



基于 A-ELM 的移动视觉搜索方法

胡海洋^{1,2}, 许军^{1,2}, 胡华^{1,2}

(1.杭州电子科技大学计算机学院, 浙江 杭州 310018;

2.杭州电子科技大学复杂系统建模与仿真教育部重点实验室, 浙江 杭州 310018)

摘要: 计算机智能技术在图像领域已经得到广泛的应用。极限学习机(ELM)作为一种新兴技术, 克服了其他传统智能技术所面临的一些问题, 吸引了越来越多研究人员的关注。首先对 ELM 算法的性能进行了分析验证, 并将其延伸到图像分类搜索上。在此基础上, 提出了基本视觉搜索(BMVS)框架, 将 ELM 运用到此框架服务器端, 并进一步优化了 ELM 的分类性能。最后实验证明 ELM 在移动视觉搜索方面的可行性, 并通过和支持向量机(SVM)的实验对比验证相关方法的高效性。

关键词: 分类; 极限学习机; 移动视觉搜索

中图分类号: TP311

文献标识码: A

doi: 10.11959/j.issn.1000-0801.2016088

Mobile visual searching method based on ascending extreme learning machine

HU Haiyang^{1,2}, XU Jun^{1,2}, HU Hua^{1,2}

1.School of Computer Science and Technology, Hangzhou Dianzi University, Hangzhou 310018, China

2.Key Laboratory of Complex Systems Modeling and Simulation, Ministry of Education,

Hangzhou Dianzi University, Hangzhou 310018, China

Abstract: Computer intelligence technology had been widely used in the field of image searching. Extreme learning machine has emerged as a new technology which overcomes the problems in traditional intelligent field and it has attracted more and more researchers. The algorithm performance of ELM was analyzed firstly, extending the method to image classification field. A basic mobile visual searching (BMVS) framework was proposed which applies ELM to image searching and optimizes the performance of ELM. Finally, the experiment proves the effectiveness of the method proposed by using ELM for the mobile vision searching. Through the experiments of comparison with SVM-based methods, the efficiency of the method proposed was also confirmed.

Key words: classification, extreme learning machine, mobile visual searching

收稿日期: 2015-04-16; 修回日期: 2016-03-02

基金项目: 国家自然科学基金资助项目(No.61572162, No. 61272188); 南京大学计算机软件新技术国家重点实验室开放基金资助项目(No.KFKT2014B15); 浙江省哲学社会科学重点研究基地(信息化与经济社会发展研究中心)课题资助项目(No.14JDXX04YB)

Foundation Items: The National Natural Science Foundation of China (No. 61572162, No. 61272188), The Foundation of State Key Laboratory for Novel Software Technology of Nanjing University (No. KFKT2014B15), The Foundation of Key Research Base for Philosophy and Social Sciences of Zhejiang Province (Research Center of Information Technology & Economic and Social Development)(No. 14JDXX04YB)

1 引言

随着互联网和多媒体技术的飞速发展,声音、图像、视频等数字信息急速膨胀。图像作为一种变现直观的多媒体显示信息,受到人们的重视和青睐。在现实社会应用中有大量的图像产生,如何快速有效地在图像库中匹配检索出需要用到的信息,一直是多媒体以及图像检索研究领域的热点问题,多媒体图像检索不仅可以帮助普通用户寻找到想要的多媒体内容,还可以涉及移动视觉搜索等多个领域。

图1为移动视觉搜索示例,通过移动端的拍摄,可以在很短的时间内搜索到相同或相近的图片,并返回相关信息。图1(a)为拍摄一本软件测试书籍后通过检索反馈此书籍的相关信息;图1(b)为拍摄一张海报,最终显示出有关此海报的内容。与简单匹配不同的是,移动视觉搜索是基于算法和数据的,需要搜索框架拥有大量的图像库,之后根据图像处理方法得到图像描述特征,在特征值的基础上进行聚类分析或分类识别。它需要将图像一步步进行解析、提取特征、检索模型、识别等,最后根据检索算法得出和原图类似或相关的图像。

移动视觉搜索是指将移动终端获取的真实世界中的图像或视频作为查询对象,通过移动互联网去搜索视觉对象的关联信息的检索方式^[1]。移动视觉搜索是基于图像模式识别的基础,其中移动视觉搜索应用系统包括两个基本的内容:图像匹配和图像检索。在众多多媒体图像检索中,基于内容的图像检索技术^[2]成为了近年来图像检索的热点,其中内容有使用颜色特征、纹理特征、形状特征等。传统的一般使用前馈神经网络来训练图像特征数据,其中单隐层前馈神经网络(single-hidden layer feed forward network, SLFN)被广泛地使用。但是在之前的理论研究工作中,输入权重和隐含层阈值总是需要调节的。且

传统的SLFN主要使用基于梯度下降方法(如BP(back propagation)神经网络^[3])、标准的优化约束方法(如SVM(support vector machine)^[4])、最小二乘法(RBF(radial basis function)网络^[5])等方法来训练数据,需要多次迭代,且容易陷入局部最优解。

极限学习机(extreme learning machine, ELM)^[6-14]是由Huang等人最初提出适用于SLFN的一种简单的学习算法;不同于传统的学习算法(如BP学习算法),ELM不仅能达到最小的训练错误,同时致力于达到最小的标准权重;它避免了传统的SLFN(single-hidden layer feedforward)学习方法收敛速度慢以及容易陷入局部最优解的可能。本文在深入了解ELM算法的基础上,对ELM算法思想和算法公式进行了仔细的研究分析,并且验证分析了ELM的分类效果,然后设计ELM分类器使之进一步运用到图像搜索方面;同时在给出ELM基本概念的基础上,提出了一个基本移动视觉搜索(basic mobile visual searching, BMVS)框架来解决视觉搜索中对于网络带宽的依赖,该框架采用ELM分类器进行训练测试数据。通过这个框架上传下载图片相关信息时候,不用频繁地依赖网络,只需交互一次即可针对搜索的图像进行处理。

