

文章编号:1007-2780(2018)04-0317-09

基于深度卷积神经网络的输电线路 可见光图像目标检测

周筑博^{1*}, 高 佼², 张 巍³, 王晓婧¹, 张 静¹

(1. 天津航天中为数据系统科技有限公司(天津市智能遥感信息处理技术企业重点实验室), 天津 300301;

2. 济南汤尼机器人科技有限公司, 山东 济南 250101;

3. 南方电网科学研究院有限责任公司, 广东 广州 510080)

摘要:为了检测输电线路可见光图像中的塔材、玻璃绝缘子和复合绝缘子, 本文采用了一种基于深度卷积神经网络的技术。通过有人直升机搭载高清相机拍摄 19 条不同的输电线路近 600 张图片, 对图片中的背景、塔材、玻璃绝缘子和复合绝缘子目标进行人工标注及分块, 采用数据扩展生成包含 15 万个样本的输电线路图像库。构造 5 层深度卷积神经网络, 首先用 Cifar-100 数据集对网络进行预训练, 然后用输电线路图像库进行网络调优。本文方法在检测真阳率为 90% 时, 假阳率低于 10%, 明显优于传统方法, 可用于输电线路可见光图像中的塔材、玻璃绝缘子和复合绝缘子检测, 检测结果可用于诊断参考或进一步的目标状态分析。可对输电线路可见光图像中的塔材和绝缘子目标进行检测, 并可扩展到其它类型目标的检测。

关键词:输电线路图像; 绝缘子; 目标检测; 深度学习; 卷积神经网络

中图分类号: TP751.1 **文献标识码:** A **doi:** 10.3788/YJYXS20183304.0317

Object detection of transmission line visual images based on deep convolutional neural network

ZHOU Zhu-bo^{1*}, GAO Jiao², ZHANG Wei³, WANG Xiao-jing¹, ZHANG Jing¹(1. *Tianjin ZhongWei Aerospace Data System Technology Co., Ltd, Tianjin Key Laboratory of Intelligent Information Processing in Remote Sensing, Tianjin 300301, China;*2. *Jinan Tony Robotics Co., Ltd., Jinan 250101, China;*3. *Electric Power Research Institute, China Southern Power Grid Co., Ltd., Guangzhou 510080, China*)

Abstract: A deep convolutional neural network based method is adopted to detect objects such as tower, glass insulator and composite insulator in visible images of transmission lines. About 600 visible images of 19 different transmission lines are captured by manned helicopter with high-definition camera. All of the images are then annotated manually and segmented into blocks with 4 different labels: background, tower, glass insulator and composite insulator. These blocks are then augmented to

收稿日期: 2017-12-04; 修订日期: 2017-12-31.

* 通信联系人, E-mail: zhouzhuo90@163.com

around 150 000 training samples which comprise the transmission line image dataset. A five-layer deep convolutional neural network is designed and pre-trained by using Cifar-100 dataset, the trained network is then fine-tuned by using transmission line image dataset. The experimental results show that when detection true positive rate is 90%, the false alarm rate is less than 10%, which is obviously superior to the traditional methods. It can be used for the detection of tower, glass insulator and composite insulator in visible images of transmission lines. The detection result can be used as reference for diagnosis or state analysis of transmission lines. This method can be used to detect tower and insulator in visible images of transmission lines, and can be extended to detect other typical objects.

Key words: transmission line image; insulator; object detection; deep learning; convolutional neural networks

1 引言

近年来,随着输电线路有人直升机/无人机巡检技术的逐渐推广,使得基于可见光图像的输电线路目标分析成为了现实。输电线路本身是一种复杂的人造设施,其主要组成部件尺度差异较大。大尺度部件主要有:各类型的杆塔塔体、各类导线、安装有各种类型及方位的绝缘子、连接杆塔两端的引流线等。小尺度部件主要有:防震锤、间隔棒、均压环、线夹等。要实现输电线路目标状态自动分析,首先需要检测目标在图像中的位置。

输电线路图像目标的检测与普通图像目标检测是相同的,都是通过特定的算法对原始图像进行分析,确定目标在图像中的位置。通常,目标检测结果为目标的外接矩形框(Bounding Box),而输电线路图像中存在各种角度的“条”状目标(如绝缘子),矩形框无法描述其精确位置,因此需要以掩码的形式确定目标在图像中的位置。

