

基于 AlexNet 的迁移学习在流程工业图像识别中的应用 *

Transfer Learning of AlexNet on Image Recognition Application in Process Industry

周 曼 刘志勇 陈梦迟 赵浴阳 杨鲁江 (浙江中控技术股份有限公司,浙江 杭州 310053)

摘要:流程工业生产中存在着大量视频智能监控的需求,通过对生产相关的画面进行智能识别,提高工厂自动化水平,并降低人为出错的概率。传统图像识别的方法在实际应用中存在维护复杂、环境适应性较差等问题。近年来深度学习技术在图像分类识别上的表现突出,越来越多地被应用到各个领域。基于深度卷积神经网络 AlexNet 对流程工业视频图像进行迁移学习训练,取得了良好的识别效果,并成功应用在工业现场环境中。

关键词:图像分类,卷积神经网络,AlexNet,迁移学习,微调

Abstract:There are a large demand for intelligent monitoring by vision in process industry. Through intelligent recognition of production-related images, the industrial automation level is improved and the probability of human factor error is reduced. Traditional image recognition method has the problems of complex maintenance and poor environmental adaptability in practical applications. In recent years, deep learning has achieved outstanding performance in image classification tasks and has been widely applied to various fields. Based on a kind of deep convolutional neural network-AlexNet, this paper uses transfer learning technique to train images which comes from a process industry factory. The result shows that the model performs well on the application.

Keywords:image classification, CNN, AlexNet, transfer learning, fine-tune

随着工业生产对检测的要求越来越高,投入的成本也逐步加大,传统接触式检测已无法满足客户要求及质量保证。这就促使机器视觉技术井喷式发展,工业检测项目也越来越多地开始采用非接触式光、机、电、视觉、化工等综合性技术手段。

在工业现场中,虽然通过网络摄像头进行远程监控能够解决人员现场作业问题,但人们工作状态的客观性和持久性存在一定的缺陷,工业视频图像智能识别技术就能够很好地代替人做一些识别工作。同时在大批量工业生产过程中,用人工视觉检查产品质量效率低且精度不高,而机器视觉检测系统的检测精度高,可以快速获取大量信息且自动处理,大大提高了生产效率和生产自动化程度。因此,机器视觉检测系统取代人工检测将是现代化进程的必然趋势。

目前流程工业现场中,存在大量需要通过视觉进行监控分类判断的应用场景,例如传送带物料承载状态监测、生产仓位位情况变化监测等,监控的结果往往作为下一步生产操作的依据。传统的机器视觉及图像处理技术对观测对象的环境要求、摄像头安装、画质等要求较高,近年来以深度学习为代表的图像识别技术取得了较大的进步,尤其是在图像分类识别应用研究中达到了甚至超过人眼的识别精度,并具有良好的适应性。因此在流程工业现场图像识别任务中,可以采用深度学习的方法进行有效的识别。

1 网络模型

1.1 卷积神经网络 CNN

卷积神经网络(Convolutional Neural Network, CNN)是一种深度学习的神经网络模型,通过前向传播计算输出值,反向传播调整连接权值。它最重要的特点是稀疏连接和权值共享,每层神经元和与它相邻层的神经元之间不再是全连接,而是部分连接,即某个神经元的感知区域仅受其上层的部分神经元的影响,

这样的网络结构降低了模型的复杂度。CNN 的基本架构通常包含卷积层、池化层和全连接层,对原始输入多次进行卷积(conv)操作和池化(pooling)操作得到越来越复杂的特征图。针对这个特征计算输入图片的卷积值,其中数值最大的区域与特征匹配,得到合适的特征,并将其输出到下一层结构中。

常见的基于深度神经网络的算法主要有:8 层的 AlexNet、19 层的 VGGNet、22 层的 GoogleNet,还有最深达 200 层的 ResNet 等等。理论上来说网络层数越深,对多类识别检测效果越优异,但是 AlexNet 模型在图像分类任务上取得了显著的效果,且模型隐含层相对来说较少,从而能够更快地完成训练,因此本文采用基于 AlexNet 的模型进行迁移学习。

1.2 AlexNet

AlexNet^[1-2]是 Hinton 和他的学生 Alex Krizhevsky 在参加 2012 年的 ImageNet 竞赛时所设计的神经网络,并赢得了该届图像识别大赛的冠军。AlexNet 有八层结构、五个卷积层、相应的池化层和三个全连接层,再加上最后的 Softmax 层所组成,能使分类的 Logistic 回归的目标值最大化,该方法能最大化预测在分布条件下的训练样本中的正确标签的对数概率平均值。

