

基于 QPSO-MIL 算法的图像标注

李大湘¹ 彭进业^{1,2} 卜起荣¹

(西北大学信息科学与技术学院 西安 710069)¹ (西北工业大学电子信息学院 西安 710072)²

摘 要 在多数现有图像标注图像库中,关键字只标注在图像级而非区域级,使有监督学习方法在图像标注中难以应用。基于量子粒子群优化算法(quantum-behaved particle swarm optimization, QPSO)提出了一种新的多示例学习(multi-instance learning, MIL)算法——QPSO-MIL 算法,在多示例学习的框架下将基于区域的图像标注问题描述成一个有监督的学习问题。该方法将图像当作包,分割的区域当作包中的示例,利用多样性密度(DD)函数,定义了粒子的适应度向量。在示例空间,利用 QPSO 方法在各个维度上同时搜索 DD 函数的全局极大值点,作为关键字的概念点,然后根据 Bayesian 后验概率最大准则(MAP)对图像进行标注。通过 ECCV 2002 图像库的实验结果表明,QPSO-MIL 算法是有效的。

关键词 多示例学习, 图像标注, 量子粒子群优化

中图法分类号 TP311 **文献标识码** A

QPSO-based Multi-instance Learning for Image Annotation

LI Da-xiang¹ PENG Jin-ye^{1,2} BU Qi-rong¹

(School of Information Science and Technology, Northwest University, Xi'an 710069, China)¹

(School of Electronics Information, Northwestern Polytechnical University, Xi'an 710072, China)²

Abstract In most existing training data set for image annotation, keywords are usually associated with images instead of individual regions, so it is difficult to use supervised learning methods for image annotation. In this paper, based on quantum-behaved particle swarm optimization algorithm(QPSO), a novel multi-instance learning (MIL) algorithm was presented(QPSO-MIL), we formulated image annotation as a supervised learning problem under Multiple-Instance Learning framework. This algorithm regards every image as a bag, and the feature vectors of the segmented regions in this image as instances. We defined a fitness vector for each particle based on the diversity density(DD) function. In the instance feature space we used QPSO to search the global maxima of DD function in each dimension simultaneously, and took the result as a concept point of the keyword, finally assigned corresponding key words to a test image according to the Bayesian maximum a posteriori probability criteria. Experimental results on ECCV 2002 data set indicated that the QPSO-MIL method is effective.

Keywords Multi-instance learning(MIL), Image annotation, Quantum-behaved particle swarm optimization(QPSO)

1 引言

图像自动标注就是采用一定的机器学习算法建立图像视觉特征与高层语义的联系,自动为图像分配相应语义的关键字。随着数字图像数量的爆炸式增长,仅凭人工或半自动的方式获得图像的标注信息,不但非常费时费力,还带有很强的主观偏差,已不能满足人们的需要。由于“语义鸿沟”的存在,使得自动图像标注极具挑战性,成为近几年来来的一个研究热点,具有广阔的应用前景与研究价值。

通常,一幅图像都包含多个局部区域,每个区域都具有相同的视觉特征且代表一个明确的语义概念。所以,一个非常直观的图像标注思路就是首先将图像分割成多个区域,并且提取每个区域的视觉特征,然后利用统计模型或机器学习的

方法,从训练集中找到视觉特征与关键字之间的关系,从而对未标注图像进行标注^[1]。遗憾的是,在大部分图像标注实验的训练图像集中,关键字只标注在图像级(image level),而不是区域级(region level)。如图 1 所示,左边一列是两幅“tiger”类图像,中间一列为其分割区域,右边一列是人工为每幅图像标注的关键字。我们只知道这两幅图像中都包含有“tiger”区域,但具体是哪个区域则不知道。如果用传统的有监督学习方法(例如 SVM 方法)进行图像标注,则需要人工重新将关键字标注到区域级,不但工作量巨大,而且非常繁琐。

在传统的机器学习框架中,训练样本与概念标号是一一对应的。而在多示例学习(Multi-instance Learning, MIL)^[2]问题中,训练样本称之为包(bag),每个包含有多个示例(instance),示例没有概念标号。只有包具有概念标号,若包中

到稿日期:2009-07-31 返修日期:2009-11-18 本文受教育部新世纪优秀人才支持计划项目(NCET-07-0693)资助。

李大湘(1974—),男,博士,工程师,主要研究方向为图像检索、图像标注与图像分类, E-mail: www_ldx@163.com; 彭进业(1964—),男,教授,博士生导师,主要研究方向为数字图像处理; 卜起荣(1976—),男,博士,主要研究方向为模式识别与数据挖掘。

