

基于多分类器协同学习的卷积神经网络训练算法

陈文¹ 张恩阳² 赵勇²

(四川大学计算机学院 成都 610065)¹ (长虹技术中心基础技术研究所 成都 610094)²

摘要 卷积神经网络(CNN)是一类重要的深度神经网络,然而其训练过程需要大量的已标记样本,从而限制了其实际应用。针对这一问题,分析了CNN分类器的协同学习过程,给出了基于迭代进化的分类器协同训练算法CAMC。该算法结合了CNN和多分类器协同训练的优势,首先采用不同的卷积核提取出多种样本特征以产生不同的CNN分类器;然后利用少量的已标记样本和大量的未标记样本对多个分类器进行协同训练,以持续提高分类性能。在人脸表情标准数据集上的实验结果表明,相对于传统的表情特征识别法LBP和Gabor,CAMC能够在分类过程中利用未标记样本持续实现性能提升,从而具有更高的分类准确率。

关键词 机器学习,卷积神经网络,协同训练,图像识别,多分类器

中图分类号 TP301 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2016.9.044

CNN Training Algorithm Based on Co-studying of Multiple Classifiers

CHEN Wen¹ ZHANG En-yang² ZHAO Yong²

(College of Computer Science, Sichuan University, Chengdu 610065, China)¹

(Basis Technology Institution, Changhong Technology Center, Chengdu 610094, China)²

Abstract Convolutional neural network (CNN) is a kind of important deep neural network. However, amounts of labeled samples are needed in the training process of CNN, which largely limits its applications. For the problem, the co-training process of CNN was analyzed, and a co-training algorithm CAMC based on the iterative evolution of classifiers was given. The advantages of CNN and co-training of multiple classifiers are combined in CAMC. Firstly, multiple features are drawn using different convolutional kernels to build up different CNN classifiers. Then to continually improve the classification performance, some labeled samples together with many unlabeled samples are employed to co-train the multi classifiers. The experimental results based on the standard data sets of human expression demonstrates that, compared with traditional expression recognition methods LBP and Gabor, the performance of CAMC can be continually improved, thus it has higher classification accuracy.

Keywords Machine learning, Convolutional neural network, Co-training, Image recognition, Multiple classifiers

1 引言

卷积神经网络(Convolutional Neural Network, CNN)是一种受生物视觉神经机制的启发而设计的多层感知器。CNN由多个卷积层和次采样层组成,具备自动提取样本特征的能力。CNN的局部感受野保证了原始信号的空间结构关系,从而可以将图像直接作为神经网络的输入,避免了传统识别算法中复杂的特征提取和数据重建过程^[1]。此外,CNN通过权值共享网络结构使之更类似于生物神经网络,降低了网络模型的复杂度,减少了权值的数量,因此在模式识别等许多领域取得了更好的效果。目前,CNN已成为图像识别和语音分析等领域的研究热点^[2,3]。

然而,CNN作为一种典型的有监督的学习算法,依赖于大量有标记的(labeled)样本对模型进行训练^[4,5]。一方面,有

标记的训练样本往往较难获取,尤其是在图像识别领域,靠人工进行样本类别标注的代价过大,从而限制了CNN的广泛应用^[6]。另一方面,网络中存在大量的未标记样本,因为缺少标记而无法被使用。因此,为了利用大量的未标记样本对深度卷积神经网络进行训练,本文将研究CNN图像分类器的协同训练算法。

协同训练算法(co-training)利用少量已标记样本和大量未标记样本对多个分类器进行协同训练,以提高分类器性能^[7]。co-training算法^[8]假设数据集有两个充分冗余(sufficient and redundant)的属性特征视图(view),然后利用不同的视图对分类器进行交互式迭代训练,直至分类精度达到期望值。然而,实际应用中往往较难得到充分冗余的样本特征视图,因此S. Goldman和Y. Zhou^[9]使用不同的决策树算法,从同一个属性集上训练出不同的分类器;虽然该算法避免了

到稿日期:2015-08-21 返修日期:2015-11-13 本文受国家自然科学基金:基于云的人工免疫检测器生成算法及其中网络安全中的应用研究(61402308)资助。

