多层卷积神经网络深度学习算法可移植性分析

肖堃
(电子科技大学 计算机科学与工程学院,四川 成都 611731)

摘 要:在现实环境下,出现恶意用户或攻击者对机器学习算法的攻击;在应用过程中,机器学习算法也会受到物体形状、位移、尺度、光照、背景等因素的影响。针对这些使用过程中所产生的安全性问题,本文提出了基于多层卷积神经网络深度学习算法的图像识别方法,并对其可移植性进行分析,通过对抗性训练提高模型泛化能力来防御对抗样例攻击。针对可用性攻击,在前向传播过程中,采用训练好的多层卷积神经网络深度学习模型自动提取输入图像特征,并利用模型权值共享、更新、下采样等操作对输入图像做降采样处理,降低计算复杂度;在反向传播过程中,利用 delta 法则和 Fisher 准则,以及基于类内距离和类间距离的能量约束函数实时调整多层卷积神经网络深度学习模型参数,计算模型输出层各个输出单元的残差,使模型权值能够更加快速收敛到有利于图像识别的最优值。测试结果表明:多层卷积神经网络深度学习算法在图像识别领域的应用具有识别准确率和鲁棒性较高,耗时较短的优点,从理论和实验 2 方面证明了算法的可移植性。

关键词:多层卷积神经网络;深度学习算法;可移植性;分析;图像识别;拟合效果;delta 法则;Fisher 准则
DOI: 10. 11990/jheu.201810076
网络出版地址:http: //www.cnki.net/kcms/detail/23. 1390.u.20200507. 1018. 004.html
中图分类号:TP783 **文献标志码**:A **文章编号**:1006-7043(2020) 03-0420-05

Portability analysis of multilayer convolutional neural network´s deep learning algorithm

XIAO Kun
(School of Computer Science and Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China)

Abstract: In the real environment, malicious users or attackers attack the machine learning algorithm; in the application process, the machine learning algorithm will also be affected by the shape, displacement, scale, light, background and other factors. In view of the safety problems in the use process. In this paper, an image recognition method based on multi-level convolutional neural network deep learning algorithm is proposed and its portability is analyzed. Defense against sample attacks by improving the generalization of models through confrontational training. For the usability attack, in the forward propagation process, the trained multi-layer convolutional neural network deep learning model is used to automatically extract the input image features, and the input image is downsampled by using the model weight sharing, updating, and downsampling operations, to reduce the computational complexity; in the backpropagation process, use the delta rule and the Fisher criterion, and the energy constraint function based on the intra-class distance and the inter-class distance to adjust the parameters of the multi-level convolutional neural network deep learning model in real time, and calculate the model output layer. The residuals of the individual output units enable the model weights to converge more quickly to the optimal value for image recognition. The test results show that the application of multi-level convolutional neural network deep learning algorithm in the field of image recognition has the advantages of high recognition accuracy, high robustness and short time-consuming. It proves the portability of the algorithm in both theoretical and experimental aspects.

Keywords: multi-layer convolutional neural network; deep learning algorithm; portability; analysis; image recognition; fitting effect; delta rule; fisher criteria

收稿日期:2018-10-24. **网络出版日期**:2020-05-07.
基金项目:国家科技重大专项项目(2014ZX03002001).
作者简介:肖堃,男,高级实验师,硕士研究生.
通信作者:肖堃,E-mail: danni445566@ 163.com

深度学习从大类上可以归入神经网络,不过在具体实现中有许多变化。深度学习的核心是特征学习,旨在通过分层网络获取分层次的特征信息,从而解决以往需要人工设计特征的重要难题[1-2]。深度

学习是一个框架,包含卷积神经网络、稀疏编码器、自动编码器等多个重要算法。针对不同问题,需要选取的网络模型达到的处理效果也各不相同[3]。

由于图像特征数目过少,可能无法精确地实现分类,即欠拟合,也不会由于提取的特征数目过多,导致分类过程中过于注重某个特征出现分类错误,即过拟合。卷积神经网络不需要做大量的特征提取和特征选择工作,只需要在其训练完成后,将需要处理的问题输入,即可得到一个较好的拟合效果[4-5]。

