

深度置信网络模型及应用研究综述

刘方园, 王水花, 张煜东

LIU Fangyuan, WANG Shuihua, ZHANG Yudong

南京师范大学 计算机科学与技术学院, 南京 210023

School of Computer Science and Technology, Nanjing Normal University, Nanjing 210023, China

LIU Fangyuan, WANG Shuihua, ZHANG Yudong. Survey on deep belief network model and its applications. *Computer Engineering and Applications*, 2018, 54(1): 11-18.

Abstract: This paper firstly introduces the development of Deep Belief Network (DBN) based on theory foundation. Afterwards, the difference between deep network structure and shallow network structure is analyzed. Finally, the literature makes a study and analysis of DBN, in the field of text detection, facial and expression recognition, and remote sensing image classification by quoting multiple representative documents. Through a comprehensive introduction to the deep learning model DBN and deeply understanding the construction and practical application of DBN, it provides researchers with the idea of improving DBN and applying it to a wider emerging field in the future.

Key words: Deep Belief Network(DBN); text detection; facial and expression recognition; remote sensing image field

摘 要: 介绍深度置信网络(DBN)理论基础的发展, 对比分析深层结构DBN与浅层网络结构的差异, 最后引用多篇文献分析研究DBN在文字检测、人脸及表情识别领域和遥感图像领域的应用效果。全面介绍了深度学习模型DBN, 深入分析DBN的构建与实际应用, 为研究人员提供改进DBN的思路, 以期在未来将其运用到更宽广的新兴领域中。

关键词: 深度置信网络; 文字检测; 人脸及表情识别; 遥感图像领域

文献标志码: A **中图分类号:** TP391 **doi:** 10.3778/j.issn.1002-8331.1711-0028

1 引言

现今, 深度学习在计算机视觉、模式识别、人工智能、图像处理以及数据挖掘等领域掀起的研究热潮引人深思, 研究人员提出的自编码器、卷积神经网络、循环神经网络以及深度置信网络等深度学习模型越来越值得深入探索和挖掘。本文以深度置信网络为例, 从理论到实际, 由浅入深地分析其应用意义。

深度置信网络(DBN)发展伊始, 为解决快速、自动学习特征的研究问题, Hinton等^[1]提出一种应用于DBN的快速学习算法, 提供了理解分析DBN模型的新途径, 打开了深度学习模型扩展的大门, 并将其应用于实际生活的各个领域。

本文针对目前流行的应用领域——文字检测、人脸及情感识别、遥感图像地物分类, 对DBN进行详细分析。在上述应用领域中, DBN改进的相同之处, 一是通

过预处理方式(特征融合、去噪)提高输入样本的质量; 二是利用各种优化策略(提前终止、Dropout)对DBN模型参数, 进行深度改进或选取不同的分类器。DBN改进的不同之处在于, 文字检测领域中, DBN结合计算设备架构、文字笔画宽度; 人脸及情感领域中, DBN结合人体识别所需的关键特征(局部纹理特征); 遥感图像地物分类领域中, DBN结合高光谱图像包含的光谱信息和空间信息。

2 深度置信网络简介

深度置信网络(DBN)是根据生物神经网络的研究及浅层神经网络发展而来的, 为概率生成模型, 通过联合概率分布推断出数据样本分布。DBN生成模型通过训练网络结构中的神经元间的权重使得整个神经网络依据最大概率生成训练数据, 形成高层抽象特征, 提升

基金项目: 国家自然科学基金(No.61602250, No.61503188); 江苏省自然科学基金(No.BK20150983, No.BK20150982); 江苏省高校自然科学研究面上项目(No.16KJB520025, No.15KJB470010)。

作者简介: 刘方园(1993—), 女, 博士生, CCF会员, 主要研究方向: 计算机视觉; 王水花(1985—), 女, 博士, 讲师; 张煜东(1985—), 通讯作者, 男, 教授、博导, 主要研究方向: 人工智能与医学图像处理, E-mail: yudongzhang@ieee.org。

收稿日期: 2017-10-30 **修回日期:** 2017-12-05 **文章编号:** 1002-8331(2018)01-0011-08

模型分类性能。

2.1 DBN模型的构建——RBM

首先, DBN是基于人工神经网络发展而来, 由两部分神经元组成: 显性神经元(显元)——用于接收输入数据; 隐性神经元(隐元)——用于高层特征提取。DBN的关键组成元件是受限玻尔兹曼机(RBM), 通过将多层RBM组合并结合最终分类器对输入数据进行检测、识别以及分类。

RBM组成结构中含有两层神经元(显元、隐元), 且每一层可用一个向量表示, 向量的维数由每层神经元的个数决定, 具体结构图如图1所示。

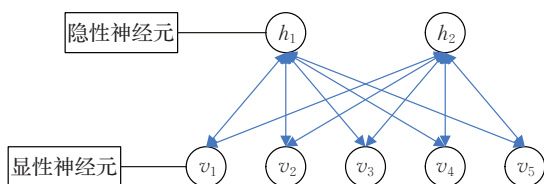


图1 RBM结构图(v 代表显元, h 代表隐元)

从图1可明显发现, RBM结构图中层内的神经元之间无连接, 层间的神经元之间双向连接。该结构保证层内神经元无互连的条件独立性, 即在给定显元的取值时所对应的隐元的取值是互不相关的, 同样在给定隐元值时显元也保留该特性。

DBN的组成元件RBM需通过训练优化特征提取能力, 其目的是求得一个最接近训练样本的联合概率分布从而能够更准确、抽象地提取或者还原特征, 即求得决定训练样本最大概率产生分布的影响因素——权值。训练RBM的过程简单来说就是寻找可视层节点和隐藏层节点之间连接的最优权值, 以经典的学习算法——对比散度(CD)算法为例更新权值, 基于对比散度算法的RBM权重更新步骤如下。

步骤1 对训练样本集进行采样, 假设每一采样记录记为 X 。

步骤2 将 X 输入到可视层 $V^{(0)}$, 计算该记录使得隐元开启的概率如公式(1):

$$P(h_j^{(0)} = 1 | v^{(0)}) = \sigma(W_j v^{(0)}) \quad (1)$$

步骤3 重构显层, 从上述所计算的概率分布中抽取隐藏层的一个样本如公式(2):

$$h^{(0)} \sim P(h^{(0)} | v^{(0)}) \quad (2)$$

步骤4 计算隐元激活概率, 基于显层中抽取样本并运用重构后显层神经元进行计算如公式(3):

$$P(v_i^{(1)} = 1 | h^{(0)}) = \sigma(W_i^T h^{(0)}), v^{(1)} \sim P(v^{(1)} | h^{(0)}) \quad (3)$$