陈文(1983-),男,博士,副教授,主要研究方向为机器学习、网络安全, E-mail: wenchen@scu.edu.cn; 张恩阳(1965-),男,硕士,研究员,主要研究方向为数据挖掘、智能系统; 赵勇(1963-),男,博士,研究员,主要研究方向为数据挖掘、信号处理。

对冗余特征视图的需要,但是对分类器的种类进行了限制。此外,分类的过程中需要频繁地进行交叉验证,导致算法的时间开销过大。针对这一问题,我国学者周志华等^[10]提出了 tri-training 模型。tri-training 使用了三分类器,可以简便地处理标记置信度估计问题以及对未知样本的预测分类问题。

在 CNN 的训练过程中,可以很容易地通过设置不同的卷积核矩阵和激励函数获取到不同的图像局部感受野,从而得到多分类器协同训练所需要的冗余特征视图。因此,CNN 适合于多分类器协同训练过程。本文提出卷积神经网络的多分类器协同训练算法(Co-training Algorithm for Multiple CNN, CAMC),利用大量未标记的图像样本对多个 CNN 分类器进行协同训练;首先采用不同的卷积核提取不同的图像特征来初始化训练 3 个不同的 CNN 图像分类器;然后基于多分类器协同训练算法,利用少量已标记的样本和大量未标记的样本对多个分类器进行周期性的协同训练,以提高分类准确度;采用投票机制就多个分类器对未知样本的最终分类结果进行融合,算法的详细描述请见第 3 节。

2 卷积神经网络

如图 1 所示,卷积神经网络包含两种特殊的网络结构:卷积层和次采样层。卷积层中含有多个特征平面,特征平面由神经元构成,同一个特征平面的所有神经元共享相同的连接权值,特征平面上的每个神经元只接收其相应感受野所传输的信号。每个特征平面可以与前一层的几个特征图的卷积建立关系,代表上一层某方面的特征。

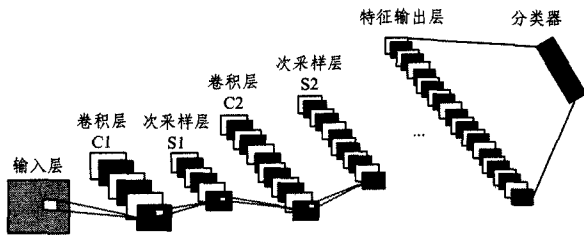


图 1 卷积神经网络结构

在卷积层,前一层的特征图与一个可学习的核进行卷积,卷积的结果经过激励函数后的输出形成这一层的特征图。一般卷积层的输出形式如式(1)所示:

$$x_j^l = f\left(\sum_{i \in M_j} x_i^{l-1} \cdot k_{ij}^l + b_j^l\right) \quad (1)$$

其中, l 代表层数, k 是卷积核, M_j 代表输入特征图的一个选择。每个输出图有一个偏置 b 。由于输入的样本通过卷积层在特征空间重构映射到了高维空间,因此要通过次采样层进行降维,否则容易造成过拟合和维数灾难^[11]。

次采样层的神经元对卷积层上相应的感受野进行抽样,以降低特征维数。次采样层上的每个神经元共享同一链接权重,其输出的一般形式如式(2)所示:

$$x_j^l = f(\beta_j^l \cdot \text{down}(k_{ij}^l) + b_j^l) \quad (2)$$

其中, $\text{down}(\cdot)$ 表示次采样函数。次采样函数一般对该层输入图像的一个 $n \times n$ 大小的区域进行求和、取极值或均值运算,因此输出图像的大小是输入图像大小的 $1/n$ 。每一个输出的特征图有自己的 β 和 b 。卷积神经网络通过卷积层和次采样层的相互配合来学习原始图像的特征,并且通过经典的 BP(Back Propagation)算法来调整参数,完成权值的更新,最终完成学习任务。BP 网络更新权值公式为^[12]:

$$w(t+1) = w(t) + \eta \delta(t) \cdot x(t) \quad (3)$$

其中, $x(t)$ 为神经元的输出, η 代表学习速率, $\delta(t)$ 表示该神经元的误差项,该误差项对于输出层神经元和隐藏层神经元有不同的表达,关于 CNN 权值调整的过程请参见文献^[12],此处不再赘述。