至少有一个示例是正例,则该包被标记为正;若包中所有示例都是反例,则该包被标记为负。多示例学习算法就是通过对训练包的学习来得到一个能对未知包或示例进行预测的分类器。例如,要训练一个关键字为“tiger”的分类器,若将图像当作包(bag),分割区域对应的视觉特征当作包中的示例(instance),对任一幅标注有“tiger”图像对应的包为正包,其它图像对应的包为负包,则可用 MIL 的方法来训练一个“tiger”分类器。所以,本文在 MIL 的框架下,将图像标注问题描述成一个有监督学习问题,并且针对图像标注这个具体的应用问题,基于量子粒子群优化(QPSO)算法,提出了一个新的多示例学习算法,称为 QPSO-MIL 算法。在图像标注的标准测试集 ECCV 2002 上的实验结果表明,QPSO-MIL 算法是有效的。

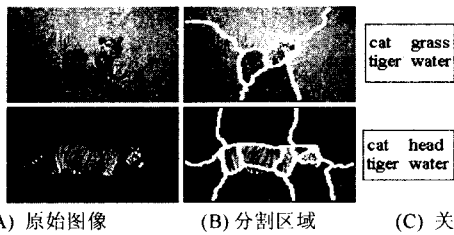


图 1 部分图像示例

2 相关工作

2.1 基于区域的图像标注

近年来,利用机器学习、概率统计等方法提出了很多图像标注方法。Claudi 等^[3]提出了基于 SVM 的图像标注方法,首先将图像分割成不同区域并识别出其中的显著区域,然后手工为每个显著区标注相应的语义关键字,作为训练集。对每一个关键字均训练一个对应的 SVM 分类器。当系统遇到一幅新的图像时,只要利用这些分类器来判断每一个关键字是否应该作为该图像的标注,就可以完成图像的标注。Fan 等^[4]结合本体、显著区检测,提取区域的颜色、纹理和 SIFT 等特征,设计了多核 SVM 分类器进行图像标注。但是,这类有监督的学习方法在训练每个分类器时,需要大量标注到区域级的训练样本,难以推广到类别数较多的情况。

图像标注的另一类方法是建立图像区域视觉特征和语义概念的统计概率模型,对图像进行自动标注。Duygulu 等^[5]提出的翻译模型,将图像标注过程视为从视觉关键字到文本关键字之间的翻译过程,通过寻找标注词和图像特征之间的关系对待标注图像进行标注;Jeon 等^[6]提出的跨媒体相关模型(Cross-Media Relevance Model, CMRM)方法,将图像标注问题看作跨语言检索问题,通过计算 blobs 和语义概念的联合概率进行图像标注。因为 CMRM 在标注过程中,要用聚类的方法产生 blobs,聚类过程往往都会造成视觉特征内容的损失,从而影响了标注效果。在 CMRM 模型的基础上,出现了连续相关模型(CRM)^[7]和多伯努利相关模型(MBRM)^[8]等改进方法。这类无监督的图像标注方法,都要使用聚类的方法产生视觉 blobs,所以聚类质量的好坏直接影响最后的标注精度。

其它的还有 LSA&PLSA 模型和传播模型等图像标注方法。最近,利用领域本体、WordNet 或训练集中词汇的共生关系,对标注结果进行改善的图像标注方法不断出现,并且取

得了更高的标注精度^[9]。

2.2 多示例学习

1997 年, Dietterich 等^[2]在药物活性预测(drug activity prediction)的研究工作中提出了多示例学习(Multi-instance Learning, MIL)的概念,及 6 种不同的轴平行矩形(APR)MIL 算法。Maron 首先对 MIL 进行扩展,定义了多样性密度(Diverse Density, DD)函数^[10],并提出了 DD 算法,用于图像分类;Andrews 等^[11]通过将 MIL 的约束引入到 SVM 的目标函数中,提出了两种基于 SVM 的 MIL 算法:mi-SVM 算法与 MI-SVM 算法;Kwok 等^[12]则将边缘核的定义方法用于度量包之间的相似性,将 MIL 问题转化成传统的 SVM 学习问题;Chen 等通过空间转化的方法,先后提出了 DD-SVM^[13]与 MILES^[14]两种 MIL 算法,均用于图像分类。在 DD-SVM 算法中,基于多样性密度(DD)函数,在示例属性空间寻找正负样本的“概念点”,构造“包属性空间”,再利用一个非线性映射,将每个训练包投影成“包属性空间”中的一个点,将 MIL 问题转化成一个有监督的学习问题。因采用 DD 函数寻找概念点非常费时,在 MILES 算法中则简单地利用训练包中所有的示例构造投影空间。显然,通过这样的方式得到的投影特征维数将特别高,存在大量冗余的、与分类无关的信息。Chen 则采用了 1 范数 SVM,它在 SVM 学习的同时具有特征选择的功能,但其分类性能往往不如 2 范数 SVM。