2 相关工作

参考文献[1]对图像检索的过去、现在和将来进行了分析,对基于内容的图像检索的主要研究就似乎进行了详细和全面的论述,介绍了典型的基于内容的图像检索系统。最后指出了现存的问题和今后的研究方向。参考文献[6-14]提出了一个新的学习算法——ELM(极限学习机),该算法由新加坡南洋理工大学的黄广斌教授提出,它是一种主要针对SLFN(单隐层前馈神经网络)的监督型学习算法。参考文献[12]在讲解了ELM前对传统的神经网络算法进行了分析,如SVM、PSVM、LS-SVM等,并且在目标函数、稳定



(a)



(b)

图1 移动视觉搜索示例



性、相似度匹配、学习速度、泛化性能等方面进行了实验对比分析。参考文献[13]重点对单隐层前馈神经网络的通用逼近性能进行了分析,并且在此基础上提出了一种 I-ELM 的单隐含层神经网络的增量学习算法。之后重点介绍了黄广斌在 ELM 的基础上进行的一些推广,如将 ELM 推广到复数域,提出基于 ELM 的在线时序算法等。参考文献[15]提出了一个新的多任务特征选择提取算法。它是以批处理的方式选取每个相关的特征,取代原先要单独评估每个特征。该算法基于一个假设:不同相关的任务拥有共同的结构,不同任务的多特征功能在连接架构中能够被同时学习,这使得该算法可以将多任务的共同理论作为补充信息来促进决策,同时提出了一个有效的迭代算法来促进优化从而保证收敛性。参考文献[16]提出了一种新的模式识别算法,称为遗传主成分分析 (genetic principal component analysis, GPCA),它是结合了遗传算法和主成分分析,初始由特征值和特征向量来构成特征空间。通过特征空间,使用欧式距离来对输入图片进行分类。最后通过在 ORL (olivetti research laboratory, 人脸数据库) 中进行实验来证明提出的方法优于之前的人脸识别方法。

基于以上分析,传统的方法如神经网络或者 SVM 等,都面临一些挑战问题,如学习速度慢、容易陷入局部最优解和计算顽健性低等。而极限学习机作为一个新兴的技术,克服了上述传统技术所面临的一些问题,在性能和可扩展性方面均能达到很好的效果。

3 极限学习机

极限学习机最初提出应用于单隐藏层前馈神经网络,然后延伸到了泛化 SLFN。给定 N 个不同的数据样本 $\{(x_i, t_i)\}_{i=1}^N \subset R^n \times R^m$, 且 L 个隐层神经元的单隐层神经网络可以零误差逼近这 N 个互异的数据样本,则网络模型为:

$$f_L(X_j) = \sum_{i=1}^L \beta_i h_i(x) \\ = \sum_{i=1}^L \beta_i G(a_i \cdot x_j + b_i) = t_j, j=1, \dots, N \quad (1)$$

其中, a_i, b_i 分别为输入权值和隐层阈值, β 为第 i 个隐层节点到 ELM 网络输出节点之间的权值, $h(x_i) = [G(a_1, b_1, x_i), \dots, G(a_L, b_L, x_i)]^T$ 为隐藏层关于输入 x_i 的输出向量,它的作用是将样本 x_i 由 d 维输入空间映射到 L 维的特征空间。

式(1)可简记为: $H \cdot \beta = T$, 其中 $H(a_1, \dots, a_L, b_1, \dots, b_L, x_1, \dots,$

$$x_L) = \begin{pmatrix} G(a_1, b_1, x_1) & \cdots & G(a_L, b_L, x_1) \\ \vdots & \ddots & \vdots \\ G(a_1, b_1, x_N) & \cdots & G(a_L, b_L, x_N) \end{pmatrix}_{N \times L}$$

为神经网络隐层输出矩阵。矩阵中第 i 行为输入向量 x_i 在各个隐层节点的输出,第 j 列为第 j 个隐层节点对应的输入向量 $x_1, x_2, \dots,$

$$x_L \text{ 的输出向量, } \beta = \begin{pmatrix} \beta_1^T \\ \vdots \\ \beta_L^T \end{pmatrix}_{L \times M}, T = \begin{pmatrix} t_1^T \\ \vdots \\ t_N^T \end{pmatrix}_{N \times M}。然而由于多$$

数情况下, $L \ll N$ 。从而使得训练样本的网络输出和实际输出之间的误差也随之产生: $H\beta = T + E$, 即 $E = \sum_{j=1}^L (\beta_j G(a_j, b_j, x_j)) - t_j, j=1, \dots, N$ 。

即网络参数的训练问题转化成了最小化平方损失函数的问题,即寻找最小二乘解 $\hat{\beta}$, 使得: $\|H\hat{\beta} - T\| = \min_{\beta} \|H\beta - T\|$, 利用广义逆可得:

$$\beta = H^+ T \quad (2)$$

其中, $H^+ = (H^T H)^{-1} H^T$ 。

4 基本移动视觉搜索框架

为了更好地分析和测试 ELM 在多分类情况下的泛化性,同时能更加符合移动视觉搜索的应用所需,本文在移动视觉领域常用三大框架模型的基础上,构建了自己的移动视觉搜索框架模型——BMVS 框架,其模型如图 2 所示。

在图 2 所示的框架中,移动端采集图像,不需要对图像进行相关处理,只需收集整理图像并使用网络传输图像到服务器端数据库,服务器数据库端会自动完成图像的相关处理任务(图片接收、特征提取、特征训练测试、显示结果),最后移动端接收识别结果即可。此外,移动端也可下载该分类器和配置信息,便能对相关图像进行匹配和检索。整个模型中,只需上传下载过程中需要网络链接,大大减少了对网络带宽等的需求。在这一过程中,ELM 分类器构造识别包括两个阶段:离线阶段,将图像库特征提取后,同类别标签一起输入 ELM 分类器中,构造一个 ELM 分类器;在线阶段,首先提取图像的特征,然后送入 ELM 分类器中进行判别,输出识别结果和信息、资源等。由于服务器端要对图片完成特征训练阶段,所以使用 ELM 算法来训练分类器。为了提高训练准确度,本文对提取的特征经过相关预处理过程,包括 PCA 降维、归一化、数据标准化等 4 个步骤。

