

# 基于自编码算法的深度学习综述<sup>①</sup>

崔广新, 李殿奎

(佳木斯大学 信息电子技术学院, 佳木斯 154007)

通讯作者: 李殿奎, E-mail: datutu2002@sina.com

**摘要:** 深度学习是机器学习的一个分支, 开创了神经网络发展的新纪元. 自编码算法作为深度学习结构的重要组成部分, 在无监督学习及非线性特征提取过程中起到了至关重要的作用. 首先介绍自编码算法的基本概念及原理, 然后介绍基于自编码算法的改进算法, 最后列举了自编码算法在若干领域应用的知名案例和发展趋势.

**关键词:** 机器学习; 深度学习; 自编码算法; 无监督学习; 神经网络

引用格式: 崔广新, 李殿奎. 基于自编码算法的深度学习综述. 计算机系统应用, 2018, 27(9): 47-51. <http://www.c-s-a.org.cn/1003-3254/6542.html>

## Overview on Deep Learning Based on Automatic Encoder Algorithms

CUI Guang-Xin, LI Dian-Kui

(College of Information Science and Electronic Technology, Jiamusi University, Jiamusi 154007, China)

**Abstract:** Deep learning is a branch of machine learning, creating a new era in the development of neural networks. As an important part of deep learning structure, self-coding algorithm plays a crucial role in unsupervised learning and nonlinear feature extraction. Firstly, the basic concepts and principles of self-encoding algorithm are introduced. Then, the improved algorithm based on self-encoding algorithm is presented. Finally, the well-known cases and development trends of self-encoding algorithm applied in several fields are elaborated.

**Key words:** machine learning; deep learning; automatic encoder; unsupervised learning; neural networks

随着大数据时代的到来, 互联网的海量资源比较容易获得. 但是, 由于“维数灾难”的影响, 就使得高维数据处理变得十分困难, 基于深度学习的自编码算法是数据降维的一个新的热点. 自编码算法横跨计算机科学、数学和工程技术等其他多个学科, 并且它广泛的应用于经济、政治、天文、地理等社会生活领域中. 自编码算法的最大优点是模型的表达能力强, 它可以很好的处理具有高维稀疏特征的数据, 通过无监督学习可以减少无关和冗余数据, 从而降低维数.

### 1 自编码算法

自动编码器是 Rumelhart 于 1986 年提出的, 是一

个典型的三层神经网络结构, 包括输入层、隐藏层和输出层<sup>[1]</sup>.

设输入层和输出层的维度为  $n$ , 隐藏层的维度为  $m$ , 样本数据集  $S = \{x^{(i)}\}_{i=1}^N$ .

设编码函数为  $f$ , 编码过程指的是从输入层到隐藏层; 设解码函数为  $g$ , 解码过程指的是从隐藏层到输出层.

自动编码器的编码部分是利用非线性映射函数将输入数据映射到隐藏层单元, 设  $h$  表示隐藏层的神经单元激活, 则其数学表达式为:

$$h = f(x) = s_f(wx + p) \quad (1)$$

其中,  $w$  表示的是连接输入层和隐藏层之间的权值矩阵,  $s_f$  表示的是编码器激活函数, 通常取 Sigmoid 函数,

① 基金项目: 2017 年黑龙江省教育科研专项 (2017-0001)

Foundation item: Specific Foundation for Education and Scientific Research of Heilongjiang Province in 2017 (2017-0001)

收稿时间: 2018-01-09; 修改时间: 2018-01-31, 2018-02-28; 采用时间: 2018-03-08; csa 在线出版时间: 2018-07-26

即  $f(x) = 1/(1 + e^{-x})$ .

解码部分计算原理与编码部分类似, 利用编码得到的隐藏层重构出原始输入数据. 其解码函数的数学表示如下:

$$y = g(h) = s_g(\tilde{w}h + q) \quad (2)$$

其中,  $y$  表示的是解码器对输入数据的重构,  $s_g$  表示的是解码器激活函数, 通常取 Sigmoid 函数或者恒等函数,  $w$  表示的是隐藏层与输出层之间的权值矩阵, 并且, 自动编码器的参数  $\theta = \{w, p, q\}$ .