输电线路可见光图像的目标检测主要包括对塔材、绝缘子串、导线、防震锤等典型设备目标的检测。以绝缘子串检测为例,分别有利用点特征^[1]、纹理特征^[2]、形态学特征^[3]以及多特征结合^[4]的检测方法。但每种方法都有其局限性,利用颜色特征的检测方法通常用于玻璃绝缘子的检测,对瓷绝缘子和复合绝缘子则不适用。利用特征点的方法对图像清晰度要求很高。利用纹理特征的方法则主要用于复合绝缘子的检测,易受背景噪声影响。

随着计算机硬件性能的提升以及相关算法的提出,深度学习技术^[5]逐渐取代了经典的机器学习方法,在图像目标识别与检测、语音识别、自然语

言处理、自动驾驶等领域取得了显著的效果。其中,在目标检测领域,包括日常物体检测^[6]、人脸检测^[7]、文字检测^[8]、医学图像检测^[9]乃至行为检测^[10]等方面,基于深度学习的方法效果远远超过了传统方法。深度学习算法通过收集大量训练数据,学习描述力更强的模型,利用巧妙的训练技巧来提高算法精度。在现实场景中,由于拍摄角度、拍摄距离、光照等因素的差异,同一类目标在图像中的表现形式是多种多样的。因此利用机器学习来进行目标检测时,需要大量的训练样本,能够涵盖不同角度、不同尺度、不同光照的情况。目前,公开的带标记图像库中,十万张数量级的有 Cifar-100^[6]、NORB^[11]和 Caltech-101^[12]等,百万张数量级的有 ImageNet^[13]、LabelMe^[14],包含了不同场景下(室内、街道、野外等)不同种类物体(人造物、自然物、动物、植物等)的图像及对应的类别标签。

在众多类型的深度神经网络中,卷积神经网络(Convolutional Neural Networks, CNN)是以哺乳动物的视觉感受为启发而提出的,在计算机视觉领域用于图像的分类与识别。LeCun^[15]在1989年就将其用于手写数字识别。由于当时训练数据集较少,计算机运算能力不足,没有应用于自然场景下物体的识别。随着公开图像数据集不断增加、硬件计算能力的提升, Alex 等^[16]提出了一个8层的深度卷积神经网络“AlexNet”,以 ImageNet 图像库中的120万张带标记数据为训练样本,在训练过程中加入特殊的训练技巧防止过拟合,最终准确率大大超越了传统算法,从而在学术界和工业界掀起了一轮深度学习的热潮。之后,研究者们提出了新的网络结构及训练方法,正确率得到了进一步提升。

本文采用一个5层的深度卷积神经网络,检测输电线路可见光图像中的背景、杆塔、绝缘子区域。构造了包含4类目标的输电线路数据集:背景、杆塔、玻璃绝缘子和复合绝缘子。用Cifar-100数据集对网络预训练,然后利用输电线路数据集进行网络调优。检测时,将图像分为图像块,利用神经网络识别每个图像块的类别,将相同类别图像块所在区域合并后即为此类别目标在图像中的区域。实验结果表明,该方法明显优于传统方法,可用于输电线路可见光图像中典型目标检测,检测结果可用于诊断参考或进一步的目标状态分析。

2 原理

输电线路可见光图像数据数量相对较少,其标记任务也需要大量人工完成。为此,本文从两方面解决输电线路图像数据量缺乏的问题,第一,采用数据扩展(Data Augmentation)方法对已有数据进行扩充,包括图像的平移、旋转、尺度和颜色变换;第二,利用公开数据集对模型进行预训练,得到初步的模型,再用输电线路图像数据对模型进行调优。

算法流程主要分为3个阶段:样本构造阶段、模型训练阶段和目标检测阶段。

1)样本构造阶段。对输电线路原始图像中的玻璃绝缘子、复合绝缘子、塔材和背景4类目标进行标注,根据标注结果生成4类目标的样本。

2)模型训练阶段。首先定义深度神经网络模型,并用Cifar-100图像数据库训练;调整网络模型输出层个数,并保持其它层结构及参数不变;最后,利用样本构造阶段生成的数据对修改后的网络进行微调,得到最终的网络参数。