3 CNN 协同训练算法 CAMC

如图 2 所示,CAMC 包含 3 个基本 CNN 分类器: h_1, h_2, h_3 。 $h_i (i=1, 2, 3)$ 使用不同的卷积核提取局部感受野来形成不同的特征视图,以满足协同训练过程对于差异化样本特征的需要。训练过程使用少量已标记的样本 L 以及大量未标记的样本 U 对 CNN 进行协同训练。

在每一轮协同训练过程中,分类器 h_k 和 h_j 从 U 中随机选择样本进行标记;若 h_k 和 h_j 对来自 U 中的样本 x 的标记相同: $label_k(x) = label_j(x)$,则将样本 $(x, label_k(x))$ 添加到新标记样本集 L_i 。 L_i 连同已标记样本集 L 用于对分类器 h_i 进行更新训练,其中 $i, j, k \in \{1, 2, 3\}$,且 $j, k \neq i$ 。注意,每一轮新标记的样本在分类器更新训练后重新被放回 U 中,再开启新一轮迭代训练,直至所有分类器不再发生变化。分类器训练结束后,在未知样本的检测与分类过程中,分类器采用投票机制对样本的类别进行融合判断,如式(4)所示。

$$h(x) = \operatorname{argmax}_{y \in \text{label}} \sum_{i, h_i(x)=y} 1 \quad (4)$$

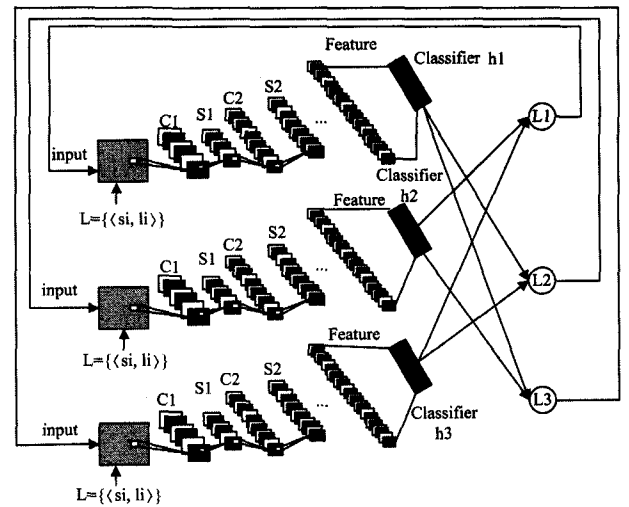


图 2 CNN 协同训练过程 CAMC

根据 Angluin 和 Laird 提出的噪声学习理论^[13],假设 $\sigma_i = \{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\}$ 表示抽取出的 m 个样本序列,其中 x_i 为样本数据, y_i 为样本标记,若样本数量 m 满足 $m \geq \frac{2}{\epsilon^2 (1-2\eta)^2} \ln(\frac{2N}{\delta})$,则有

$$\Pr[d(h_j, h^*) \geq \epsilon] \leq \delta \quad (5)$$

其中, ϵ 是分类模型的误差上限, δ 是置信度参数, η 是训练集分类错误率, N 是类别数目, $d(h_j, h^*)$ 是假设(分类器) h_j 和真实情况 h^* 间的均衡差别(symmetric difference)的概率之和。因此,在给定分类误差上限 ϵ 和置信度参数 δ 的情况下,为满足式(5),有式(6)成立。

$$(1-2\eta)^2 > \frac{2}{\epsilon^2 m} \ln(\frac{2N}{\delta}) \quad (6)$$

对式(6)的左侧进行展开,得到

$$1 - \frac{2}{\epsilon^2 m} \ln(\frac{2N}{\delta}) \geq 4(\eta - \eta^2) \geq 4\eta \quad (7)$$

对于给定的期望误差上限 ϵ ,训练集的分类误差 η 需满足

$$\eta \leq \frac{1 - \frac{2}{\epsilon^2 m} \ln(\frac{2N}{\delta})}{4} \quad (8)$$