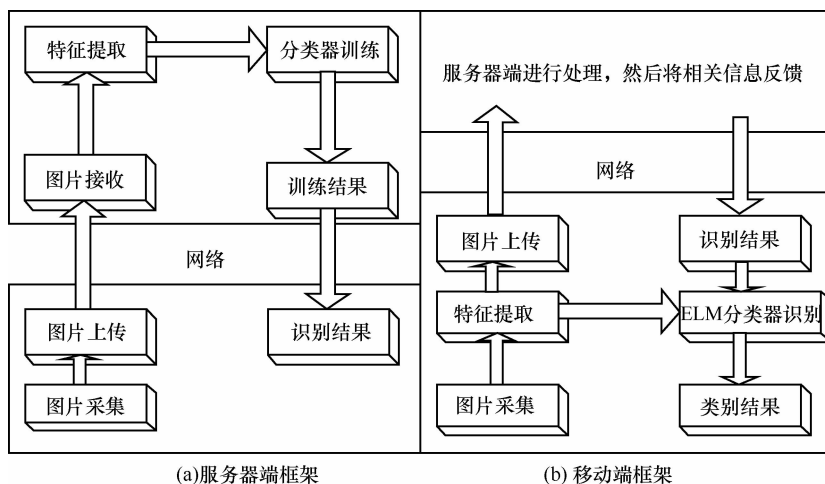


图2 基本移动视觉搜索框架

本文首先对接收的图像进行 RGB 提取,接着进行 HSV 转化,即对转化的 3 种特征通道进行 HSV 特征量化(H 通道量化 16 级,S 通道和 V 通道量化成 4 级),再将 HSV 分量转换成一个一维数组,并对其进行特征直方图提取。在此基础上本文使用 PCA 方法^[16]来对提取的图像特征进行降维操作,其原理就是将原来的数据投影到一个新的空间,新的空间下数据是原样本的一个最大线性无关组特征值所对应的空间数据,本文所实现的相关算法伪码如下。

步骤 1 对于 m 个训练样本图片进行上面的特征提取转换,第 i 副图像的特征数据为: $\{x_{i1}, x_{i2}, x_{i3}, \dots, x_{in}\}$, n 为每张图片的特征维数,可以构建一个 $A=m \times n$ 的样本矩阵。

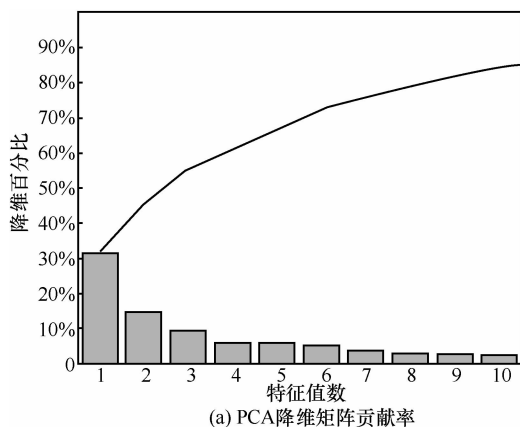
步骤 2 求取样本矩阵 A 的平均值: $\theta = \frac{1}{m} \sum_{i=1}^m x_{i0}$ 。

步骤 3 求样本矩阵与平均值的差值: $D_i = x_i - \theta$, $i = \{1, 2, 3, \dots, n\}$ 。

步骤 4 计算样本矩阵的协方差矩阵,它的维度与每个样本的维度相等: $B = \frac{\sum_{i=1}^m D_i \cdot D_i^T}{m-1}$ 。

步骤 5 求协方差矩阵 B 的特征值和特征向量,并且根据特征值从大到小进行排列,得到给出成分的重要性级别,忽略重要性低的成分,最后根据得到的最大特征值的特征向量构成一个转换矩阵 $n \cdot t (t < n)$,之后用原来样本数据 $m \cdot n \times n \cdot t$ 得到原样本数据在新空间的数据,即降维后的矩阵数据。

如图 3 所示的实验测试中,对特征图像进行 PCA 降维后,图 3(a)是计算特征值的贡献率,一般会选前 85%~95%的特征值对应的特征向量,其对应的特征值数量为降维的列数。图 3(b)表示的是测试图在 PCA 降维后,通过重构与原特征数据进行对比,可以发现 PCA 重构的误差较小。



(a) PCA降维矩阵贡献率



(b) PCA降维重构图片对比

图3 PCA降维与重构示例



由于 ELM 算法中使用的 sigmoid 函数的取值是 0~1, 所以要对图像提取的 PCA 降维后的特征数据进行归一化处理。归一化映射函数如下: $y = \frac{(y_{\max} - y_{\min}) \cdot (x - x_{\min})}{x_{\max} - x_{\min}} + y_{\min}$, 其中, $y_{\max}, y_{\min}$ 为参数, 默认为 -1 和 1, $x_{\min}, x_{\max}$ 代表数据矩阵中的最大值和最小值, x 代表要归一化的矩阵每一行中的每个数据。

根据第 3 节 ELM 算法思想和 ELM 分类的分析, 本文设计 ELM 分类器来训练图像特征集, 其中对于输入层和隐含层之间的连接权值 w 和隐含层阈值 b 进行随机选取, 同时选择激活函数为 sigmoid 函数, 相关程序分类器训练的程序伪码如下表所示。

步骤 1 初始化 $N=150$ 幅图像训练样本矩阵 $A=\{(x_i, t_i) | x_i \in R^n, t_i \in R^m, i=1, \dots, N\}$ (即对图像进行上述的提取、降维、归一化操作), 其中 $T=\{t_1, t_2, \dots, t_i\}, i=1, \dots, N$ 为训练样本的类别矩阵。