输出数据  $y$  可以看作是对输入数据  $x$  的预测, 如果输出数据  $y$  与输入数据  $x$  的接近程度是我们接受的, 那么该自动编码器就保留了原始数据的大部分信息, 该自动编码器也就训练好了<sup>[2]</sup>.

用重构误差函数  $L(x, y)$  来刻画  $y$  与  $x$  的接近程度.

当  $s_g$  为恒等函数时:

$$L(x, y) = \|x - y\|^2 \quad (3)$$

当  $s_g$  为 Sigmoid 函数时:

$$L(x, y) = - \sum_{i=1}^n [x_i \log y_i + (1 - x_i) \log(1 - y_i)] \quad (4)$$

当训练样本集为  $S = \{X^{(i)}\}_{i=1}^N$  时, 自动编码器的整体损失函数为:

$$J_{AE}(\theta) = \sum_{x \in S} L(x, g(f(x))) \quad (5)$$

最后重复使用梯度下降算法迭代计算  $J_{AE}(\theta)$  的最小值, 就可以求解自编码网络的参数  $\theta$ , 也就完成了自动编码器的训练.

## 2 基于自编码算法的改进算法

### 2.1 稀疏自编码算法

稀疏自编码算法与传统自编码算法降维特性不同, 稀疏自编码算法试图学习过完备数据, 通过适当的稀疏正则化可以达到过完备的效果<sup>[3]</sup>.

一般采用稀疏性限制确定神经元是否被激活, 如果神经元输出接近 1, 则说明被激活. 对于 Sigmoid 函数, 如果输出接近 0 则认为神经元被抑制; 对于 tanh 函数, 输出接近 -1 则认为被抑制, 通过使大量神经元处于被抑制的条件限制下从而使自动编码器达到稀疏的目的<sup>[4]</sup>.

### 2.2 降噪自编码算法

降噪自编码算法 (Denoising Automatic Encoder,

DAE) 是自编码算法的一个变种, 其实就是在自编码算法上加入了噪声鲁棒性约束.

降噪自编码算法的过程是: 首先将原始输入数据  $x$  随机的添加噪声干扰, 就会得到受干扰的输入数据  $\tilde{x}$ . 自动编码器通过编码函数  $f$  将  $\tilde{x}$  映射到隐藏层, 从而就会获得隐藏层特征表达式  $h$ , 然后自动编码器通过解码函数  $g$  将隐藏层的特征表达式  $h$  再映射到输出层, 就会获得输入数据的重构  $y$ . 如果最终得到的重构数据  $y$  能够很好地保持原始输入数据  $x$  的信息, 那么我们就可以说这个网络具有很好的鲁棒性.

降噪自编码算法最重要的就在于为输入数据添加干扰噪声<sup>[5]</sup>. 最常见的方法有两种, 一种是加入 additive isotropic Gaussian noise, 令  $\tilde{x} = x + \varepsilon$ ,  $\varepsilon \sim N(0, \sigma^2 I)$ , 即噪声干扰部分  $\varepsilon$  服从标准方差为  $\sigma^2$  的标准正态分布. 另外一种方法是 binary masking noise, 该方法只需要以概率  $P(A)$  将输入向量  $x$  的部分分量赋值为 0, 以  $1 - P(A)$  的概率分量保持不变, 这样就做到了将输入向量  $x$  添加干扰变成了干扰向量  $\tilde{x}$ , 参数  $\sigma$  和  $P(A)$  都是用来控制噪声强度的.

### 2.3 收缩自编码算法

收缩自编码算法 (Contractive Automatic Encoder, CAE) 是自编码算法的一个变种, 其实就是在自编码算法上加入了一个规则项. 当训练样本的输入存在微小扰动时, CAE 被提出来保持中间层表示的鲁棒性, 它主要通过惩罚输入数据的敏感性来保持中间层表示的鲁棒性<sup>[6]</sup>. 该惩罚因子是自动编码器编码部分的 Jacobian 矩阵的 Frobenius 范数.

