

8-20-2020

Fast Training Model for Image Classification Based on Spatial Information of Deep Belief Network

Gao Qiang

Institute of Electrical and Electronic Engineering, North China Electric Power University, Baoding 071003, China;

Yang Wu

Institute of Electrical and Electronic Engineering, North China Electric Power University, Baoding 071003, China;

Li Qian

Institute of Electrical and Electronic Engineering, North China Electric Power University, Baoding 071003, China;

Follow this and additional works at: <https://dc-china-simulation.researchcommons.org/journal>



Part of the [Artificial Intelligence and Robotics Commons](#), [Computer Engineering Commons](#), [Numerical Analysis and Scientific Computing Commons](#), [Operations Research, Systems Engineering and Industrial Engineering Commons](#), and the [Systems Science Commons](#)

This Paper is brought to you for free and open access by Journal of System Simulation. It has been accepted for inclusion in Journal of System Simulation by an authorized editor of Journal of System Simulation.

Fast Training Model for Image Classification Based on Spatial Information of Deep Belief Network

Abstract

Abstract: With the exponential growth of data and the complexity of algorithm, efficient computing of DBN (Deep Belief Network,) has become an important issue. A fast training model for image classification of DBN was built according to sample images of DBN have nothing to do with spatial information, an improved algorithm simply as LSMI (Linear Superposition Multiple Images) was proposed for classifying images based on the idea that linear combination of multiple images. The characteristics of training images have nothing to do with spatial information was proved via information entropy theory, meanwhile, the characteristics was identified to be correct based on ORL database. According to ergodic theory, an algorithm simply as LSMI was proposed according to the ergodic theory. The LSMI algorithm was contrasted to other improved algorithms, using COREL and MIT database, judging whether the LSMI algorithm is effective from the correct recognition rate, time complexity and other quotas. The simulation results show that LSMI algorithm can ensure recognition rate, decrease training time greatly and achieve the goal of fast learning.

Keywords

deep belief network, spatial information, image classification, fast learning, LSMI algorithm

Recommended Citation

Gao Qiang, Yang Wu, Li Qian. Fast Training Model for Image Classification Based on Spatial Information of Deep Belief Network[J]. Journal of System Simulation, 2015, 27(3): 549-558.

基于空间信息的 DBN 图像分类快速训练模型

高强, 阳武, 李倩

(华北电力大学电气与电子工程学院, 保定 071003)

摘要: 数据的指数级增长及算法本身的复杂性使深度信念网络 (DBN) 面临着学习效率问题。根据 DBN 的样本图像与空间信息无关的特点, 建立了 DBN 图像分类快速训练模型, 提出了基于多幅样本图像线性叠加合成思想的 DBN 图像分类算法—LSMI 算法。利用信息熵理论, 证明了样本图像与空间信息无关的特点, 并以 ORL 库为依据进行了验证。根据正态历经性, 提出了 LSMI 算法, 并以 COREL 库和 MIT 库为仿真对象, 与其他改进算法进行对比, 从正确识别率和算法时间复杂度等指标, 判断该算法的有效性。仿真结果表明 LSMI 算法在保证识别率不变的同时, 大幅度降低了算法的训练时间, 达到快速学习的目的。

关键词: 深度信念网络; 空间信息; 图像分类; 快速学习; LSMI 算法

中图分类号: TP18

文献标识码: A

文章编号: 1004-731X (2015) 03-0549-10

Fast Training Model for Image Classification Based on Spatial Information of Deep Belief Network

Gao Qiang, Yang Wu, Li Qian

(Institute of Electrical and Electronic Engineering, North China Electric Power University, Baoding 071003, China)

Abstract: With the exponential growth of data and the complexity of algorithm, efficient computing of DBN (Deep Belief Network,) has become an important issue. A fast training model for image classification of DBN was built according to sample images of DBN have nothing to do with spatial information, an improved algorithm simply as LSMI (Linear Superposition Multiple Images) was proposed for classifying images based on the idea that linear combination of multiple images. The characteristics of training images have nothing to do with spatial information was proved via information entropy theory, meanwhile, the characteristics was identified to be correct based on ORL database. According to ergodic theory, an algorithm simply as LSMI was proposed according to the ergodic theory. The LSMI algorithm was contrasted to other improved algorithms, using COREL and MIT database, judging whether the LSMI algorithm is effective from the correct recognition rate, time complexity and other quotas. The simulation results show that LSMI algorithm can ensure recognition rate, decrease training time greatly and achieve the goal of fast learning.

Keywords: deep belief network; spatial information; image classification; fast learning; LSMI algorithm

引言

图像分类系统主要由图像采集器、预处理模



收稿日期: 2014-06-30 修回日期: 2014-10-13;
作者简介: 高强(1960-), 男, 河北, 博士, 教授, 研究方向为电力通信、深度学习以及图像处理; 阳武(1975-), 男, 湖南, 博士生, 研究方向为图像处理、电力与通信系统; 李倩(1990-), 女, 内蒙, 硕士, 研究方向为深度学习及图像处理。

块、分类模型和分类器等部分组成, 分类模型的建立是图像处理领域非常重要的部分, 对基于内容的图像检索结果的准确性至关重要。目前, 人们逐渐将重点转移到图像的自动分类上并提出了各种解决方案。人工神经网络就是一种可用于图像自动分类的模仿动物神经网络行为特征的模型, 已成功用于惯导系统传递对准^[1]、输电线路故障检测^[2]、电

