

基于多任务学习的自然图像分类研究^{*}

刘 成, 彭进业

(西北大学 信息科学与技术学院, 西安 710127)

摘要: 针对自然图像内容结构复杂、难以区分的实际情况,提出了一种基于多任务学习的自然图像分类方法。通过额外任务来辅助主任务的学习,构造了衡量任务间相关性大小的相关性矩阵,提出了主任务联合额外任务共同决策的学习模式;通过额外任务与主任务的相关性来控制额外任务参与主任务决策的程度,以提高主任务的分类准确率。实验结果表明,与传统的单任务学习相比,尤其是在已知样本较少的情况下,多任务学习机制能够明显地改善分类器的泛化性能。

关键词: 多任务学习; 自然图像; 相关性

中图分类号: TP181 **文献标志码:** A **文章编号:** 1001-3695(2012)07-2773-03

doi:10.3969/j.issn.1001-3695.2012.07.102

Research of classification method for natural images based on multitask learning

LIU Cheng, PENG Jin-ye

(School of Information Science & Technology, Northwest University, Xi'an 710127, China)

Abstract: Considering it's difficult to distinguish different images because the content and structure of natural images is very complex, this paper proposed a classification method for natural images based on multitasks learning. It designed some extra tasks to assist the main task's learning, and constructed correlation matrix that measured the correlation between tasks. It proposed a study model that main task jointed extra tasks to make decision. And controlled the degree of extra tasks involved main task's decision through the correlation between main task and extra tasks. Experiment results show that the proposed method can markedly improve the generalization capability of classifier comparing with single task learning, especially in the case of lacking enough prior knowledge.

Key words: multitask learning; natural images; correlation

随着信息采集和存储技术的发展,人们获取信息的能力也在不断增强。互联网上的数据以指数方式增长,而图像在这些数据中占有很大的比例。如何让计算机对这些图像进行有效的分类管理,以达到方便查询和检索的目的,机器学习便成为首选方案。

目前的图像分类研究大多集中于遥感和医学等专业图像^[1-3],这类图像大多是灰度图像,并且背景纹理简单,只需要将一幅图像分成几个区域或将感兴趣的区域提取出来就可以了。而自然图像多是彩色图像,图像中包含的对象五花八门,有意义的形状特征不易提取。文献[4,5]提取了自然图像的颜色和纹理特征,分别使用支持向量机和BP神经网络对自然图像进行了单任务分类操作。由于特征的区分能力不是很强,这种传统的单任务学习只专注于最终的分类结果,对任务间的相关性信息没有作进一步的挖掘,从输入样本中所获取的信息有限,单任务学习训练出来的分类器很难具有较高的泛化性能。

多任务学习是一种迁移学习方法,它可以通过使用包含在其他相关任务中的信息来改善单任务学习的性能^[6],将多个相关任务放在一起同时学习,加入的任务会改变网络权值更新的动力特性^[7],使得分类器输出的结果相互牵制,尤其是在已

知样本较少的情况下,多任务学习更能改善分类器的泛化性能^[8-10]。文献[11]提出了改造SVM分类器,增加代表所有任务间共有分类面的方法来实现多任务分类,但是将所有任务的相关性都视为均匀的、同一的。文献[12]提出了一种基于多任务学习的垃圾邮件过滤系统,获得了较好的过滤效果。淘宝网也对它的客户进行了分类,根据同一类人对某些商品有共同的喜好这一事实,推荐了人们可能喜欢的商品。因此,充分利用任务间的相关性,建立多任务学习机制,会对某些实际问题产生很好的分类识别效果。本文将自然图像的分类问题看成是一个多任务学习问题,构建了自然图像分类的多任务学习的BP神经网络模型,将最终所需要的类别看成是主任务,同时引入一些额外任务来辅助主任务的学习。

1 BP神经网络与多任务学习

1.1 BP神经网络

BP(error back propagation)神经网络是一种多层前馈神经网络,如图1所示。 $(x_1, x_2, \dots, x_n)$ 为某一个样本的特征向量, $(y_1, y_2, \dots, y_m)$ 为网络对该样本的判决输出,通过不断地调整