从几何角度来说, 特征表示的鲁棒性可以看作是投影到特征空间的输入数据空间, 尤其是数据生成分布的样本领域空间的压缩. 从解析的角度来说, 惩罚因子的限制提高了 CAE 对于输入数据中的微小扰动的鲁棒性, 数据表示的准确性较高, 且重构特征并不受惩罚因子的影响. 可以平衡误差代价函数的重构误差与鲁棒性.

### 2.4 栈式自编码算法

栈式自编码是一种由多层自编码构成的神经网络模型. 它通常能得到输入的“层次型分组”或者“部分-整体分解”结构. 栈式自编码神经网络的第一层会学习到原始输入的一阶特征 (像图片里的边缘), 第二层会学习到二阶特征, 该特征对应一阶特征里包含的一些模式 (像在构成轮廓或者角点时, 什么样的边缘会出现),

更高层还会学到更高阶的特征,但耗时较长。

该深度学习模型的训练过程为:首先,网络中的每一层依次通过逐层贪婪训练法来训练整个深度神经网络,即原始输入用于训练第一层网络参数,然后把该隐藏层的输出作为下一层的输入,如此反复,依次逐层训练学习各层网络的参数。在无监督预训练之后,我们在最后一个自动编码器的编码层上添加一个分类器,然后使用监督学习通过有标记样本来调整分类器的参数。此外,多层神经网络的监督学习也可以用来微调所有层的参数,同时改善结果。

### 2.5 边缘降噪自编码算法

边缘降噪自编码算法是 Chen 为克服降噪自编码算法计算强度高、处理时间长等缺点提出的一种衍生算法<sup>[7]</sup>,但控制噪声强度的参数不易确定。

### 2.6 稀疏降噪自编码算法

通过研究传统自编码算法,提出了一种将稀疏性限制加在降噪自动编码器的隐藏神经元上的方法,可以让较少隐藏层的激活单元表示最有效的特征,但没有考虑输入与输出的局部近似性。

### 2.7 稀疏边缘降噪自编码算法

稀疏边缘降噪自编码算法 (Sparse Margin Denoising Auto-Encoder, SMDAE) 是自编码算法的一个变种,其实就是将稀疏性约束条件表达式和边缘降噪约束条件表达式添加在隐藏层上面。同时结合了稀疏自编码算法和边缘降噪自编码算法的特点,噪声干扰分布和重构误差函数还可以继续改进。

### 2.8 卷积自编码算法

卷积自编码算法 (Convolutional Auto-Encoder) 是在自编码算法基础上,将普通的矩阵内积操作替换为卷积操作。这种操作更适合处理二维输入数据,但耗时较长。

### 2.9 卷积稀疏自编码算法

在自编码算法中加入卷积与反卷积操作构成 CAE,再加入稀疏化操作,构成卷积稀疏自编码器。使用 CSAE-S 对输入数据进行预处理,尽可能实现初步无损特征提取,然后使用 CSAE-V 对数据进行降维,实现低级特征抽象,再加入 Pooling 层实现高级特征抽象以及提高网络泛化性。将 CSAE 中的卷积编码部分按 CSAE-S、CSAE-V、Pooling 顺序栈式堆叠,最后加入 Full 连接层和 Softmax 分类层,则形成 CSAENN。

整个 CSAENN 的耗时主要在于权值稀疏化部分,

因此整个网络的时间复杂度达到了  $O(n^3)$ 。CSAENN 的另外一个优点就是准确率有一定的提高。经过稀疏化后的权值提取出的特征相似性更少,即代表性更强,因此可以更好地表示原图像。增加迭代次数、增加样本数量可获得更高的准确率。