步骤5 最终依据隐藏层神经元和显层神经元之间的相关性差异来更新权重 W , 更新如公式(4):

$$P(h_j^{(1)} = 1 | v^{(1)}) = \sigma(W_j v^{(1)}),$$

$$W \leftarrow W + \lambda (P(h^{(0)} = 1 | v^{(0)}) v^{(0)T} -$$

$$P(h^{(1)} = 1 | v^{(1)}) v^{(1)T}) \quad (4)$$

公式中, v 代表显元, h 代表隐元, m 、 n 代表显元和隐元的个数, 公式中的上标代表取样步骤, $\langle v^{(0)} h_j^{(0)} \rangle$ 代表基于显层和隐藏层的第一次取样, W 代表层间连接权重, 其训练好之后可确定输入显层的一条新记录所对应隐元的状态。

2.2 DBN模型的构建——整体

本文以两层受限玻尔兹曼机和Softmax分类层为例, 构建最终能够实现实际应用的DBN模型, 模型构建如图2所示。 v 代表显层神经元, h 代表隐藏层神经元, y 代表标签已知的样本, o 代表分类结果输出。DBN模型构建在于: 固定第一个训练好的RBM的权重和偏置, 将其隐元所处的状态作为第二个RBM的输入, 对第二个RBM进行训练后堆叠于第一个RBM上, 对于多层RBM重复上述过程。若训练数据集中含有带标签的样本, 在第二个RBM训练时需加入, 最终采用Softmax对数据进行分类。

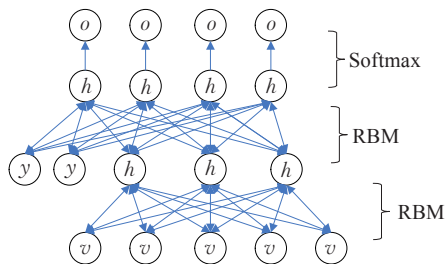


图2 DBN模型结构简图

DBN的微调步骤也尤其重要, 由于先前构建的每一层RBM都只能确保自身层内的权值对该层的特征映射提取达到最优, 为保证整体结果的最优性, 常见的是在分类前设置BP网络, 监督训练学习, 两者结合保证参数不易陷入局部最优, 且在一定程度上弥补训练时间长的劣势。

3 深度置信网络应用分析

一个模型的提出其实就是为了实现它在实际生活的各个领域中的作用, 深度置信网络(DBN)更加抽象地学习高层特征, 并且适用于高维复杂的海量数据的特性。本章主要分析DBN在文字检测领域、人脸及表情识别领域和遥感图像分类领域中的方法和性能。

3.1 文字检测领域

陆军建等^[2]采用深度置信网络(DBN)与统一计算设备架构(CUDA)相结合的方法, 分类包含字迹相似和字迹模糊的手写字符图片, 所采用的DBN由两层受限玻尔兹曼机和一层反向传播网络构成。实验过程中, 将

DBN-CUDA 与 DBN、SVM、KNN 通过识别时间和准确率的定量分析比较, DBN-CUDA 既能保证识别精度又能大幅度提高识别速率, 其分类准确率达到 96.5%; 其不足在于理论分析 SVM、KNN 与 CUDA 结合的不可取之处, 未在实验中进行充分证明。

Mleczo 等^[3]提出一个基于深度置信网络(DBN)并应用 Pawlak 粗糙集理论的粗糙深度置信网络(RDBN), 经理论分析其适用于不完全文字信息的分类。将 RDBN 识别系统用于分类具有缺失值的手写数字样本中, 同时与传统 DBN 方法进行比较, 实验结果证实随着缺失样本的增加, RDBN 错误识别的样本数目远远低于 DBN。为进一步扩展其应用范围, 可通过改进将其用于具有错误值或包含噪声的不完全模式下。

张董等^[4]分析在手写数字识别研究中深度学习和浅层学习算法的性能, 确定深度置信网络模型的可行性和有效性, 在手写数字 0~9 中, DBN 方法在各类别中识别准确率良好。通过文中待测样本图片的展示, 其实验过程中确定各个待识别数字的位置, 消除对识别过程无贡献的区域, 进一步提高识别效率, 提升幅度为 10% 左右。

Liu^[5]利用一种将深度置信网络(DBN)和支持向量机(SVM)相结合的文本分类方法, 同时从两个方面提高方法的性能: 一是采用不同的微调策略; 二是采用不同的网络结构。利用 DBN 构建的深层架构获取高效的文本信息特征表示, 将所学习的深度特征训练 SVM 使其对所输入文本进行有效分类, 实验结果表明性能良好, 但对于中文文本分类可考虑在所 DBN+SVM 方法中改进特征降维策略以提升性能指标。

翟文洁等^[6]提出一种混合深度置信网络(HDBN)的分类方法并将其用于文本分类和检索中, 通过实验比较验证其有效性。HDBN 模型是基于两层深度玻尔兹曼机(DBM)和两层深度置信网络(DBN)实现的, DBM 在此过程中相当于一个优化过程, 一方面对样本特征进行初降维从而提取更准确的特征信息, 另一方面确定模型的各个最优参数。从实验结果来看其性能达到 98.01%, 但该混合模型必然增加训练的时间。

王贵新等^[7]在中文短信分类问题中采用 word2vec 结合深度置信网络方法, DBN 利用双隐层反向传播(BP)神经网络对模型参数进行微调。DBN+双隐层 BP 对六大类短信进行分析识别, 与传统方法 BP(单隐层)、BP(双隐层)、贝叶斯、SVM 方法相比, 其具备更优的分类性能, 但由于中文短信文本受发送、接收等因素的影响, 其学习速率可尝试通过改进预处理方式进一步提升。

Jiang 等^[8]采用一种混合的深度置信网络模型, 该模型将 DBN 和 Softmax 回归分类器相结合, 并将其应用于稀疏高维矩阵表示的文本数据的分类中。实验过程中, DBN 用于提取文本信息的特征, Softmax 对所学习特征空间中的文本进行回归分类, 其关键在于 Softmax