AlexNet 网络中采用了 ReLU 激活函数替代常规的 Sigmoid 函数,能更快地训练并解决由于较深网络导致的梯度弥散问题;采用 Dropout 随机失活部分神经元来防止模型过拟合(overfitting);采用重叠的最大池化层避免了平均池化层的模糊化的效果,并且步长比池化核的尺寸小,因此池化输出层之间有重叠和覆盖进而提升了特征的丰富性;采用 LRN (Local Response Normalization)局部响应归一化,对局部神经元的活动创建了竞争机制,使得其中响应比较大的值变得更大,并抑制反馈较小的神经元,从而增强了模型的泛化能力;采用具有强大并行能力的 GPU 来加速网络训练;采用从原始图片中随机截取指

* 2016 年工业转型升级专项“智能工厂解决方案”资助

定大小区域(以及水平翻转)来实现数据增强。

1.3 迁移学习

深度学习模型需要训练大量的标注数据,而对于人工重新搜集并进行标注训练的任务,需要耗费大量的人力物力。另一方面,完全丢弃已经获得的标注数据是十分可惜的。迁移学习(Transfer Learning)是用从一个环境中学到的知识来帮助新环境中的学习任务。一般的迁移学习是将训练好一个网络(我们称它为 base network)的前 n 层复制到目标网络(target network)的前 n 层,对目标网络剩下的其他层随机初始化,然后开始训练目标任务。其中,在做反向传播(backpropagate)的时候,有两种方法可以选择:①把迁移过来的这前 n 层冻结(frozen)起来,即在训练目标任务的时候,不改变这 n 层的权重值;②不冻结这前 n 层,而是不断调整它们的权重值,称为微调(fine-tune)。这主要取决于目标数据集的大小和前 n 层的参数个数,如果目标数据集很小,而参数个数很多,为了防止过拟合,通常采用冻结的方法;反之,采用微调。

Jason Yosinski^[3]发现深层神经网络在第一层提取的特征具有一般性并不是某个数据集或任务所特有的。在神经网络最高层的分类层,网络所学到的特征与数据集密切相关,对于有 N 个神经元的分类层,每个神经元的输出都会是一个特定的类别,因而具有特殊性。从一般性的特征到具有特殊性的特征,所以神经网络的某处一定存在着这种特征的过渡。正是由于深度学习的特征是逐层提取的,低层的特征具有一般性,所以这也就意味着可以在已有数据集上训练的模型用于新的数据集,通过对模型的微调(fine-tune)从而获得用于对新的数据集进行分类的网络模型。

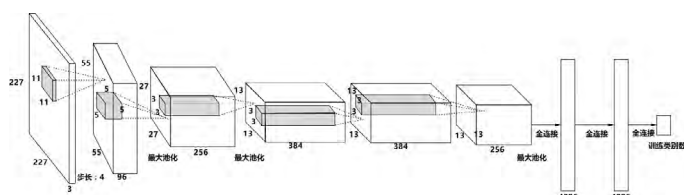


图1 AlexNet 模型示意图

本文通过迁移已经训练好的模型参数,对新的数据集进行训练。原始训练数据集采用 ImageNet 2012 分类数据集,该模型有 6 千万个参数数据集分 1 000 个类别,包含 120 万张训练图像,5 万张验证图像和 10 万张测试图像^[1]。将在 ImageNet 数据集上训练好的 AlexNet 模型参数迁移到新的数据集上进行训练,保留前五个卷积层和相应池化层以及三个全连接的参数,其它参数采用随机初始化,最后一层为用于分类的 Softmax 层,其神经元个数为被分类任务的数据集的图像类别。

2 工业视频图像处理与训练模型结果分析

2.1 工业视频图像处理

对工业传送带以及生产仓料位视频图像进行视频采集、图片截取、标签分类,然后采用基于 AlexNet 的迁移学习,修改最后一个全连接层为输出的类别数目,在训练生成的模型上进行验证测试。

本文中对摄像头采集的视频按照一定采样间隔进行图片截取,从而获得图片数据集,然后对图片数据集进行筛选,选取特征明确、图像清晰的数据进行人工分类标注,对于标注的每个类别中的数据采用无放回随机采样的方法选取 20% 的图片数据作为验证集。然后进行任务模型的训练和验证测试。对于传送带模型我们将其划分为 3 类,依次标注为无料、少料和正常料,例图如图 2 所示;对于生产仓料位模型我们将其划分为 4 类,依次