令 $c = 2\mu \ln(\frac{2N}{\delta})$, 其中 μ 为令式(8)取等号的系数, 则可得

$$\frac{c}{\epsilon^2} = m(1 - 4\eta) \quad (9)$$

由式(9)可知, 误差上限 ϵ 的平方项与 $m(1 - 4\eta)$ 成反比。设第 t 轮训练的样本规模为 $m^t = |L| + |L^t|$, 样本检测错误率为 η^t , 若满足式(10), 即第 t 轮训练结果的误差上限 ϵ 较第 $t-1$ 轮更低, 则表明增加 $|L^t|$ 个未知样本对分类器进行更新能够持续改进分类性能; 否则放弃本轮新标记的样本, 从 U 中重新进行采样训练。

$$(|L| + |L_u^t|)(1 - 4\eta^t) > (|L| + |L_u^{t-1}|)(1 - 4\eta^{t-1}) \quad (10)$$

其中, 样本检测错误率 $\eta^t = \frac{e_L^t |L| + e_u^t |L_u^t|}{|L| + |L_u^t|}$, 其中 e_L^t 和 e_u^t 分别代表第 t 轮训练过程中 CNN 在已标记样本集 L 和未标记样本集 L_u^t 上的错误率。将 η^t 的表达式代入式(10), 为了使该式成立, 即训练过程能够保证持续降低分类误差上限 ϵ , 需满足

$$|L_u^t| - 4(e_L^t |L| + e_u^t |L_u^t|) > |L_u^{t-1}| - 4(e_L^{t-1} |L| + e_u^{t-1} |L_u^{t-1}|) \quad (11)$$

通常模型在已标记样本上的分类错误率 $e_L^t \ll e_u^t$, 式(11)可以进一步简化为:

$$|L_u^t| (1 - 4e_u^t) > |L_u^{t-1}| (1 - 4e_u^{t-1}), \text{ 即} \quad (12)$$

$$\frac{|L_u^t|}{|L_u^{t-1}|} > \frac{1 - 4e_u^{t-1}}{1 - 4e_u^t}$$

因此, 若相邻两轮新标记样本规模和错误率满足式(12), 则将 h_j, h_k 分类标记相同的新标记样本提交给 h_i 进行训练, 其中 $i, j, k \in \{1, 2, 3\}$, 且 $j, k \neq i$; 否则舍弃本轮新标记的样本, 重新从 U 中选择未标记样本进行分类器训练。实际应用中, 由于无法直接计算未标记样本的分类错误率 e_u^t , 因此采用十折交叉验证法, 从 L 中随机选择 $1/10$ 的已标记样本作为估计 e_u^t 的测试集, L 中的其余样本连同 U 一起对分类器进行训练。协同训练算法 CAMC 如表 1 所列。

表 1 多分类器协同训练算法

CAMC ($L, U, \delta, N, \epsilon_e, \sigma$)	
Input:	L : 已标记样本集; U : 未标记样本集; δ : 置信度参数 N : 类别数目; ϵ_e : 期望误差上限; σ : 错误率阈值
Output:	h_1, h_2, h_3
Step1	从已标记样本集 L 中随机选择 $1/10$ 的样本作为测试集 T , 其余已标记样本 $L' = L - T$ 用于初始化分类器 h_1, h_2, h_3 , 迭代训练次数 $t = 1, e_u^0 = 1$;
Step2	对于分类器 $h_i, i \in \{1, 2, 3\}$, 采用 h_j, h_k 对 U 中的每个样本进行标记, 对于任意 $x \in U$, 若 $h_j(x) = h_k(x) (j, k \neq i)$, 则将 x 并入 h_i 的更新训练集 L_i ;
Step3	基于测试集 T 和 L 计算 CNN 的分类错误率 η^t , 若 $\eta^t > 1/4 - \frac{1}{2\epsilon_e^2 m} \ln(\frac{2N}{\delta})$, 违反式(8), 则舍弃 L_i 转 Step2;
Step4	计算 e_u^t , 若 $\frac{ L_i^t }{ L_i^{t-1} } \leq \frac{1 - 4e_u^{t-1}}{1 - 4e_u^t}$, 违反式(12), 则舍弃 L_i 转 Step2;
Step5	采用训练集 $L' \cup L_i$ 对 h_i 进行更新 $h_i = \text{update}(h_i, L', L_i), i \in \{1, 2, 3\}$;
Step6	若对于分类器 h_1, h_2, h_3 , 均有 $ e_u^t - e_u^{t-1} < \sigma$, 表明分类器性能不再发生改变, 转 Step7, 否则转 Step2;
Step7	分类器 h_1, h_2, h_3 采用投票法对 $L \cup L_1 \cup L_2 \cup L_3$ 进行分类测试, 若分类错误率 $\eta < \epsilon_e$, 则中止分类器训练过程; 否则转 Step1, 重新开始迭代训练。

