

基于分阶段深度神经网络的施工违章识别

刘思雨, 薛劲松, 景栋盛

(国网江苏省电力有限公司苏州供电分公司, 江苏 苏州 215004)

✉lsy2801@163.com; 6802569@qq.com; jds19810119@163.com



摘要: 目前很多施工场地仍然使用人工方式检测施工人员是否佩戴安全帽。针对此, 设计并实现了一个基于分阶段深度神经网络的施工违章识别系统, 用以检测施工人员是否佩戴安全帽。系统利用深度神经网络, 通过在视频中采样获得图片, 然后将其分割成若干子区域, 接着利用预处理后的数据训练模型, 不断优化提升识别精度, 然后将训练好的模型应用到系统中。在室内、室外和红外线三个场景中进行测试。实验结果表明, 系统具有良好的实时检测能力, 总体平均正确检出率达86.79%。

关键词: 违章识别; 深度学习; 神经网络; 物体识别; 视频监控

中图分类号: TP18 **文献标识码:** A

Construction Safety Violation Recognition based on Staged Deep Neural Network

LIU Siyu, XUE Jingsong, JING Dongsheng

(Suzhou Power Supply Branch, State Grid Jiangsu Electric Power Limited Company, Suzhou 215004, China)

✉lsy2801@163.com; 6802569@qq.com; jds19810119@163.com

Abstract: In many building sites, people still manually monitor whether the construction crew wear safety helmets or not. To solve this problem, this paper designs a safety-violation identification system based on phased deep neural network in order to locate the construction crews who appear without safety helmets. By using deep neural network, this proposed system samples pictures from videos, divides them into sub-sections, and then uses processed data to train the model. When the identification accuracy is high enough, the model is then applied to practical use. The model is tested respectively in indoor scene, outdoor scene and infrared scene. The experiment results show that this system can achieve good real-time monitoring, with an average correct detection rate of 86.79%.

Keywords: violation recognition; deep learning; neural network; object recognition; video surveillance

1 引言(Introduction)

计算机在众多领域得到了迅速的发展, 其中包括计算机视觉技术、图像技术等。不少工作研究如何将计算智能应用于各种领域^[1,2]。通过视频监控设备实时监控检测施工场地中人员是否有违章操作就是很重要的一类应用场景。然而, 目前大部分还是通过人工方式来查看施工人员是否违章, 这种方法比较耗时耗力, 而且由于人的工作时间也有限, 无法实现全天候实时监控。随着工作时间的增加, 效率也会降低。因此, 人工方式来检测违章的情况存在一定的局限性。

针对此, 本文设计通过深度学习来自动地检测是否存在施工违章情况, 生成并保存相应的违章。这种方法和人

工方法相比, 在识别准确率方面可能会下降, 但是这种方式更加高效, 并且大大节省了人工的成本。

2 相关研究(Related studies)

2.1 图像处理技术

通过计算机来对图像进行一些操作包括对图像进行去噪, 提取特征等类似方法的技术就是图像处理技术。计算机软硬件的发展、离散数学的发展, 以及各种需求推动了图像技术的发展。数学的对图像处理技术的贡献主要是在图像处理的过程中需要将图像信号转换成数字信号, 转化成数字信号后, 计算机就可以处理这种信号。图像处理技术在农林、气象和水利等众多领域得到广泛的应用。

图像处理有几个目的。第一个目的是提高图像的视

感质量,主要是让图像的质量变得更好;第二个目的是提取图像中特征,在计算机分析图像时这些特征至关重要;第三个目的是对图像中数据进行处理,处理的方法主要是变换、编码和压缩,从而让图像的存储和传输更为快捷。图像处理技术有很多,而图像分类是其中的一种,其主要内容是将图像预处理,然后提取特征,根据特征来进行分类。图像分类的方法有很多种,而人工神经网络模式分类在图像技术中的表现和以前的方法相比,更为准确,因此这种方法就越来越受到重视。

2.2 人脸识别技术

人脸识别技术是近几年兴起的一种识别技术。这种识别技术是基于输入的图片的特征来识别,对于输入的一张图像,首先要做的就是判断输入的图像中是否存在人脸。如果存在的话,则需要进一步给出脸的具体信息。这些信息中提取这张脸所对应的身份特征。然后将这个身份特征与已知的情况,即可知道这张脸的身份情况。广义的说法是构建一个人脸识别系统,这个系统比较广泛,包括人脸的定位,人脸的预处理等等;而只通过人脸来进行身份上的确认和广义的人脸识别相比就是狭义上的人脸识别。

