

基于堆叠降噪自动编码器的胶囊缺陷检测方法

王宪保^{1,2} 何文秀¹ 王辛刚¹ 姚明海¹ 钱运涛²

(浙江工业大学信息工程学院 杭州 310023)¹ (浙江大学计算机科学与技术学院 杭州 310027)²

摘 要 目前医用胶囊生产过程中的缺陷检测主要由人工完成,费时费力,容易受主观因素的影响。提出一种基于堆叠降噪自动编码器的胶囊表面缺陷检测方法,该方法首先建立深度自动编码器网络,并根据缺陷样本进行降噪训练,获取网络的初始权值;然后通过 BP 算法进行微调,得到训练样本到无缺陷模板之间的映射关系;最后利用重构图像与缺陷图像之间的对比关系,实现测试样本的缺陷检测。实验表明,堆叠降噪自动编码器较好地建立了上述映射关系,能快速、准确地进行缺陷检测,对噪声具有很强的鲁棒性和稳定性。

关键词 堆叠降噪自动编码器,缺陷检测,深度学习

中图法分类号 TP391

文献标识码 A

DOI 10.11896/j.issn.1002-137X.2016.2.014

Capsule Defects Detection Based on Stacked Denoising Autoencoders

WANG Xian-bao^{1,2} HE Wen-xiu¹ WANG Xin-gang¹ YAO Ming-hai¹ QIAN Yun-tao²

(College of Information Engineering, Zhejiang University of Technology, Hangzhou 310023, China)¹

(College of Computer Science and Technology, Zhejiang University, Hangzhou 310027, China)²

Abstract At present defects of capsules are detected mainly by manual operation, which is time-consuming and needs high labor costs, besides, it is easily misled by subjective factors. This paper proposed a method of detection of capsules surface defects based on stacked denoising autoencoders (SDAE). Our method firstly establishes deep autoencoders networks and trains using a denoising criterion according to the defect samples to obtain the initial weights at first. Then, BP algorithm fine-tunes the network parameters to get the mapping relationship between the training sample and defect-free template. Finally, defect detection of the testing samples is finished by comparing the reconstruction image and defect image. Experimental results show that SDAE perfectly establishes the mapping relationship, which is robust and stable to noise, and can quickly detect defects with high accuracy.

Keywords Stacked denoising autoencoders, Defect detection, Deep learning

1 前言

深度学习(Deep Learning, DL)是机器学习研究的新领域,其自 Hinton 在 2006 年^[1]提出以来,由于在学习时间^[2]、特征提取^[3]等方面表现出来的优越特性,已在图像^[4]、声音和文本^[5]等领域的信号处理与识别中得到了广泛应用,并取得了良好的实验效果,引起了众多学者的研究兴趣。深信度网络(Deep Belief Networks, DBN)是 DL 算法的主要实现结构之一,其主要思想是在每一深度下尽量保持输入信号的信息,从而达到特征提取的目的,因此一个 DL 结构又被称为编码器^[6]。但在实际应用中,系统的输入信号往往带有各种类型的噪声,一个理想的编码器应该能从带噪信号中鲁棒地获取特征并能恢复出原有的纯净信号^[7],这显然是 DBN 没有考虑的。

堆叠降噪自动编码器(Stacked Denoising Autoencoders, SDAE)把降噪作为网络学习准则,训练时可在输入信号中加

入不同类型、不同强度的噪声,因此其编码过程具有良好的稳定性和鲁棒性,在特征提取和模式识别领域表现出了超越其它结构的特性。文献[8]把 SDAE 用于人脸识别中,实现了人脸姿态的自动校正,简化了识别模型,大大提高了正确识别率。文献[9]基于 SDAE 实现了遮挡人脸的自动恢复,在脸部有饰物或部分被遮挡的情况下,仍能得到正确的识别。文献[10]利用 SDAE 对噪声鲁棒性强的特点,对深度图像进行学习,并与 SVM 识别算法相结合,实现了 3 维动作识别。文献[11]把多个 SDAE 组合在一起,并分别由一种类型噪声样本训练,工作时,样本首先经过一个预测网络估计出样本的噪声类型,然后再计算出对应网络应有的权值,从而达到自适应降噪的目的。