4 实验

本节通过实验验证 CAMC 的性能。实验采用卡耐基梅隆大学发布的人脸表情标准数据集 CMU Face Images Data Set^[14] 作为已标记训练样本集 L , 其中包含 640 幅已标记表情的灰度图片, 如图 3 所示, 该样本集拥有 Neutral, Sad, Happy, Angry 4 种典型的表情特征。此外, 从 Cohn-Kanade 人脸库^[15] 中随机选择了 90 个表情序列, 每个序列选取 20 幅包含 4 种不同表情的图像共 1800 幅人脸表情样本图像作为未标记样本集 U 。实验过程中, 首先将全部未标记图像转换为 48×48 的灰度图像, 然后采用 CAMC 对分类器进行协同训练。测试过程采用日本表情数据库 JAFFE^[16] 作为测试集对模型进行分类能力进行验证。需要指出的是, 实验过程使用的已标记样本训练集、未标记样本训练集以及测试样本集均来自于不同的数据集, 且拥有不同的初始分辨率和拍摄角度, 因此实验环境设置接近于真实应用场景。

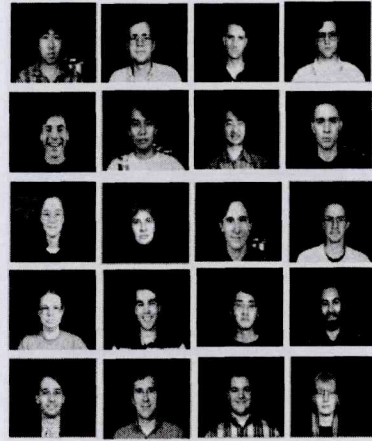


图 3 人脸表情数据集 CMU Face Images Data Set

实验过程中 CAMC 包含 3 个 CNN 分类器 h_1, h_2, h_3 , 它们各自拥有 5 个网络层 C_1, S_2, C_3, S_4, O_5 , 其中 C_1, C_3 为卷积层, S_2, S_4 为次采样层, O_5 为输出层; h_1, h_2, h_3 的 O_5 层输出特征分别连接到 BP, SVM 和 RBF 分类器。 h_1, h_2, h_3 在卷积层和次采样层采用不同的卷积核和激励函数, 通过不同的局部感受野提取差异化的特征视图用于多分类器协同训练。实验过程首先基于差异化的样本特征对 h_1, h_2, h_3 进行初始化; 然后利用 CAMC 对 3 分类器进行持续的协同训练。训练完毕后, 在测试集上对 Neutral, Sad, Happy, Angry 4 种表情进行测试, 并与传统的特征识别法 LBP^[17] 和 Gabor^[18] 进行比较。

在不同表情上的实验结果如图 4—图 7 所示, 其中 h_1, h_2, h_3 代表单个分类器的分类准确率, Fusion 代表融合分类准确率。从图中可以看出, 协同训练后的 3 个分类器 h_1, h_2, h_3 在 4 种表情上的分类准确率随着无标记训练样本的增加而增加。对于图像表情识别问题, 在训练样本较少 (小于 600) 时, 由于可供 CNN 权值调整训练的已标记样本特征较少, 深度学习模型对表情的正确识别率较低, 单个分类器的分类准确率均低于 75%, 随着参与训练的未标记样本的增加, CNN 网格权重得到持续的优化调整, 表情分类精度逐渐提