人脸识别技术主要采用的是特征分析算法,这个算法用到了两个方面的技术,一个是计算机图像技术,需要用图像技术来提取图像中的特征点,第二个是生物统计学原理,利用这个来分析特征点来建立对应数学模型,这个模型就是人脸特征模板。这个模板需要与被测的人的图像来进行特征分析,特征分析会得到一个相似值。判断是否为同一个人就是根据这个值推理出来的。

2.3 分类问题

机器学习是一门交叉学科,主要研究怎样让机器更加智能化,是目前人工智能领域的一个重要方向,从诞生到现在经历了若干个阶段。机器学习的一些方法的应用逐渐从一些理论转化为落地的产品,应用范围更为逐渐变广。

分类问题是机器学习中的基础问题。目前很多落地的产品当中好很多就是分类问题的应用。现阶段机器的计算能力很强,使得机器学习可以对海量的数据进行处理,由于数据的规模比较大,因此数据内部存在着有用的信息,因此机器学习可以通过这些数据可以挖掘出很多有用的信息。监督学习和无监督学习是机器学习中的两种数据处理方式^[3]。机器学习中的分类算法有很多,其中具有代表性方法包括决策树^[4]、支持向量机^[5]、关联规则^[6]和AdaBoost算法^[7]等。这些算法的步骤大致一致:首先,建立一个训练模型,根据模型所需要的训练数据和标签,然后构建相应的模型;然后,验证数据的输入,模型建立好之后有对应的输入和输出,将装备好的数据放入模型之中,模型会输出结果;最后,将结果与数据对应的正确标签进行比对,从而评价算法的效果。数据的验证是算法准确率的体现,评价一个算法好坏最直接的标准就是准确率,因此需要测试数据来进行验证数据。随着精确率的不断提高,模型就变得越来越好,模型就可以运用到现实之中。自深度学习提出以后,深度学习也可以运用到分类问题中,并取得了很好的效果。

2.4 深度学习

深度学习与机器学习领域中的神经网络有着特别的关联。深度学习也是学习海量数据中的潜在的信息,学

到的信息对于对应的数据的解释有着很大的帮助。深度学习的目的是使计算机更智能,可以识别各种问题,看懂各种图像,听懂各种语音。深度学习近两年在各种方面的取得较好的成就,其中在文本、图像、语音和生成式对抗网(Generative Adversarial Network, GAN)等方面取得了很好的效果。

深度学习与神经网络有着很强的关联主要是因为深度学习来源于人工神经网络的研究。深度学习可以通过组合底层特征来发现数据的特征表示,形成更高级的特征。神经网络是模拟人脑而建立的。深度学习就是通过这种方式来分析数据。深度学习中有许多网络结构,其中包括卷积神经网络(Convolutional Neural Networks, CNN)、循环神经网络(Recurrent Neural Network, RNN)和长短期记忆网(Long Short-Term Memory, LSTM)等。卷积神经网络是比较常见的一种网络,第一个卷积神经网络^[8]计算模型是在Fukushima D的神经认知机中提出的,基于神经元之间的局部连接得到一种平移不变神经网络结构形式。之后,有研究人员在此基础上,用误差梯度设计并训练卷积神经网络,在一些任务上得到优越的性能^[9]。

3 方法设计(Method design)

3.1 训练模型

训练模型采用的卷积神经网络所训练的模型。输入层、隐藏层和全连接层构成了卷积神经网络^[10]。神经网络的输出是输入层,本文中所使用的神经网络结构中的卷积神经网络,该网络的输入是一个像素矩阵,这个像素矩阵是图片所对应的。卷积神经网络示意图,如图1所示。