胶囊缺陷检测是药品生产领域的重要环节,目前主要由人工来完成,不仅生产效率低、漏检率高,而且人工检测的标准差异较大,造成产品质量参差不齐,不能满足日益严格的生产要求。随着计算机视觉技术的发展,机器视觉在胶囊检测

到稿日期:2015-05-03 返修日期:2015-06-27 本文受浙江省自然科学基金(LY14F030009, LZ14F030001)资助。

王宪保(1977—),男,博士,讲师,CCF 会员,主要研究领域为神经网络、模式识别、图像处理等, E-mail: wxb@zjut.edu.cn; 何文秀(1979—),女,博士生,讲师,主要研究领域为人工智能、无线传感网络等; 王辛刚(1974—),男,博士,副教授,主要研究领域为多项式优化理论、并行计算等; 姚明海(1963—),男,博士,教授,博士生导师,主要研究领域为模式识别、智能控制、嵌入式系统和无线传感器网络等; 钱运涛(1968—),男,博士,教授,博士生导师,主要研究领域为机器学习、模式识别和图像处理。

中也逐步得到应用。文献[12]通过提取胶囊接口处的直线特征和胶囊的颜色,利用 BP 神经网络实现了缺陷样本的识别。文献[13]提取了胶囊的宽度、高度和面积 3 个特征,然后利用最小贝叶斯分类器对胶囊完整性进行检测,从而判断胶囊是否异常。文献[14]提取了胶囊投影面积并与 6 个形态参数组成特征向量,然后利用 BP 神经网络进行分类识别。文献[15]针对胶囊的凹陷缺陷进行不变矩分析,然后提取偏心率和圆弧度作为特征变量,最后基于极端学习机对胶囊的缺陷进行检测与识别。可以看出,目前胶囊缺陷检测方法都是通过“特征表达”+“分类器”的框架来进行缺陷分类检测,一般先对采集到的胶囊图像进行处理,然后提取相应的特征,最后利用分类器进行分类检测。这些方法对胶囊的外观缺陷没有通用性,而且胶囊的特征表达是人为设计的,参数复杂,缺乏扩展性。

本文提出了一种基于 SDAE 的胶囊缺陷自动检测方法。该方法首先通过 SDAE 对大量缺陷样本进行学习,得到训练样本与无缺陷模板之间的映射关系;然后利用重构图像与缺陷图像之间的对比,实现样本的缺陷检测。该方法避免了传统检测方法中复杂的图像处理环节,大大提高了检测效率。

2 自动编码器及改进算法

2.1 自动编码器

一个自动编码器(Autoencoders, AE)对输入样本进行学习时,其训练目标是将输入信号从目标表达中重构出来,因此,训练时往往把输出设置为输入自身。AE 结构可分为两部分:编码器和解码器。

编码器完成从输入向量 X 到输出表征 Y 的映射转换,其典型表达形式为:

$$Y = f_{\theta}(X) = s(WX + b) \quad (1)$$

模型参数 $\theta = \{W, b\}$, 其中 W 为 $d' \times d$ 维权值矩阵, b 为 d' 维偏移向量。

解码器实现输出表征 Y 逆向映射回 d 维输入空间, 重构向量 Z , 其典型形式为:

$$Z = g_{\theta'}(Y) = s(W'Y + b') \quad (2)$$

模型参数为 $\theta' = \{W', b'\}$, 其中 W' 为 $d \times d'$ 维权值矩阵, b' 为 d 维偏移向量。

一般情况下, Z 不是输入变量 X 的精确重构, 它只是在满足一定分布的条件概率 $p(X|Z)$ 下, 最大程度地接近 X , 所以目标函数可转化为对下面重构误差函数的优化:

$$L(X, Z) \propto -\log p(X|Z) \quad (3)$$

对于式(3)函数形式的选择, 可分为两种情况考虑:

(1) X 为实数, 即

$$X \in \mathbb{R}^d; X|Z \sim N(Z, \sigma^2 I)$$

这时误差函数可选择为:

$$L(X, Z) = L_2(X, Z) = C(\sigma^2) \|X - Z\|^2 \quad (4)$$

其中, $C(\sigma^2)$ 为由 σ^2 决定的常数, 在优化时可忽略。这种平方差的形式在传统 AE 中是最常用的。

(2) X 为二进制数, 即

$$X \in \{0, 1\}^d; X|Z \sim B(Z)$$

这时误差函数经常选用扁平状的非线性形式(例如 s 型函数), 可写为:

$$\begin{aligned} L(X, Z) &= L_H(X, Z) \\ &= -\sum_j [x_j \log z_j + (1 - x_j) \log(1 - z_j)] \end{aligned}$$

$$= H(B(X) \| B(Z)) \quad (5)$$

这种形式由于像两个独立变量的伯努利交叉熵, 因此称为交叉熵损失函数, $B(X)$ 和 $B(Z)$ 分别表示二者的均值。这种函数形式也可用在非严格的二进制形式下, 即 $X \in [0, 1]^d$ 。

另外, 根据 AE 的定义, 容易得到进一步的约束条件, 即

$$W' = W^T \quad (6)$$

AE 的训练过程就是最小化重构误差函数, 即求取如下最优:

$$\arg \min_{\theta, \theta'} E_{q^0(X)} [L(X, Z(X))]$$

结合式(3), 最优函数可重写为:

$$\arg \min_{\theta, \theta'} E_{q^0(X)} [\log p(X|Z = g_{\theta'}(f_{\theta}(X)))] \quad (7)$$

或

$$\arg \min_{\theta, \theta'} E_{q^0(X)} [\log p(X|Y = f_{\theta}(X); \theta')] \quad (8)$$

2.2 DAE

降噪自动编码器(Denoising Autoencoders, DAE)与传统的 AE 具有相同的结构, 只是在样本输入时加入了某种类型的噪声, 其学习目标是从受污染的输入中重构出纯净的输入。其计算过程如下。

设输入样本为 X , 污染后的随机映射样本为 $\tilde{X} \sim q_D(\tilde{X}|X)$, 则 DAE 的编码输出为:

$$Y = f_{\theta}(X) = s(W\tilde{X} + b) \quad (9)$$

重构后的解码输出为:

$$Z = g_{\theta'}(Y) = s(W'Y + b') \quad (10)$$

其参数、结构与基本的 AE 完全相同, 不同点在于损失函数:

$$L_2(X, Z) = C(\sigma^2) \|X - Z\|^2, \text{线性解码器}$$

或

$$L_H(X, Z) = H(B(X) \| B(Z)), \text{非线性解码器}$$

损失函数虽然仍旧是最小化纯净输入信号 X 和重构信号 Y 之间的误差, 或者说是最大化纯净输入信号和重构信号之间的共有信息, 但与传统 AE 不同的是, 重构信号 Y 是由受污染信号 $\tilde{X}$ 重构得到的。具体计算过程如图 1 所示。

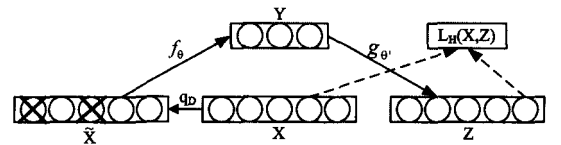


图1 降噪自动编码器流程

需要说明的是, 样本 X 每训练一次, 都会根据 $q_D(\tilde{X}|X)$ 产生一个不同的 $\tilde{X}$ 。

2.3 SDAE

为了得到更高级的特征表达, 用深度网络结构的形式, 把 DAE 逐层堆叠起来, 形成一个由 DAE 上下连接而成的模型结构, 即为 SDAE。在训练时, 前一层的输出作为后一层的纯净输入, 逐层进行训练。其学习过程如图 2 所示。

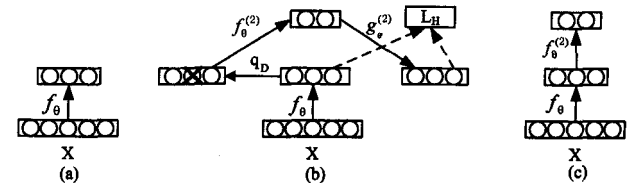


图2 SDAE 计算过程

图 2(a)表示第一层 DAE, 训练过程如图 1 所示, 得到的

编码函数 f_{θ} 用来对输入 X 进行降噪编码。图 2(b) 表示第一层的结果输出作为第二层的样本输入,并训练得到第二层的编码函数 $f_{\theta}^{(2)}$ 。整个深度网络的训练过程以此重复,如图 2(c) 所示。