收稿日期: 2011-11-11; **修回日期:** 2011-12-19 **基金项目:** 国家自然科学基金资助项目(61075014);西北大学硕士研究生创新基金资助项目(10YZZ17)

作者简介: 刘成(1985-),男,陕西西安人,硕士研究生,主要研究方向为人工智能、模式识别(570015061@163.com);彭进业(1964-),男,陕西西安人,教授,博导,主要研究方向为图像处理、量子通信等。

输入层到隐层的权值 w_{ij} 、偏置 a_j 和隐层到输出层的权值 w_{jk} 、偏置 b_k 来建立从输入到输出的映射。首先计算隐层的输出 $H: H_j = f(\sum_{i=1}^n w_{ij}x_i - a_j) (j=1, 2, \dots, l)$, 其中 l 为隐层节点数, f 为隐层激励函数, 常选用 S 函数 $f(x) = 1/(1 + \exp(-x))$; 再计算输出 $O: O_k = \sum_{j=1}^l H_j w_{jk} - b_k (k=1, 2, \dots, m)$, 根据预测输出 O 和期望输出 Y 计算网络预测误差 $e: e_k = Y_k - O_k (k=1, 2, \dots, m)$; 最后根据预测误差 e 更新网络连接权值 w_{ij} 、 w_{jk} 和偏置 a_j 、 b_k : $w_{ij} = w_{ij} + \eta H_j (1 - H_j) x(i) \sum_{k=1}^m w_{jk} e_k$, $w_{jk} = w_{jk} + \eta H_j e_k$, $a_j = a_j + \eta H_j (1 - H_j) \sum_{k=1}^m w_{jk} e_k$, $b_k = b_k + e_k (i=1, 2, \dots, n; j=1, 2, \dots, l)$ 。当误差小于某一限定值或迭代次数超出规定的最大值时, 停止更新网络权值, 训练完成。

1.2 自然图像分类的多任务 BP 神经网络

按照不同的分类标准, 同一幅图像可以被分为不同的类别。例如, 根据图像中所包含的内容可以将图像分为非洲生活、美食、花、汽车四类, 而这些图像按照另一种分类标准可以分为室内、室外、风景、非风景等。如图2所示, 任务1是主任务, 它的输出 B_1 及其所代表的类别是本文想要的类别; 而任务2~5属于额外任务, 它的输出 $B_2 \sim B_5$ 及其所代表的类别不是本文所关心的。但是额外任务与主任务是相关的, 额外任务可以帮助主任务决策得更好, 每个任务的输出及其对应的类别如表1所示。

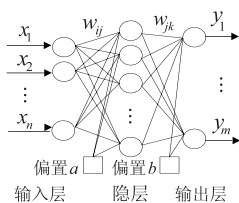


图1 BP神经网络结构示意图

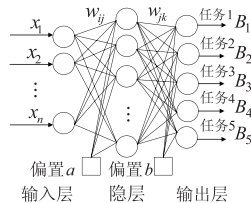


图2 多任务BP神经网络

表1 各个任务输出的值及其所代表的含义

输出符号	B_1	B_2	B_3	B_4	B_5
输出值	1	2	3	4	0~1
对应类别	非洲生活	美食	花	汽车	室内

这五个任务分别从不同角度和不同层次对同一幅图像进行了描述。在对一幅未知图像进行分类时, 与只获得任务1的输出相比, 在不增加样本特征的前提下, 同时获得这五个任务的输出, 将会得到对一幅图像更完全更精准的表达。当任务1判定一幅图像到底是非洲生活还是美食而犹豫不决时, 可以参考任务2~5的结果。如果这幅图像是室外的风景, 那么有理由认为它应该属于非洲生活; 如果是一幅室内的非风景图, 那么就认为它应该属于美食。额外任务参与主任务决策的程度可以通过每个额外任务与主任务的相关性大小来控制。关于每一个额外任务与主任务的相关性大小的定义, 以及额外任务如何辅助主任务决策的数学模型, 在1.3节中将给出明确的定义。