3)目标检测阶段。对于任意图像,首先进行分块,然后利用训练好的模型判断每一个子块的类别标签。

2.1 数据源

2.1.1 Cifar-100 图像数据库

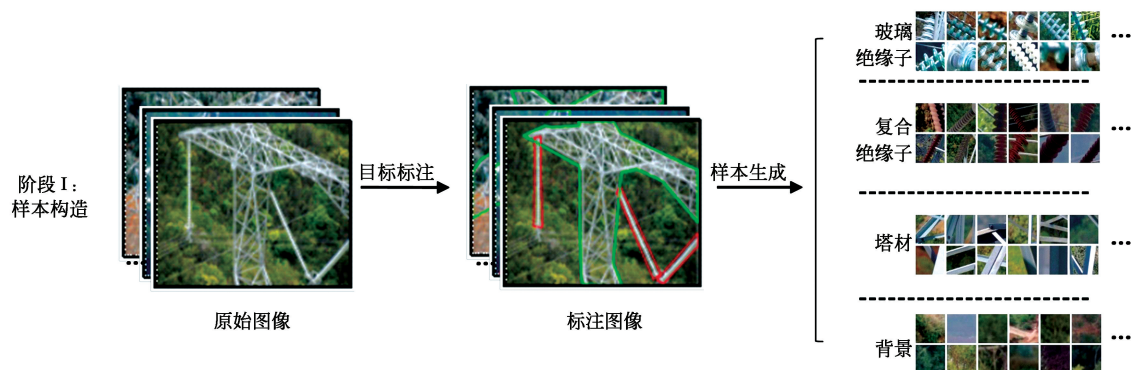
Cifar-100是一个用于普适物体识别的数据集,共有60 000张大小为32 pixel×32 pixel的彩色图像,包含了100类常见物体(如家具、交通工具、建筑、动物等),每类600个样本。其中50 000个样本为训练集,10 000个样本为测试集。Cifar-100广泛用于计算机视觉领域目标分类与识别算法的开发和测试。

2.1.2 输电线路图像库

输电线路图像库原始数据是南方电网公司在输电线路巡检过程中,由直升机搭载高清摄像机拍摄获得,包含19条不同线路(220 kV与500 kV)共600余张图片。每张图片中包含杆塔、玻璃绝缘子、复合绝缘子中的一类或多类目标。由于拍摄时背景、光照的变化以及拍摄角度、距离、焦距的不同,各类目标在不同图片中的色彩、角度、尺度会存在差异。

2.2 样本构造

样本构造的目标是输电线路图像库。如图1(a)所示,样本构造分为目标标注和样本生成两个步骤。目标标注即将输电线路原始图像中的背景、杆塔、玻璃绝缘子和复合绝缘子4类区域用多边形框标记,均由人工完成。如图2所示,其中黄色区域表示背景,绿色区域表示杆塔,红色区域表示玻璃绝缘子,蓝色区域表示复合绝缘子。完成标注后,对图像进行分块,根据标注结果,指定每个图像块所属的类别标签:背景—1、塔材—2、玻璃绝缘子—3、复合绝缘子—4。