采用不同的参数优化算法, Softmax(1)基于梯度下降法, Softmax(2)基于 L-BFGS 算法。通过分析实验结果得出, 其提出的方法优于 SVM、KNN, 且 DBN+Softmax(2)分类海量文本数据时准确率最优达到 86.88%, 性能优于小型文本数据。

宋春静等^[9]将基于三层 RBM 和一层 BP 的深度置信网络模型应用于集装箱上的字符识别研究中, 采用两个 DBN 模型进行字符匹配, 这是考虑到部分字母与数字的高度相似性影响识别精度。最终实验结果 DBN 相较于 BP、模板匹配方法准确率高, 但其实验样本为 200 幅未能在较大程度上体现 DBN 方法处理海量数据的优势, 且其检测准确度可通过优化模型权重和偏置等参数提高。

Xu 等^[10]对现实生活中自然场景图像中的文本信息进行识别, 综合考虑文本语言环境因素的影响, 提出笔画宽度变换(SWT)与深度置信网络(DBN)结合的检测方法。基于 SWT 能消除文本字符间与字符内的错误, 基于 DBN 可高效学习字符与非字符的差异性, 两者结合能有效识别不同场景下的文本内容且保证结果的鲁棒性。所提出方法的识别率仅为 75%, 需进一步提高, 可尝试通过优化 DBN 模型的参数和改善网络结构来实现。

周世超等^[11]将由两层受限玻尔兹曼机和三层反向传播网络构成的深度置信网络(DBN)应用于语音文本的识别过程中, 从而对语音通讯的服务质量进行分类。实验结果表明, 所采用的 DBN 方法在各类别中分类效果为 96.1%, 均优于传统分类模型: 朴素贝叶斯、KNN、SVM、BP。该方法在语音服务文本取得良好效果, 但在实际应用中, 其语音文本化的预处理步骤较很大程度上影响了分类准确度。

张庆庆等^[12]通过改进优化深度置信网络模型的结构和参数, 提高 DBN 在文本情感分类检测中的识别率, 达到 83%。实验中采用来自携程网中的酒店的正面和负面评论, 对文本进行向量化预处理后, 利用信息增益选择特征以降低时间成本, 模型结构为 4000-600-300、2000-600-300、1000-300-100, 结果表明第二种结构分类性能相对较优。

Ray 等^[13]研究的关键在于比较分析基于不同训练算法的深度置信网络(DBN)应用在场景文本的识别问题中的精度, 寻找较优的检测方法。将 DBN1(结合对比散列算法)、DBN2(结合去噪分数匹配算法)和 DBN3(结合监督学习算法)进行对比研究, 在相同文本数据集下, DBN1 取得最优检测效果。实验证明了 DBN1 方法在识别和分析文本信息的可行性与有效性, 但需研究如何将计算复杂度降低。

翟继友^[14]利用深度置信网络进行各个文本中语义相关度计算模型的构建, 表明深度学习方法在识别计算文本语义相似度方面具有明显优势。实验分为两部分:

一是改变DBN模型的深度,观察计算结果;二是在相同实验条件下对比分析不同识别方法。文中得出层数增加结果较好,但若层数增加到一定程度不一定保证算法的性能且时间损耗大;同时,将DBN_3与其他算法比较后,各个指标体现了DBN方法的优越性。

Sarikaya等^[15]研究前期改进深度置信网络(DBN)并将其用于自然语言识别检测中。实验从两方向扩展,将DBN学习的特征和原始输入结合使用,利用SVM对其分类;将额外的未标记数据用来改进模型性能。将文中采用SGD微调策略和提前终止优化的DBN方法与传统文本分类算法SVM、Boosting和最大熵进行比较,其分类精度最高达到90.8%,同时有效降低了计算时间。

深度置信网络在文字检测领域中的应用效果分析如表1。文字检测领域中,样本数据主要为手写数字、短信文本、网络评论、场景文本以及语音文本等。表1中,文献[11]采用的传统DBN方法用于识别语音文本信息,未结合预处理技术导致方法的抗噪声能力差,未结合优化策略导致方法的性能低,但其优势是计算时间成本低;其余为改进方法,将深度置信网络结合文本识别的框架或从自身进行改进,最后比较分析研究人员所提出的不同方法,总结其优势与劣势,提出进一步的改进思路。

3.2 人脸及表情识别领域

Li等^[16]首先考虑到深度置信网络(DBN)提取人脸局部特征信息的不完全性提出LBP+DBN,进而发现LBP对于多姿态的人脸纹理特征作用有待加强,最终提出基于中心对称原理的CS-LBP。CS-LBP可对四个不同方向的人脸图像编码,极大提升DBN特征学习的能力,在多姿态以及正面人脸识别中,CS-LBP+DBN的性能为92.78%和97.5%,均优于DBN、LBP+DBN。今后需从

DBN自身改进特征学习能力,降低整体方案的时间成本。

李鑫等^[17]提出一种新的人脸识别技术,考虑到计算时间成本和识别的准确率从而在深度置信网络(DBN)中加入多任务学习机制,通过分析预测误差指标值验证方法的可行性。DBN提取特征后多个任务并行处理,利用权值竞争策略需找最优参数,避免陷入局部最优和过学习现象,相较于DBN+单任务机制的方法其预测误差降低了50%,达到8.75%,但该方法的抗噪声的能力弱,实验过程中去除了残缺的人脸样本,可将DBN结合粗糙理论进行改进。

Fatahi等^[18]提出一个基于深度置信网络改进(SDBN)的人脸识别系统,该系统利用两层具有Siegert单元的RBM来构建,并且通过调整Siegert方程设置DBN中LIF神经元参数。SDBN方法关键在于,改进自身网络结构的组成元件(RBM)和调整模型参数来提高所提出方法的分类识别率。

杨瑞等^[19]对人脸样本提取Gabor小波特征并在卷积融合后输入到构建好的深度置信网络(DBN)中,DBN微调策略采用交叉熵方法确定最优模型参数,利用Softmax回归分类测试样本。实验的分类精度虽然达到92.7%,但样本不充足,未能在海量数据上证明深度学习方法在人脸识别的可靠性,且训练时间成本高,建议改进DBN结构或微调策略或更合理地划分图像块以改善时间复杂度和学习效率。