标注为无料、少料、正常料和满料,例图如图 3 所示。



图2 传送带无料、少料和正常料例图



图3 生产仓料位无料、少料、正常料和满料例图

2.2 训练模型及结果分析

2.2.1 传送带模型

对于传送带模型,选取了无料、少料和正常料 3 种类别的原始图片一共 144 119 张,然后对这 3 种类别的图片每种每次随机选取 100 张进行数据增广,包括随机缩放、上下翻转、左右翻转、旋转 90°、旋转 270°、沿着中心位置随机截取 60% 的区域图片、随机截取全图位置 60% 的区域图片、随机灰度处理、随机饱和度处理以及随机对比度处理等 10 种数据增广的方法来增广数据集,进而得到 3 000 张增广图片,对这一共 147 119 张图片进行模型训练,其中设置学习率为 0.0005,Dropout 值为 0.5。如图 4:传送带训练步数与准确率图中可以看出经过 20 步训练后模型的准确率达到 93.02%,到 40 步就升高到了 98.19%,170 步时达到最大值 99.61%,此后在 97.22%到 99.41%之间徘徊,说明该模型训练到 170 步就已经达到较好的准确度,再训练容易导致过拟合,此时应该停止训练了。

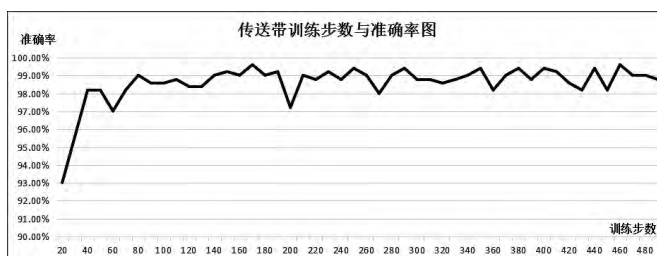


图4 传送带训练步数与准确率图

对于传送带模型采用无放回随机采样的方法选取 20% 左右的图片数据作为验证集,选取了共 36 000 张图片,对验证集进行回归验证,得到 99.68% 的准确率。

2.2.2 生产仓料位模型

对于生产仓料位模型,选取了无料、少料、正常料和满料 4 种类别 22 135 张原始图片,然后对这 4 种类别的图片每种每次随机选取 100 张进行数据增广,包括随机缩放、上下翻转、左右翻转、旋转 90°、旋转 270°、沿着中心位置随机截取 60% 的区域图片、随机截取全图位置 60% 的区域图片、随机灰度处理、随机饱和度处理以及随机对比度处理等 10 种数据增广的方法来

增广数据集,进而得到 4 000 张增广图片,对这 26 135 张图片进行模型训练,其中学习率设置为 0.0005,Dropout 值设置为 0.5。如图 5:生产仓料位训练步数与准确率图(学习率为 0.0005)中可以看出经过 20 步训练后模型的准确率达到 83.98%,到 40 步升高到了 90.58%,490 步时达到最大值 99.22%,说明学习率设置得太小导致训练到结束时准确率还在上升过程中,因此需要增加训练步数继续进行训练或者可以将学习率改大来加快训练。

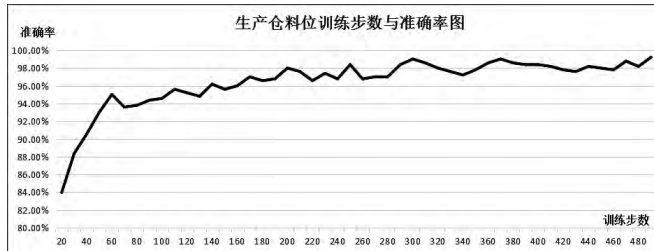


图5 生产仓料位训练步数与准确率图(学习率为 0.0005)

采用将学习率改为 0.001,其他不变,如图 6:生产仓料位训练步数与准确率图(学习率为 0.001)中可以看出经过 20 步训练后模型的准确率达到 88.62%,到 30 步达到 92.56%,470 步时达到最大值 99.80%,此后又下降为 480 步的 98.58%,到 490 步的时候下降为 98.39%,说明该模型训练到 470 步就达到了理想的准确率,此时可以停止训练了。