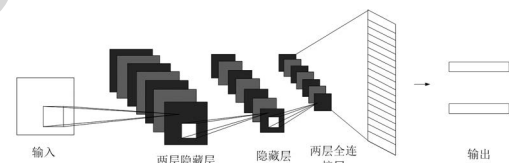


图1 卷积神经网络示意图

Fig.1 Schematic diagram of convolutional neural network

在图1中,最左侧的输入是一个三维矩阵,矩阵长宽代表图像大小,深度代表色彩通道。卷积神经网络从输入层通过不同的神经网络结构将上层的三维矩阵转化为下一层的三维矩阵,生成的三维矩阵作为下一层的输入,直到最后一个全连接层。图中的隐藏层中包括卷积层和池化层。而卷积网络结构中最重要的一部分就是卷积层。卷积层的输入是上一层神经网络的输出,卷积层再次进行卷积操作从而获得更抽象的特征。池化层的功能是缩小矩阵的大小,将高分辨率的图像转化为低分辨率的图像。通过池化操作,最后一个全连接层的节点数就会减少。在经历多隐藏层的操作后,最后通过全连接层来输出结果。隐藏层的操作是图像抽象提出的过程,在完成了特征提取后,通过最后的全连接层就完成分类。

在此基础上,本文把一些符合安全条件的施工图片和不符合安全的施工图片放入此神经网络中进行训练,最后准确度高于90%的情况下,将此时的参数进行保存,得到了训练后的模型。只要将施工中得到的图片放入此模型中便可以得到施工中的人是否符合安全施工的条件。

3.2 视频处理

由于摄像头获取到数据是以视频的方式呈现的,因此

系统的输入是视频。然而视频数据量很大，因此需要将视频处理一下，以便做之后的处理。将视频处理成需要的图片是必不可少的。

目前的图像处理技术很容易做到将视频分割成图片^[11,12]。视频中有帧率属性，本文就是在帧率的基础之上对视频进行，按照15帧获取一次图片，在所使用的视频中，每15帧为视频中的1秒，如果间隔帧率比15小的话，获取到的图片就会更多，反正，获得的图片就会更少，本文采用1秒钟截图一次，尽量保存必要的信息。采用OpenCV技术来讲视频处理为若干图片，第一步就是获取视频，这一步相当于输入，获取到对应的视频后，然后获取视频的总帧数，然后设置一个标记，并且标记此标记为真，在标记为真的情况下，持续读取视频，按照每15帧的情况截图视频中当前的画面，并保存在对应的文件下，直至视频结束，在视频结束后，让标记变为假。获取视频总帧数的情况是保证获取到的图片数量不会太多，由于系统使用视频帧率是已知的，因此规定每15帧取一张图，在未知的情况下，可以通过总帧数和规划获取的图片总量来决定多少帧获取一张图片。

3.3 人脸识别

在图片中，有一部分图片中是包含人脸的，有一部分是不包含人脸的，因此需要将包含人脸的图片挑选出来。本系统主要是来识别施工人员中是否穿戴安全帽的情况，因此需要将人脸和戴安全帽一起进行检测。因此，本文中做工作首先就是先将每个图中的人脸部分获取到，然后将获取到的人脸进行检测，检测人脸是否穿戴安全帽，穿戴安全帽的标记为真，反之，标记为假。

本文中采用的人脸识别技术主要是采用深度学习的方法，根据图像中的内容，来检测图像中是否存在人脸，通过已经训练好的检测人脸的模型就可以做到，这种模型和其他的神经网络模型一致，其中最重要的部分是数据。本文采用的模型是谷歌训练好的模型。只要将一张图片放入模型的输入之中，就可以得到该图所对应的人脸，根据输入的图片的名称来以此命名对应输出的图片的名称，这是因为一张图中对应的人脸可能不止一个，因此需要这个操作。在得到这些人脸之后，我们可以把这些图片放入检测是否佩戴安全帽的模型之中，并根绝对应的情况来输出未佩戴安全帽的图片对应的原始图片。

3.4 分阶段处理

由于系统比较复杂，因此需要采取分阶段处理的方法去将复杂的问题转化为若干个简单的问题并分别处理，直至系统的完成。

第一个阶段就是从视频处理为图片的阶段，需要将视频转化为图片，根据视频中的帧率情况将视频转化为若干图片，用以保证产生的图片不会太多也不会太少，更加保证产生的图片中包含脸的图片不会被遗漏。第二个阶段主要是针对产生的图片的处理，由于视频清晰度的原因，产生的图片可能会由于视频的原因而大小不一。对图片处理的速度受到了图片清晰度的影响，图片越清晰，对图片的处理的速度就越快。基于此，需要先对图片进行处理，无关图片的清晰度，都需要对其进行处理，从而保证后续的处理更加快捷。第三个阶段就是将处理后的图片进行人脸的识别，由于有图片的预处理过程，人脸识别的部分的效