Yang 等^[15]基于 DD 函数,提出了 SPWDD 多示例学习算法,首次将 MIL 用于图像标注。该法首先选择一个 DD 值最大的示例,作为相应关键字的代表示例,再使用梯度搜索法得到这个代表示例对应的尺度因子 w ,然后采用后验概率最大原则进行图像标注;之后, Yang 等^[16]通过在 SVM 的目标函数中引入一个损失函数,提出了一种 ASVM-MIL 算法,用于图像标注。最近,还提出基于半监督思想的^[17]和多示例多标签^[18]的多示例学习算法,用于图像分类和图像检索。

3 基于 QPSO-MIL 算法的图像标注

3.1 问题描述

对于基于区域的图像标注,首先用图像分割的方法将图像分割成多个区域(Region),然后提取每个区域的颜色、纹理和形状等底层视觉特征,设其为 $X_i \in R^D, i=1, 2, \dots, n$,其中 D 为视觉特征向量的维数, n 为图像分割的区域数。

设 $TST = \{J_1, J_2, \dots, J_M\}$ 表示由 M 幅未标注图像组成的测试集,其中 J_i 用它 m_j 个分割区域对应的视觉特征表示,记作 $J_j = \{X_{j1}, X_{j2}, \dots, X_{jm_j}\}, j=1, 2, \dots, M$; 设 $TRN = \{I_1, I_2, \dots, I_N\}$ 表示由 N 幅已标注图像组成的训练集,其中 I_i 用它对应的 n_i 个分割区域的视觉特征和手工标注的关键字集进行表示,记作 $I_i = \{X_{i1}, X_{i2}, \dots, X_{in_i}\}, i=1, 2, \dots, N$ 和 $W_i = \{w_1, w_2, \dots, w_{l_i}\} \subseteq V$, 其中 l_i 表示该图像标注关键字的个数, $V = \{w_1, w_2, \dots, w_l\}$ 表示关键字词汇表。对于给定的任一未标注图像,自动图像标注的任务就是要从 V 中选择一些关键字来准确地描述图像中所包含的高层语义。如果将每个关键字 w_i 当作一个类的标号,图像标注可转化成图像分类问题,采用有监督的学习方法来解决。训练集中的关键字只标注到图像级,没有标注到区域级,限制了传统的有监督学习方法的直接应用。本文在 MIL 的框架下,用有监督学习的方法进行图像标注。

3.2 图像标注与 MIL

对于任意关键字 $w_i \in V$, 首先在训练图像集 TRN 中, 将所有标注有 w_i 的图像对应的包标为正包, 其它没标注 w_i 的图像标为负包, 组成 MIL 的训练集。按照 MIL 训练包的定义, 在理想的情况下, 正包中所有示例的交集减去负包中所有示例的并集, 就会得到真正的正例。这些真正的正例只会出现在正包中, 不会出现在任何负包中。但是在实际应用过程中, 由于样本中存在噪声, 这样严格的交、并、差运算, 往往得到的最后结果是空集。1998 年, O. Maron 等^[10] 提出了多样性密度 (Diverse Density, DD) 算法。在他们的定义中, 在示例属性空间中的某点附近出现的正包数越多, 而反包中的示例都离它越远, 则该点的 DD 值越大, 把它作为目标概念点, 用一种软交的、并、差方法求解 MIL 问题。

设 B_i^+ 表示第 i 个正包, 其中包含 n_i^+ 个示例 $X_{ij}^+ \in R^D, j=1, 2, \dots, n_i^+, x_{ijk}^+, k=1, 2, \dots, D$ 表示第 i 个正包的第 j 个示例的第 k 个属性值; 同理, B_i^- 表示第 i 个负包, 其中包含 n_i^- 个示例 $X_{ij}^- \in R^D, j=1, 2, \dots, n_i^-, x_{ijk}^-$ 表示第 i 个负包的第 j 个示例的第 k 个属性值, 正包(负包)个数表示为 $l^+ (l^-)$ 。令 h 代表 DD 值最大的点, 则可通过最大化 $\Pr(x=h | B_i^+, \dots, B_i^+, B_i^-, \dots, B_i^-)$ 确定 h 。假设各包互相独立, 根据 Bayes 理论, 可通过式(1)确定 h ^[10]:

$$\operatorname{argmax}_h \left(\prod_i P(h | B_i^+) \prod_i P(h | B_i^-) \right) \quad (1)$$