## 3 自编码算法的应用案例

### 3.1 图像识别领域

利用人体生物学特征已成为获取相关信息的重要途径,徐文瀚等人提出一种运用栈式自编码网络进行面部年龄识别的方法<sup>[8]</sup>;针对传统人脸表情识别算法鲁棒性差,李江等人提出一种降噪自编码神经网络的方法<sup>[9]</sup>;余涛将稀疏自编码应用在手写字的识别中,与传统人工神经网络相比,参数少,对图像的平移具有一定的容忍度,不易造成过拟合<sup>[10]</sup>;针对网络上大量重复视频的问题,王飞飞将稀疏自编码用于近重复视频检索<sup>[11]</sup>,取得了较好的检索精度;王勇将稀疏自编码应用在森林火灾的图像分类中<sup>[12]</sup>,该方法直接从图像块的像素中提取图像数据的特征,提高了森林火灾与类似物体的分类准确率。

### 3.2 语音识别领域

马勇等人提出了一种基于稀疏神经网络的说话人分割方法<sup>[13]</sup>,为了避免网络在训练中出现过拟合问题,还采用了 dropout 技术提高网络的鲁棒性;针对中文语音情感识别的问题,朱芳枚同时结合了降噪自编码算法和稀疏自编码算法的优点,提取了具有鲁棒性和稀疏性的语音情感二次特征<sup>[14]</sup>;针对在语音识别中,传统径向基随机初始化的问题,黄丽霞等人将深度自编码应用在语音识别噪声鲁棒性的研究中<sup>[15]</sup>;王辉等人采用稀疏自编码算法来提取模仿听觉神经的稀疏触动信号,将其作为表征语音信号<sup>[16]</sup>;在声学模型方面,也不断地尝试新的深度神经网络,王山海利用深度自编码算法研究出它的模型参数对于孤立词语音识别产生的影响<sup>[17]</sup>。

### 3.3 故障诊断领域

针对浅层网络特征提取能力有限的问题,陈仁祥等人提出采用深度自编码神经网络的方法对轴承进行识别<sup>[18]</sup>;温博文将深度自编码算法应用于飞机操纵面故障诊断当中<sup>[19]</sup>,该方法无需精准的飞机模型,诊断速度快、精度高;为了实现风机齿轮箱的故障诊断检测分析,刘辉海提出了一种深度自编码网络模型<sup>[20]</sup>;针对



在复杂纹理背景噪声下传统路面裂缝检测系统检测效率低,易造成漏检、错检等问题,钱彬等人提出了一种运用稀疏自编码进行裂缝自动检测的方法<sup>[21]</sup>;朱煜奇将栈式降噪自编码应用在故障诊断中,有效的克服了局部极值与梯度消失的问题<sup>[22]</sup>。

3.4 其它领域

袁静等人将稀疏去噪自编码应用在异常行为检测中,该方法比传统的自编码网络在全局异常行为检测中更加有效<sup>[23]</sup>;针对在入侵检测中 SVM 受时间与空间复杂度的约束,高妮等人提出一种自编码神经网络应用于此<sup>[24]</sup>;杨辉华等人采用稀疏降噪自编码对真假药

品近红外光谱数据进行建模,可以做到真实有效的鉴别药品的真伪性<sup>[25]</sup>;针对常规毒理学实验方法的周期长、耗资高等问题,黎红等人提出采用降噪自编码提取隐含毒性化合物的特征,使得有毒化合物的化合物毒性预测和活性预测都能被实现;魏扬威将自动编码器应用于英语作文自动评分系统,相比于传统自动评分方法皮尔森相关系数高出 9.7%,具有非常好的应用价值<sup>[26]</sup>。