1.3 任务相关性与任务判决机制

如何寻找能够很好地辅助主任务决策的额外任务, 这是多任务学习的第一步。额外任务的选取应该遵循以下原则: a) 额外任务必须能够从输入的样本特征中反映出来; b) 额外任务应该能够把握主任务类别间微妙的差异性, 在主任务很难决策时, 能够帮助主任务决策。比如, 非洲生活多属于室外风景图像, 能够从输入图像颜色特征的亮度上反映出是室内还是室

外, 风景图的颜色分布较为均匀, 非风景图的颜色分布较为密集, 这一点也能从输入获得。1.2节中额外任务的选取就是按照这一原则选取出来的。

在获得了额外任务之后, 重点就在于如何度量额外任务与主任务的相关性, 本文使用 $R_{i,j}$ 来代表额外任务与主任务的相关性。其中: $i=1, 2, 3, 4$ 代表主任务的四个类别; $j=2, 3, 4, 5$ 代表四个额外任务; $R_{1,2}$ 就代表室内与非洲生活的相关性大小, 可以通过式(1)来计算 $R_{1,2}$ 的值。

$$R_{1,2} = \frac{\text{在非洲生活的图像中室内图像的数量}}{\text{类别为非洲生活图像的总数量}} \quad (1)$$

$$R_{i,j} = \frac{\text{在第 } i \text{ 类图像中属于 } j \text{ 类图像的数目}}{\text{第 } i \text{ 类图像的总数量}} \quad (2)$$

根据式(2)可以计算出其他主任务与额外任务的相关性大小, 进而得到描述任务间相关性大小的相关性矩阵。

根据与主任务相关性越大的额外任务越多参与主任务的决策和与主任务相关性越小的额外任务越少参与主任务的决策这一原则, 本文构造了如式(3)所示的任务判决机制, C_i 代表了对一幅图像最终的判决结果。

$$C_i = B_1 - \frac{1}{4}(B_2 R_{i,2} + B_3 R_{i,3}) + \frac{1}{4}(B_2 R_{i+1,2} + B_3 R_{i+1,3}) - \frac{1}{4}(B_4 R_{i,4} + B_5 R_{i,5}) + \frac{1}{4}(B_4 R_{i+1,4} + B_5 R_{i+1,5}) \quad (3)$$

对于式(3)中参数的选取可以遵循式(4)的约束:

$$\begin{cases} \text{if } (B_1 < 2), i = 1 \\ \text{if } (2 \leq B_1 < 3), i = 2 \\ \text{if } (3 \leq B_1 < 4), i = 3 \\ \text{if } (B_1 \geq 4), i = 4 \\ \text{if } (B_j \leq 0.5) \quad B_j = 0, \text{ else } B_j = 1, \quad j = 2, \dots, 5 \end{cases} \quad (4)$$

如式(3)所示, 在判决中, 主任务的输出 B_1 对最终的判决起决定性作用, 额外任务的输出 B_2, B_3, B_4, B_5 修正主任务的输出, 它的修正量不会超过0.5。式(3)中第2、4项的修正会使最终的输出类别更接近于 i , 而第3、5项的修正会使最终的输出类别更接近于 $i+1$, 当 $B_1 \geq 4$ 时, $C_i = B_1$ 。为了使额外任务能够更好地在两个难以区分的类别间作修正, 让两类相似的图像的类别号相邻, 在额外任务的辅助下, 主任务的决策将会更加合理。

2 实验结果与分析

2.1 实验图像与特征选取

如1.2节所述, 本实验选取 Corel 图像集中的非洲生活、美食、花、汽车共四类图像进行实验, 每类100幅, 总计400幅, 训练集每类20幅, 共80幅, 测试集每类80幅, 共320幅。提取27维的空间颜色直方图 $H = (H_1, H_2, \dots, H_{27})$ 和40维的 Gabor 纹理特征 $T = \{\mu_1, \sigma_1, \dots, \mu_{20}, \sigma_{20}\}$, 组合成67维的综合特征 (H, T) 。将非洲生活、美食、花、汽车四个类别作为主任务, 将室内、室外、风景、非风景四个类别作为额外任务进行多任务学习。