(a)样本构造阶段

(a) Sample construction phase

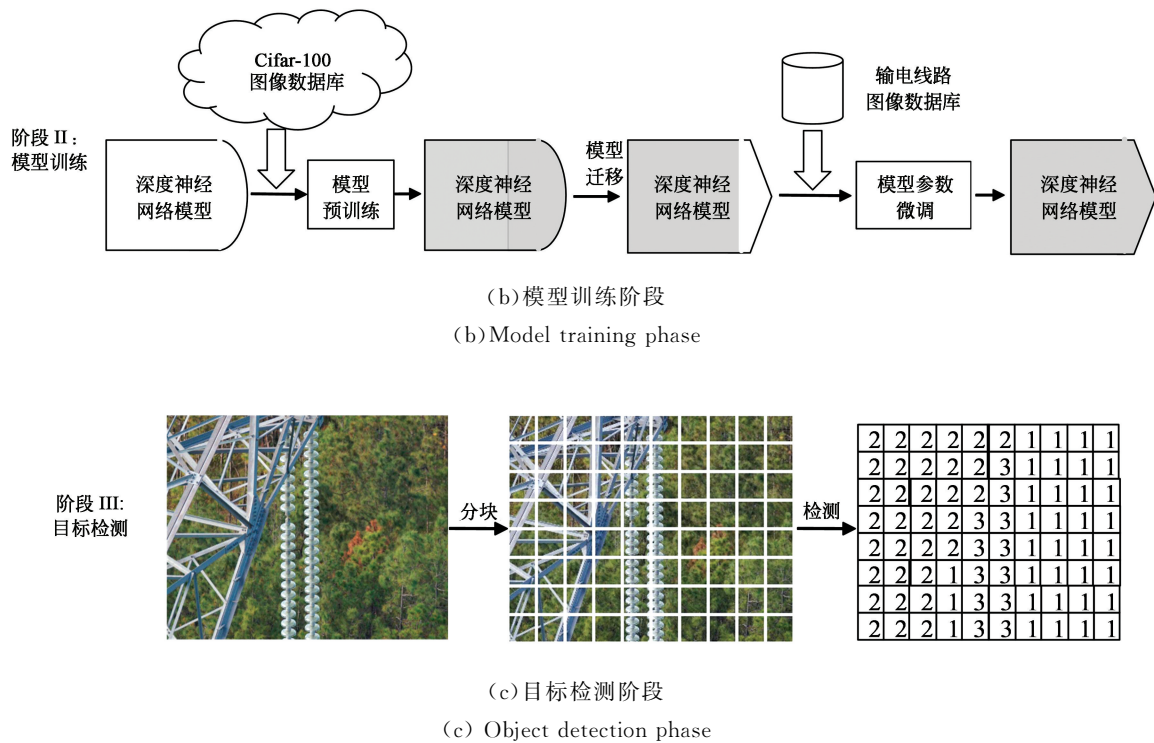


图 1 深度卷积神经网络目标检测算法流程

Fig.1 Algorithm procedure of deep neural network based object detection

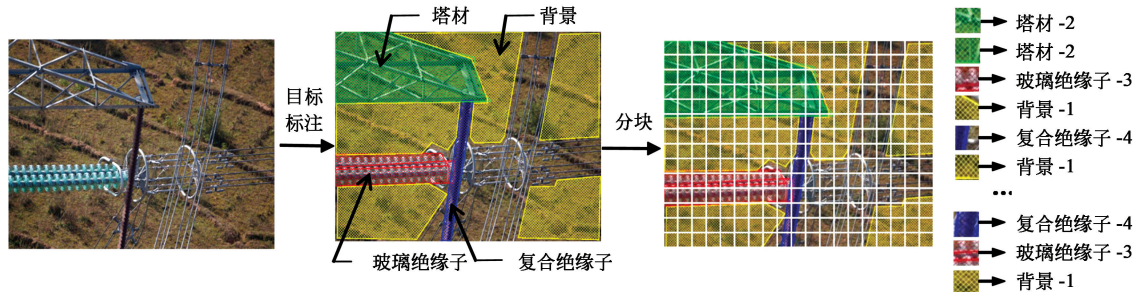


图 2 输电线路图像库目标样本构造

Fig.2 Sample construction of transmission line images dataset

为了获取更多样本,采用数据扩展方法来增加样本的数量。数据扩展通过对图像块的旋转、平移、尺度、颜色 4 种变换实现,这样有两个好处:第一,训练样本数量大大增加,防止训练过程中出现过拟合。第二,使算法在进行目标识别时具有更好的平移不变性、旋转不变性、尺度不变性和颜色不变性。如图 3 所示,分别展示了 4 种变换形式,其中实线代表原始图像块,划线代表在原始图片的基础上扩展的图像块,数量分别扩展了 5 倍、3 倍、3 倍和 4 倍,因此图像块数量总共扩展为原来的 180 倍。

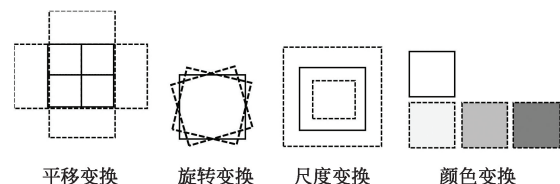


图 3 样本数据扩展

Fig.3 Data augmentation of samples

2.3 卷积神经网络结构

参照 AlexNet 网络结构,定义 5 层深度卷积神经网络,如图 4 所示:分别为 1 个输入层、3 个卷积(Conv)/池化层和 1 个全连接(Fully Connected,

FC)输出层。输入为归一化的、大小为 64×64 、RGB 三通道的图片;3 个卷积池化层的卷积核个数分别为 64、128 和 256,卷积核的大小分别为 11×11 、 7×7 和 5×5 ,每一层卷积核的通道数分别由其前一层输出的通道数决定,即分别为 3、64 和 128,池化均采用最大值池化(Max Pooling, MP),且窗口大小均为 2×2 ;全连接层神经元个数与输出类别数相同,在此 C 代表网络输出标签的个数,通常定义输出值为集合 $\{1, 2, \dots, C\}$ 中元素。因此,整个卷积神经网络可以看作一个内部带参数的黑盒,输入为相同大小的图片,输出为图片的类别标签,即

$$c = F_{\text{CNN}}(s | P), c \in \{1, 2, \dots, C\}, \quad (1)$$

式中: s 为输入图片, P 为神经网络模型的参数, c 为图像的输出类别标签。 $F_{\text{CNN}}()$ 表示深度卷积神经网络前向运算,即已知模型参数 P ,计算图片 s 的类别标签 c 。