步骤 2 随机产生输入层与隐含层间的连接权值 I 与隐含层神经元的阈值 B , 取激活函数 TF 为 sig 函数。

步骤 3 根据第 3.1 节相关计算式计算隐含层输出矩阵 $H = \frac{1}{1 + e^{-(I \cdot A + B)}}$ (计算式忽略维数相关转变)。

步骤 4 计算输出权值矩阵: $L = \text{pinv}(H') \cdot T'$ (H', T' 均为转置矩阵)。

步骤 5 计算输入矩阵的预测输出类型矩阵: $P = (H' \cdot L)'$ 。

如图 4 所示, 本文对训练样本进行分类器的训练并对预测结果与原类别矩阵进行对比。当神经元的个数向训练样本逼近时, 可以发现 ELM 分类器训练精度会越来越高。

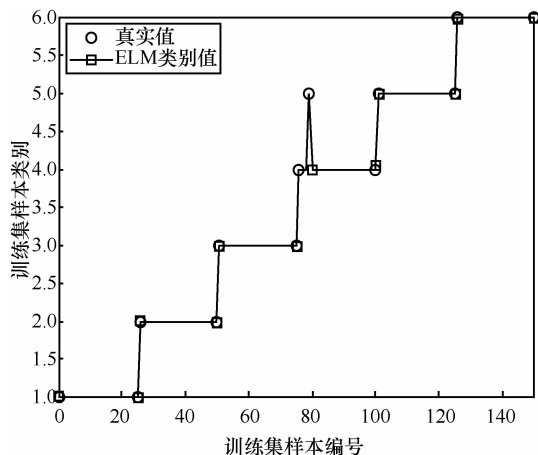
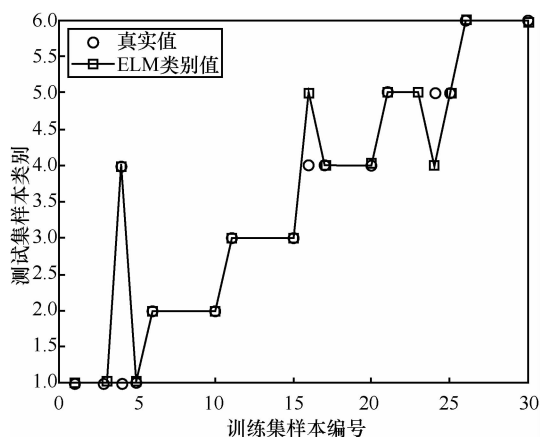


图 4 ELM 分类器训练结果 (正确率 $Accuracy=99.333\ 3\%$)

进一步对训练好的 ELM 分类器进行泛化性能研究, 需要取测试数据样本来进行测试, 其中测试数据的相关处理与训练数据一致, 测试分为两部分: 对样本数据测试, 取样本数据中测试部分进行分类测试, 由测试结果可以得知分类器的测试性能; 对非样本中数据的测试, 随机选取相关图像特征取相关数据进行分类预测, 考察分类器的泛化性能。相关实验如图 5 所示, 其中图 5(a)是对样本数据的测试结果, 可知训练好的 ELM 分类器测试性能良好; 图 5(b)是对非样本数据中的随机相关图像数据进行的分类预测结果, 可知对于非样本相关数据, 本 ELM 分类器能准确的预测出相关类别并输出相关信息。



(a) 样本数据的测试结果 (正确率 $Accuracy=90\%$)

图片输出类别为: 第4类 相关信息为:
盆栽系由中国传统的园林艺术变化而来。
至于对观赏植物的栽培, 起源于古代园林街景,
以摹仿自然山水景色营造园林。夏有瑶参、
商有鹿台、周有灵台。汉武帝在建章宫
筑太液池, 叠在蓬莱、方丈、瀛台三山。
北魏中期, 宣武皇帝元恪所建景明寺和瑶光寺,
寺里“青林垂影, 绿叶为文,”青台紫阁,
浮道相通, 虽外有四时, 内无寒暑。
房檐之外, 皆是山池, 遍布崔、蒲、菱、藕;
紫甲黄鳞, 出没于繁藻; 青凫白雁, 浮沉于绿水
参考: <http://baike.baidu.com/view/2413013.htm?fr=aladdin>

(b) 非样本数据的预测结果

图 5 ELM 分类器预测效果

5 性能优化方法 A-ELM

由于 ELM 的一些阈值和输入权值为 0 导致隐含层节点无效。为了保持 ELM 的高效性, 可以设置隐含层节点数目接近训练样本数目, 此时样本能达到最好的训练效果, 但此时 ELM 的泛化测试能力大大降低。

为了解决上述训练样本与测试样本分类准确率均衡

的问题, 本文提出 A-ELM (ascending-extreme learning machine) 一种类似增量学习的方法来训练优化分类器的性能, 从而解决隐含层节点需要手动随机设置的问题, A-ELM 算法的基本模型如图 6 所示。

由第 3 节可知, 对于含有 L 个隐含层的神经网络的函数为: $f_L(X) = \sum_{i=1}^L \beta_i h_i(x)$ 。同时, 理论与实验表明^[9], 含有 L 个隐含层节点的输出函数为:

$$f_L(x) = f_{L-1}(x) + \beta_L g_L(x) \quad (3)$$

A-ELM 优化方法的主要思想为: 在图中的隐含层逐一添加隐含层神经元节点数 (最大为训练样本数量 N)。每次添加节点时, 计算出添加节点前后的模型残差 $E_n = f - f_n^{[1]}$ (其中 f 为任意连续的目标函数, 初始化为训练样本的目标向量), 命名为 φ_1 , 将此时的隐含层节点数和模型参数带入测试样本中, 计算比较出测试样本的残差 φ_2 , 并对训练样本和测试样本的残差和求二范数 φ_{sum} , 最后依次比较求出 $\varphi_{\text{min}} = \min(\varphi_{\text{sum}})$ 即可求出此分类器均衡的隐含层节点数, 从而进一步对未知样本进行测试实验。相关的算法伪码如下。