网区分故障诊断^[3]以及交流伺服系统^[4]等领域中。

经过 Hinton 教授与其他几位学者的共同努力,于 2006 年提出了深度神经网络的概念,开启了深度学习在学术界和工业界的浪潮。

深度神经网络是一种自动学习输入样本特征的方法,且学习到的特征对数据具有更本质的刻画。它包括深度信念网络^[5](DBN)、自动编码器^[6]以及深度卷积网络^[7-8]三种多层结构模型。其中,深度信念网络由于其具有无监督学习的特征受到广泛的关注^[9],目前已成功应用到了语音识别^[10],手写字体识别^[11]、交通标志识别、人脸识别^[12]等领域中,显示出优越的学习性能。但是,由于 DBN 算法自身需要大量的训练样本,且训练过程中存在收敛速度慢、易陷入局部最小等缺点,导致其需要大量的精力和时间,学习效率低^[13]。针对以上问题,研究者采取了 2 方面的措施,一方面是通过硬件 GPU 和 FPGAs 等进行并行处理进而加快学习速度^[14],缺点是并不是所有人都具备这样的硬件条件。另一方面是对算法本身进行改进, Hinton 提出了 RBM 的快速学习算法—对比散度(CD)算法^[15]。进一步,文献[16]提出蒙特卡罗方法 PT 算法代替 CD 算法, Tijmen Tieleman 修正了 CD 算法无法极大似然的缺陷,提出了 PCD 算法^[17]。许多学者^[18-20]还提出利用自适应步长的方法代替传统的固定步长,提高 PT 算法和 CD 算法的性能,进而提升 RBM 的学习效率。James Martens 通过仿真研究,提出了一种二阶优化算法—Hessian-free 算法取代梯度

下降法^[21]。而文献[22]从另外一个角度提出了基于云平台的 RBM 推荐算法,将 RBM 过程分解为若干个 Hadoop 任务的循环,实现并行计算。

综上, DBN 存在收敛速度慢、学习效率低等缺点。本文基于 DBN 训练样本图像无空间信息的特点,提出了一种线性叠加的快速学习算法 LSMI (Linear Superposition Multiple Images),用于实现图像分类。首先通过基于 DBN 图像分类系统输入样本的信息熵证明了训练样本图像空间位置改变对训练效果无影响,即深度信念网络训练时,图像与空间信息无关。结合此性质,根据各态历经性理论,借鉴了多幅同类别训练样本图像线性叠加处理的思想,提出了深度信念网络的快速学习改进算法 LSMI。最后以 COREL 场景分类库和 MIT 人脸库为训练样本,与传统算法和基于自适应步长的 DBN 改进算法进行了仿真对比验证。仿真结果表明 LSMI 算法可以在保证图像有效识别率的基础上,降低时间复杂度,提高学习效率。

1 深度信念网络

1.1 基于 DBN 的图像分类系统

图像分类系统包括图像采集、存储、预处理、分类和输出等模块。其中,最核心的模块是图像分类,可用的算法有很多种,本文采用深度信念网络(DBN)分类算法。系统组成如图 1 所示。

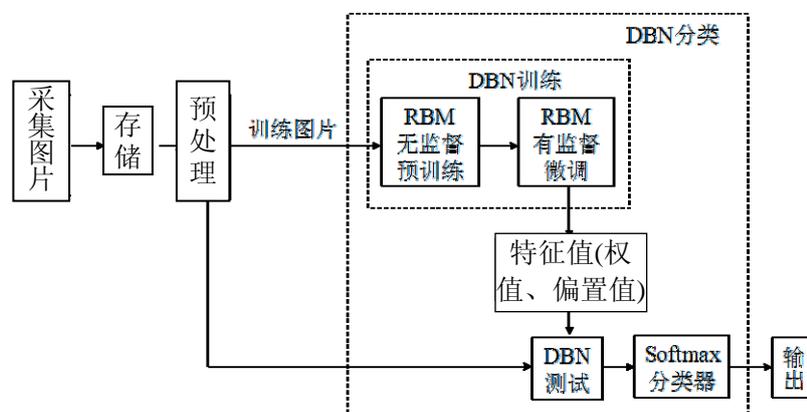


图 1 基于 DBN 的图像分类系统框图

其中, 分类器模型的选择是系统分类性能好坏的关键。基于 DBN 的分类器需要通过大量的训练学习, 才能使分类效果达到最佳, 当训练图片数量达到成千上万甚至达到百万级时, 训练时间就会非常长。

1.2 深度信念网络的模型

DBN 作为深度神经网络的一种模型, 由多层受限玻尔兹曼机(RBM)堆栈构成。RBM 是一种两层的神经网络模型, 由可视层和隐含层组成, 每层有若干个节点, 层内节点无连接, 不同层之间的节点依靠权值系数矩阵来连接, 根据不同的样本确定不同的权重值。DBN 的结构如图 2 所示。

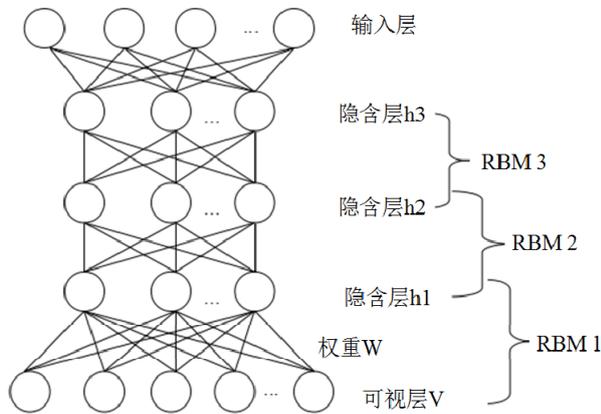


图 2 深度信念网络模型

RBM 是一种能量模型, 其可视层和隐含层的联合组态能量表示为:

$$E(v, h) = -\sum_i a_i v_i - \sum_j b_j h_j - \sum_i \sum_j w_{ij} v_i h_j \quad (1)$$