果会更好，减小了识别过程中的误差。第四个阶段就是识别当前的人脸的图片中是否佩戴了安全帽，如果没有带的话，此图片就保留下来用以输出。经过这种分阶段的处理过，保证了实施性更高，速度更快，效果更好，准确率更高。

3.5 算法描述

根据方法设计中所提到的各种技术，在提前训练好两个模型的基础之上提出一个算法用以开发这个系统，其中第一个模型用以识别该张图片中是否存在人脸，而第二个模型用以识别该张图片中是否存在安全帽，图片符合第一个模型识别下在第二个模型中得到了一个否的情况就是我们最后得到的情况，即该张图片中存在人脸，但是并不存在安全帽。

算法1基于分阶段深度神经网络的安全违章识别

输入：实时视频

输出：带有安全帽的人员图片

 读取输入的视频

 根据帧率情况将视频分解为合适的图片集

 For $i \leftarrow 1$ to 图片集末尾 do:

 调整当前图片的大小

 将图片进行灰度化处理

 If 当前图片包含人脸:

 圈出人脸并生成人脸图片并放入一个新的图片集

 标记生成人脸的原始图片

 End For

 根据人脸图片集的标记情况去除原始图片集中不包含人脸的图片

 For $j \leftarrow 1$ to 人脸图片集末尾 do:

 If 当前图片包含安全帽:

 根据当前图片的标记，去除原始图片集中对应的图片

 End For

 返回处理后的原始图片集

4 系统效果展示(System effect display)

本文系统主要是通过对视频进行处理，得到最后的未戴安全帽的人员的图片并进行展示。系统监控的展示效果如图2所示。其中，图2(a)是室内场景效果图，图2(b)是室外场景效果图，图2(c)是红外线场景效果图。在本文中，对效果图的涉及敏感的隐私信息部分进行了处理。



(a)室内场景效果图 (b)室外场景效果图 (c)红外线场景效果图

图2 系统监控展示效果

Fig.2 Demonstration of system monitoring

本文使用三种场景的视频文件中所获取的图像作为训练数据。在完成了训练之后，本文分别在室内、室外和红外线三个场景中对模型进行测试，统计分析正确率和误报率，其中室内场景下得到图像13张，检测出未戴安全帽的图像13张，正确率100%，误报率0%，室外场景下得到图像10张，检测出未戴安全帽的图像7张，正确率70%，误报率30%，红外场景下得到图像30张，检测到未戴安全帽的图像26张，正确率86.67%，误报率13.33%。通过分析可以发

现，由于室外场景相对较为复杂，因此正确率较低。在三种场景中，总体正确率为86.79%，误报率达13.21%，可以满足识别需求。三种场景检测结果统计如表1所示。

表1 三种场景检测结果

Tab.1 Testing results of three scenes

场景	准确率(%)	误报率(%)
室内	100	0
室外	70.00	30.00
红外线	86.67	13.33
总体	86.79	13.21

5 结论(Conclusion)

本文提出了一个基于分阶段深度神经网的施工违章识别系统，该系统可以面向视频实现实时监控，识别出未按要求佩戴安全帽的人员。本系统提供了模型学习的功能，能利用监控视频数据自动学习获得检测模型。用户只要输入所需要的视频文件，就可以得到检测模型。学习完成后，系统即可在此基础上进行检测。在本文实验中，学习了两个模型，并将其运用到系统中，实现了从视频当中检测出未佩戴安全帽的情况，获得了较高的正确识别率。

参考文献(References)

- [1] 孙卓晟,彭来湖,倪利明,等.基于虚拟现实技术的机房动力环境监控系统设计[J].软件工程,2020,23(3):28-31.
- [2] 王昌海,申红雪,张王卫,等.一种基于人脸识别的课堂教学监控系统[J].软件工程,2020,23(1):48-50.
- [3] 吕永标,赵建伟,曹飞龙.基于复合卷积神经网络的图像去噪算法[J].模式识别与人工智能,2017,30(2):97-105.
- [4] 汤鲲,蒋炳南,彭艳兵,等.基于决策树的多维属性自动推理识

(上接第38页)