与 DBN 一样,SDAE 也可以在顶层加入逻辑分类器,把得到的高级特征作为分类器的输入,形成有监督学习。网络所有层的参数可利用梯度下降法进行微调,从而达到最优的识别效果。

3 胶囊缺陷检测可行性分析

若把缺陷看作噪声,把完整胶囊看作共有特征,则利用 SDAE 特征提取的鲁棒性和稳定性,用缺陷样本库训练得到缺陷图像到完整胶囊的映射关系,然后利用重构图像与缺陷图像之间的对比,就可以实现胶囊的缺陷检测。基于 SDAE 的胶囊缺陷检测方法的可行性还在于:

(1) 胶囊表面具有简单的纹理图案,在检测背景不变、检测位置变化不大的情况下,无缺陷图像具有一致的纹理特征,而有缺陷图像只是在局部存在灰度差异,可看作某种类型的噪声,这些特点完全适用于 SDAE 鲁棒的特征提取能力;

(2) SDAE 把图像处理和特征提取结合为一体,大大缩短了缺陷检测时间,能更好地满足在线产品检测;

(3) SDAE 能快速地完成权值训练,即使在产品背景、位置、型号发生改变时,也能通过短时训练完成新映射关系的重新学习,满足灵活的生产需要;

(4) SDAE 与传统识别网络相结合,不但可实现图像处理、特征提取,还可以完成模式识别,为后续的缺陷识别提供实现途径。

4 缺陷检测实现

SDAE 的输入层是 4096 维,4 个隐藏层维度分别是 1000、500、200、50。首先按图 2 的形式连接网络,并依 SDAE 算法得到相应的预训练权值,如图 3(a) 所示。然后在此权值的基础上,按图 3(b) 的形式展开。最后把输入样本作为网络的理想输出,用 BP 算法对权值做进一步的微调,如图 3(c) 所示。

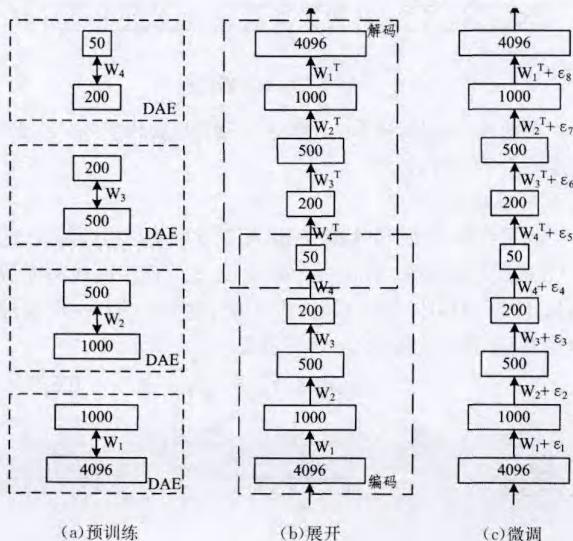


图 3 缺陷检测训练过程

(1) 首先训练第一个 DAE 网络,即输入层 4096 维和第一

个隐藏层 1000 维构成的网络。采用最优梯度下降法优化,计算得到训练样本在隐藏层的输出值。

(2) 将(1)的结果作为第 2 个网络训练的输入值,同样用最优梯度下降法来优化,并计算网络的输出值。用同样的方法训练第 3 个 DAE 网络和第 4 个 DAE 网络。

(3) 将上面 4 个 DAE 网络展开连接成新的网络,且分成编码和解码两部分,用步骤(1)和步骤(2)得到的权值给网络赋初值。

(4) 将输入值作为 SDAE 网络的理论输出值,同样用最优梯度下降法来优化,得到最终的网路权值。

5 实验及结果分析

本次训练是在 Intel Core i7 4600M 2.9GHz、内存为 8GB 的 CPU 平台上,利用 MATLAB 工具软件完成的,其中的缺陷包括裂纹、孔洞和胶囊帽缺失等。实验输入样本为 64×64 灰度图像,包括有缺陷图像和无缺陷图像。图 4 是待检测胶囊图像,其中图 4(a) 为无缺陷图像,图 4(b) 为胶囊帽缺失图像,图 4(c) 为孔洞图像,图 4(d) 为裂纹图像。