针对当前方法在图像识别领域应用过程中容易受到物体形状、位移、尺度、光照、背景等因素的影响导致识别准确率不高,鲁棒性较低、方法可移植性较差的缺点,本文提出了基于多层卷积神经网络深度学习算法的图像识别方法,并分析了该方法在图像识别领域的可移植性。

1　多层卷积神经网络深度学习算法下的图像识别

1.1　模型训练过程

多层卷积神经网络深度学习模型是由多层感知机演变而来的,主要包括输入图像局部感受野、图像权值共享及图像时间/空间下采样 3 个部分,这 3 部分操作的最终目的是为了保证多层卷积神经网络能够学习到输入图像的更多细节特征,大大减少了神经元个数,使得输入图像具有形状、尺度及位移不变性,同时又能够减少输入图像的分辨率,降低计算复杂性,缩小搜索空间,更有利于图像识别[6-7]。

本文使用了 6 层卷积神经模型,包括卷积层、池化层、卷积层、池化层、下采样层、全连接层以及输出层,如图 1 所示。并采用 3×3 的卷积核进行了10 次迭代。

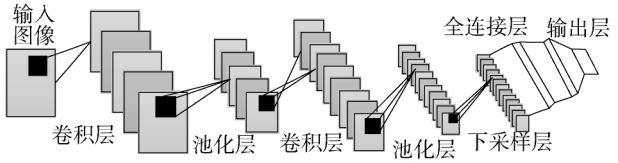

图 1　卷积神经网络结构

Fig.1　Structure of convolutional neural network

训练过程包括数据训练和对应的类别标签训练。在模型初始化设置过程中,多层卷积神经网络深度学习模型的参数多为一些较小的、各部分相同的随机数,不仅能够保证模型不会进入过于饱和状态,而且还能保证模型具有较强的学习能力。多层卷积神经网络深度学习模型的训练方法与 BP 神经网络(BP neural network)模型的训练方法类似,主要分为 2 步: 1) 模型的前向传播,将训练图像数据中的任意一个样本 (X_p, Y_p) 输入模型中,通过多层卷积神经网络深度学习,计算得到模型输出层的实际输出值,即最终的识别结果[8-9],具体计算公式描述为:

$$O_p = F_n(L(F_2(F_1(O_p W_1) W_2)) LW_n) \tag{1}$$

式中: F_1、F_2、LF_n 和 W_1、W_2、LW_n 分别表示多层卷积神经网络深度学习模型中的一组滤波器及其对应权值; n 为模型中包含的训练数据个数。

根据式(1)计算得到模型输出层的实际输出值 O_p 与期望输出值 Y_p 之间的误差大小,采用极小化规则反向传播方法调整多层卷积神经网络深度学习模型中的对应参数。具体过程如下:

1) 极小化规则反向传播方法。

①前向传播。

采用平方误差代价函数计算模型输出层的实际输出值 O_p 与期望输出值 Y_p 之间的误差大小。假设训练的输入图像数据有 N 个,对应的类别标签有 C 个,则可得模型训练整体误差 E^n 的计算公式:

$$E^n = \frac{1}{2}\sum_{n=1}^{N}\sum_{k=1}^{C}(t_k^n - y_k^n)^2 \tag{2}$$

式中: t_k^n 为多层卷积神经网络深度学习模型中第 n 个训练数据期望输出的第 k 维数据; y_k^n 为多层卷积神经网络深度学习模型中第 n 个训练数据实际输出的第 k 维数据。

如果只考虑单个训练数据,则可得多层卷积神经网络深度学习模型中第 n 个训练数据的误差为:

$$E^n = \frac{1}{2}\sum_{n=1}^{N}(t_k^n - y_k^n)^2 = \frac{1}{2}\|\boldsymbol{t}^n - \boldsymbol{y}^n\|_2^2 \tag{3}$$