式(1)就是 DD 函数的一般定义。在实际应用中, 对乘积项采用 noisy-or 模型进行例化:

$$P(x | B_i^+) = 1 - \prod_j (1 - \exp(-\|X_{ij} - x\|^2)) \quad (2)$$

$$P(x | B_i^-) = \prod_j (1 - \exp(-\|X_{ij} - x\|^2))$$

式中, $\|X_{ij} - x\|^2 = \sum_{k=1}^D (x_{ijk} - x_k)^2$ 。现在不妨假设 h 就是关键字 w_i 的概念点(即 DD 值最大的点), 则关键字 w_i 的类条件概率定义为:

$$\Pr(X | w_i) = \exp(-\|X - h\|^2) \quad (3)$$

式中, X 表示任意示例。

设测试图像 J 分割成 m 个区域, 对应的视觉特征记作 X_j , 其中 $j=1, 2, \dots, m$ 。如果至少存在一个区域具有关键字 w_i 的语义, 则可用 w_i 来标注该图像。根据 Bayesian 后验概率最大准则, 设 $p(w_i)$ 表示 w_i 的先验概率, 则测试图像 J 与 w_i 的后验概率为:

$$\Pr(w_i | J) = \operatorname{argmax}_j \{P(w_i | X_j) * p(w_i)\} \quad (4)$$

对于词汇表 $V = \{w_1, w_2, \dots, w_L\}$ 中的每个关键字, 通过相同的方法为其在示例空间找到一个概念点, 记为 $H = \{h_1, h_2, \dots, h_L\}$, 自动进行图像标注, 则可以用式(4)来计算每个关键字与它的后验概率, 然后选择概率最大的几个关键字作为图像的标注信息。

为了搜索 DD 函数的最大值点 h , 常用的方法有梯度下降法^[12] 与拟牛顿搜索法^[16]。因为 DD 函数是连续且高度非线性的, 存在很多局部极大值点, 这些算法的搜索过程中要将每一个正包中的示例作为初始起点进行搜索, 不但非常费时, 而且常导致算法收敛到局部最优, 使得到的概念点带有很大的偏差, 影响算法最后的性能。

3.3 QPSO-MIL 算法

1995 年, Eberhart 博士和 Kennedy 博士模拟鸟类的觅食

行为, 基于群体的随机搜索算法提出了粒子群优化算法 (Particle Swarm Optimization, PSO)^[19], 系统初始化一组随机解, 通过不断地迭代寻找最优解, 可以很好地求解非线性、不可微和多峰等问题。传统的 PSO 算法存在两个严重的缺点: 易发生过早收敛、在搜索后期效率较低。这使得最终搜索得到的结果可能不是全局最优解, 而是局部最优解。为了提高算法性能, 已有学者提出多种改进方案^[20]。

3.3.1 QPSO 算法

J. Sun 在研究了 Clerc 等人关于粒子收敛行为的研究成果后, 从量子力学的角度提出了一种新的 PSO 算法模型。这种模型以 δ 势阱为基础, 认为粒子具有量子行为, 提出了量子粒子群优化算法 (Quantum-behaved Particle Swarm Optimization, QPSO)^[21]。该算法显著的特点是控制参数少, 设置简单, 搜索能力强, 具有更好的全局搜索能力。在 QPSO 算法中, 粒子只带有位置信息, 粒子的进化过程如下:

$$\begin{cases} x_i^{t+1} = p + \beta \cdot |Mbest - x_i^t| \cdot \ln(1/u), & \text{if } k \geq 0.5 \\ x_i^{t+1} = p + \beta \cdot |Mbest - x_i^t| \cdot \ln(1/u), & \text{if } k < 0.5 \end{cases} \quad (5)$$

式中, x_i^t 是 t 时刻第 i 个粒子的位置, u, k 为 $(0, 1)$ 之间的随机数; $Mbest$ 是群体的最优中心, 定义为所有 M 个局部最优粒子的中心位置, 计算公式如下:

$$Mbest = \frac{1}{M} \left(\sum_{i=1}^M Pbest_{i1}, \dots, \sum_{i=1}^M Pbest_{iD} \right) \quad (6)$$