2.4 模型训练

神经网络的学习过程即求解模型参数 P 的过程。为防止过拟合,采用以下步骤训练:首先,按照 2.3 节的网络结构,定义 $C=100$,即 100 类

输出类别,采用 Cifar-100 图像数据库进行初步训练。算法收敛后,将模型最后一层改为 $C=4$,分别代表背景、塔材、玻璃绝缘子和复合绝缘子 4 类目标,其余各层参数不变。然后利用输电线路图像库进行训练,直至收敛。Cifar-100 图像数据库数据类型丰富,可以避免算法陷入局部最优解。网络的前 4 层特征提取,最后的第 5 层用作分类器。利用 Cifar-100 图像数据库训练得到的特征提取参数(即网络的前 4 层参数),可转化为输电线路图像的特征提取参数。在此基础上进行网络参数调优,有效地避免了过拟合现象。

定义好网络结构后,采用最速梯度下降法训练神经网络模型参数,训练过程由开源深度卷积神经网络学习工具 Caffe^[17] 完成。预训练阶段的学习参数为:

- 1) 训练数据分块(Batch)大小为 128。
- 2) 所有权值都初始化为均值为 0,标准差为 0.01 的高斯噪声。
- 3) 系数动量(momentum)为 0.9,权值衰减系数为 0.000 5。
- 4) 对于第 2、3、4 卷积层和全连接层的神经元

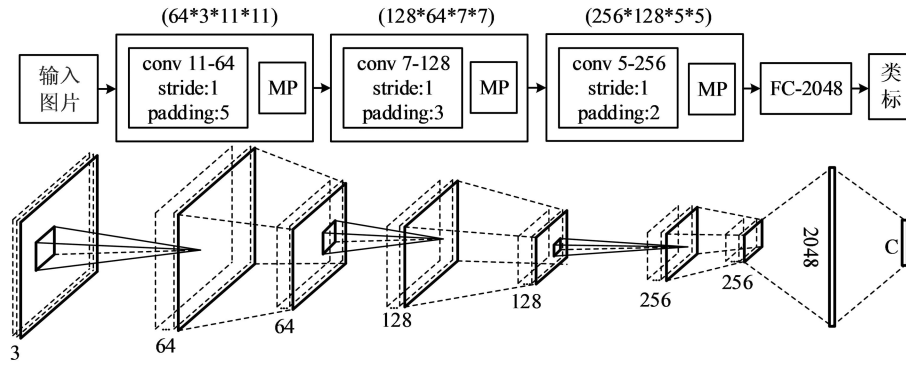


图4 深度卷积神经网络结构

Fig.4 Structure of deep convolutional neural network

偏置,初始化为 1,第 1 层的神经元偏置初始化为 0。

5) 每一层的学习速率保持一致,初始学习速率是 0.01。在训练过程中采用启发式方法:在当前学习速率下,如果验证集错误率不变,则将学习速率减小至当前的 1/10。

调优阶段,学习速率固定为 0.001,其它训练参数与预先训练阶段相同。最终算法收敛得到网络的模型参数 P 。

2.5 目标检测

进行目标检测时,首先将原始图像分为 $M \times N$ 个子块,其中 M 为纵向子块个数, N 为横向子块个数,定义 $s_{m,n}$ 为原始图像 I 的第 m 行、第 n 列个子块,则:

$$s_{m,n} = \begin{bmatrix} I_{\frac{m \times h}{M}+1, \frac{n \times w}{N}+1} \dots I_{\frac{m \times h}{M}+1, \frac{n \times w}{N}+w} \\ I_{\frac{m \times h}{M}+2, \frac{n \times w}{N}+1} \dots I_{\frac{m \times h}{M}+2, \frac{n \times w}{N}+w} \\ \dots \\ I_{\frac{m \times h}{M}+h, \frac{n \times w}{N}+1} \dots I_{\frac{m \times h}{M}+h, \frac{n \times w}{N}+w} \end{bmatrix}, \quad (2)$$

式中: h 和 w 分别为原始图像 I 的高度和宽度。这种子块划分方式存在的问题是, 目标有可能被划分到两个相邻子块中, 从而导致无正确识别, 尤其是像绝缘子这种细长的目标。为了解决这一问题, 将子块划分方式进行扩展, 生成 2 批子块: 第 1 批以原始图片左上角为起始点, 用 $s_{m,n}^1$ 表示; 第 2 批将起始点的横坐标和纵坐标进行平移, 平移量为分别为子块宽度和高度的一半, 用 $s_{m,n}^2$ 表示, 如图 5 所示。利用 2.4 节训练得到的模型参数 P , 分别计算每个子块的类别标签 $c_{m,n}^1$ 和 $c_{m,n}^2$ 。