根据上述计算可得当前模型第 l 层的输出值为:

$$X^l = f(u^l) \tag{4}$$

$$u^l = \boldsymbol{W}^l X^{l-1} + b \tag{5}$$

式中: $f(g)$ 为模型输出激活函数,用于获得待识别图像特征图; $\boldsymbol{W}^l$ 为模型第 l 层训练数据的权值矩阵; X^{l-1} 为当前模型第 $l-1$ 层的输出; b 为模型偏置项。

②模型反向传播过程中,模型中每个神经元对于误差偏置项的灵敏度计算公式:

$$\frac{\partial E}{\partial b} = \frac{\partial E}{\partial u}\frac{\partial u}{\partial b} = \delta \tag{6}$$

$$\frac{\partial u}{\partial b} = 1 \tag{7}$$

式中: $\partial E/\partial u$ 为模型误差对于输入图像数据 u 的导数。则可得模型反向传播过程中第 l 层神经元偏置项灵敏度计算公式为:

$$\delta^l = (\boldsymbol{W}^{l+1})^{\mathrm{T}}\delta^{l+1} \odot f(u^l) \tag{8}$$

式中: δ^{l+1} 为模型第 $l+1$ 层神经元偏置项灵敏度; $\odot$ 表示模型中每个元素相乘。

对于误差函数,模型输出层的偏置项灵敏度与上述计算略有不同,表示为:

$$\delta^l = f(u^l) \odot (y^n - t^n) \tag{9}$$

在上述计算基础上，采用 delta 法则，实时更新多层卷积神经网络深度学习模型的各神经元权值，计算公式为：

$$\Delta \boldsymbol{W}^l = -\eta \frac{\partial E}{\partial \boldsymbol{W}^l} \tag{10}$$

$$\frac{\partial E}{\partial \boldsymbol{W}^l} = \boldsymbol{X}^{l-1} (\boldsymbol{\delta}^l)^{\mathrm{T}} \tag{11}$$

式中 η 为模型的学习率。

2) 模型中下采样层权值更新。

对于模型中的下采样层，输入的特征图数目与输出的特征图数目相等，即：

$$x_j^l = f(\beta \mathrm{down}(x_j^{l-1}) + b) \tag{12}$$

下采样层神经元偏置项灵敏度，即对输入图像做降采样处理：

$$\delta_j^l = f(\beta \mathrm{down}(x_j^{l-1}) + b) \tag{13}$$

式中 $\mathrm{down}(\cdot)$ 表示下采样函数。

模型中下采样层权值更新方法与卷积层权值更新方法类似，计算公式为：

$$\frac{\partial E}{\partial \beta} = \sum_{u,v} (\delta_j^l \mathrm{down}(x_j^{l-1}))_{u,v} \tag{14}$$

1.2 多层卷积神经网络深度学习

在上述多层卷积神经网络深度学习模型训练完成基础上，为了提高该模型在图像识别过程中的准确率，引入 Fisher 准则，提出了基于类内距离和类间距离的能量函数[10-11]，计算公式为：

$$J = R + \gamma J_1 - \eta J_2 \tag{15}$$

式中：R 为多层卷积神经网络深度学习模型的代价函数；γ 为能量函数调节参数；J_1 为待识别图像样本之间的相似度函数；J_2 为待识别图像类别标签之间的相似度函数。

在充分考虑类内距离 J_1 和类间距离能量函数 J_2 的基础上，计算模型输出层各个输出单元的残差，使得模型权值能更加快速收敛到有利于图像识别的最优值[12]，计算公式为：

$$\begin{cases} \varepsilon = \dfrac{\partial J_1}{\partial \delta_j^l} = \dfrac{\partial}{\partial \delta_j^l} \dfrac{1}{2} \displaystyle\sum_{i=1}^{m} \sum_{j=1}^{n} \| \boldsymbol{Y}_p - \boldsymbol{M}^{(i)} \|^2 \\ \varepsilon = \dfrac{\partial J_2}{\partial \delta_j^l} = \dfrac{\partial}{\partial \delta_j^l} \dfrac{1}{2} \displaystyle\sum_{i=1}^{m} \sum_{j=1}^{n} \| \boldsymbol{Q}_p - \boldsymbol{M}^{(i)} \|^2 \end{cases} \tag{16}$$

$$M^{(i)} = \frac{\sum_{j=1}^{n} Y_p}{n} \tag{17}$$

式中 $M^{(i)}$ 为待识别图像第 i 类样本的均值。根据式(17)即可找到有利于图像识别的最优权值，将该权值计算结果代入式(17)中，即可获得图像识别结果，该理论分析结果证明多层卷积神经网络深度学习算法具有可移植性。