高;在未标记样本数量达到 1800 时,融合分类准确率 Fusion 均超过 85%。实验结果表明将未标记样本用于 CNN 分类器训练能够有效提高分类准确率。

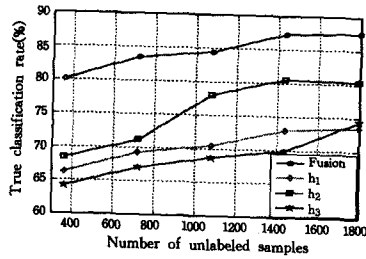


图 4 Neutral 上的分类结果

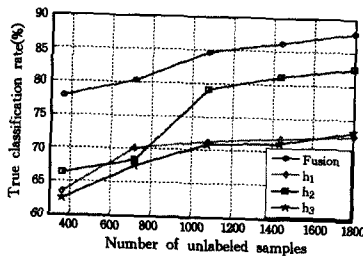


图 5 Sad 上的分类结果

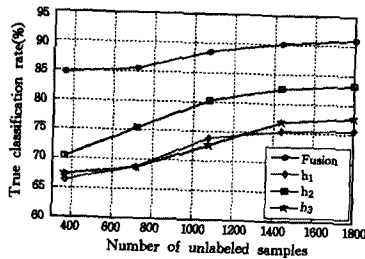


图 6 Happy 上的分类结果

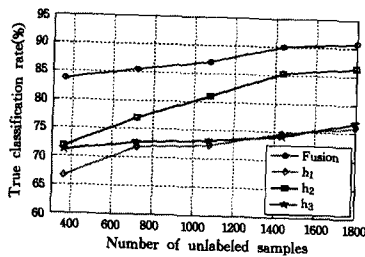


图 7 Angry 上的分类结果

为了比较分类器的性能,对比分析了分类器训练结束后 CAMC 与传统的表情特征识别方法 LBP 及 Gabor 在同样本集上的平均识别准确率,实验结果如表 2 所列。从表中可以看出,CAMC 在表情识别率上优于传统的表情特征识别法 Gabor 和 LBP;尤其在 Sad 和 Neutral 两种表情上,识别率远高于 Gabor 和 LBP。这是因为不同的人脸悲伤或正常表情差异较大,且两种表情带有的纹理信息变化较少,导致样本分类特征的区分度较弱。若仅通过少量已标记样本对分类器进行建模训练,则易导致分类器训练不足,且对部分训练样本特征存在过度拟合,导致分类器泛化能力较差,分类准确率较低。而 CAMC 同时通过 640 幅已标记样本和 1800 幅未标记样本对分类器进行协同训练,一方面,从不同的局部感受野提取到的样本特征有利于描述不同的人脸表情差异;另一方面,CAMC 获得的冗余的特征视图有利于使用不同的分类器基于不同的特

征进行协同训练,持续优化卷积核权重,最终提升分类准确率。

表 2 表情识别准确率

expression	LBP(%)	Gabor(%)	CAMC(%)
Neutral	76.05	77.86	85.3
Sad	81.05	83.6	91.5
Happy	87.7	86.2	93.06
Angry	88.3	90.1	93.2

结束语 本文提出了一种基于多分类器协同学习的卷积神经网络训练算法 CAMC。首先采用多个不同的卷积核提取出多种样本特征以满足协同训练算法多特征视图的需要;然后将不同的特征视图提交给 3 个 CNN 分类器进行协同训练,分类器两两协同以标记新的训练样本;最后新的已标记样本被提交给第 3 个分类器用于周期性的更新训练,从而实现了 CNN 分类器整体性能的持续改进。CAMC 结合了 CNN 多特征视图的提取能力和协同学习算法利用未标记样本进行分类器训练的优势,实现了利用大量未标记样本进行分类器训练的目的,降低了 CNN 对人工标记样本的依赖,提高了卷积神经网络的实用性。