初始化 对于给定的训练样本集合 $\psi = \{(x_i, t_i) | x_i \in R^n, t_i \in R, i=1, \dots, N\}$, $t=[t_1, t_2, \dots, t_N]$ 为样本目标向量, 隐含层神经元个数 $L=0$, 样本残差 $E=t$, 最小残差和 $\varphi_{\text{min}}=0$ (N 为训练样本个数)。

步骤 1 当隐含层节点数 L 小于最大样本数时候, 逐一添加隐含层神经元个数: $L=L+1$ 。

步骤 2 对于添加的新的隐含层节点 $j(j < L)$ 随机生成

输入权值 a_j 和阈值 b_j 。

步骤 3 计算新添加的隐含层节点的输出权重 β_j 为:

$$\beta_j = \frac{\langle e_{j-1}, g_j \rangle}{\|g_j\|^2} = \frac{E \cdot H_j^T}{H_j \cdot H_j^T}, \text{ 其中 } E \text{ 添加节点前的残差,}$$

$$H_j = \begin{pmatrix} G(a_j, b_j, x_1) \\ \vdots \\ G(a_j, b_j, x_N) \end{pmatrix} \text{ 为第 } j \text{ 个隐含层节点与 } N \text{ 个训练样}$$

本相关的隐含层输出矩阵。

步骤 4 计算添加新的隐含层节点后的残差: $E_L = E_L - f - f_L$, 结合式 (3) 有: $E_L = f - f_L = f - f_{L-1} - \beta_L \cdot g_L = E_{L-1} - \beta_L \cdot g_L$ (其中 E_{L-1} 表示添加新节点之前的模型残差, E_L 表示添加之后的新的残差)。

步骤 5 将 a_j, b_j 等模型相关信息带入到测试集框架中, 重复步骤 3、步骤 4 求出测试集合残差 E_2 , 将

$$\varphi_{\text{sum}} = \left(\sum_{i=1}^N E_1^2 \right)^{\frac{1}{2}} + \left(\sum_{i=1}^N E_2^2 \right)^{\frac{1}{2}} \text{ 与 } \varphi_{\text{Min}} \text{ 比较, 重复实现步骤}$$

1~步骤 5 直到求出 $\varphi_{\text{min}} = \min(\varphi_{\text{sum}})$ 。

步骤 6 对训练样本和测试样本之外的模型样本进行测试, 进一步提高分类器的泛化性能。

6 实验部分

在本节中, 通过实验来验证上述所有理论, 采用“MATLAB2010b + VC++”编程, 具体实验环境: Asus A43s 2.50 GHz CPU、4 GB 内存、500 GB 硬盘。实验主要程序采用 MATLAB2010 语言编写。首先第 6.1 节中对 ELM 分类器训

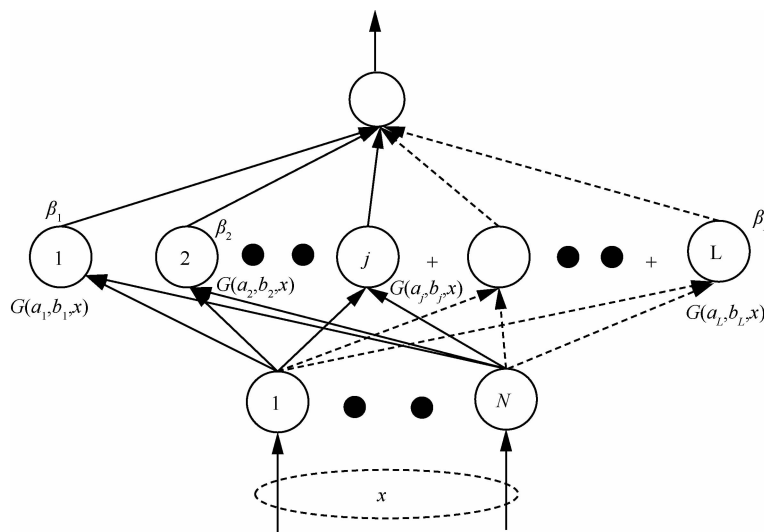


图6 A-ELM 算法模型



练,通过对图片集中种类的采集和处理来训练分类器。

6.1 服务器端 ELM 分类器实验

对 ELM 分类器训练,通过对图片集中种类的采集和处理来训练分类器。

图 7 中实验所用图像是从 101_ObjectCategories 图片集中随机选取 6 种类图片集,每种 30 样本(其中,每种选取 25 样本进行训练,5 样本进行测试,共计 $6 \times 25 = 150$ 个训练样本, $6 \times 5 = 30$ 个测试样本)进行 ELM 分类器的训练,由于随机产生输入和隐含层连接权值 w 和隐含层阈值 b ,所以分类准确率会有一个波动范围。所以本文采取多训练求平均值的方式来衡量训练的结果,由图 6 可知对于训练分类器准确率达到了 98%,此时测试准确率也能达到 90%可以证明对于此分类器的训练测试效果较好。