4 结论与展望

对各算法作对比分析,如表 1。

表 1 各算法对比

年份	模型	构成	优点	缺点
1986	自编码	典型三层	无监督学习	训练时间长
2007	稀疏自编码	在 AE 隐藏层神经元加稀疏性	降维,学习稀疏的特征表达	对随机数据,压缩学习会很困难
2007	降噪自编码	AE 上加噪声鲁棒性	降低噪声,提高鲁棒性	计算度高、处理时间长
2010	收缩自编码	AE 上加一个规则项	平衡重构误差、鲁棒性	处理时间较长
2009	栈式自编码	多层自编码	强大的非线性拟合能力	耗时较长
2014	边缘降噪自编码	对 DAE 添加的噪声做边缘化处理	减少重构误差,降低计算开销	MDAE 的控制噪声强度的参数
2016	稀疏降噪自编码	对 DAE 的隐藏神经元加入稀疏性限制	较少隐藏层的激活单元表示最有效的特征	没有考虑输入与输出的局部近似性
2016	稀疏边缘降噪自编码	结合 SAE、MDAE 的约束条件	好的噪声鲁棒性;好的稀疏性约束	重构误差函数还可以继续改进
2011	卷积自编码	AE 引入卷积操作	可直接处理二维图像	耗时较长
2017	卷积稀疏自编码	AE 中加入卷积与反卷积,再加稀疏化	耗时少,较少隐藏层的激活单元表示最有效的特征	增加迭代次数、增加样本数量等获得更高的测试准确率

自编码的最大优点是模型的表达能力强,它可以很好的处理具有高维稀疏特征的数据,通过无监督学习可以减少无关和冗余数据,从而降低维数。虽然自编码算法已经在很多领域都取得了令人瞩目的成绩,但其仍处于起步阶段,还面临诸多问题:像如何判定网络的隐藏层层数和隐藏层节点数;如何改进梯度下降法,使得网络可以获得更好的局部极值点,甚至是可以获得全局最优点;如何评价通过自动编码器所学到的特征是优劣的,并且如何解释该特征;如何处理大数据的大样本量,高维属性和多样性的数据类型等问题。

未来自编码的发展趋势:在自然语言处理领域,期待利用 AE 模型的更多新算法,可以更好地理解自然语言中的句子或整个文档;在信息检索领域仍有待突破,期待涌现新的深度结构及其算法;在计算机视觉领域,期待更多的深层结构和算法能够更好地处理目标和行为识别这类复杂的问题。

参考文献

1 Rumelhart DE, Hinton GE, Williams RJ. Learning representations by back-propagating errors. *Nature*, 1986, 323(9): 533–536.

2 Baldi P, Guyon G, Dror V, *et al.* Autoencoders, unsupervised learning, and deep architectures. *Journal of Machine Learning Research*, 2012, 27: 37–50.

3 Luo Y, Wan Y. A novel efficient method for training sparse auto-encoders. 2013 6th International Congress on Image and Signal Processing (CISP 2013). Hangzhou, China. 2013. 1019–1023. [doi: 10.1109/CISP.2013.6745205]

4 史彩娟, 阮秋琦. 基于增强稀疏性特征选择的网络图像标注. *软件学报*, 2015, 26(7): 1800–1811. [doi: 10.13328/j.cnki.jos.004687]

5 Vincent P, Larochelle H, Bengio Y, *et al.* Extracting and composing robust features with denoising autoencoders. *Proceedings of the 25th International Conference on Machine Learning*. ACM. New York, NY, USA. 2008. 1096–1103. [doi: 10.1145/1390156.1390294]