吴进等^[20]为改善深度置信网络(DBN)学习人脸样本特征时忽略局部特征的学习,提出将DBN结合多尺度局部二值模式(LBP)算法,LBP用于提取更关键的人脸局部特征。同等实验条件下,LBP+DBN算法的检测效果为95.2%,相较于Gabor+PCA、Gabor+SVM的分类

表1 文字检测领域优劣分析

作者	方法	适用领域	优势	劣势
陆军建等 ^[2]	DBN+CUDA	字符相似或模糊的手写字符图片	识别准确率高且计算速度快	缺乏实验结果的对比分析
Mleczo等 ^[3]	RDBN	信息不完全的文本数据	方法的抗噪性强	分类的准确率需进一步提高
张董等 ^[4]	区域分割+DBN	无用区域大的字符图片	大幅度提高识别率	时间成本较高
Liu ^[5]	DBN+SVM	分类中文文本	大大降低分类过程中预测误差	需改进样本的特征降维策略
翟文洁等 ^[6]	HDBN	文本分类检索	分类性能良好	训练时间长
王贵新等 ^[7]	DBN+word2vec	用于文字通讯行业短信分类拦截	采用合适的微调策略	对样本质量有较强依赖性
Jiang等 ^[8]	DBN +Softmax	大型文本数据	选择最有效的参数优化算法	对于小型文本数据分类效果差
宋春静等 ^[9]	DBN+DBN	集装箱字符识别	利用字符分离提高识别效果	采用两个特征学习模型分类时间较长
Xu等 ^[10]	DBN+SWT	包含较多错误字符的文本数据	保证结果鲁棒性	识别准确率较低
周世超等 ^[11]	DBN	语音文本内容	大幅度提高分类精度	分类前需对样本集进行内容整理,去噪
张庆庆等 ^[12]	改进DBN	文本情感分类	优化模型结构和参数,节约训练时间	实验分类性能低
Ray等 ^[13]	DBN+不同优化算法	场景文本检测	验证识别方法的可行性和有效性	计算复杂度较高
翟继友 ^[14]	不同深度的DBN	语义相似的文本	有效区分相似文本内容	实验文本具有特定性,应扩大文本量
Sarikaya等 ^[15]	改进DBN+SVM	自然语言信息	计算时间成本低	需增强其普适性

精度有明显优势,但时间损耗需进一步改进,若能提出方法与 Gabor+DBN、LBP+SVM 比较,更能够比较分析出方法的有效性和可靠性。

Wang 等^[21]提出一种基于深度学习的混合模型算法解决人脸检测问题,该模型为卷积池深度置信网络(CPDBN)是对传统 DBN 的组成元件 RBM 进行改进构建 CPDBN 的。CPDBN 的人脸检测算法提高了局部面部识别的精度和多姿态人脸的鲁棒性,可用于研究和预测闭塞的特征位置。实验结果表明,CPDBN 方法正确检测率优于传统 DBN(91.6%)和卷积神经网络(CNN,90.8%),达到 94.2%。

黄寿喜等^[22]利用改进后的受限玻尔兹曼机(GB-RBM)构建新的深度置信网络(GB-DBNs),来学习人脸表情特征,并将其作为堆叠自动编码器(SAE)的输入,采用七折交叉验证方式对数据样本进行识别分类。GB-DBNs+SAE 方法增强了分类的鲁棒性,识别精度达到 92.46%,优于传统分类和模板匹配法。

柴瑞敏等^[23]基于压缩感知理论改进传统的稀疏深度置信网络(DBN),增强 DBN 对于人脸特征的学习能力,其关键在于优化稀疏受限玻尔兹曼机模型中的稀疏项(tan-sigmoid 正则项)。改进的 DBN 可基于不同问题进行多个任务学习而非相同的稀疏水平,实验表明改进后的稀疏 DBN 分类多个人脸数据库时效果良好,平均识别率达到 98%,克服了高维特征表示的样本数据的冗余性。

Liu 等^[24]采用一个新颖的增强型深度置信网络(BDBN)应用于人脸面部表情识别中,在统一的循环框架中迭代执行训练过程中的特征学习、特征选择和分类器构建的任务。BDBN 方法学习选择一组有效表示相关面部外观/姿态变化的特征,并且采用联合微调的参数优化策略(BTD-SFS)加强所选特征的区分力。实验结果表明,BDBN 完成特征集和分类器的最优组合,采用 8 折交叉验证,识别性能达到 96.7%。

施徐敢等^[25]采用多层感知器(MLP)结合深度置信网络(DBN)的方式对人脸表情进行分类,分类精度达到 90.95%。通过 DBN 学习不同尺寸下的人脸表情图片的

特征信息,之后将其用于训练 MLP,初步确定模型参数,最终将测试样本输入 DBN+MLP 中,实验过程中通过改变参数设置提升性能。经比较,DBN+MLP 方法优于 NN、MLP、SVM、NS、SRC,但在处理相似人脸表情样本时,预测误差增加。

Wu 等^[26]针对演讲人情感的识别分类问题提出一种利用深度置信网络(DBN)对多特征进行融合后学习的方案,很好地对演讲情感类别的概率分布进行预测。该方法的新颖之处在于提出子带光谱中心加权小波包倒频谱系数(W-WPCC)特征,即组合子带能量与子带光谱并通过加权策略产生噪声鲁棒的声学特征计算而来。W-WPCC 特征组合方式提高了识别性能,平均识别率达到 86.6%,可尝试将 DBN 与其他特征结合以获取更高效的特征表示了从而提升分类准确率。

Ali 等^[27]采用一种特征结合的方式解决说话人识别的任务,将深度置信网络(DBN)学习的特征与 Mel 频率倒频谱系数(MFCC)特征组合来训练分类器 SVM,提高识别能力。通过比较分析,MFCC+DBN 的方法优于 DBN,且 DBN 的网络结构同样影响分类精度,其中 DBN-1+DBN-2+MFCC 性能最佳为 92.6%。