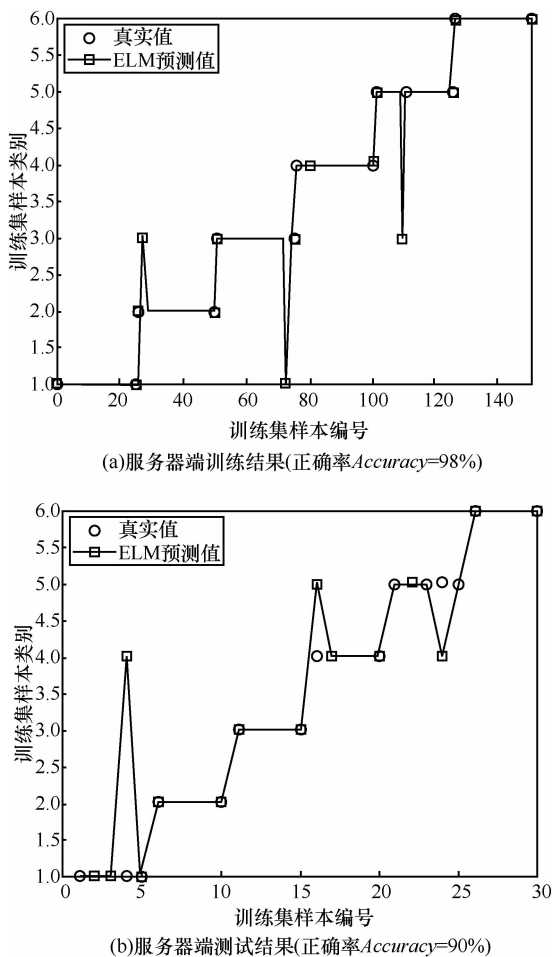


图 7 服务器端训练测试效果

6.2 对比实验

通过使用传统的分类算法 SVM 来进行比较,从而确定 ELM 分类器的优势。

如图 8 所示,系统通过使用支持向量机(SVM)来处理图片特征,最后返回训练样本的准确率。为了更好地体现对比性,本文对不同训练样本进行对比实验如图 9 所示。

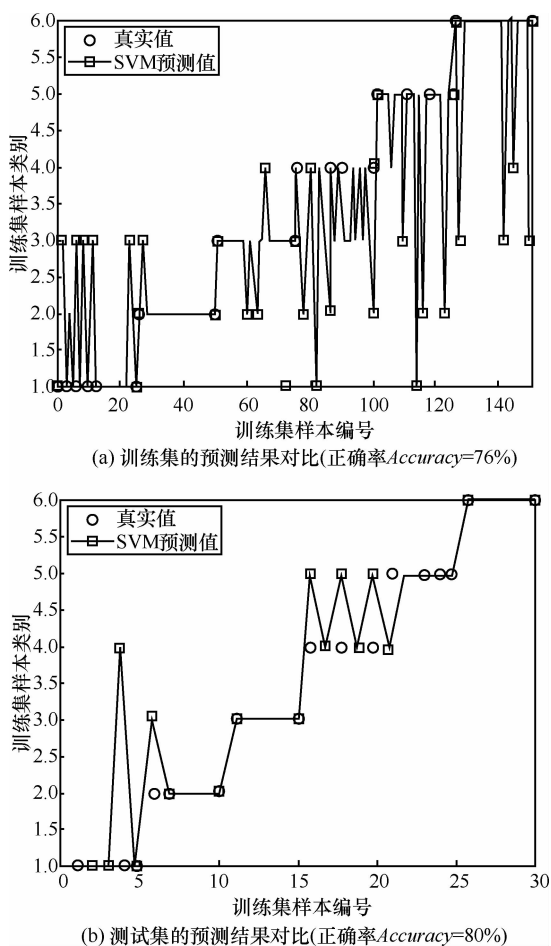


图 8 SVM 训练测试结果

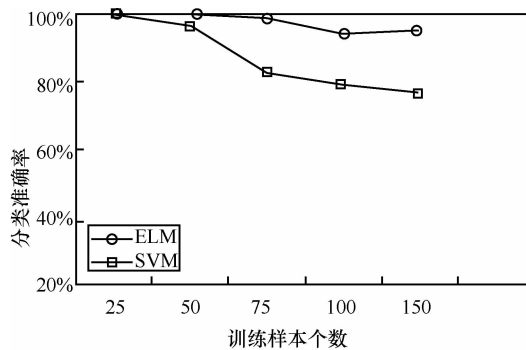


图 9 SVM 和 ELM 对比

由于随机产生 w 和 b 对测试实验结果有起伏的影响,本文对 5 次分类器测试实验取平均值,见表 1。

表 1 训练平均值

测试正确率	1	2	3	4	5	平均值
ELM	86.667%	90.0%	86.667%	83.333%	86.667%	86.667%
SVM	80.0%	80.0%	80.0%	80.0%	80.0%	80.0%

通过图 9 和表 1、表 2 可以发现,当样本数量不多时,ELM 和 SVM 的分类准确率和分类时间差不多,当数量增大时候,分类器的准确率就有了比较大的差距,训练时间更是相差比较大。因此,在图像特征检索方面,ELM 比 SVM 更有效率性和准确性。

表 2 ELM 和 SVM 对比总结

对比项	ELM	SVM
训练样本平均正确率 <i>Accuracy</i>	97.28%	85.72%
测试样本平均正确率 <i>Accuracy</i>	86.667%	80%
分类器训练时间/s	0.574 4	1.667 3
分类器测试时间/s	1.294 035	1.540 635

6.3 ELM 的性能优化实验

本节对 ELM 方法进行性能优化实验,使用优化性能的 ELM 进行训练测试,进一步提高 ELM 的分类训练测试能力和泛化性能。主要分为两部分:A-ELM 的优化结果及性能优势;对相关非样本数据的预测。

如图 10 所示,当逐一增加隐含层节点 L 时,训练样本残差 φ_1 与 φ_2 测试样本的残差的范数和成下降趋势,之后趋于一个最小点,此时即可求出无需人工设置的隐含层节点数。

如图 11 所示,当没有采取优化时候,训练样本的

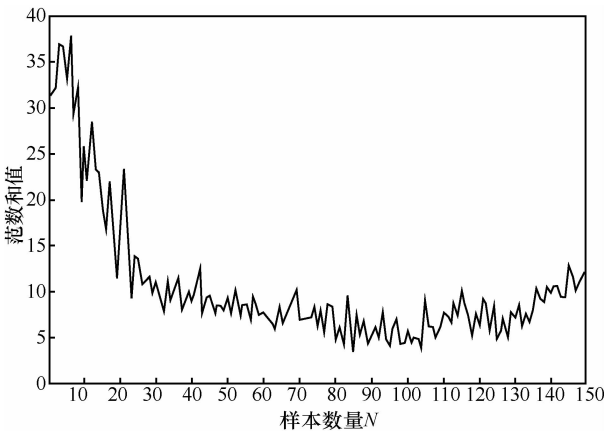


图 10 A-ELM 优化实验

神经元个数是随机生成,要想达到很好的效果,只能人工不断手动设置隐含层节点来慢慢调节,且如需达到最好的训练效果,可设置神经元个数 L =训练样本数目,但此时测试样本的 $\|\varphi_2\|$ 非常大(也即误差非常大),如果想达到一个均合适的分类效果,需要不停手动设置 L 。采取优化后,无需手动设置,通过上述的方法,可以自动调节训练样本和测试样本的误差,从而确定一个合适的神经元个数,达到一个均衡的效果,以提高分类器的稳定性。