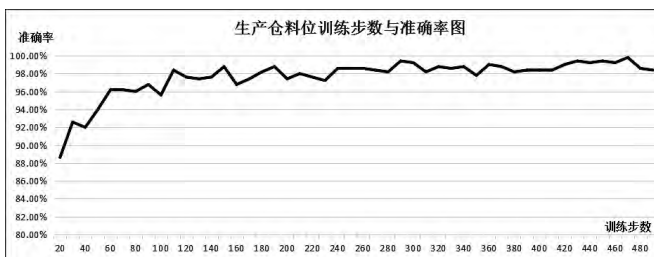


图6 生产仓料位训练步数与准确率图(学习率为 0.001)

对于生产仓料位模型采用无放回随机采样的方法选取 20%左右的图片数据作为验证集,我们选取了共 6 920 张图片,对验证集进行回归验证,得到 98.84%的准确率。

2.3 工业现场应用

将上述训练生成的模型应用于工业现场,将现场摄像头采集的视频图片输入到模型中进行识别检测,查看识别结果是否与实际相符合。在某水泥厂同时对传送带和生产仓料位进行识别,从现场识别结果图 7、图 8 可以看到模型能很好地滤除卡车进行倒料的干扰而正确识别对象,效果良好。

3 结束语

基于深层卷积神经网络的深度学习技术已经在图像分类上取得了较好的效果,本文基于 AlexNet 模型,对实际工业现场的



图7 现场识别结果(正常-满料)



图8 现场识别结果(少料-少料)

大量实拍视频图片进行迁移学习,并将参数微调后的模型用于处理实际工业现场的视频图像,取得了较好的分类结果。基于深度学习方法的图像分类识别对摄像头安装位置、角度、现场环境的光照、震动、人员走动干扰等影响具有较强的适应能力。

将深度学习方法与传统图像处理方法相结合,能够有效解决并满足各种工业现场视觉识别任务。随着深度学习技术的逐渐完善以及工程应用的不断推进,这一技术将越来越成为工业智能制造中不可或缺的一部分。

参考文献

- [1]Krizhevsky A, Sutskever I, Hinton G E. ImageNet lassification with deep convolutional neural networks [C]// International Conference on Neural Information Processing Systems. Curran Associates Inc, 2012:1097-1105
- [2]Hinton G E,Salakhutdinov R R.Reducing the dimensionality of data with neural networks[J].Science,2006,13(5786):504-507
- [3]J Yosinski,J Clune,Y Bengio,et al. How transferable are features in deep neural networks[J].Neural Information Processing Systems, 2014, 3320-3328
- [4]党宇,张继贤,邓喀中,等.基于深度学习 AlexNet 的遥感影像地表覆盖分类评价研究[J].地球信息科学学报,2017,19(11):1530-1537
- [5]自琮,黄玲,陈佳楠,等.面向大规模图像分类的深度卷积神经网络优化[J].软件学报,2018,29(4):1029-1038
- [6]苏思悦,付莹,来林静.结合单目标多窗口检测器和 AlexNet 的螺母检测[J].中国科技论文,2018,13(4):414-419

[收稿日期:2018.8.20]

(上接第 79 页)

实现稳定悬浮,同时在这种情况下,永磁悬浮轴承内部的辅助轴承理论上不再承受载荷,亦可达到长寿命的目标。

根据以上计算可知,对于永磁轴承内部的辅助轴承,其不仅起到保护轴承的作用,同时还起到辅助定位的作用,其所受到的载荷非常小,理论分析认为能够实现辅助轴承长寿命的目标。故在安装磁轴承时,应该尽量保持中心线的重叠,这样不仅可以实现轴向承载的最大化,同时产生的径向力为 0,偏转力矩为 0,对辅助轴承产生较小的附加载荷,有利于实现辅助轴承长寿命的目标。

参考文献

- [1]张钢.立式风力发电机用永磁轴承的设计及其力学性能研究[J].轴承,2015,5,1-6,12
- [2]张钢.磁悬浮轴承-转子系统的机电耦合动力学研究[D].西安:西安交通大学,1999
- [3]严道发.电主轴技术综述[J].机械研究与应用,2006,19(6):1-3
- [4]孟庆涛.永磁悬浮电机支承关键技术研究[D].上海:上海大学,2015
- [5]张钢,孟庆涛,钟永彦,等.五自由度全永磁轴承系统的稳定悬浮特性分析[J].机械工程学报,2015,51(5):56-63

[收稿日期:2018.7.1]