2 可移植性测试

硬件环境：3.4 GHz 双核处理器，4 GB 内存，M40 显卡，Ubuntu14.04 操作系统。

软件环境：C++编译器。

为检验多层卷积神经网络深度学习算法在图像识别领域应用的可移植性，随机选取 PASCAL VOC2007 数据库中的 2 750 张图片，6 个类别标签作为测试集，如表 1 所示。

表 1 实验测试数据集

Table 1 Experimental test data set

类别标签	数目/张
数字	100
车辆	150
植物	200
建筑物	500
人脸	800
指纹	1 000

选取识别准确率、鲁棒性、以及识别耗时 3 项指标作为评价指标，识别表 1 中的 6 种图像类型，测试基于多层卷积神经网络深度学习算法的图像识别方法的有效性，测试结果如表 2 所示。

表 2 测试结果统计

Table 2 Test results statistics

类别标签	识别准确率/%	鲁棒性/%	识别耗时/ms
数字	100	100	12
天气	99	99	15
植物	96	97	17
建筑物	96	98	18
人脸	95	94	21
车辆	92	93	24

从表 2 中可以看出，采用多层卷积神经网络深度学习算法识别 6 种不同类型图像的平均准确率为 96.3%；识别 6 种不同类型图像的平均鲁棒性为 96.8%；识别 6 种不同类型图像的平均耗时为 17.8 ms。这是由于所提方法在训练完多层卷积神经网络深度学习模型基础上，引入 Fisher 准则，并提出了基于类内距离和类间距离的能量函数，使得模型权值能够更加快速收敛到有利于图像识别的最优值，大大提高了 6 种不同类型图像识别准确率和鲁棒性，同时加快了识别速度，平均识别耗时均达到毫秒级。假设道路上发生交通事故，所提方法具有毫秒级处理速度，有利于决策者及时作出应对措施，避免交通事故造成的经济损失和人员伤亡。

为了进一步检验所提方法的有效性和可移植性，选取比较复杂的人脸图像和容易受光照、背景、位移等影响的车辆图像作为训练样本，训练样本中喜欢、愤怒、悲伤、快乐 4 种表情各 200 张，受背景干

扰、光照干扰、噪声干扰和雾霾干扰的车辆各 250 张，其中部分样本示例如图 2 和图 3 所示，采用所提方法识别 4 类人脸图像和 4 类不同类型干扰车辆图像，测试结果如图 4 和图 5 所示。

图 2　人脸图像训练样本

Fig.2　Face image training sample

图 3　车辆图像训练样本

Fig.3　Vehicle image training sample

分析图 4 和图 5 的实验结果可以发现，采用所提方法无论是识别具有复杂表情的人脸图像训练样本，还是具有各种干扰因素影响的车辆图像训练样本均取得了较好的成果，识别准确率和鲁棒性均较高，识别耗时较少。这是由于所提方法构建的识别模型中包括输入图像局部感受野、图像权值共享以及图像时间/空间下采样 3 个部分，这 3 部分操作保证了多层卷积神经网络能够学习到输入图像的更多细节特征，大大减少神经元个数，使得输入图像具有形状、尺度以及位移不变性，同时又能够减少输入图像的分辨率、降低计算复杂性、缩小搜索空间，更有利于识别。通过多层卷积神经网络深度学习算法对样本数据进行多次迭代训练，通过分类器输出训练后的实验样本，降低了原始数据的偏置程度，以提高实验结果的准确性。