深度置信网络在人脸及表情识别领域中的应用效果分析如表 2。人脸及表情识别领域中,数据样本主要包括人脸数据库、人脸多表情数据库以及识别说话人数据集。文献[25]采用传统的深度置信网络方法识别人脸表情,未融合人脸表示特征导致分类性能差,未采用降低误差的优化策略导致难以区分相似度大的样本,但其优势是验证深度学习模型在人脸及表情识别分类领域中的可行性且优于浅层网络模型;其余为改进方法,即针对人脸识别对于局部纹理特征的需求,将 DBN 与 LBP、Gabor 等特征相结合且改进上述特征;针对多类人脸表情的分类问题,为提升性能,结合使用多任务自学习模式,基于上述不同分类识别方法,分析其长处与不足之处。

3.3 遥感图像分类领域

Lin 等^[28]利用高光谱图像特征信息融合的方式,通过深度置信网络(DBN)对其进行分类。研究者组合使

表2 人脸及表情识别领域优劣分析

作者	方法	适用领域	优势	劣势
Li 等 ^[16]	CS-LBP+DBN	不同方向的人脸图片	有效识别多姿态下的人脸	需进一步提高特征学习能力
李鑫等 ^[17]	DBN+多任务机制	信息完整的人脸样本数据	降低计算时间且保证学习率	抗噪声能力弱
Fatahi 等 ^[18]	SDBN	正面人脸	从自身改进提高识别精度	缺乏人脸样本的多样性
杨瑞等 ^[19]	Gabor +DBN	小数据量的人脸样本	提高特征学习的能力	样本不充分且模型训练时间长
吴进等 ^[20]	LBP+DBN	多尺度人脸图片	分类效果具有明显的提升	实验方法对比的一致性需完善
Wang 等 ^[21]	CPDBN	多姿态人脸数据	分类性能优于其他深度学习模型	有效获取闭塞位置的特征信息
黄寿喜等 ^[22]	GB-RBM	特征明显的人脸表情识别	保证分类识别效果的鲁棒性	时间成本高
柴瑞敏等 ^[23]	压缩感知理论+DBN	大型高维数据	克服样本特征信息的冗余性	尝试其他优化改进策略缩短时间
Liu 等 ^[24]	BDBN	面部多表情分类	采用联合微调的优化策略	需在多个数据集上进一步验证可靠性
施徐敢等 ^[25]	DBN	人脸表情差异明显的样本数据	分类多种人脸表情的可行性	分类相似度大的样本时效果不好
Wu 等 ^[26]	W-WPCC+DBN	分类演讲人情感	多种特征融合的方式增强分类稳定性	预测效果性能需提高
Ali 等 ^[27]	MFCC+DBN	识别说话人	分类精度高	训练时间过长,稳定性差

用邻域马赛克谱和空间信息,通过缝合邻域(NIS),整合光谱信息和空间信息(SSIS)并用加权平均方式获得最优加权平均法的空谱联合,两种方法的分类精度良好为98.62%和98.71%。最终分类结果表明SSIS+DBN相对于NIS+DBN性能有小幅提升,且优于SAR-LR、RBF-SVM,但若考虑到计算复杂度,该方法的样本预处理方式需改进。

Chen等^[29]基于深度置信网络(DBN)加入Dropout策略以解决高光谱遥感图像识别和分类过程中的问题。Dropout策略是在传统方法上改进,该策略选择局部数据清除每次的权重变量,衰减至0,即在模型学习过程中增加一个可以防止过拟合问题的随机化过程,使得随机化更彻底,以便提升分类准确率。

高鑫等^[30]采用改进的深度置信网络(DBN)结合去噪处理技术的方法检测地物遥感图像中的多类别物种。通过调整网络深度、迭代次数、权重参数来确定最优模型,通过添加自适应扩散系数的PM模型消除样本图像的噪声信息的同时保留边缘信息,自适应方式不但利于有效特征的提取同时能去除冗余信息。实验过程中,对八类物种分类的性能缺乏鲁棒性,主要由于相似地物种类的难区分性影响分类总精度。

吕启等^[31]着重分析深度学习模型DBN在分类遥感图像研究中与传统模型分类效果的比较,虽然整体实验分类结果表明DBN性能优于SVM、NN,但是其性能提高幅度不大且未能较好解决区别性不大的物种之间的分类,只证实了DBN的可行性。建议尝试利用特征融合方式增强DBN特征学习的能力或将DBN与其他深度分类模型相结合,从而进一步提高模型分类识别的能力。

Zhao等^[32]建立一个深度置信网络分析合成孔径雷达SAR遥感图像的三级变化,即图像的正变化像素、负变化像素以及不变像素。首先,利用改进的对数比操作生成SAR差分图像;其次,建立DBN提取鲁棒特征,对散斑噪声具有较强抗噪性;最终,对所输入数据样本进行检测分类,DBN_FLICM分类正确率达到98%,证明DBN在海量遥感图像数据集上的分类可靠性。

徐丽坤等^[33]将深度置信网络(DBN)用于分类地物信息的遥感图像,重点对DBN的模型参数进行改进调优,分析了DBN的分类精度随着网络内部结构、迭代次数改变的结果。实验分类的总精度为92.23%,证明了DBN+Softmax方法的可行性和有效性。

Ayhan等^[34]采用正确分类精度指标进行定量分析,所提出方法应用于高光谱图像的识别任务,实验分为两大步骤:一是DBN+光谱信息/组合光谱空间信息分析DBN模型的最佳参数和训练时间;二是DBN和SVM在相同特征输入下的性能比较。基于具体的高光谱数据,发现仅使用光谱信息的效果优于组合信息,这也许是过拟合现象导致的性能下降,建议通过预处理方式去除图像中无用的特征信息。

李新国^[35]利用深度置信网络的多隐层多层感知器(MLP)的深层网络结构对地物遥感图像进行识别检测,该方法的总体分类准确度良好,DBN+MLP相较于传统的NN和SVM具有明显优势。实验过程中充分利用DBN的降维和表示能力以较少维度很好表示样本特征信息,但所采用的样本包含较少的相似物种,对于所提出模型的区分能力判定的可靠性需加强。

Le等^[36]分析深度学习架构在识别检测复杂数据中的高效性,提出将深度置信网络(DBN+LR)用于分类远程感测高光谱数据。通过研究DBN网络宽度和观察区分多类别的预训练时间来改进DBN,在反复训练学习的过程中确定最佳宽度。同时,在光谱信息中融合空间信息作为原始输入,从而大幅度提升分类性能,相比于SVM,文中提出方法性能提升约15%,达到98.123 3%。