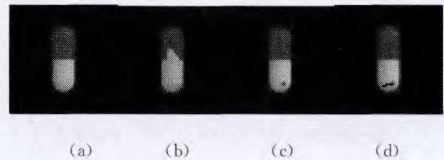


图 4 胶囊的无缺陷图像和缺陷图像

选取 2000 幅图像为训练样本,200 幅为测试样本,对网络进行训练。隐层激活函数为 s 函数,学习速率为 0.5,输入噪声为零掩蔽噪声。

实验结果如图 5 所示。左侧为输入图像,中间为经过 SDAE 计算后的重构图像,右侧为二者差补后的缺陷图像。

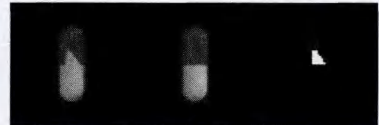


图 5 输入图像、重构图像与检测结果

部分测试图像的实验结果如图 6 所示。其中无缺陷图像的检测结果为第 1、5、7、12 和 16 幅图像(横向计数),其余为缺陷图像。需要说明的是,显示图像二值化的阈值是根据直方图动态生成的。整个实验训练时间如表 1 所列,同时给出了在同样方差($\sigma < 0.2$)情况下 BP 算法的训练时间。200 幅测试图像的所需时间共 1.803s,平均每幅检测时间为 9ms。

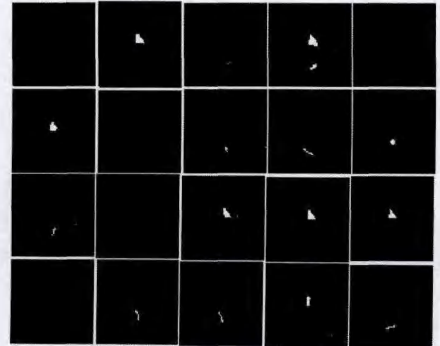


图 6 测试图像缺陷检测结果

表1 SDAE 学习时间及与 BP 算法对比

学习方式	深度学习		BP 学习
	逐层学习	BP 微调	
循环次数	20	10	119
学习时间(s)	48.4	93.8	1118

实验结果分析:

(1)本算法对文中所提 3 种缺陷都有很好的检测能力,具有很好的适用性;

(2)对未参与学习的 200 个测试样本集,本算法具有 100% 的缺陷检出率;

(3)虽然由于条件的限制,只能得到相对于网络规模较小的样本库,但由于本实验只要求得到有限缺陷样本到无缺陷模板之间的映射关系,而无需知道缺陷样本的流形分布,也不需对不同缺陷进行分类,因此仍能得到很好的实验效果;

(4)本算法在检测缺陷的同时,能给出缺陷的物理位置;

(5)本文所建立的神经网络运用了 SDAE 学习算法,与 BP 算法相比大大缩短了学习时间。

传统的胶囊缺陷检测往往首先对待测图像进行灰度归一化、形态学和中值滤波处理,得到缺陷图像,再通过图像增强、边缘检测等将缺陷图像分割提取出来,然后选取相应的特征参数如周长、面积、中心矩等作为判别缺陷类型的依据。文献[15]提出的一种基于极端学习机的胶囊缺陷检测与识别的方法就是按照上述步骤处理的。它在图像灰度归一化、图像增强的基础上,提取胶囊的偏心率和圆形度作为特征变量,用极端学习机实现了胶囊缺陷的检测与分类。在相同的输入图像分辨率及实验环境的情况下,本文算法与文献[15]所提算法相比,具有如下特点:

(1)本算法不需根据缺陷类型设定不同的图像处理方法和特征参数,具有广泛的适用性;

(2)文献[15]只有 90% 的检出率,与本算法相比,存在很大的差距;