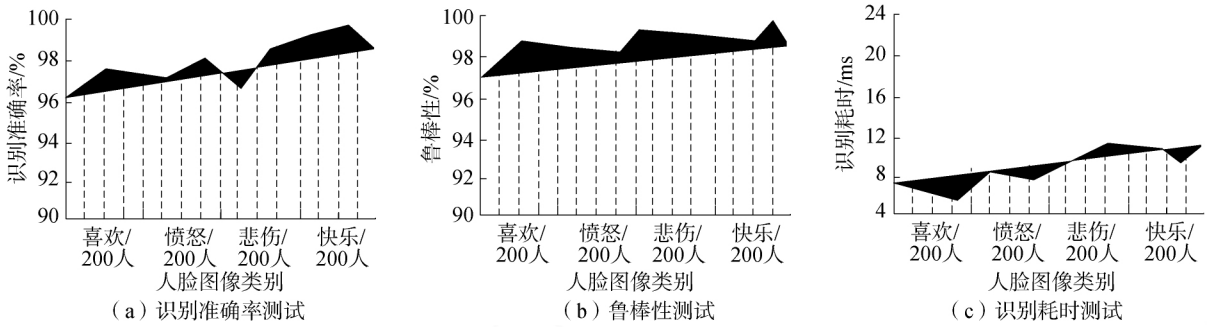

图 4　人脸图像训练样本的识别结果

Fig.4　Recognition results of face image training samples

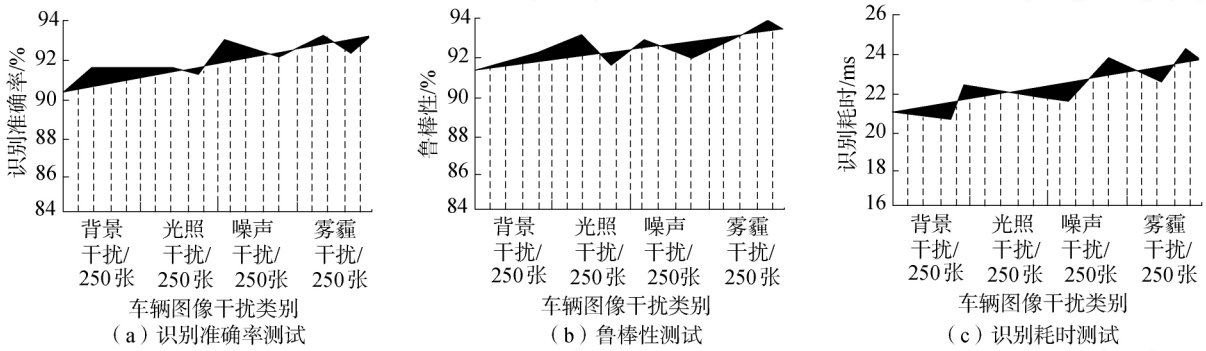

图 5　车辆图像训练样本的识别结果

Fig.5　Recognition results of vehicle image training samples

上述实验结果表明，多层卷积神经网络深度学习算法应用在图像识别领域具有可移植性，这与上述理论分析结果一致。

3　结论

1) 本文提出了基于多层卷积神经网络深度学习算法的图像识别方法，利用模型权值共享、更新、下采样等操作对输入图像做降采样处理，有效降低了计算复杂度。

2) 利用 delta 法则、Fisher 准则以及基于类内距离和类间距离的能量约束函数实时调整多层卷积神经网络深度学习模型参数，计算模型输出层各个输出单元的残差，让模型权值能够更加快速收敛到有利于图像识别的最优值。

3) 本文实现了所提出的算法，并随机选取 PASCAL VOC2007 数据库中的图片，从理论和实验两方面证明了所提方法的可移植性。

参考文献:

[1] 刘万军，梁雪剑，曲海成. 不同池化模型的卷积神经网络学习性能研究 [J]. 中国图象图形学报，2016，21(9)：1178-1190.
LIU Wanjun，LIANG Xuejian，QU，Haicheng. Learning performance of convolutional neural networks with different pooling models [J]. Journal of image and graphics，2016，21(9)：1178-1190.