深度置信网络在遥感图像分类领域中的应用效果分析如表3。遥感图像分类领域中,样本集基本为高光谱图像,文献[31]采用的是传统DBN方法,未优化模型的网络结构和参数导致分类性能低,但其优势是在遥感图像领域中证明其处理复杂、海量数据的可行性;其余为所采用的改进方法:一是结合样本自身的特征信息;二是结合不同的模型优化策略;三是结合多种解决分类任务的分类器。从分类精度、计算复杂度等方面分析DBN应用于高光谱图像的优势与劣势。

表3 遥感图像分类领域优劣分析

作者	方法	适用领域	优势	劣势
Lin等 ^[28]	SSIS+DBN	空间信息完整的高光谱图像	选择最利于分类的特征信息	完善样本图像的预处理方式
Chen等 ^[29]	DBN+Dropout	物体信息少的样本	有效避免过拟合学习缩减时间	缺少全局数据的学习和分析
高鑫等 ^[30]	DBN+去噪技术	含噪声的光谱图像	自适应学习样本特征信息	鲁棒性低且难以区分相似物种
吕启等 ^[31]	DBN	分类光谱差异度大的高光谱图像	证实可行性	未优化特征学习模型未结合不同分类器
Zhao等 ^[32]	DBN_FLICM	SAR遥感图像	有效分类海量遥感数据且抗噪声能力强	方法具有一定程度的特定性
徐丽坤等 ^[33]	改进DBN	地物信息丰富的样本	选择最优的模型参数,提升分类性能	可尝试提高模型参数的自学习能力
Ayhan等 ^[34]	特征信息+DBN+SVM	特征显著的遥感图像数据	实验完整分析特征学习和分类器的性能	组合信息易导致过学习现象
李新国 ^[35]	DBN+MLP	高维样本数据	特征降维后的表示能力强	扩增样本以保证完整性和多样性
Le等 ^[36]	DBN+LR	远程感测高光谱数据	选择最佳网络结构并降低预训练时间	增加样本提高方法的适用性

3.4 其他新兴领域

张善文等^[37]设计一种针对大棚中冬枣是否会发生病虫害的预测模型——改进的MDBN,关键步骤:首先,由于样本预测收环境、气候、时间等因素影响,在DBN组成元件RBM中加入样本先验信息(MDBN);其次,改变MDBN模型的各项参数,对其进行优化。

陈亮等^[38]将传统深度置信网络(DBN)应用于在线视频热度的预测,进一步分析视频上映的价值程度。通过分析影响视频热度的各项因素所起的作用,选择最有效的几项作为输入向量,实验结果表明,DBN方法的预测误差低,但视频样本的噪声多,预处理过程所需时间成本高。

Hu等^[39]采用基于深度置信网络(DBN)的参数语音合成隐马尔可夫模型(HMM)进行频谱特征的有效表示:首先,转换代表高维光谱包络的DBN为二进制码——DBC;其次,将DBC作为HMM的光谱特征。实验结果表明,该方法避免过平滑效应,通过全面考虑到参数语音合成过程中产生的全局变异提升性能。

Yeh等^[40]提出一种新型的识别检测篮球比赛中裁判动态手势的系统,即利用深度置信网络和时域特征分析动态手势带来的信号——sEMG和ACC。DBN模型下,采用5折交叉验证法手势识别的准确率为97.9%,同时保证实时检测手势与得分表交流。

刘瑞兰等^[41]通过比较分析DBN-BP与DBN-PLS的预测误差与测量精度,证明DBN-BP方法在工业生产中对于决定产品质量的4-CBA含量具有良好的预测效果。实验的不足之处在于根据经验确定模型参数,未能保证参数最优。

Zheng等^[42]分析改进的深度置信网络(DBN)模型在预测汇率应用中的效果,改进DBN基于两个关键部分:模型的结构设计和参数的学习规则。实验过程中,基于多方差分析的方法选择改进显著的节点以提高预测效率,与前馈神经网络(FNN)比较,改进DBN提升近50%的准确率。

李楠^[43]提出一种改进的自适应深度置信网络算法(DBN)与社会网络分析法(SNA)相结合分析经济发展

迅速的城市物流效率并将核心城市分类,自适应DBN依赖于样本数据知识集,自主学习模型的各项参数。自适应DBN+SNA方法在输入数据集的平均分类精度为98.21%,但在实验中该方法所设定的性能评价指标依赖数据集,普适性弱,并且以研究经验选择物流效率为评估指标。

李本威等^[44]在传统DBN中添加一个自适应系数动态更新学习率,模型自适应调整权值和偏置,形成自适应修正权值的ad_DBN方法。该方法取得了三种有益效果:第一,与传统DBN相比,提升发动机故障诊断准确率,达到97.84%;第二,增强故障分类识别过程中的泛化能力,节约了时间;第三,模型的抗噪性强,但其是以人为增加故障种类样本进行验证的,缺乏一定的可靠性。

Huang等^[45]提出一种将三隐层结构的深度置信网络(DBN)结合背景减法的车辆识别方法,样本数据取自雪山隧道。该方法的实验结果表明检测车辆的准确率达到96.59%,但实验样本数据量少,且为改善隧道环境影响需进行较长时间样本预处理。

深度置信网络在新兴领域中的应用效果分析如表4。表中所涉及到的车辆识别、汇率预测、手势分析、自然灾害以及视频热度预测等新兴领域应用表明,DBN在未来的发展前景值得期待。

4 总结与展望

本文以深度学习模型为背景,分析深层网络结构在模式识别、计算机视觉、图像处理等领域的发展基础,同时引入深度学习模型——深度置信网络(DBN)。首先,本文介绍DBN组成元件RBM的构建过程;接着,介绍DBN与分类器相结合的整体模型构建;最后,在此基础上展现本文关键,即对DBN在当今较为流行的各个领域中的应用效果进行研究讨论。今后,研究人员可从以下方面提升DBN性能:

(1)对DBN组成元件RBM进一步改进,从简化网络结构,降低计算复杂性以及模型参数学习规则的优化方向进行深入研究;

