

文章编号: 2095-2163(2021)03-0215-03

中图分类号: TP391.41

文献标志码: A

深度学习在影像中的算法研究

肖行

(广州市第一人民医院, 广州 510180)

摘要:深度学习技术的运用正日趋广泛,深度学习自身的高效性和智能性受到研究者的青睐。通过对深度学习影像分类的剖析,进一步探究深度学习在影像识别方向的应用,介绍了主要用于影像分类识别的基于深度学习的医疗影像检测算法,可作为开展深度学习技术运用于医学影像检测研究工作的有益参考。

关键词:深度学习;影像分类;影像识别

Research on algorithms of deep learning in images

XIAO Hang

(Guangzhou First People's Hospital, Guangzhou 510180, China)

[Abstract] The use of deep learning technology is becoming more widespread, and the efficiency and intelligence of deep learning itself are favored by researchers. Through the analysis of deep learning image classification, the paper further explores the application of deep learning in the direction of image recognition, introduces the medical image detection algorithm based on deep learning, which is mainly used for image classification and recognition. The fruits could be used as a beneficial reference for the application of deep learning technology to medical imaging detection research work.

[Key words] deep learning; image classification; image recognition

0 引言

在复杂的背景中,一幅图像会具有许多对象。如何识别这些对象,辨识其中的主要对象,并理解主对象与其他对象之间的关系已然成为目前的研究热点。对象识别有多种方法,但大部分都不能标记图像的主要对象。研究可知,可以运用改进的RCNN^[1]网络来检测和识别图像中的多个对象,由此即提出了增强型的目标评分系统来标记图像的主要对象。实验结果表明,该算法不仅保持了RCNN的优越性,而且能检测到图像的主要对象。

近年来,基于计算机的图像识别技术获得迅猛的发展。与之相适应,在医学领域就已对人工神经网络的影像学识别展开大量研究。例如,在糖尿病视网膜病变的诊断中,有较高的诊断预测能力。在乳腺癌的诊断中,在乳腺癌中淋巴结转移的病理诊断中,在胸部X射线的分类中,在食管胃十二指肠镜检查中分类等^[2],均取得了可观成果。需要指出的是,尽管人工智能技术被应用于预测manikins的气道图像的glottic开放,但在临床实践中获得的患者的喉部图像的语言开放或语言位置的研究却仍不多见。

气管插管是一项重要的医疗程序,在自发性呼

吸、气道维护上存在困难,究其原因就是麻醉、全身麻醉和心肺问题。气管插管是一种生命保存的程序,可在心脏、呼吸阻塞等情况下进行,具体来说就是当病患处于高危的呼吸、缺氧、通风不足、气道阻塞的情况下。glottis是2个声带间的一个开口。在执行插管时,应将塑料插管插入气管中。然而,气管导管有时插入不适当的结构,如食道,而不是glottis。如果气管插管插入结构出现错误,就会导致严重的并发症,如低氧血症和心脏骤停^[3]。因此,在插管时准确地识别glottis的位置是非常重要的。

视频喉镜是一种将相机插入刀片的较为尖端的技术方法。视频喉镜可提高glottic的视觉化,减少食管插管突发事件的发生概率。关于视频喉镜检查是否会增加第一次尝试成功率,不同的研究结果仍存在差异。当前的一项研究中显示,在视频喉镜检查时,使用前医院插管的成功率较低。此外,即使在视频喉镜检查时,食管插管也会发生。单食管插管增加了脱气、吸气和心脏骤停的风险。

1 ANNs的深度学习算法与反向传播的区别

在使用ANNs的临床实践中,尝试探究了在此应用中去获得气道图像的理论位置的预测模型。一

作者简介:肖行(1988-),男,硕士,中级工程师,主要研究方向:商务智能。

通讯作者:肖行 Email:hank0526@foxmail.com

收稿日期:2020-12-24

个模仿动物神经元结构的 ANN 可应用于实现类似于大脑的功能。动物的实际学习过程永远不会是非自然的数学理论,即如在反向传播方法中一样,也会有类似基于 trial-and-error 过程的生物进化的技术内容。因此,在本次研究中开发并应用了基于蒙特卡罗模拟的 ANNs 的一种新的深度学习算法。ANN 包含了成千上万个或更多的未知变量、权重因子和偏置值。这个新的深度学习算法是应用蒙特卡罗模拟的优化过程,用来确定权重因子和偏差值,使学习数据的平均训练误差最小化^[4]。