参考文献

- [1] Liu Cong, Xu Wei-sheng, Wu Qi-di. Spatiotemporal convolutional neural networks and its application in action recognition [J]. Computer Science, 2015, 42(7): 245-249 (in Chinese)
- [2] 刘琮, 许维胜, 吴启迪. 时空域深度卷积神经网络及其在行为识别上的应用[J]. 计算机科学, 2015, 42(7): 245-249
- [3] Evgeny A S, Denis M T, Serge N A. Comparison of Regularization Methods for ImageNet Classification with Deep Convolutional Neural Networks [J]. Aasri Procedia, 2014, 6(1): 89-94
- [4] Dong Zhen, Wu Yu-wei, Pei Ming-tao, et al. Vehicle Type Classification Using a Semisupervised Convolutional Neural Network [J]. IEEE Transactions on Intelligent Transportation Systems, 2015, 16(4): 2247-2256
- [5] Wu Hai-bing, Gu Xiao-dong. Towards dropout training for convolutional neural networks [J]. Neural Networks, 2015, 71: 1-10
- [6] Shi He-huan, Xu Yue-lei, Ma Shi-ping, et al. Convolutional neural networks recognition algorithm based on PCA [J]. Journal of Xidian University, 2016, 43(3): 175-180 (in Chinese)
- [7] 史鹤欢, 许悦雷, 马时平, 等. PCA 预训练的卷积神经网络目标识别算法[J]. 西安电子科技大学学报, 2016, 43(3): 175-180
- [8] Liu Fa-yao, Lin Guo-sheng, Shen Chun-hua. CRF learning with CNN features for image segmentation [J]. Pattern Recognition, 2015, 48(10): 2983-2992
- [9] Deng Chao, Guo Mao-Zu. Tri-Training and Data Editing Based Semi-Supervised Clustering Algorithm [J]. Journal of Software, 2008, 19(3): 663-673
- [10] Blum A, Mitchell T. Combining labeled and unlabeled data with co-training [C]//Conference on Computational Learning theory. ACM Press, 1998: 92-100
- [11] Goldman S, Zhou Y. Enhancing supervised learning with unlabeled data [C]//The 17th International Conference on Machine Learning. Morgan Kaufmann Publishers, 2000: 327-334
- [12] Zhou Z H, Li M. Tri-Training: Exploiting unlabeled data using three classifiers [J]. IEEE Trans. on Knowledge and Data Engineering, 2005, 17(11): 1529-1541

能力以进一步满足实际需要,并对模型的可行性和有效性进行了算例和实验验证。在实验分析中,从UCI数据集中选取了几组数据集,说明了所提出的方法明显优于传统优势关系下的变精度粗糙集模型。另外,由表4观察到两类问题的测试精度较高,而多类问题如glass与Heart Disease的测试精度较低,可以进一步研究分类数目对测试精度的影响。不过与传统模型相比,所提出的扩展模型在这两组数据上仍能明显提高测试精度。从图5中发现SPECTF在 $\alpha=0.95$ 处的值小于 $\alpha=1$ 时的值,但是其它位置的值均比较大。未来可以进一步探讨类别数及参数的取值对测试精度的影响。

参 考 文 献

[1] Pawlak Z. Rough sets[J]. International Journal of Information and Computer Sciences, 1982, 11(3): 341-356

[2] Pawlak Z. Rough sets; theoretical aspects of reasoning about data[M]. Boston: Kluwer Academic Publishers, 1991

[3] 苗夺谦, 李道国. 粗糙集理论、算法与应用[M]. 北京: 清华大学出版社, 2008

[4] Ziarko W. Variable precision rough set model [J]. Journal of Computer and System Sciences, 1993, 46(1): 39-59

[5] 张文修, 吴伟志, 梁吉业, 等. 粗糙集理论与方法[M]. 北京: 科学出版社, 2003: 123-125

[6] Xie F, Lin Y, Ren W. Optimizing model for land use land cover retrieval from remote sensing imagery based on variable precision rough set[J]. Ecological Modeling, 2011, 222(2): 232-240

[7] Xie G, Yue W, Wang S, et al. Dynamic risk management in petroleum project investment based on a variable precision rough set model [J]. Technological Forecasting & Social Change, 2010, 77(6): 891-901

[8] Chen Hang, Tao Jun, Zhang Jian-de. Intrusion detection research based on variable precision rough set[J]. Journal of Qingdao University of Science and Technology (Natural Science Edition), 2014, 35(2): 196-199 (in Chinese)
陈行, 陶军, 张建德. 基于变精度粗糙集入侵检测研究[J]. 青岛科技大学学报(自然科学版), 2014, 35(2): 196-199

[9] Greco S, Matarazzo B, Slowinski R. Rough Approximation of a preference relation by dominance relations[J]. European Journal of Operational Research, 1999, 117(1): 63-83

[10] Greco S, Matarazzo B, Slowinski B. Rough approximation by dominance relations[J]. International Journal of Intelligent Systems, 2002, 17(2): 153-171

[11] Li S Y, Li T R. Incremental update of approximations in dominance-based rough sets approach under the variation of attribute values[J]. Information Sciences, 2015, 294: 348-361