[2] 朱威，屈景怡，吴仁彪. 结合批归一化的直通卷积神经网络图像分类算法 [J]. 计算机辅助设计与图形学学报，2017，29(9)：1650-1657.
ZHU Wei，QU Jingyi，WU Renbiao. Straight convolutional neural networks algorithm based on batch normalization for image classification [J]. Journal of computer-aided design & computer graphics，2017，29(9)：1650-1657.

[3] 常玲玲，马丙鹏，常虹，等. 深度网络结构在行人检测任务中的性能对比 [J]. 计算机仿真，2017，34(7)：373-377，411.
CHANG Lingling，MA Ruipeng，CHANG Hong，et al. Comparative study on deep network architectures in pedestrian detection task [J]. Computer simulation，2017，34(7)：373-377，411.

[4] 王晨，汤心溢，高思莉. 基于深度卷积神经网络的红外场景理解算法 [J]. 红外技术，2017，39(8)：728-733.
WANG Chen，TANG Xinyi，GAO Sili. Infrared scene understanding algorithm based on deep convolutional neural network [J]. Infrared technology，2017，39(8)：728-733.

[5] 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.
LYU Yongbiao，ZHAO Jianwei，CAO Feilong. Image denoising algorithm based on composite convolutional neural network [J]. Pattern recognition and artificial intelligence，2017，30(2)：97-105.

[6] 卢宏涛，张秦川. 深度卷积神经网络在计算机视觉中的应用研究综述 [J]. 数据采集与处理，2016，31(1)：1-17.
LU Hongtao，ZHANG Qinchuan. Applications of deep convolutional neural network in computer vision [J]. Journal of data acquisition & processing，2016，31(1)：1-17.

[7] 伍家松，达臻，魏黎明，等. 基于分裂基-2/(2a) FFT算法的卷积神经网络加速性能的研究 [J]. 电子与信息学报，2017，39(2)：285-292.
WU Jiasong，DA Zhen，WEI Liming，et al. Acceleration performance study of convolutional neural network based on split-radix-2/(2a) FFT algorithms [J]. Journal of electronics & information technology，2017，39(2)：285-292.

[8] 厉智，孙玉宝，王枫，等. 基于深度卷积神经网络的服装图像分类检索算法 [J]. 计算机工程，2016，42(11)：309-315.
LI Zhi，SUN Yubao，WANG Feng，et al. Clothing image classification and retrieval algorithm based on deep convolutional neural network [J]. Computer engineering，2016，2016，42(11)：309-315.

[9] 周林腾. 基于神经网络算法的大数据分析方法研究 [J]. 电子设计工程，2018，26(9)：19-22，27.
ZHOU Linteng. Study on the analysis method of data based on neural network algorithm [J]. Electronic design engineering，2018，26(9)：19-22，27.

[10] 刘海龙，李宝安，吕学强，等. 基于深度卷积神经网络的图像检索算法研究 [J]. 计算机应用研究，2017，34(12)：3816-3819.
LIU Hailong，LI Baoan，LYU Xueqiang，et al. Image retrieval based on deep convolutional neural network [J]. Application research of computers，2017，34(12)：3816-3819.

[11] 余萍，赵继生. 基于矩阵2-范数池化的卷积神经网络图像识别算法 [J]. 图学学报，2016，37(5)：694-701.
YU Ping，ZHAO Jisheng. Image recognition algorithm of convolutional neural networks based on matrix 2-norm pooling [J]. Journal of graphics，2016，37(5)：694-701.

[12] 王华利，邹俊忠，张见，等. 基于深度卷积神经网络的快速图像分类算法 [J]. 计算机工程与应用，2017，53(13)：181-188.
WANG Huali，ZOU Junzhong，ZHANG Jian，et al. Fast image classification algorithm based on deep convolutional neural network [J]. Computer engineering and applications，2017，53(13)：181-188.

本文引用格式:

肖堃. 多层卷积神经网络深度学习算法可移植性分析 [J]. 哈尔滨工程大学学报，2020，41(3)：420-424.
XIAO Kun. Portability analysis of multilayer convolutional neural network's deep learning algorithm [J]. Journal of Harbin Engineering University，2020，41(3)：420-424.