对于反向传播方法,采用梯度下降方法,通过多次使用全部或部分学习数据来确定结果数值,直到根据给定的学习速率训练误差达到最小为止。在此条件下,节点的偏压值作为节点的附加权重因子,将其输入值设定为 1.0,而并不考虑偏差值是否为负的变化机制。本次研究的新算法具有与反向传播方法完全不同的结构。在新的深度学习算法的情况下,当前 ANN 的所有学习数据的平均训练误差都是在训练误差达到最小的情况下反复计算的结果,而随机选择的权重因子和 ANN 的偏差值正在根据给定范围内随机选择的 delts 值进行调整。当然,该算法无需计算训练错误的梯度,或者在训练阶段使用所有或一部分学习数据的权重因素和偏差值来做调整,按小随机量调整一个 ANN 的权重因子和偏压值,而不是通过计算密集型梯度下降法分别应用于所有权重因子,是与反向传播方法中在训练期间的主要差异。因此,该算法简单而高效,对计算资源的要求也不高。

2 算法原理与实现

ANN 是一个连接的简单计算元素集,称为节点。通常,是一个多层的组织。总的来说,可分为:一个输入层、多个隐藏层和一个输出层。对于深度学习方法而言,就代表着有些 ANNs 可能还会包括数百个或更多的隐藏层。输入层中的每个节点,即输入节点,将接收来自外部的输入,并将其传送给第一个隐藏层的所有节点。在本次研究中,考虑到气道图像大小和纵横比的不同,先把图像转换成正方形,再将其分辨率降到 100×100 、 70×70 、 50×50 、 45×45 、 40×40 、 35×35 、 30×30 和 25×25 像素,并且每组应用于 ANN 模型的输入结构。ANN 的输入节点的数量就等于原始气道图像减少分辨率后的总数。每个输入节点的输入值是相应的像素值,即减少分辨率的气道图像。可以通过计算来获得黑白彩色转换过程中的像素值,此处会用到的数学公式可写为:

$$value = red \times 0.299 + green \times 0.587 + blue \times 0.114. \quad (1)$$

在此基础上,将该值除以最大值 255,即将其转换为 0~1.0 之间的值。相同的过程应用于标记气道的图像,以获得像素值,再将图像分为 7 个横截面和 7 个垂直截面。该分区共生成了 49 个具有标记气道图像的单元格。这些图像被分为 49 个单元格,用来预测在气道图像中 49 个 glottic 的位置。

3 实验结果

通过将训练集和测试集划分为 2 种类型的视图:好视图和差视图,以预测选择模型的气道位置的训练和测试精度,实验结果见表 1。

表 1 不同视图的预测率

Tab. 1 Forecast rate of different views %

图像类型	数据集类型	准确预测率	相邻预测率	不准确预测率
好视图	训练集	83.3	14.3	2.4
	测试集	78.6	19.0	2.4
差视图	训练集	55.6	30.5	13.9
	测试集	53.1	34.4	12.5

表 1 中,准确预测率表示被预测的位置与 glottis 重叠;相邻预测率表示 8 个相邻的相等大小的正方形,被预测的位置与 glottis 重叠;不准确预测率表示被预测的位置与 glottis 不相邻。

对于训练集,好视图图像数量是 813,占 81.3%;差视图的图像数量是 187,占 18.7%。对于测试集,好视图的图像数量为 168,占 84.0%;差视图的图像数量为 32,占 16.0%。对于好视图而言,仿真得到的测试集的精确预测率为 83.3%,测试集的预测率为 78.6%。对于差视图来说,测试集的准确预测率为 55.6%,测试集的准确预测率为 53.1%。

4 结束语

提出了一种新的基于人工神经网络(ANNs)的深度学习算法。该方法与反向传播方法完全不同。研究中随机选择一个 ANN 的权重因子和偏置值,并在训练期间通过小随机数来调整相应数值,不需要计算训练误差的梯度来调整权重因子。该算法应用于通过视频气道装置获得的气道图像中 glottis 的位置。在 1 200 个气道图像中,使用 GlideScope R 和纤维镜检查。对于随机选取的 1 000 个训练集数据,利用上述算法训练了 84 个 ANN 模型。寻求一个 ANN 模型,通过减少输入图像分辨率,将所有训练

(下转封三)