(2)将DBN与深度分类器相结合以提高最终分类

表4 新兴领域优劣分析

作者	方法	适用领域	优势	劣势
张善文等 ^[37]	改进MDBN	预测植物病虫害	融合先验信息提升预测率	未选择性过滤环境信息易丢失重要信息
陈亮等 ^[38]	DBN	在线视频热度预测	分析影响因素作用选择最有效特征	计算时间成本高
Hu等 ^[39]	DBC+HMM	频谱特征的表示	减轻过平滑效应	性能评价指标单一
Yeh等 ^[40]	DBN+时域特征	手势识别	准确率高实时性好	选择过少手势特征信号
刘瑞兰等 ^[41]	DBN-BP	分析产品质量	含量预测效果良好	未保证模型参数最优
Zheng等 ^[42]	改进DBN	预测汇率	选择最有效的样本节点	时间损耗大稳定性差
李楠 ^[43]	自适应DBN+SNA	分析城市物流效率	自主调整模型参数	较依赖数据集质量
李本威等 ^[44]	ad_DBN	发动机故障诊断以及分类故障级别	动态更新学习率泛化能力强且时间成本低、抗噪能力强	人为添加样本种类,实验性能缺乏可靠性
Huang等 ^[45]	DBN	车辆识别	检测性能良好可行性高	预处理时间长且实时性有待提高

准确率;

(3)扩展DBN应用范围,将其运用在航空识别或医学识别领域中,增强其普适性。

针对其对高层特征学习和提取的优势与良好的性能效果可知DBN的改进之路不止于此,其发展应用前景依然广阔。

参考文献:

- [1] Hinton G E, Osindero S, Teh Y W. A fast learning algorithm for deep belief nets[J]. Neural Comput, 2006, 18(7): 1527-1554.
- [2] 陆军建, 林家骏. 基于CUDA和深度置信网络的手写字符识别[J]. 华东理工大学学报: 自然科学版, 2015, 41(2): 210-215.
- [3] Mleczko W K, Kapuściński T, Nowicki R K. Rough deep belief network—application to incomplete handwritten digits pattern classification[C]//Proceedings of 21st International Conference on Information and Software Technologies, 2015: 400-411.
- [4] 张董, 游福成, 王惠华, 等. 受限制 Boltzmann 机深度置信网络与手写数字识别[J]. 北京印刷学院学报, 2016, 24(4): 56-58.
- [5] Liu T. A novel text classification approach based on deep belief network[C]//Proceedings of Neural Information Processing Theory and Algorithms International Conference, Sydney, Australia, 2010: 314-321.
- [6] 翟文洁, 闫琰, 张博文, 等. 基于混合深度信念网络的多类文本表示与分类方法[J]. 情报工程, 2016, 2(5): 30-40.
- [7] 王贵新, 郑孝宗, 张浩然, 等. 利用深度置信网络的中文短信分类[J]. 现代电子技术, 2016, 39(9): 37-40.
- [8] Jiang M, Liang Y, Feng X, et al. Text classification based on deep belief network and softmax regression[J]. Neural Computing and Applications, 2016(1): 1-10.
- [9] 宋春静, 丁香乾, 郭乙运. 基于深度信念网络的集装箱字符识别方法[J]. 计算机工程与设计, 2016, 37(3): 742-745.
- [10] Xu H, Xue L, Su F. Scene text detection based on robust stroke width transform and deep belief network[C]//Proceedings of the 12th Asian Conference on Computer Vision, Singapore, 2015: 195-209.
- [11] 周世超, 张沪寅, 杨冰. 基于深度信念网络的语音服务文本分类[J]. 计算机工程与应用, 2016, 52(21): 157-161.
- [12] 张庆庆, 刘西林. 基于深度信念网络的文本情感分类研究[J]. 西北工业大学学报: 社会科学版, 2016, 36(1): 62-66.
- [13] Ray A, Rajeswar S, Chaudhury S. Scene text analysis using deep belief networks[C]//Proceedings of the 2014 Indian Conference on Computer Vision Graphics and Image Processing, Bangalore, India, 2014: 1-8.
- [14] 翟继友. 基于深度置信网络的语义相关度计算模型[J]. 科学技术与工程, 2014, 14(32): 58-62.
- [15] Sarikaya R, Hinton G E, Deoras A. Application of deep belief networks for natural language understanding[J]. IEEE/ACM Trans on Audio, Speech and Lang Proc, 2014, 22(4): 778-784.
- [16] Li C, Wei W, Wang J. Face recognition based on deep belief network combined with center-symmetric local binary pattern[C]//Advanced Multimedia and Ubiquitous Engineering: FutureTech & MUE, Singapore, 2016: 277-283.
- [17] 李鑫, 张俊. 基于深度信念网络和多任务学习的人脸识别[J]. 电脑知识与技术, 2016, 12(17): 165-168.
- [18] Fatahi M, Ahmadi M, Ahmadi A, et al. Towards an spiking deep belief network for face recognition application[C]//Proceedings of International Conference on Computer and Knowledge Engineering, 2016.
- [19] 杨瑞, 张云伟, 苟爽, 等. Gabor特征与深度信念网络结合的人脸识别方法[J]. 传感器与微系统, 2017, 36(5): 68-70.
- [20] 吴进, 严辉, 王洁. 采用局部二值模式与深度信念网络的人脸识别[J]. 电讯技术, 2016, 56(10): 1119-1123.
- [21] Wang D, Li M, Li X. Face detection algorithm based on convolutional pooling deep belief network[C]//Proceedings of International Conference on Electrical, 2017.
- [22] 黄寿喜, 邱卫根. 基于改进的深度信念网络的人脸表情识别[J]. 计算机工程与设计, 2017, 38(6): 1580-1584.
- [23] 柴瑞敏, 曹振基. 基于改进的稀疏深度信念网络的人脸识别方法[J]. 计算机应用研究, 2015, 32(7): 2179-2183.
- [24] Liu P, Han S, Meng Z, et al. Facial expression recognition via a boosted deep belief network[C]//Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition, 2014: 1805-1812.
- [25] 施徐敢, 张石清, 赵小明. 融合深度信念网络和多层感知器的人脸表情识别[J]. 小型微型计算机系统, 2015, 36(7): 1629-1632.
- [26] Wu A, Huang Y, Zhang G. Feature fusion methods for robust speech emotion recognition based on deep belief networks[C]//Proceedings of the Fifth International Conference on Network, Communication and Computing, Kyoto, Japan, 2016: 6-10.
- [27] Ali H, Tran S N, Benetos E, et al. Speaker recognition with hybrid features from a deep belief network[J]. Neural Computing and Applications, 2016(1): 1-7.
- [28] Lin L, Dong H, Song X. DBN-based classification of spatial-spectral hyperspectral data[C]//Proceedings of the 12th International Conference on Intelligent Information Hiding and Multimedia Signal Processing, Kaohsiung, Taiwan, China, 2017: 53-60.
- [29] Chen G, Li X, Liu L. A study on the recognition and classification method of high resolution remote sensing image based on deep belief network[C]//Proceedings of the 11th International Conference on Bio-inspired Computing—Theories and Applications, Xi'an, China, 2016: 362-370.