式(1)中: v_i, h_j 分别是可视层和隐含层的节点状态; a_i, b_j 分别为可视层和隐含层节点对应的偏置值; w_{ij} 为可视层与隐含层之间的连接权重值。隐含层节点的输出为:

$$h_j = \sum_i v_i w_{ij} + b_j \quad (2)$$

可视节点和隐藏节点的联合概率为:

$$p(v, h) = \frac{e^{-E(v, h)}}{Z} \quad (3)$$

其中: Z 为归一化因子, 即配分函数, 表达式为:

$$Z = \sum_{v, h} e^{-E(v, h)} \quad (4)$$

$p(v, h)$ 对 h 的边缘分布为:

$$p(v) = \sum_h p(v, h) = \frac{\sum_h e^{-E(v, h)}}{Z} \quad (5)$$

一般采用最大化似然函数 $L(\theta)$ 的方法求解 RBM 的参数, 表达式为:

$$L(\theta) = \prod_{i=1}^n p(v) \quad (6)$$

式(6)中: n 为训练样本的数目。最大化似然函数常用的数值方法是梯度上升, 参数 $\theta = (w, a, b)$ 的迭代公式为:

$$\theta = \theta + \eta \cdot \frac{\partial \ln L}{\partial \theta} \quad (7)$$

式(7)中: η 为学习率, 也称步长。采用对比散度算法(CD)的方法对重构数据进行近似采样, 得:

$$\frac{\partial \ln p(v)}{\partial w_{ij}} = \langle v_i h_j \rangle_0 - \langle v_i h_j \rangle_k \quad (8)$$

式(8)中: $\langle \rangle_0$ 是数据集的期望值; $\langle \rangle_k$ 是模型中定义的期望值。则新的更新准则如下:

$$\Delta w_{ij} = \eta (\langle v_i h_j \rangle_0 - \langle v_i h_j \rangle_k) \quad (9)$$

$$\Delta a_i = \eta (\langle v_i \rangle_0 - \langle v_i \rangle_k) \quad (10)$$

$$\Delta b_j = \eta (\langle h_j \rangle_0 - \langle h_j \rangle_k) \quad (11)$$

DBN 的预训练就是利用以上 RBM 的自训练实现自底向上无监督的逐层训练所有的 RBM 结构。通过使输出和输入差别最小, 即系统能量最小训练第一层的参数, 从而得到比输入更具有表示能力的特征, 在学习得到第 $n-1$ 层的输出后, 将第 $n-1$ 层的输出作为第 n 层的输入, 一层一层进行训练。由于基于 DBN 的图像分类所需的训练样本数目较大, DBN 训练过程中解决算法时间复杂度成为了一大难题。

2 深度信念网络的快速训练算法设计

2.1 DBN 提取图像空间信息的分析

研究表明, 深度信念网络应用于模式识别等应用中, 具有鲁棒性, 对伸缩、旋转、平移等处理后,

仍能有效地识别^[23]。为什么将图像进行旋转平移后,图像中的空间位置关系发生了变化,但仍能准确识别?深度信念网络训练图像时,是不是与图像的空间信息无关,只与自动学习到的特征点有关呢?

设 DBN 的可视层节点数为 p 个,第一层隐含层的节点数为 q 个,输入的训练样本图像为 \mathbf{X} ,像素为 $m \times p$,可视层节点的偏移量为 a ,第一层隐含层节点的偏移量为 $b = (b_1, b_2, \dots, b_q)$ 。下面结合熵的表示进行分析图像的空间信息。

根据统计力学,自由能量的定义为

$$FreeEnergy(x) = -\ln \sum_h e^{-E(x,h)} \quad (12)$$

根据式(12), $p(x)$ 可以重新写成

$$p(x) = \frac{e^{-FreeEnergy(x)}}{Z} \quad (13)$$

对于确定的系统,归一化因子 Z 是一个常数。式(13)两边取对数,得

$$\ln p(x) = -FreeEnergy(x) - \ln Z \quad (14)$$

式(14)两边同乘以 $-p(x)$, 并做累加,得

$$-\sum_x p(x) \ln p(x) = \sum_x p(x) FreeEnergy(x) + \sum_x p(x) \ln Z \quad (15)$$

式(15)简化为

$$H_x(p) = \sum_x p(x) FreeEnergy(x) + \ln Z \quad (16)$$

由式(16)可知,等式左边为可视层输入样本图像的信息熵,用 $H_x(p)$ 表示,等式右边的第 2 项中 $\ln Z$ 是一个常数,求概率平均后仍然为 $\ln Z$,熵函数只与 x 的概率分布有关,与 x 的空间位置无关,即系统收敛后的结果与空间无关。

当样本图像旋转平移后与原始样本图像进行相比,虽然空间位置发生了变化,但总的图像内容是不变的,每个像素所占的比例 $p(x)$ 是确定的,样本图像的信息熵是一定的,系统的自由能量也没有发生变化,目标函数一致。因此,基于深度神经网络的图像不具有空间结构信息,旋转平移后的训

练样本图像与原始样本图像有相同的目标函数值,仍能有效地识别,具有鲁棒性。

2.2 基于图像空间信息的快速学习改进算法

设 \mathbf{X} 是 DBN 系统中训练样本图像中的 n 幅同类图像,像素为 $m \times p$,可视节点的偏移量 a ,隐含层节点的偏移量 b 。在 T 时间段内,读取 n 幅训练图像,并将其一起输入 DBN 系统时,所有图像的像素点都放入一个大矩阵中。这个图像矩阵记为:

$$\begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix} = \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,mp} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,mp} \\ \vdots & \vdots & \vdots & \vdots \\ x_{n,1} & x_{n,2} & \cdots & x_{n,mp} \end{bmatrix} \quad (17)$$

n 幅图像矩阵可以看作是一条马尔科夫链,且既没有周期性又不可约,因此是各态遍历的,即不论初始时矩阵元素取各值的概率是多少,随着转移次数的增多,取值分布最终会收敛于唯一的平稳分布。将式(14)展开为

$$\ln p(x) = \ln \left(\sum_h e^{-E(x,h)} \right) - \ln \left(\sum_{x,h} e^{-E(x,h)} \right) \quad (18)$$