对相关非样本数据的预测方面,为了检验调节 L 对分类器的泛化性能的影响,如图 12 中(a)~(d)随机下载 4 幅图片进行测试,图 13 为测试结果。

表 3 中给出了 A-ELM 和未优化的 ELM 对于非样本数据的测试,可以看到优化后,预测准确度并没有降低,但是由于需要重复调节神经元个数,花费时间增加。但是手动调节来说,A-ELM 有很大的优势。

6.4 移动计算设备的测试

本文对移动计算设备进行了视觉搜索测试,移动计算设备为拥有安卓系统的智能手机或者拥有安卓模拟器的电脑,并且只需要下载分类器和相关信息即可。实验整体流程如下:首先在服务器端和移动端建立连接,设置好端口号和服务器端的 IP 地址,同时在移动端上安装一个安卓 App。打开此应用程序程序,移动计算设备拍摄图像完成采集,上传到服务器进行处理,上传过程中使用序列化方式传输,将图像组织成二进制流的方式进行图像的接收和传输。与此同时,服务器端监听事先设置好的端口并接收图像数据流,之后反序列化成图像显示送入服务器端的分类器中进行分类预测。根据本文所提出的模式框架,服务器端将预测的图像进行相关处理并传输到网络中,移动端接收此图像检测进行显示,同时能返回相关信息。实验展示效果如图 14 所示。

7 结束语

本文在阐述 ELM 的理论基础上对 ELM 进行了进一步的理论与实验分析,并将之运用到移动视觉搜索方面;本文首先对 ELM 算法的性能进行了分析验证,在此基础

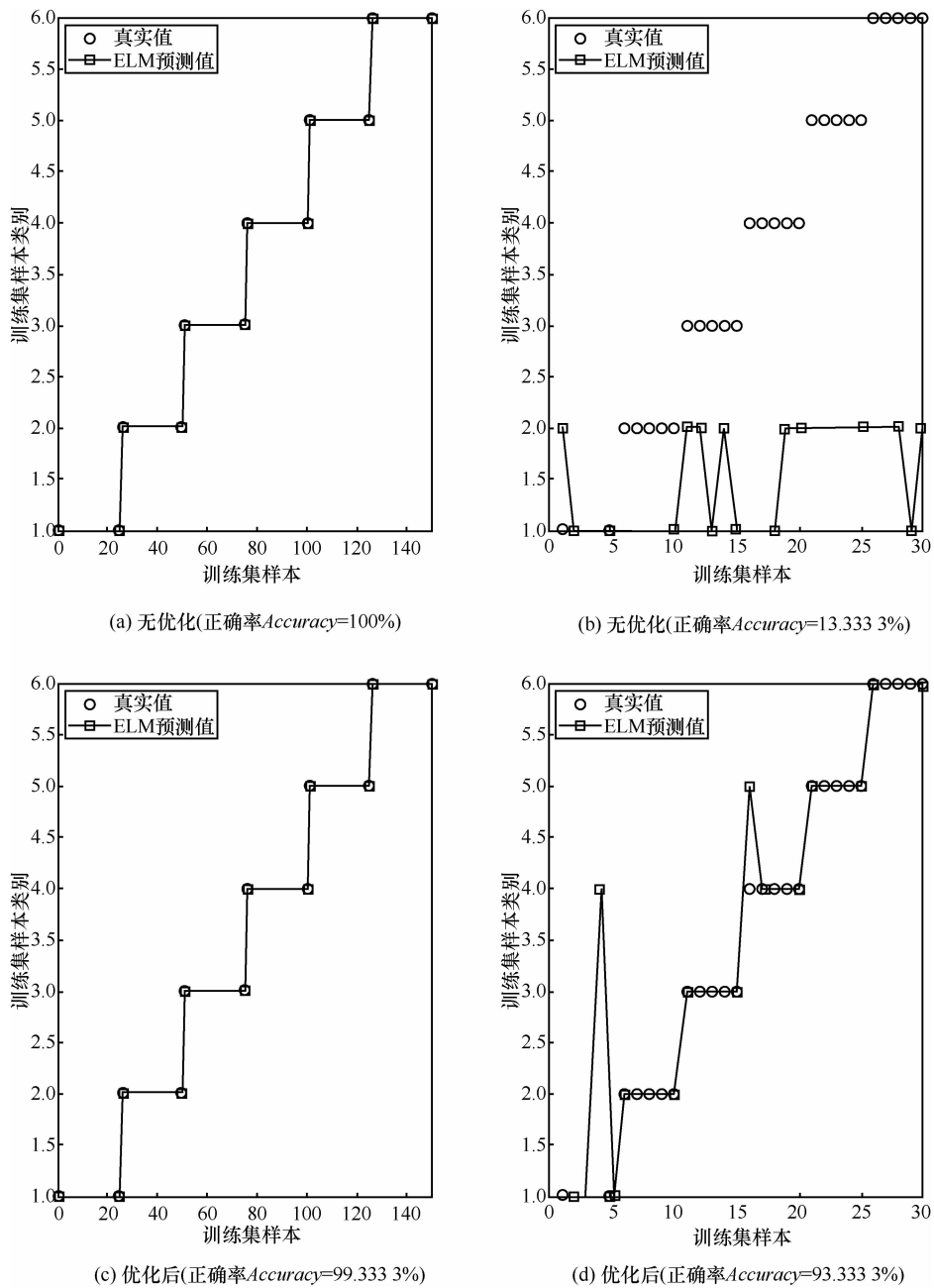
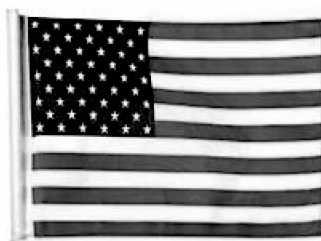