(下转47页)

- [15] Du Y, Pearlson G D, Yu Q, et al. Interaction among sub-systems within default mode network diminished in schizophrenia patients: A dynamic connectivity approach[J]. Schizophrenia Research, 2015, 170(1):55-65.
- [16] Suk H I, Wee C Y, Lee S W, et al. State-space model with deep learning for functional dynamics estimation in resting-state fMRI[J]. Neuroimage, 2016, 129:292-307.
- [17] Reich D, Price A L, Patterson N. Principal component analysis of genetic data[J]. Nature Genetics, 2008, 40(5):491-492.
- [18] Donoho D L, Grimes C. Hessian eigenmaps: Locally linear embedding techniques for high-dimensional data[J]. Proceedings of the National Academy of Sciences, 2003, 100(10):5591-5596.
- [19] Hagmann P, Cammoun L, Gigandet X, et al. Mapping the structural core of human cerebral cortex[J]. Plos Biology, 2008, 6(7):159.
- [20] Hinton G, Roweis S. Stochastic neighbor embedding[J]. Advances in Neural Information Processing Systems, 2010, 41(4):833-840.
- [21] Leon P S, Woodman M, McIntosh R, et al. The virtual brain: A neuroinformatics platform for simulating large-scale brain network models[J]. Bmc Neuroscience, 2013, 14(1):1-2.
- [22] Rana K D, Vaina L M, Hämäläinen M S. A fast statistical significance test for baseline correction and comparative analysis in phase locking[J]. Frontiers in Neuroinformatics, 2013, 7(3):3.
- [23] Torben-Nielsen B. An efficient and extendable python library to analyze neuronal morphologies[J]. Neuroinformatics, 2014, 12(4):619-622.
- [24] Brain Imaging & Analysis Center. Python/FSL resting state pipeline[DB/OL]. [2016-07-15]. https://wiki.biac.duke.edu/biac:analysis:resting_pipeline.
- [25] Darvas F, Pantazis D, Kucukaltun-Yildirim E, et al. Mapping human brain function with MEG and EEG: Methods and validation[J]. Neuroimage, 2004, 23(1):289-299.
- [26] Wang X, Luo C. Researches on chaos phenomenon of EEG dynamics model[J]. Applied Mathematics & Computation, 2006, 183(1):30-41.
- [27] Wang X, Luo C, Meng J. Nonlinear dynamic research on EEG signals in HAI experiment[J]. Applied Mathematics & Computation, 2009, 207(1):63-74.

(上接18页)

- [30] 高鑫, 欧阳宁, 袁华. 基于快速去噪和深度信念网络的高光谱图像分类方法[J]. 桂林电子科技大学学报, 2016, 36(6):469-476.
- [31] 吕启, 窦勇, 牛新, 等. 基于DBN模型的遥感图像分类[J]. 计算机研究与发展, 2014, 51(9):1911-1918.
- [32] Zhao Q, Gong M, Li H, et al. Three-class change detection in synthetic aperture radar images based on deep belief network[C]//Proceedings of the 10th International Conference of Bio-Inspired Computing—Theories and Applications, Hefei, China, 2015:696-705.
- [33] 徐丽坤, 刘晓东, 向小翠. 基于深度信念网络的遥感影像识别与分类[J]. 地质科技情报, 2017(4):244-249.
- [34] Ayhan B, Kwan C. Application of deep belief network to land cover classification using hyperspectral images[C]//Proceedings of the 14th International Symposium, Sapporo, Hakodate, 2017:269-276.
- [35] 李新国. 一种基于DBN的高光谱遥感图像分类方法[J]. 电子测量技术, 2016, 39(7):34-41.
- [36] Le J H, Yazdanpanah A P, Regentova E E, et al. A deep belief network for classifying remotely-sensed hyperspectral data[C]//Proceedings of the 11th International Symposium on Advances in Visual Computing, Las Vegas, NV, USA, 2015:682-692.
- [37] 张善文, 张传雷, 丁军. 基于改进深度置信网络的大棚冬枣病虫害预测模型[J]. 农业工程学报, 2017, 33(19):202-208.
- [38] 陈亮, 张俊池, 王娜, 等. 基于深度信念网络的在线视频热度预测[J]. 计算机工程与应用, 2017, 53(9):162-169.
- [39] Hu Y J, Ling Z H. DBN-based spectral feature representation for statistical parametric speech synthesis[J]. IEEE Signal Processing Letters, 2016, 23(3):321-325.
- [40] Yeh C W, Pan T Y, Hu M C. A sensor-based official basketball referee signals recognition system using deep belief networks[C]//Proceedings of the 23rd International Conference, Iceland, 2017:565-575.
- [41] 刘瑞兰, 毛佳敏. 基于深度置信网络的4-CBA软测量建模[J]. 计算机工程与应用, 2017, 53(6):227-230.
- [42] Zheng J, Fu X, Zhang G. Research on exchange rate forecasting based on deep belief network[J]. Neural Computing and Applications, 2017(1):1-10.
- [43] 李楠. 城市物流效率分析自适应DBN算法研究[J]. 计算机工程与应用, 2017, 53(12):9-15.
- [44] 李本威, 林学森, 杨欣毅, 等. 深度置信网络在发动机气路部件性能衰退故障诊断中的应用研究[J]. 推进技术, 2016, 37(11):2173-2180.
- [45] Huang B J, Hsieh J W, Tsai C M. Vehicle detection in hsuehshan tunnel using background subtraction and deep belief network[C]//Proceedings of the 9th Asian Conference on Intelligent Information and Database Systems, 2017:217-226.