
- [17] ZHANG M M, CHOI J, DANILIDIS K, et al. VAIS: A dataset for recognizing maritime imagery in the visible and infrared spectrums[C]//Computer Vision and Pattern Recognition Workshops. IEEE, 2015:10-16.
- [18] LIU F, SHEN T S, MA X X, et al. Ship recognition based on multi-band neural network[J]. *Optics and Precision Engineering*, 2017, 25(11):2939-2946. (in Chinese)
刘峰, 沈同圣, 马新星, 等. 基于多波段深度神经网络的舰船目标识别[J]. *光学精密工程*, 2017, 25(11):2939-2946.
- [19] LIU F, SHEN T S, MA X X. Convolutional Neural Network Based Multi-Band Ship Target Recognition with Feature Fusion[J]. *Acta Optica Sinica*, 2017(10):240-248. (in Chinese)
刘峰, 沈同圣, 马新星. 特征融合的卷积神经网络多波段舰船目标识别[J]. *光学学报*, 2017(10):240-248.
- [20] HAJHASSAN H, CHADDAD A, HARKOUSS Y, et al. Classifications of Multispectral Colorectal Cancer Tissues Using Convolution Neural Network[J]. *Journal of Pathology Informatics*, 2017, 8(1):1-7.
- [21] WANG J, LUO C, HUANG H, et al. Transferring Pre-Trained Deep CNNs for Remote Scene Classification with General Features Learned from Linear PCA Network[J]. *Remote Sensing*, 2017, 9(3):225.
- [22] LOWE D G. Distinctive image features from scale-invariant keypoint[J]. *International Journal of Computer Vision*, 2004, 60(2):91-110.
- [23] LEUTENEGGER S, CHLI M, SIEGWART R Y. BRISK: Binary Robust invariant scalable keypoints[C]//IEEE International Conference on Computer Vision. IEEE, 2012:2548-2555.