[12] Li S. Y, Li T R, Liu D. Incremental updating approximations in dominance-based rough sets approach under the variation of attribute set[J]. Knowledge-Based Systems, 2013, 40(1): 17-26

[13] Du W, Hu B. Approximate distribution reducts in inconsistent interval-valued ordered decision tables[J]. Information Sciences, 2014, 271(7): 93-114

[14] Wong S K M, Ziarko W. Comparison of the probabilistic approximate classification and the fuzzy set model[J]. Fuzzy Sets and Systems, 1987, 21(3): 357-362

[15] Yao Y Y, Wong S K M. A decision theoretic framework for approximating concepts[J]. International Journal of Man-machine Studies, 1992, 37(6): 793-809

[16] Dubois D, Prade H. Rough fuzzy sets and fuzzy rough sets[J]. International journal of general system, 1990, 17(2): 191-209

[17] Parthalaïn N M, Shen Q, Jensen R. A distance measure approach to exploring the rough set boundary region for attribute reduction[J]. IEEE Transactions on Knowledge and Data Engineering, 2010, 22(3): 306-317

[18] Hu Jie, Zhao Hui, Huang Chang-qiang, et al. The variable precision dominance rough set in the application of UCAV in threat estimation[J]. Journal of Air Force Engineering University (Natural Science Edition), 2009, 10(5): 27-31 (in Chinese)
胡杰, 赵辉, 黄长强, 等. 优势变精度粗糙集在UCAV威胁估计中的应用[J]. 空军工程大学学报(自然科学版), 2009, 10(5): 27-31

[19] Luo Gong-zhi, Yang Xiao-jiang. Variable precision dominance rough set attribute reduction preferential algorithm[J]. Chinese Journal of Management Science, 2009, 17(2): 169-175 (in Chinese)
骆公志, 杨晓江. 变精度优势粗糙集属性约简择优算法[J]. 中国管理科学, 2009, 17(2): 169-175

[20] Inuiguchi M, Yoshioka Y, Kusunoki Y. Variable-precision dominance-based rough set approach and attribute reduction[J]. International Journal of Approximate Reasoning, 2009, 50(8): 1199-1214

[21] Karami J, Alimohammadi A, Seifouri T. Water quality analysis using a variable consistency dominance-based rough set approach[J]. Computers, Environment and Urban Systems, 2014, 43(1): 25-33

[22] Zhai L, Khoo L, Zhong Z. A rough set based decision support approach to improving consumer affective satisfaction in product design[J]. International Journal of Industrial Ergonomics, 2009, 39(2): 295-302

[23] 张文修, 仇国芳. 基于粗糙集的不确定决策[M]. 北京: 清华大学出版社, 2005

(上接第226页)

[11] Li Hai-feng, Li Chuan-guo. Note on deep architecture and deep learning algorithms [J]. Journal of Hebei University (Natural Science Edition), 2012, 32(5): 538-544 (in Chinese)
李海峰, 李纯果. 深度学习结构和算法比较分析[J]. 河北大学学报(自然科学版), 2012, 32(5): 538-544

[12] Hinton G E, Salakhutdinov R R. Reducing the dimensionality of data with neural networks [J]. Science, 2006, 313 (5786): 504-507

[13] Angluin D, Laird P. Learning from noisy examples[J]. Machine Learning, 1988, 2(4): 343-370

[14] CMU face images data set [OL]. <http://archive.ics.uci.edu/ml/datasets>

[15] Kanade T, Cohn J, Tian Y. Comprehensive Database for Facial Expression Analysis[C]//Proceedings of the Fourth IEEE In-

ternational Conference on Automatic Face and Gesture Recognition. EE Computer Society, 2000: 46-53

[16] Bashyal S, Venayagamoorthy G K. Recognition of facial expressions using Gabor wavelets and learning vector quantization [J]. Engineering Applications of Artificial Intelligence, 2008, 1(7): 1056-1064

[17] Ojala T, Pietikainen M. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(7): 971-987

[18] Bashyal S, Venayagamoorthy G K. Recognition of Facial Expressions Using Gabor Wavelets and Learning Vector Quantization [J]. Engineering Applications of Artificial Intelligence, 2008, 21(7): 1056-1064