对式(18)左右两边进行求导,有

$$\frac{\partial \ln p(x)}{\partial \theta} = -\sum_h p(h|x) \frac{\partial E(x,h)}{\partial \theta} + \sum_{x,h} p(x,h) \frac{\partial E(x,h)}{\partial \theta} \quad (19)$$

由式(1)和式(19)得

$$\begin{aligned} \frac{\partial \ln p(x)}{\partial w_{ij}} &= -\sum_h p(h|x) \frac{\partial E(x,h)}{\partial w_{ij}} + \sum_{x,h} p(x,h) \frac{\partial E(x,h)}{\partial w_{ij}} = \\ &= \sum_h p(h|x) x_i h_j - \sum_{x,h} p(x,h) x_i h_j \\ &= \sum_h p(h|x) x_i h_j - \sum_x p(x) \sum_h p(h|x) x_i h_j = \\ &= p(h_j = 1|x) x_i - \sum_x p(x) p(h_j = 1|x) x_i \end{aligned} \quad (20)$$

要求式(20)的第二项,必须遍历所有可能的 x

值, 为了方便, 一般对每个训练样本 x_i , 都用某种抽样方法(常用的是 Gibbs 采样)抽取一个它对应的样本 y_i , 进而计算 RBM 的参数值 w 。那么式(20)近似为

$$\frac{\partial \ln p(x)}{\partial w_{ij}} = p(h_j = 1 | x)x_i - \frac{1}{n} \sum_{i=1}^n p(h_j = 1 | y_i)y_i \quad (21)$$

根据概率论, 一个各态历经性的平稳随机过程, 多个样本函数在某一时间点的统计平均值等于一个样本函数在整个时间轴上的平均值, 即统计平均值=时间平均值。即

$$\sum_{i=1}^n p(x_i)x_i = \frac{1}{T} \int_0^T x(t)dt \quad (22)$$

式(21)中的第 2 项等价于求解 n 幅采样后的样本图像的统计平均值。利用时间平均值代替统计平均值, 进而实现深度信念网络的快速学习。

设 $0-T$ 时间段内共输入 n 幅训练图像, 则可以

计算出所有样本图像的时间平均值。下面将式(22)代入式(21), 表达式近似为

$$\frac{\partial \ln p(x)}{\partial w_{ij}} = p(h_j = 1 | x)x_i - \frac{1}{nT} \int_0^T y(t)dt \quad (23)$$

积分等价于求和, 则式(23)化简为

$$\frac{\partial \ln p(x)}{\partial w_{ij}} = p(h_j = 1 | x)x_i - \frac{1}{n} \sum_{i=1}^n y_i \quad (24)$$

其中, $y_i(i=1,2,\dots,n)$ 为输入 DBN 系统的 n 幅训练样本经过采样后的样本图像。根据 2.1 节分析可知, 改变式(17)中像素的位置, 不影响训练结果。我们称这种利用时间平均代替统计平均, 基于图像空间信息的线性叠加合成同类样本的快速学习改进算法为 LSMI(Linear Superposition Multiple Images) 算法。

在 LSMI 算法的基础上, 建立基于空间信息的 DBN 图像分类系统, 如图 3 所示。

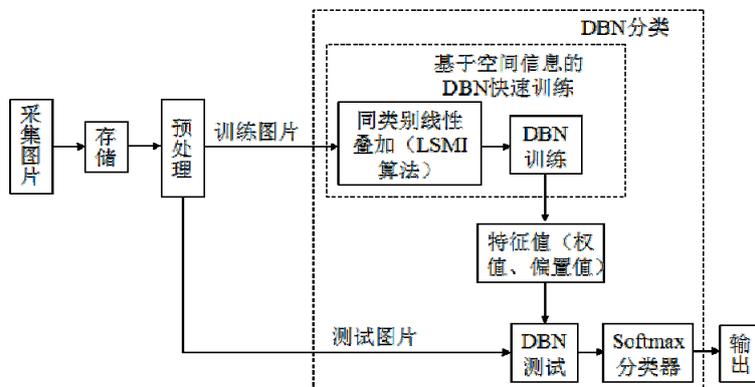


图 3 基于空间信息的 DBN 图像分类系统

针对不同的训练样本, 利用图 3 的模型, LSMI 算法一般是通过实验的方法选择恰当的幅数, 进而在保证有效识别的基础上, 加快 DBN 的收敛速度。

假设同类训练样本图像共有 m 张, 训练过程中需要对 m 张图像进行处理, 但如果利用 LSMI 算法, 将其中的 n 张图像进行线性叠加合成, 则训练过程相当于是对 m/n 张图像进行处理。这样的话, 训练时间会大大缩减。因此, 利用 LSMI 算法可以实现深度信念网络的快速学习。

2.3 LSMI 算法收敛的衡量指标

很明显, 训练时间复杂度是 LSMI 算法相比于其他算法, 速度收敛快慢的衡量指标。下面从中心极限定理角度证明 LSMI 算法收敛速度的另一个衡量指标, 即权重矩阵收敛于正态分布的趋势。

假设深度信念网络第 i 层的输入节点数是 p , 输入的待训练图像记为 X , 大小为 $m \times p$, 第 $i+1$ 层的输出节点数是 q , 输出的矩阵记为 Y , 大小为