(3)文献[15]的方法需要大量的图像预处理过程,包括去噪、滤波等,流程繁琐复杂,而本算法不需要图像处理工作,可直接得到缺陷图像。

结束语 SDAE 技术是一种深度学习方法,它在输入样本中人为地加入噪声,通过运用降噪的方法,实现样本特征的提取鲁棒性和稳定性。本文把 SDAE 技术应用于胶囊图像的缺陷检测中,通过对缺陷样本的学习,建立了带噪图像与纯净模板之间的映射关系,然后对比原始图像和重构图像,实现了胶囊的缺陷检测。本方法不需要复杂的图像处理,也无需考虑缺陷的类别,具有广泛的适用性。但本方法也存在一些不足,只能处理 64×64 维的图像,如何处理高精度的缺陷图像将是本方法今后研究的主要问题。

参考文献

- [1] Hinton G E, Salakhutdinov R R. Reducing the dimensionality of data with neural networks[J]. Science, 2006, 313(28): 504-507
- [2] Hinton G E, Osindero S, Teh Y-W. A fast learning algorithm for deep belief nets[J]. Neural Computation, 2006, 18(7): 1527-1554
- [3] Bengio Y. Learning deep architectures for AI[J]. Foundations and Trends in Machine Learning, 2009, 2(1): 1-127
- [4] Dong Yu, Li Deng. Deep learning and its applications to signal and information processing[J]. IEEE Signal Processing Magazine, 2011, 28(1): 145-154
- [5] Karnowski T P, Arel I, et al. Deep spatiotemporal feature learning with application to image classification[C]//Proceedings of the Ninth International Conference on Machine Learning and Applications. Washington, USA, 2010: 883-888
- [6] Schulz H, Behnke S. Deep learning: layer-wise learning of feature hierarchies[J]. Künstl Intell, 2012, 26: 357-363
- [7] Vincent P, Larochelle H, Lajoie I, et al. Stacked denoising autoencoders: learning useful representations in a deep network with a local denoising criterion [J]. Journal of Machine Learning Research, 2010, 11(12): 3371-3408
- [8] Kang Y, Lee K T, Eun J, et al. Stacked Denoising Autoencoders for Face Pose Normalization [C]//20th International Conference on Neural Information Processing. Berlin: Springer Press, 2013: 241-248
- [9] Zhang Ying, Liu Rui, Zhang Sai-zheng, et al. Occlusion-Robust Face Recognition Using Iterative Stacked Denoising Autoencoder[C]//20th International Conference on Neural Information Processing. Berlin: Springer Press, 2013: 352-359
- [10] Sang Ruo-xin, Jin Pei-quan, Wan Shou-hong. Discriminative Feature Learning for Action Recognition Using a Stacked Denoising Autoencoder[M]//Intelligent Data Analysis and Its Applications, Volume 1. 2014: 521-531
- [11] Agostinelli F, Anderson M R, et al. Adaptive multi-column deep neural networks with application to robust image denoising[C]//Proceedings of the Advances in Neural Information Processing Systems 26 (NIPS'12). Cambridge: MIT Press, 2013: 1-9
- [12] Wei Lin, Hu Rong-qiang. The Automatic Inspection System of Capsule Product Based on BP Neural Network[J]. Application Research of Computers, 2002, 19(4): 52-53 (in Chinese)
- 韦琳, 胡荣强. 基于 BP 网络的胶囊药片自动检测系统[J]. 计算机应用研究, 2002, 19(4): 52-53
- [13] Zuo Qi, Shi Zhong-ke. General design for capsule integrity detection system based on machine vision[J]. Journal of Xi'an Jiaotong University, 2002, 36(12): 1262-1265 (in Chinese)
- 左奇, 史忠科. 基于机器视觉的胶囊完整性检测系统研究[J]. 西安交通大学学报, 2002, 36(12): 1262-1265
- [14] Feng Shan-shan, Chen Shu-yue. Research on identification method for capsule-grain of real and false medicine based on image analyzing[J]. Transducer and Micro system Technologies, 2008, 27(8): 54-56 (in Chinese)
- 冯姗姗, 陈树越. 基于图像分析的真假药胶囊颗粒识别方法研究[J]. 传感器与微系统, 2008, 27(8): 54-56
- [15] Lai Da-hu, Huang Yan-wei. Inspection for defected capsules based on extreme learning machine[J]. Journal of Fuzhou University (Natural Science Edition), 2012, 40(4): 489-494 (in Chinese)
- 赖大虎, 黄宴委. 基于极端学习机的胶囊缺陷检测[J]. 福州大学学报(自然科学版), 2012, 40(4): 489-494