图 11 性能优化对比



(a) 盆栽



(b) 美国国旗



(c) 雷峰塔



(d) 向日葵

图 12 测试图片示例

表 3 ELM 优化前后性能对比

对比项	ELM	A-ELM
测试准确度(是/否)	是	是
测试所花费时间/s:	0.563 23	6.230 265
手动调节 L 次数	11	0

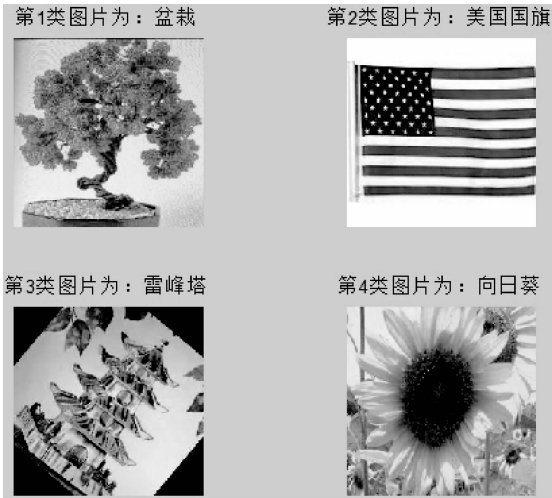


图 13 针对图 12 中图片的测试结果



(a) 图片输入



(b) 信息显示

图 14 移动计算设备的测试

上,提出了基本视觉框架模型 BMVS,将 ELM 运用到此框架服务器端,并进一步给出了优化分类性能的 A-ELM 方法。最后通过实验证明了 ELM 在移动视觉搜索方面的可行性,并通过和支持向量机 SVM 的实验对比验证相关方法的高效性。

参考文献:

[1] 段凌宇,黄铁军,高文. 移动视觉搜索技术与标准化进展[J]. 信息通信技术, 2012(6): 51-58.
DUAN L Y, HUANG T J, GAO W. Technical research and standardization in mobile visual search [J]. Information and Communications Technologies, 2012(6): 51-58

[2] 李向阳,庄越挺,潘云鹤. 基于内容的图像检索技术与系统[J]. 计算机研究与发展, 2001(3), 38(3): 344-354.
LI X Y, ZHUANG Y T, PAN Y H. The technique and systems of content-based image retrieval [J]. Journal of Computer Research and Development. 2001(3), 38(3): 344-354

[3] GURNEY K. An introduction to neural networks [M]. London: UCL Press, 1997: 20-25.

[4] HSU C W, LIN C J. A comparison of methods for multiclass support vector machines [J]. IEEE Transactions on Neural Networks, 2002, 13(2): 415-425.

[5] HHUANG G B, SIEW C K. Extreme learning machine: RBF network case[C]// 2004 8th Conference on Control, Automation, Robotics and Vision Conference, December 6-9, 2004, Kuming, China. New Jersey: IEEE Press, 2004(2): 1029-1036.

[6] HUANG G B, CHEN L. Convex incremental extreme learning machine[J]. Neuro Computing, 2007(70): 2056-2062.

[7] HUANG G B, BABRI H A. Upper bounds on the number of hidden neurons in feedforward networks with arbitrary bounded nonlinear activation functions [J]. IEEE Transactions on Neural Networks, 1998, 9(1): 224-229.

[8] FUNG G, MANGASARIAN O L. Proximal support vector machine classifiers [C] // The 7th ACM Sigkdd International Conference on Knowledge Discovery & Data Mining, August 26-29, 2001, San Francisco, CA, USA. New York: ACM Press, 2001: 77-86.

[9] HUANG G B, WANG D H, LAN Y. Extreme learning machines: a survey [J]. International Journal of Machine Learning and Cybernetics, 2011, 2(2):107-122.

[10] HUANG G B, ZHU Q Y, SIEW C K. Extreme learning machine: a new learning scheme of feedforward neural networks[J]. IEEE International Joint Conference. 2004(2): 985-990.

[11] HUANG G B, ZHU Q Y, SIEW C K. Extreme learning machine: theory and applications[J]. Neuro Computing, 2005 (70): 489-501.

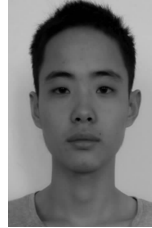


- [12] HUANG G B, ZHOU H, DING X, et al. Extreme learning machine for regression and multi-class classification [J]. IEEE Transactions on Pattern Anal Mach Intell, 2012, 42(2): 513-529.
- [13] HUANG G B, CHEN L, SIEW C K. Universal approximation using incremental constructive feedforward networks with random hidden nodes [J]. IEEE Transactions on Neural Networks, 2006, 17(4): 879-892.
- [14] HUANG G B, SIEW C K. Extreme learning machine with randomly assigned RBF kernels [J]. International Journal of Information Technology, 2005, 11(1): 16-24.
- [15] HALL M A, SMITH L A. Feature selection for machine learning: Comparing a correlation-based filter approach to the wrapper [C] // The 20th International Florida Artificial Intelligence Research Society Conference, May 1-5, 1999, Orlando, Florida, USA. California: AAAI Press, 1999: 235-239.
- [16] PRICE A L, PATTERSON N J, PLENCE R M, et al. Principal components analysis corrects for stratification in genome-wide association studies [J]. Nature Genetics, 2006, 38 (8): 904-909.
- [17] HOLTEN D, WIJK V J J, MARTENS J B. A perceptually based spectral model for isotropic textures [J]. ACM Transactions on Applied Perception, 2006, 3(4): 376-398.

[作者简介]



胡海洋(1977-),男,博士,杭州电子科技大学教授,主要研究方向为分布式系统、数据库技术、 workflow 技术。



许军(1990-),男,杭州电子科技大学硕士生,主要研究方向为数据库技术、图像检索。



胡华(1964-),男,博士,杭州电子科技大学教授,主要研究方向为分布式系统、数据库技术。