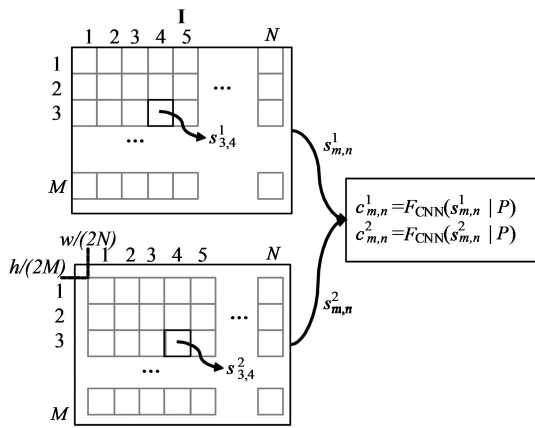


图 5 目标检测流程示意图

Fig.5 Schematic of object detection phase

3 实验结果与分析

3.1 实验方式及结果

在输电线路图像库中, 选取 15 条线路中的 420 张图片作为训练集, 剩余 4 条线路中的 100 张图片作为测试集。对图像库中所有图像的背景、塔材、玻璃绝缘子和复合绝缘子进行手动标注, 利用样本生成数据扩增方法生成样本, 4 类目标训练集与测试集样本数量如表 1 所示。

然后, 按照 2.3 节方法构造网络结构 ($C = 100$), 利用 Cifar-100 数据集进行预训练; 改变网络结构 (令 $C = 4$), 利用输电线路图像库生成的训练集样本进行调优。

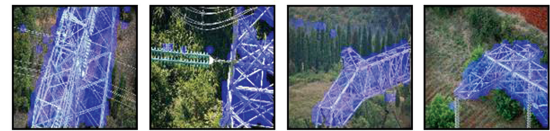
目标检测阶段, 选取测试集中的原始图像进行测试。令 $M = 24$ 、 $N = 18$, 生成 2 批子块, 总子块个数为 864。利用训练好的模型, 根据式(1)判

表 1 训练集与测试集样本数量

Tab.1 Quantities of training and testing samples

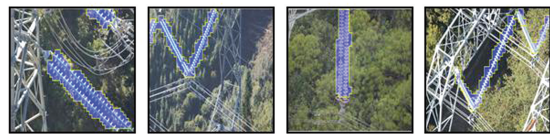
	训练集数量	测试集数量
背景	36 528	4 271
塔材	34 550	4 563
玻璃绝缘子	32 805	4 369
复合绝缘子	31 268	3 916
合计	135 151	17 119

断每个子块所属类别标签, 最终检测结果如图 6 所示。其中蓝色区域代表对应区域的子块被判定为属于该类目标, 可见除少数虚警外, 算法能够准确检测出每类目标的位置。



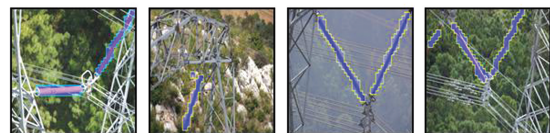
(a) 塔材检测结果

(a) Tower detection



(b) 玻璃绝缘子检测结果

(b) Glass insulator detection



(c) 复合绝缘子检测结果

(c) Composite insulator detections

图 6 输电线路目标检测结果

Fig.6 Detection results of transmission line objects

3.2 结果分析

分别对比本文方法与传统算法对各类目标的识别正确率。对塔材、玻璃绝缘子和复合绝缘子三类典型目标进行识别。每类目标包含一定数量的正样本和相同数量的负样本。测试样本选自 4 条输电线路中的 100 张图片, 对这 100 张图片进行标注、分块、扩充后形成了近 1.7 万个测试样本 (包含“背景”类型)。其中, 正样本在本类图像块中选取, 负样本在非本类图像块中随机选取。总

共对比4种算法的识别效果,如表2所示。其中,CNN+支持向量机(Support Vector Machine, SVM)算法是将预训练得到的模型最后一层替换为线性SVM^[18];SIFT(Scale-Invariant Feature Transform)+SVM算法首先提取图像SIFT特征^[19],再利用线性SVM做分类器;HOG(Histogram of Oriented Gradient)+SVM算法同理,首先提取图像HOG特征^[20],然后利用线性SVM做分类器。

对于每种算法,统计在三类目标总计真阳率(将三类目标合在一起统计)为90%左右时,各算法的总计假阳率及对每类目标的真阳率和假阳率。这样做的原因是,在调节判断阈值时,随着真阳率的增大,假阳率也会增大。因此,可以在真阳率固定的情况下(均为90%左右),统计对比每种算法的假阳率,假阳率越低则算法越优。另外,将三类目标总计的真阳率固定为90%左右时,每类目标的真阳率也各不相同,从而也可以对比不同