2.2 图像特征的空间可视化

为了验证所选取特征的有效性, 并且分析后续的分类结果, 有必要对所选取的高维特征在特征空间中进行可视化操作。使用平行坐标将67维的高维特征向量投影到一个二维的平面坐标上来观察这400个样本的空间分布情况, 判断所选取的特征是否能够将不同类别的图像区分开来。如图3所示, 横坐标1~67代表每一个特征矢量的67个分量, 纵坐标0~1代表归一化后每一个分量的值, 一个矢量的67个分量连起来就

形成一条曲线,在同一个坐标中绘出每一类的 100 个样本,便形成了一类图像的特征分布。图 3 依次演示了非洲生活、美食、花、汽车四类图像的特征矢量在平行坐标中的分布。从图中可以看出,各类图像的分布具有一定的可区分性,但同时也有一定的重叠区域,这是由两类图像太过相似造成的,也造成某些图像难以区分的现象。

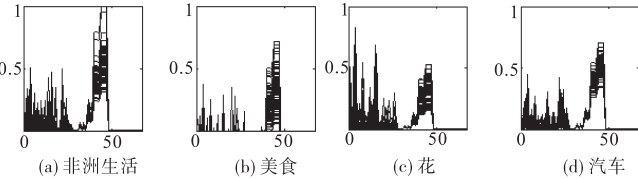


图 3 高维特征可视化

2.3 实验结果与分析

根据式(2)所定义的相关性参数,可以获得如表 2 所示的各个任务间的相关性矩阵。根据额外任务 B_2 、 B_3 、 B_4 、 B_5 和额外任务与主任务的相关性 R_{ij} ,就可以控制额外任务参与主任务决策的程度,额外任务的修正量与主任务的输出相结合便可以得到最终的决策输出。本次实验采用单隐层 BP 神经网络,输入层节点数目为 67 个,隐层节点数目为 100 个,输出节点数目为 5 个,学习率为 0.1,目标误差为 0.001。

表 2 任务间相关性矩阵

主任务(B_1)	室内(B_2)	室外(B_3)	风景(B_4)	非风景(B_5)
非洲生活	10%	90%	15%	75%
美食	95%	5%	0%	100%
花	7%	93%	30%	70%
汽车	5%	95%	20%	80%

在 320 幅测试图像中,1~80 属于第一类,类别号为 1,依此类推,241~320 属于第四类,类别号为 4。图 4 与 5 分别是单任务学习(single task learning, STL)和多任务学习(multitask learning, MTL)在 Corel 图像集上实验结果的图形化输出。图 4 (a)与图 5(a)显示的是两种学习方式下的预测输出(用 $\odot$ 表示)与期望输出(用 * 表示)的对比,横坐标代表样本序号,纵坐标代表样本类别号。期望输出表示的是图像实际的类别号,也是希望得到的分类结果,预测输出是本文训练出来的网络对测试图像所属类别的判断,没有与期望输出重合的样本都是预测出错的样本。图 4(b)与图 5(b)显示的是两种学习方式下的预测输出偏离期望输出的误差(即期望输出减预测输出的结果)。在表 3 中,对单任务学习网络与多任务学习网络对未知图像的分类准确率进行了统计。结合图 3 所示的样本特征可视化结果对单任务学习的结果进行分析可知,第一类图像的样本空间分布较为密集,并与其他三类区别较大,因而分类准确率较高。从图 4(a)与图 5(a)可以看出,第一类中被判错的图像大多错判为第二类,其他三类图像的特征重叠较多,难以区分,因而分类准确率较低。而多任务学习充分利用了任务间的相关性,选取的额外任务能够很好地修正主任务的输出,主任务从其他额外任务中获取到了更多的偏执信息,利用任务的相关性弥补了先验知识的不足,对每一类图像的表达都更为精确,因而在每一类测试图像上都获得了较高的分类准确率,使分类器的泛化性能得到了增强。