$m \times q$ 。已知 DBN 每一层中的所有节点之间都是条件独立的。根据基本的神经元模型知, 输入、输出之间的关系可用数学式表达出来:

$$Y = \varphi(XW + \Theta) \quad (25)$$

其中: W 是连接相邻两层之间的权重; Θ 为输出层的阈值。假设激活函数的逆函数表示为 $\varphi^{-1}(\cdot)$, 矩阵 X 的逆矩阵为 T , 即 $T = \text{dinv}(X)$, 则矩阵 W 为

$$W = T \cdot [\varphi^{-1}(Y) - \Theta] \quad (26)$$

设 $S = \varphi^{-1}(Y) - \Theta$, 则权值系数矩阵 W 中的每个元素可以表示为

$$w_{ij} = \sum_{l=1}^m t_{il} \cdot s_{lj} \quad (27)$$

由于 X 中的元素相互独立, 则矩阵 T 中的元素相互独立, 同理, S 中的元素也相互独立, 因此, $t_{i1} \cdot s_{1j}, t_{i2} \cdot s_{2j} \cdots t_{ip} \cdot s_{pj}$ 之间也是相互独立的随机变量, 即权值系数矩阵 W 中的每个元素 w_{ij} 是由 m 个随机独立的变量的和组成。基于 DBN 每一层的输出作为下一层的输入, 即

$$Y_1 = \varphi(YW_1 + \Theta_1) \quad (28)$$

$$W_1 = \varphi^{-1}(XW + \Theta) \cdot [\varphi^{-1}(Y_1) - \Theta_1] \quad (29)$$

随着层数的增加, DBN 越来越收敛, 权值系数矩阵中独立随机变量的个数越来越多。根据中心极限定理, 独立随机变量和的极限分布为正态分布。因此随着层数的增加, DBN 越来越收敛, 权值系数矩阵中的元素由越来越多的独立随机变量的和组成, 即权重必将越来越趋于正态分布。

因此, 通过 LSMI 算法, 将多幅图像合成后系统达到稳态的收敛速度可以通过权值系数矩阵趋于正态分布的趋势来衡量。权重趋于正态分布的趋势越来越快, DBN 系统收敛速度越快, 算法的学习速度更快。

3 算法分析与仿真结果

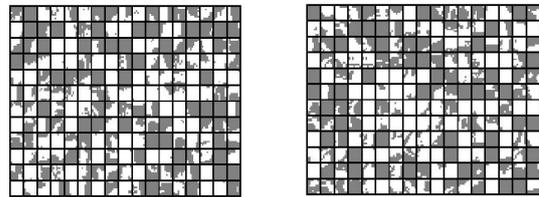
3.1 图像空间信息的仿真研究

本仿真实验以图 3 为系统, 选取 ORL 人脸库中的一幅图像作为仿真对象, 验证 2.1 节提出的深

度信念网络的训练图像与空间信息无关的特点。仿真步骤如下:

随机初始化深度信念网络的权值初始矩阵, 选取 ORL 人脸库中的一幅图像进行训练, 图 4(a) 为 204 个采样出的人脸小 patch。保存收敛后得到的 4 层权值系数矩阵 w_1, w_2, w_3, w_4 ;

旋转(1)中的人脸图像, 使人脸在图片中的空间位置发生改变, 图 4(b) 为 204 个采样出的人脸小 patch。注意保证此次训练过程中初始权重矩阵与(1)中的保持一致, 并保存收敛后得到的 4 层权值系数矩阵 $w_{11}, w_{22}, w_{33}, w_{44}$;



(a) 旋转前的小 patch (b) 旋转后的小 patch

图 4 比较两次实验的 204 个小 patch

通过比较(1), (2)两次实验中的权值系数矩阵, 来验证识别效果与图像的空间位置信息是否相关。仿真结果如表 1 所示。

表 1 比较两次实验的权值系数矩阵

权重矩阵	均值	方差
w_1-w_{11}	-1.2321×10^{-4}	2.6112×10^{-5}
w_2-w_{22}	7.0225×10^{-4}	1.0529×10^{-7}
w_3-w_{33}	6.7168×10^{-4}	8.6171×10^{-9}
w_4-w_{44}	6.5469×10^{-4}	7.2484×10^{-9}

由表 1 可知, 旋转图像前后深度信念网络的权值系数矩阵仅有微小变化, 均值接近于 0, 且识别率一致, 为 94.3996%。因此对于图像旋转平移等空间位置的改变, 不影响深度信念网络的有效识别, 即与空间信息无关。

3.2 LSMI 算法的仿真研究

3.2.1 COREL 库仿真及实验结果

本实验选取场景图像分类 COREL 库中的 300 张图像(包括 3 个主题: 花、汽车、飞机, 每个主

题包含不同背景的 100 幅图像)作为仿真对象, 以 Matlab R2010b 为运行环境, 以图 3 为系统, 对基于 DBN 的 3 类图像进行分类识别。将本文提出的 LSMI 算法与文献[2]中的传统数据降维算法以及文献[19]提出的基于自适应步长的改进算法进行比较。其中, 文献[19]的改进算法是通过误差函数的梯度变化, 自适应调节每一个权值对应的步长, 加快 DBN 的收敛速度, 进而达到快速学习的目的。调整步长的公式为

$$\eta_{ij} = \begin{cases} u\eta_{ij}^{old}, (\langle v_i h_j \rangle_0 - \langle v_i h_j \rangle_k) \\ \quad (\langle v_i h_j \rangle_0^{old} - \langle v_i h_j \rangle_k^{old}) > 0 \\ d\eta_{ij}^{old}, (\langle v_i h_j \rangle_0 - \langle v_i h_j \rangle_k) \\ \quad (\langle v_i h_j \rangle_0^{old} - \langle v_i h_j \rangle_k^{old}) < 0 \end{cases} \quad (30)$$

3 种算法的正确识别率和时间复杂度仿真结果分别如表 2、表 3 所示。

表 2 算法的正确识别率仿真表 /%

正确识别率	1 轮学习	2 轮学习	3 轮学习
传统算法	90.00%	93.333 3%	93.333 3%
自适应步长算法	93.333 3%	90.00%	93.333 3%
LSMI 算法	90.00%	90.00%	93.333 3%

表 3 算法的时间复杂度仿真表 /s

时间	图像预处理	训练图像
传统算法	8.041 372	308.120 236
自适应步长算法	8.355 101	294.642 161
LSMI 算法	10.545 219	179.266 538