- automated guided vehicle based on computer vision[J]. IEEE Robot Mag, 2007, 14(3):24-34.
- [2] Wu X., Zhang Y., Zou T., et al. Coordinated path tracking of two vision-guided tractors for heavy-duty robotic vehicles[J]. Robot Comput Integr Manuf, 2018, 53(100):93-107.
- [3] 陈敏,黎展滔,陈庆新,等.考虑有限物流运输能力的智能车间AGV调度算法研究[J].工业工程,2019,22(4):49-57.
- [4] 常建斌,王璐,莫易敏,等.基于Manhattan距离的汽车总装车间带时间窗多AGV小车调度优化[J].武汉理工大学学报(交通科学与工程版),2017,41(4):489-594.
- [5] 于滨,靳鹏欢,杨忠振.两阶段启发式算法求解带时间窗的多中心车辆路径问题[J].系统工程理论与实践,2012,32(8):1793-1800.
- [6] 梁承姬,沈珊珊,胡文辉.基于路段时间窗考虑备选路径的AGV路径规划[J].工程设计学报,2018,25(2):200-208.
- [7] 徐镇华,马殷元.基于时间窗的改进两阶段AGV路径规划研究[J].测控技术,2018,37(06):145-149.
- [8] Duinkerken MB, Lodwijks C. Routing of AGVs on automated container terminals[C]. IEEE 19th International Conference on Computer-Supported Cooperative Work in Design, Calabria, 2015.

别[J].计算机与现代化,2017(2):83-87.

- [5] 游清顺,王建新,张秀宇,等.基于支持向量机的食品安全抽检数据分析方法[J].软件工程,2019,22(2):29-31.
- [6] 李昌盛,伍之昂,张璐,等.关联规则推荐的高效分布式计算框架[J].计算机学报,2019(6):1218-1231.
- [7] 黄健.基于深度学习与二维离散小波分解特征相融合的adaboost人脸识别模型[J].软件工程,2020,23(2):43-46.
- [8] 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.
- [9] 万晓琪,宋辉,罗林根,等.卷积神经网络在局部放电图像模式识别中的应用[J].电网技术,2019(6):2219-2226.
- [10] Jin K H, McCann M T, Froustey E, et al. Deep convolutional neural network for inverse problems in imaging[J]. IEEE Transactions on Image Processing, 2017, 26(9): 4509-4522.
- [11] 周力,闵海.基于局部连接度和差异度算子的水平集纹理图像分割[J].中国图像图形学报,2018,24(1):39-49.
- [12] Ding, Changxing, Dacheng Tao. Robust face recognition via multimodal deep face representation[J]. IEEE Transactions on Multimedia, 2015, 17(11): 2049-2058.

作者简介:

- 刘思雨(1994-),女,硕士,助理工程师.研究领域:智能信息系统,计算机应用.
- 薛劲松(1977-),男,本科,高级工程师.研究领域:网络安全,智能化信息系统.
- 景栋盛(1981-),男,硕士,高级工程师.研究领域:软件智能化,信息安全.

- [9] Song L., Huang S. A hybrid metaheuristic method for dispatching antomated guided vehicles in container terminals[C]. IEEE Symposium on Computational Intelligence in Scheduling, Singapore, 2013.
- [10] 霍凯歌,张亚琦,胡志华.自动化集装箱码头多载AGV调度问题研究[J].大连理工大学学报,2016,56(3):244-251.
- [11] 彭成吉.基于时间窗算法的多AGV路径冲突的研究[A].中国烟草学会学术年会优秀论文集[C].2017:5.
- [12] 梁承姬,沈珊珊,胡文辉.基于路段时间窗考虑备选路径的AGV路径规划[J].工程设计学报,2018(25):200-208.
- [13] 许伦辉,黄宝山,钟海兴.AGV系统路径规划时间窗模型及算法[J].广西师范大学学报(自然科学版),2019(37):1-8.

作者简介:

- 邱亭秀(1999-),女,本科生.研究领域:信息管理与信息系统,软件工程,仿真处理.
- 倪欣园(1999-),女,本科生.研究领域:信息管理与信息系统.
- 于露(1999-),女,本科生.研究领域:软件工程,通信工程.
- 窦万峰(1968-),男,博士,教授.研究领域:软件工程,分布式与并行计算,大数据分析挖掘.