表 3 两种学习方法在测试集上的分类准确率统计

方法	非洲生活	美食	花	汽车
单任务学习(STL)	0.725 0	0.275 0	0.337 5	0.425 3
本文提出的多任务学习(MTL)	0.812 5	0.737 5	0.912 5	0.825 0

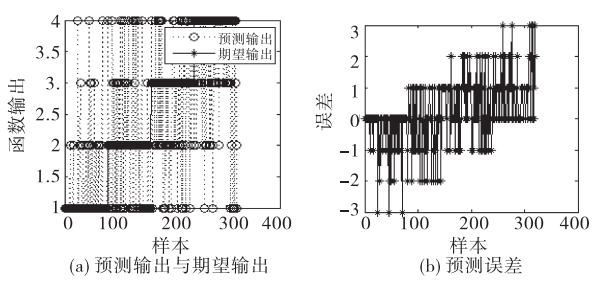


图 4 单任务学习的输出结果与预测误差

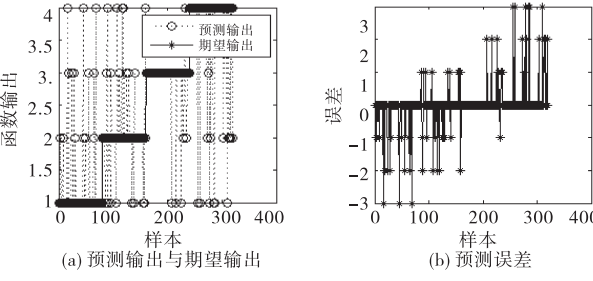


图 5 多任务学习的输出结果与预测误差

3 结束语

在现实生活中,常常遇到已知样本不足的实际问题,即缺乏足够的先验知识,人们在小样本集上训练出来的分类器很难经得起大量样本的考验,面临泛化能力不强的问题。本文所提出的基于多任务学习的自然图像分类方法,在不增加输入信息的情况下,同时关注多个任务的输出,能够从输入端得到更多的信息,训练出来的网络也具备了获取这些信息的能力,而这些信息又能够很好地辅助人们所关注的任务,进而提高了分类器的泛化性能。

参考文献:

[1] 孙锋利,何明一,高全华.基于自适应脊波网络的高光谱遥感图像分类[J].计算机科学,2011,38(8):216-236.

[2] 任军号,吉沛琦,耿跃.SOM 神经网络改进及在遥感图像分类中的应用[J].计算机应用研究,2011,28(3):1170-1172.

[3] 戴丽丽,宋余庆,陈健美,等.遗传 K-均值初始化的 t 混合模型医学图像聚类[J].计算机应用研究,2010,27(8):3151-3153.

[4] 任建峰,郭雷,李刚.多类支持向量机的自然图像分类[J].西北工业大学学报,2005,23(3):295-298.

[5] 谢文兰,石跃祥,肖平.应用 BP 神经网络对自然图像分类[J].计算机工程与应用,2010,46(2):163-166.

[6] CARUANA R. Multitask learning[D]. Pennsylvania: Carnegie Mellon University, 1997:1-4.

[7] 金凤.多视角的构建及其在单任务和多任务学习中的应用[D].武汉:华东师范大学,2009.

[8] BAKKER B, HESKES T. Task clustering and gating for Bayesian multi-task learning[J]. Journal of Machine Learning Research, 2003,4: 83-99.

[9] HESKES T. Empirical Bayes for learning to learn[C]//Proc of IC-ML. 2000:367-374.

[10] THRUN S, PRATT L. Learning to learn[M]. [S. l.]: Kluwer Academic Publishers, 1997.

[11] EVGENIOU T, MASSIMILIANO P. Regularized multi-task learning[C]//Proc of ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2004:109-117.

[12] 许隼华,王志坚.基于多任务学习的邮件过滤系统的研究[J].计算机技术与发展,2010,20(10):137-140.