算法对每类目标的识别效果。

由表2的统计结果可以看出,在三类目标总计真阳率一定的情况下,本文方法对每一类目标都能保持较低的假阳率,优于其它方法。三类目标中,塔材的假阳率高于其它两类,这是由于导线或金具(如防振锤)与塔材特征相似,很容易被误识别为塔材,产生更多假阳性样本。而玻璃绝缘子在三种类别中特征最明显,其假阳率相对较低,与实际经验相符。

为了对深度卷积神经网络算法的敏感性和特异性进行分析,在此又比较了本文方法与其它经典方法的ROC(Receiver Operating Characteristic)曲线。如图7所示,结果与表2的统计结果一致。当检测真阳率为90%时,基于SIFT+SVM算法和基于HOG+SVM算法的假阳率高达50%左右,基于CNN+SVM的算法也在20%左右,而深度卷积神经网络方法假阳率低于10%,明显优于其它方法。

表2 各算法识别率数据表

Tab.2 Recognition accuracy of each algorithm

		总计	塔材	玻璃绝缘子	复合绝缘子
测试样本	正样本个数	12 848	4 563	4 369	3 916
	负样本个数	12 848	4 563	4 369	3 916
HOG+SVM	真阳样本个数	11 637	3 995	4 070	3 572
	假阳样本个数	5 951	2 419	1 770	1 762
	真阳率	90.6%	87.5%	93.1%	91.2%
	假阳率	46.3%	53.0%	40.5%	45.0%
SIFT+SVM	真阳样本个数	11 745	4 162	4 046	3 537
	假阳样本个数	7 097	2 735	2 234	2 128
	真阳率	91.4%	91.2%	92.6%	90.3%
	假阳率	55.2%	59.9%	51.1%	54.3%
CNN+SVM	真阳样本个数	11 704	4 157	4 052	3 495
	假阳样本个数	2 832	1 580	564	688
	真阳率	91.1%	88.4%	92.7%	89.2%
	假阳率	22.0%	34.6%	12.9%	17.6%
本文方法	真阳样本个数	11 673	4 068	4 076	3 529
	假阳样本个数	1 072	581	245	246
	真阳率	90.9%	89.1%	93.3%	90.1%
	假阳率	8.3%	12.7%	5.6%	6.3%

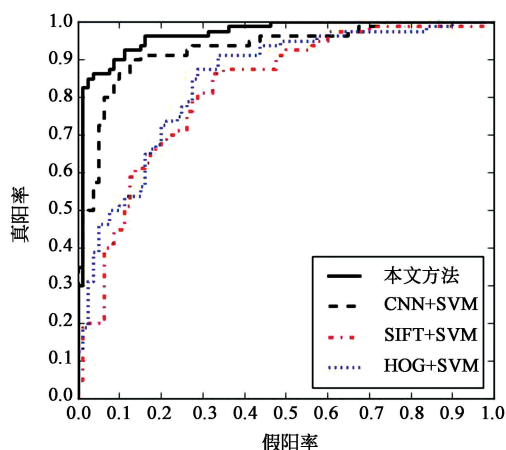


图 7 各算法 ROC 曲线

Fig.7 ROC curve of different algorithms

4 结 论

本文将深度卷积神经网络算法应用到输电线

路可见光图像目标检测,可精确检测出复杂背景图像中塔材、玻璃绝缘子和复合绝缘子。训练样本为载人直升机拍摄的 19 条输电线路、近 600 张图片。利用样本构造和数据扩展方法,生成 15 万个输电线路图像库训练样本。定义了 5 层深度卷积神经网络,先利用 Cifar-100 数据集对网络进行预训练,然后调整网络结构,利用输电线路图像库对网络参数调优。训练过程均用深度卷积神经网络训练工具 Caffe 完成。在保证检测真阳率为 90% 的条件下,本文采用的深度卷积神经网络方法假阳率低于 10%,明显低于其它方法 20% 以上,可用于输电线路可见光图像中的塔材、玻璃绝缘子和复合绝缘子检测,检测结果可进行辅助诊断或进一步的目标状态分析。本文方法可扩展到输电线路可见光图像中其它目标类型的检测。

参 考 文 献:

- [1] ZHAO Z B, LIU N, YUAN Y P. The recognition and localization of insulators based on SIFT and RANSAC [C]// *Proceedings of the 3rd International Conference on Multimedia Technology (ICMT-13)*. Atlantis Press, 2013: 699-706.
- [2] 李卫国,叶高生,黄锋,等.基于改进 MPEG-7 纹理特征的绝缘子图像识别[J].高压电器,2010,46(10):65-68.
LI W G, YE G S, HUANG F, et al. Recognition of insulator based on developed MPEG-7 texture feature [J]. *High Voltage Apparatus*, 2010, 46(10): 65-68. (in Chinese)
- [3] 韩军,张晶晶,王滨海.基于感知组织的输电线路结构识别方法[J].红外与激光工程,2013,42(12):3458-3463.
HANJ, ZHANG J J, WANG B H. Method on recognizing the structure of transmission line based on perceptual organization [J]. *Infrared and Laser Engineering*, 2013, 42(12): 3458-3463. (in Chinese)
- [4] 赵俊梅,张利平.绝缘子图像的多种特征提取技术的研究[J].电测与仪表,2013,50(12):37-41.
ZHAO J M, ZHANG L P. Research on various feature extraction technology of insulator image [J]. *Electrical Measurement & Instrumentation*, 2013, 50(12): 37-41. (in Chinese)
- [5] SZEGEDY C, TOSHEV A, ERHAN D. Deep Neural Networks for object detection [C]// *Proceedings of the 27th Annual Conference on Neural Information Processing Systems*. Lake Tahoe, Nevada: Curran Associates Inc., 2013: 2553-2561.
- [6] KRIZHEVSKY A. Learning multiple layers of features from tiny images [J]. 2009.
- [7] SUN Y, WANG X G, TANG X O. Deep learning face representation from predicting 10,000 classes [C]// *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*. Columbus, OH, USA: IEEE, 2014: 1891-1898.
- [8] WANG T, WU D J, COATES A, et al. End-to-end text recognition with convolutional neural networks [C]// *Proceedings of the 21st International Conference on Pattern Recognition*. Tsukuba, Japan: IEEE, 2012: 3304-3308.
- [9] 黄文博,王珂,燕杨.彩色视网膜眼底图像血管自动检测方法[J].光学精密工程,2017,25(5):1378-1386.
HUANG W B, WANG K, YAN Y. Automatic detection method of blood vessel for color retina fundus images [J]. *Optics and Precision Engineering*, 2017, 25(5): 1378-1386. (in Chinese)

- [10] 刘智,黄江涛,冯欣.构建多尺度深度卷积神经网络行为识别模型[J].光学精密工程,2017,25(3):799-805.
LIU Z, HUANG J T, FENG X. Action recognition model construction based on multi-scale deep convolution neural network [J]. *Optics and Precision Engineering*, 2017, 25(3): 799-805. (in Chinese)
- [11] LECUN Y, HUANG F J, BOTTOU L. Learning methods for generic object recognition with invariance to pose and lighting [C]//*Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. IEEE, 2004, 2: II-97-II-104.
- [12] LI F F, FERGUS R, PERONA P. Learning generative visual models from few training examples: An incremental Bayesian approach tested on 101 object categories [J]. *Computer vision and Image understanding*, 2007, 106(1): 59-70.
- [13] DENG J, DONG W, SOCHER R, *et al.* ImageNet: A large-scale hierarchical image database [C]//*Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*. Miami, FL, USA: IEEE, 2009: 248-255.
- [14] RUSSELL B C, TORRALBA A, MURPHY K P, *et al.* LabelMe: A database and web-based tool for image annotation [J]. *International Journal of Computer Vision*, 2008, 77(1/3): 157-173.
- [15] LECUN Y, BOSER B, DENKER J S, *et al.* Backpropagation applied to handwritten zip code recognition [J]. *Neural Computation*, 1989, 1(4): 541-551.
- [16] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks [C]//*Proceedings of the 25th International Conference on Neural Information Processing Systems*. Lake Tahoe, Nevada: ACM, 2012: 1097-1105.
- [17] JIA Y, SHELHAMER E, DONAHUE J, *et al.* Caffe: Convolutional architecture for fast feature embedding [C]//*Proceedings of the 22nd ACM International Conference on Multimedia*. Orlando, Florida: ACM, 2014: 675-678.
- [18] FAN R E, CHANG K W, HSIEH C J, *et al.* LIBLINEAR: A library for large linear classification [J]. *Journal of Machine Learning Research*, 2008, 9(8): 1871-1874.
- [19] LOWE D G. Distinctive image features from scale-invariant keypoints [J]. *International Journal of Computer Vision*, 2004, 60(2): 91-110.
- [20] DALAL N, TRIGGS B. Histograms of oriented gradients for human detection [C]//*Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. San Diego, CA, USA: IEEE, 2005: 886-893.

作者简介:周筑博(1990—),男,山东潍坊人,硕士,工程师,主要从事机器学习及电网应用方面的工作。E-mail:zhouzhu-bo90@163.com