- 6 Rifai S, Vincent P, Muller X, *et al.* Contractive auto-encoders: Explicit invariance during feature extraction. *Proceeding ICML'11 Proceedings of the 28th International Conference on International Conference on Machine Learning*. Bellevue, WA, USA. 2011. 833-840.
- 7 Chen M, Weinberger K, Sha F, *et al.* Marginalized denoising auto-encoders for nonlinear representations. *Proceedings of the 31st International Conference on Machine Learning*. Beijing, China. 2014. 1476-1484.
- 8 徐文瀚, 程石磊. 基于深度栈式自编码网络的面部年龄识别. *信息通信*, 2017, (11): 45-47. [doi: [10.3969/j.issn.1673-1131.2017.11.020](https://doi.org/10.3969/j.issn.1673-1131.2017.11.020)]
- 9 李江, 冉君军, 张克非. 一种基于降噪自编码器的人脸表情识别方法. *计算机应用研究*, 2016, 33(12): 3843-3846. [doi: [10.3969/j.issn.1001-3695.2016.12.072](https://doi.org/10.3969/j.issn.1001-3695.2016.12.072)]
- 10 余涛. 基于稀疏自编码器的手写体数字识别. *数字技术与应用*, 2017, (1): 66-68.
- 11 王飞飞. 基于稀疏自动编码器的近重复视频检索. *电子技术与软件工程*, 2017, (3): 194-196.
- 12 王勇, 赵俭辉, 章登义, 等. 基于稀疏自编码深度神经网络的林火图像分类. *计算机工程与应用*, 2014, 50(24): 173-177. [doi: [10.3778/j.issn.1002-8331.1401-0392](https://doi.org/10.3778/j.issn.1002-8331.1401-0392)]
- 13 马勇, 鲍长春. 基于稀疏神经网络的说话人分割. *北京工业大学学报*, 2015, 41(5): 662-667.
- 14 朱芳枚, 赵力, 梁瑞宇, 等. 面向中文语音情感识别的改进栈式自编码结构. *东南大学学报*, 2017, 47(4): 631-636.
- 15 黄丽霞, 王亚楠, 张雪英, 等. 基于深度自编码网络语音识别噪声鲁棒性研究. *计算机工程与应用*, 2017, 53(13): 49-54. [doi: [10.3778/j.issn.1002-8331.1611-0217](https://doi.org/10.3778/j.issn.1002-8331.1611-0217)]
- 16 王辉, 赵悦, 刘晓凤, 等. 基于深度特征学习的藏语语音识别. *东北师大学报*, 2015, 47(4): 69-73.
- 17 王山海, 景新幸, 杨海燕. 基于深度学习神经网络的孤立词语音识别的研究. *计算机应用研究*, 2015, 32(8): 2289-2291, 2298. [doi: [10.3969/j.issn.1001-3695.2015.08.011](https://doi.org/10.3969/j.issn.1001-3695.2015.08.011)]
- 18 陈仁祥, 黄鑫, 杨黎霞, 等. 加噪样本扩展深度稀疏自编码神经网络的滚动轴承寿命阶段识别. *振动工程学报*, 2017, 30(5): 874-882.
- 19 温博文, 董文瀚, 解武杰, 等. 深度自编码观测器飞机操纵面快速故障诊断. *飞行力学*, 2016, 34(6): 34-38, 44.
- 20 刘辉海, 赵星宇, 赵洪山, 等. 基于深度自编码网络模型的风电机组齿轮箱故障检测. *电工技术学报*, 2017, 32(17): 156-163.
- 21 钱彬, 唐振民, 徐威. 基于稀疏自编码的路面裂缝检测. *北京理工大学学报*, 2015, 35(8): 800-804, 809.
- 22 朱煜奇, 黄双喜, 杨天祺, 等. 基于栈式降噪自编码的故障诊断. *制造业自动化*, 2017, 39(3): 152-156.
- 23 袁静, 章毓晋. 融合梯度差信息的稀疏去噪自编码网络在异常行为检测中的应用. *自动化学报*, 2017, 43(4): 604-610.
- 24 高妮, 高岭, 贺毅岳, 等. 基于自编码网络特征降维的轻量级入侵检测模型. *电子学报*, 2017, 45(3): 730-739. [doi: [10.3969/j.issn.0372-2112.2017.03.033](https://doi.org/10.3969/j.issn.0372-2112.2017.03.033)]
- 25 杨辉华, 雒志超, 蒋淑洁, 等. 稀疏降噪自编码算法用于近红外光谱鉴别真假药的研究. *光谱学与光谱分析*, 2016, 36(9): 2774-2779.
- 26 魏扬威, 黄萱菁. 结合语言学特征和自编码器的英语作文自动评分. *计算机系统应用*, 2017, 26(1): 1-8. [doi: [10.15888/j.cnki.csa.005535](https://doi.org/10.15888/j.cnki.csa.005535)]