本次仿真实验中, LSMI 算法是将同类别的 2 幅样本图像进行线性叠加合成。文献[19]中的自适应步长算法是在保证识别率与传统算法识别率一致的前提下, 尽量减少迭代次数, 进而与传统算法及 LSMI 算法比较算法的时间复杂度。且传统算法、自适应步长算法以及 LSMI 算法的迭代次数分别设置为 60 次、27 次、60 次。表 3 中的图像预处理指的是读取样本图像并对其处理的时间。

由表 2、表 3 可知, LSMI 算法和文献[2]中的传统算法相比, 图像的识别率基本保持一致, 从算法时间复杂度角度相比, 尽管图像预处理的时间增加了 2.503 847 s, 但训练图像的时间缩短了 128.853 698 s。文献[19]中的自适应步长算法尽管迭代次数减小了一多半, 但每次迭代时计算每个权值矩阵元素的步长耗费时间长, 因此, 其和 LSMI 算法相比, 训练样本的时间也较长。

由 2.3 知, 权值系数矩阵收敛于正态分布的速度是衡量深度信念网络学习算法快慢的指标。下面对三种算法训练结束后的权值系数矩阵进行研究。传统算法、自适应步长算法和 LSMI 算法的第一层权重值的分布情况分别如图 5~7 所示。其中, (a), (b) 分别为算法对应的直方图和正态概率纸检验函数。正态概率纸检验函数指的是权重矩阵的概率密度分布(粗线条)与正态分布(细虚线)的趋近关系, 位于细虚线上的数据点越多, 权重矩阵越趋于正态分布。

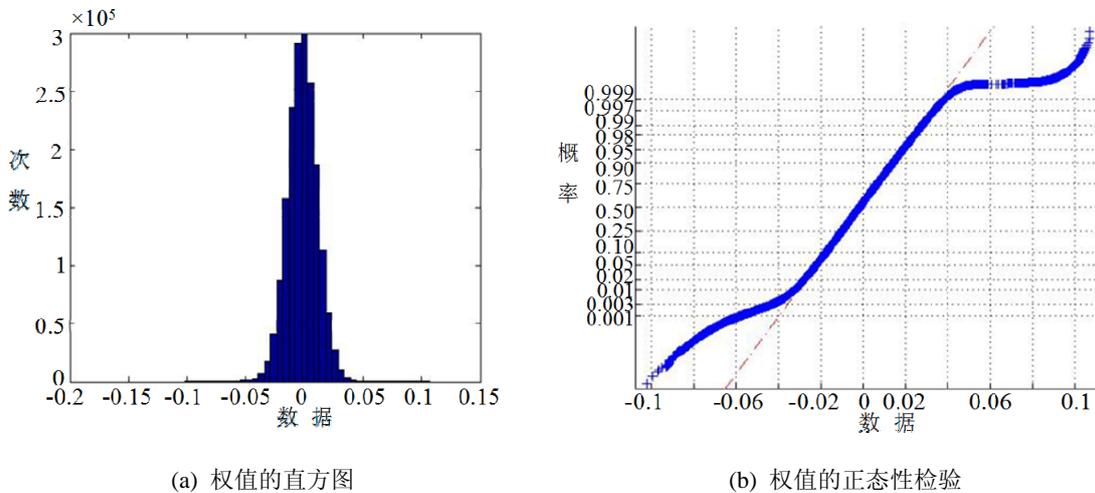


图 5 传统算法的权值分布情况

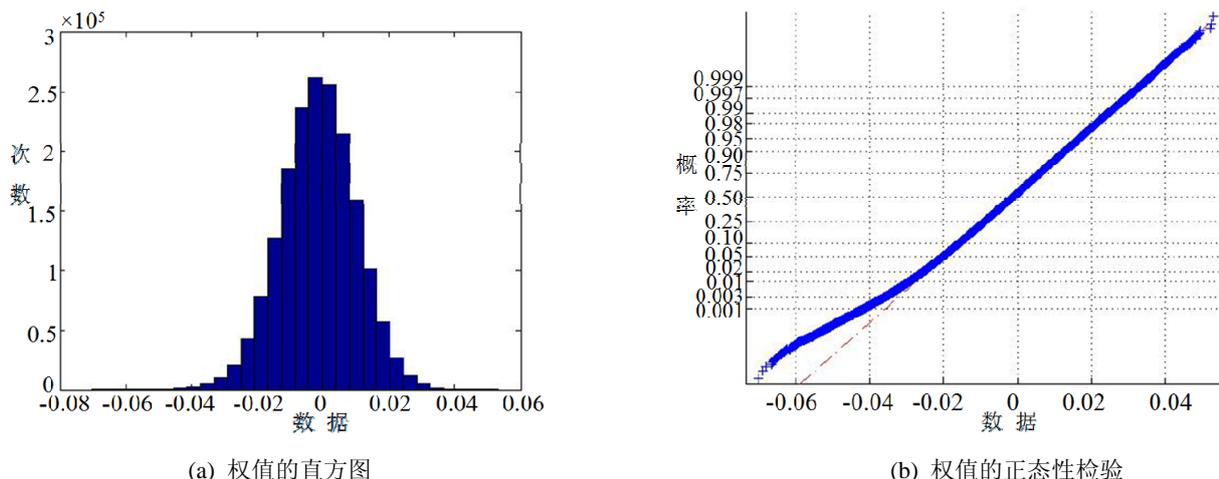


图 6 自适应步长算法的权值分布情况

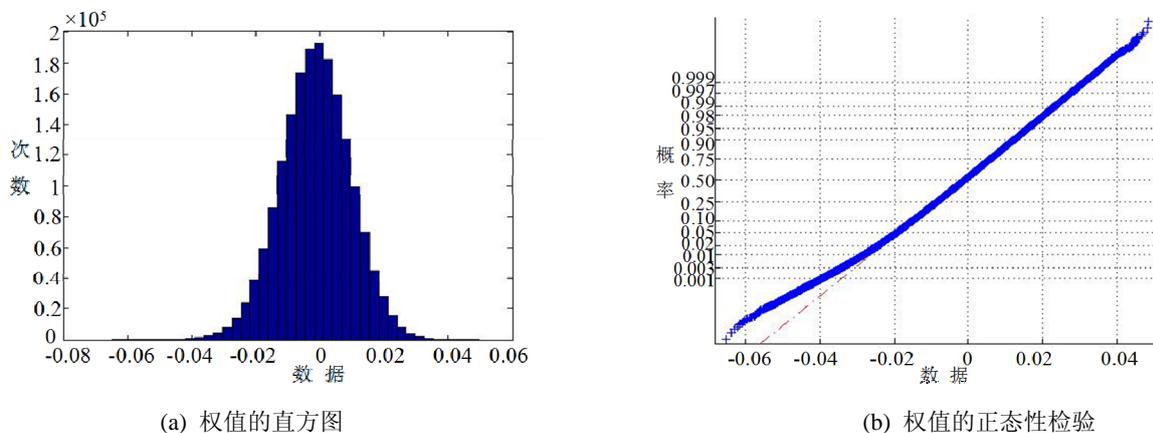


图 7 LSMI 算法的权值分布情况

由图 5、图 6、图 7 可知，不论是从直方图还是正态概率纸检验函数来看，图 6 和图 7 与图 5 相比，权重矩阵明显更趋于正态分布；图 7 和图 6 相比，权重矩阵趋于正态分布的程度稍好。因此，本文提出的 LSMI 算法和文献[19]中的基于自适应步长的 DBN 改进算法相比于传统算法，收敛速度明显加快，且 LSMI 算法效果稍好一些。

综上，以 COREL 库为依据，从识别率、时间复杂度、权重矩阵趋于正态分布的趋势等三个角度证明了本文提出的 LSMI 算法相比于传统算法和自适应步长算法的优越性。在保证正确识别率的基础上可以实现快速学习的目的。

3.2.2 MIT 人脸库仿真及实验结果

本仿真实验选取 MIT 人脸数据库中的 4 000

幅图像(包括 2 000 幅人脸图像和 2 000 幅非人脸图像)，并按训练样本与测试样本 9:1 的比例随机选取进行仿真研究。以 Matlab R2010b 为运行环境，以图 3 为系统进行分类识别，进一步验证 LSMI 算法的有效性、快速性。同样，将 LSMI 算法与文献[2]中的传统数据降维算法和文献[19]中的自适应步长算法进行比较。算法的正确识别率和时间复杂度比较，分别如表 4、表 5 所示。

正确识别率	1 轮学习	2 轮学习	3 轮学习
传统算法	83.25%	83.50%	83.75%
自适应步长算法	83.50%	83.50%	83.25%
LSMI 算法-1	82.75%	83.00%	82.75%
LSMI 算法-2	82.75%	82.50%	82.50%
LSMI 算法-3	82.50%	82.75%	83.00%

表 5 算法的时间复杂度仿真表 /s

时间	图像预处理	训练图像
传统算法	112.349 642	1125.848 530
自适应步长算法	108.460 529	876.307 486
LSMI 算法-1	133.744 099	535.263 103
LSMI 算法-2	128.395 074	226.349 869
LSMI 算法-3	121.868 517	112.888 312

表 4~5 中的 LSMI 算法-1、LSMI 算法-2 和 LSMI 算法-3 分别是将同类别的 3, 5 和 10 幅图像进行线性叠加合成。由表 4~5 可知, LSMI 算法相比于传统算法和自适应步长算法, 识别率基本一致, 但明显缩短了算法的时间复杂度。

尽管同类图像合成的幅数越多, 训练时间越短, 但并不是合成的图像幅数越多越好。如利用 LSMI 算法合成图像前后, 权重值有明显的变化, 则会降低识别率。造成识别率下降的原因有待进一步研究。

综上, 通过 COREL 库和 MIT 人脸库仿真实验表明 LSMI 算法可以加快 DBN 的收敛速度, 降低算法的时间复杂度, 提高学习效率, 从而验证了本文提出的算法的正确性。且 LSMI 算法和文献 [19] 中的自适应步长算法都有加快学习的目的, 只是从不同的角度着手。基于自适应步长的 DBN 改进算法通过减小迭代次数加快收敛, 适合于迭代次数较多的仿真实验; 而本文的 LSMI 算法是在保证特征点不变的前提下, 减少训练样本的数目进而达到快速学习的目的, 适合于大训练样本的仿真实验。随着大数据时代的到来, LSMI 算法将会显示出其优越性。

4 结论

本文利用深度信念网络分类模型构建图像分类识别系统。由于 DBN 分类模型中的 DBN 算法存在训练过程中收敛速度慢、学习效率等缺点, 本文基于 DBN 的训练样本图像无空间信息的特点, 根据各态遍历性, 提出了线性叠加合成的深度信念网络图像分类快速学习改进算法—LSMI 算法。理论分析和仿真结果表明: (1)深度信念网络的样本

图像与空间结构信息无关。(2)DBN 各层的权值矩阵趋于正态分布的程度可以作为衡量 DBN 收敛速度的指标。(3)以 COREL 场景分类库和 MIT 人脸库为仿真对象, 分析比较了文献[2]、文献[19]以及本文提出的 LSMI 算法, 从正确识别率、算法的时间复杂度以及权重趋于正态分布的趋势三个角度验证了 LSMI 算法的有效性、快速性。因此, LSMI 算法可为深度信念网络的进一步研究、发展和应用提供一定的理论依据。

参考文献:

- [1] 赵国荣, 王希彬, 高青伟. 粒子群优化小波神经网络在馈导线系统传递对准中的应用 [J]. 系统仿真学报, 2010, 22(3): 670-673.
- [2] 董天祯, 郭江鸿, 吕娟, 等. 基于神经网络的输电线路故障检测研究 [J]. 系统仿真学报, 2009, 21(15): 4903-4911.
- [3] 石东源, 熊国江, 陈金福, 等. 基于径向基函数神经网络和模糊积分融合的电网分区故障诊断 [J]. 中国电机工程学报, 2014, 34(4): 562-569.
- [4] 侯润民, 刘荣忠, 高强, 等. 粒子群小波神经网络在交流伺服系统中的应用 [J]. 系统仿真学报, 2014, 26(4): 881-886.
- [5] 刘建伟, 刘媛, 罗雄麟. 玻尔兹曼机研究进展 [J]. 计算机研究与发展, 2014, 51(1): 1-16.
- [6] Bengio Y, Lamblin P, Popovici D, et al. Greedy Layer-wise Training of Deep Networks [C]// Advances in Neural Information Processing Systems 19 (NIPS'06). Cambridge, MA, USA: MIT Press, 2007: 153-160.
- [7] Hadsell R, Chopra S, Le Cun Y. Dimensionality Reduction by Learning an Invariant Mapping [C]// IEEE Computer Society Conference on Computer Vision and Pattern Recognition (S10636919), New York, USA. USA: IEEE, 2006: 1735-1742.
- [8] 赵志宏, 杨绍普, 马增强. 基于卷积神经网络 LeNet-5 的车牌字符识别研究 [J]. 系统仿真学报, 2010, 22(3): 638-641.
- [9] Schulz H, Miiller A, Behnke S. Investigating Convergence of Restricted Boltzmann Machine Learning [C]// NIPS 2010 Workshop on Deep Learning and Unsupervised Feature Learning. Canada: NISP, 2010: 1-9.
- [10] 陈硕. 深度学习神经网络在语音识别中的应用研究 [D]. 广州: 华南理工大学, 2013.
- [11] Ji Nannan, Zhang Jiangshe, Zhang Chunxia. A

- Sparse-Response Deep Belief Network Based on Rate Distortion Theory [J]. *Pattern Recognition (S0031-3203)*, 2014, 47(9): 3179-3191.
- [12] Lin Miaozen, Xin Fan. Low Resolution Face Recognition with Pose Variations Using Deep Belief Networks [C]// 2011 4th International Congress on Image and Signal Processing. Shanghai, China: CISP, 2011: 1522-1526.
- [13] Lopes N, Ribeiro B, Goncalves J. Restricted Boltzmann Machines and Deep Belief Networks on Multi-core Processors [C]// 2012 International Joint Conference on Neural Networks. Brisbane, Australia: IJCNN, 2012: 1-7.
- [14] Raina R, Madhavan A, Ng A. Large-scale Deep Unsupervised Learning Using Graphics Processors [C]// Proceedings of the 26th International Conference on Machine Learning, ICML 2009. Montreal, Canada: ICML, 2009: 873-880.
- [15] G E Hinton. Training Products of Experts by Minimizing Contrastive Divergence [J]. *Neural Computation (S0899-7667)*, 2002, 14(8): 1771-1800.
- [16] Cho K, Raiko T, Ilin A. Parallel Tempering is Efficient for Learning Restricted Boltzmann Machines [C]// 2010 International Joint Conference on Neural Networks. New York, USA: IJCNN, 2010: 1-8.
- [17] Tieleman T. Training Restricted Boltzmann Machines Using Approximations to the Likelihood gradient [C]// Proceedings of the 25th international conference on Machine learning. Canada: ICML, 2008: 1064-1071.
- [18] Cho K, Raiko T, Ilin A. Enhanced Gradient and Adaptive Learning Rate for Training Restricted Boltzmann Machines [C]// Proceedings of the 28th International Conference on Machine Learning. New York, USA: ICML, 2011: 105-112.
- [19] Lopes N, Ribeiro B. Improving Convergence of Restricted Boltzmann Machines via a Learning Adaptive Step Size [C]// Proceedings off the 17th Iberoamerican Congress on Pattern Recognition CIARP 2012. Berlin, Germany: Springer, 2012: 511-518.
- [20] Lopes N, Ribeiro B. Towards Adaptive Learning with Improved Convergence of Deep Belief Networks on Graphics Processing Units [J]. *Pattern Recognition (S0031-3203)*, 2014, 47(1): 114-127.
- [21] Martens J. Deep Learning Via Hessian-free Optimization [C]// Proceedings of the 27th International on Machine Learning, ICML 2010. Haifa, Israel: ICML, 2010: 735-742.
- [22] 郑志蕴, 李步源, 李伦, 等. 基于云计算的受限玻尔兹曼机推荐算法研究 [J]. *计算机科学*, 2013, 40(12): 259-263.
- [23] 吴川. 基于神经网络的目标识别及定位方法的研究 [D]. 长春: 中国科学院长春光学精密机械与物理研究所, 2005.

《系统仿真学报》荣获“2014 中国最具国际影响力学术期刊”证书

由中国学术期刊（光盘版）电子杂志社与清华大学图书馆联合成立的中国学术文献国际评价研究中心，以美国汤森路透 Web of Science 收录的 1.2 万余种期刊为引文统计源，首次研制发布了 2012《中国学术期刊国际引证年报》（CAJ-IJCR 年报）。第一次给出了我国 5600 余种中外文学术期刊总被引频次、影响因子、半衰期等各项国际引证指标，并采用了新的国际影响力综合评价指标 CI 对期刊排序，发布了“中国最具国际影响力学术期刊”（排序 TOP5%）和“中国国际影响力优秀学术期刊”（排序 TOP5-10%），在国内外学术界产生了较大反响。之后，2013 年版年报，将引文统计源期刊扩展到 1.44 万多种。目前，2014 版国际、国内年报与 TOP5% 和 TOP5-10% 期刊的遴选业已完成，《系统仿真学报》被列入“2014 中国最具国际影响力学术期刊”行列。

我学报连续 2 年被列入 TOP5% 国内一流的中国最具国际影响力学术期刊，走向世界，进入国际一流，指日可待！