式中, m 是粒子的个数, $Pbest_i$ 是第 i 个粒子的历史最优位置。为了保证所有粒子向最优粒子靠拢, 定义 p 如下:

$$p = (c_1 \cdot Pbest + c_2 \cdot Gbest) / (c_1 + c_2) \quad (7)$$

式中, c_1, c_2 为 $(0, 1)$ 之间的随机数; $Gbest$ 是全局最优粒子。 β 是收缩扩张因子, 用于控制算法的收敛速度, 运行过程中动态调节方式如下:

$$\beta = 1.0 - \frac{t}{(2 * MAXITER)} \quad (8)$$

式中, $t, MAXITER$ 分别表示当前和最大迭代次数。

3.3.2 适应度函数设计

适应度函数是衡量粒子优劣的唯一标准。在求解 MIL 问题的应用中, 本文定义了粒子的适应度向量, 从示例属性空间的每个维度上对粒子的适应度进行衡量。

定义 1(粒子的适应度向量) 利用 DD 函数, 分别计算每个粒子 $x = (x_1, x_2, \dots, x_D)$ 在示例特征空间每个维度上的 DD 值, 记作 $F_d(x)$

$$F_d(x) = \prod_i P(x_d | B_i^+) \prod_i P(x_d | B_i^-) \quad (9)$$

称 $Fit(x) = [F_1(x), F_2(x), \dots, F_D(x)]$ 为粒子 x 的适应度向量, 每个分量分别从各个维度上对粒子的适应度进行衡量。类似于式(2), 其中

$$P(x_d | B_i^+) = 1 - \prod_j (1 - \exp(-(x_d - x_{jd})^2))$$

$$P(x_d | B_i^-) = \prod_j (1 - \exp(-(x_d - x_{jd})^2)) \quad (10)$$

3.3.3 QPSO-MIL 算法步骤

根据 MIL 与 DD 函数的定义, DD 值大的点往往都离正包中的示例近, 离负包中的示例远, 所以本文从正包中随机选择不同的示例, 初始化第一代粒子。本文定义了粒子的适应度向量, 从各个维度上同时搜索 DD 函数的全局极大值点, 以极大地提高粒子收敛的速度。虽然能够保证算法的全局收敛性, 但是在收敛的情况下, 由于所有的粒子都向最优解的方向

飞去,导致粒子不可避免地趋向同一化(多样性损失),丧失了粒子的全局搜索能力。同时,算法收敛到一定精度时,无法继续优化,进而陷入局部最优解。针对此问题,本文定义了粒子群在各个维度上的拥挤度函数:

$$G(d) = \sum_{i=1}^M (|x_d - Gbest_d|) \quad (11)$$

式中, $d=1, 2, \dots, D$ 。当在任一个维度上,拥挤度小于一个阈值时,对粒子群该维上的数据进行重新随机初始化,让粒子群获得新的搜索位置,从而增加了搜索更优解的机会。如此反复不已,直到迭代结束。最后, QPSO-MIL 算法步骤总结如下。

算法 QPSO-MIL 算法

1) 搜索关键字对应的概念点

输入: 带有概念标记的训练包组成的训练集 B ;

输出: 关键字 w_i 对应的目标概念点 h ;

初始化: 设置粒子种群规模 M , 最大迭代次数 $MAXITER$, 拥挤度阈值 Th ;

Step 1 初始化 M 个粒子的初始位置, 组成第一代种群 $X^1 = \{x_1, x_2, \dots, x_M\}$, 具体方法是: 将所有正包中的全部示例取出来, 组成一个集合 S , 从中随机选择 M 个示例, 作为每个粒子的初始位置; 若 S 中元素的个数小于 M , 则对剩余的粒子进行随机初始化。设置当前迭代次数 $t=1$, 初始化每个粒子的历史最优位置 $Pbest_i$ 与适应度向量 $[PF_1(x_i), PF_2(x_i), \dots, PF_D(x_i)]$ 、全局最优位置 $Gbest$ 与适应度向量 $[GF_1, GF_2, \dots, GF_D]$ 。

Step 2 ①根据式(5)一式(8)更新粒子的位置; ②由式(9)计算每个粒子的适应度向量 $F(x_i)$; ③更新每个粒子的历史最优位置与适应度; 具体是:

For $i=1$ to $M // M$ 为种群规模

For $d=1$ to $D // D$ 为数据维数

{ if $F_d(x_i) > PF_d(x_i)$

$PF_d(x_i) = F_d(x_i); //$ 更新 d 维上的适应度

$Pbest_i(d) = x_i(d); //$ 更新 d 维上的数值

end if

}

④用类似方法更新全局最优位置及适应度向量; ⑤用式(11)判断粒子群各个维度上的拥挤度。若存在小于阈值 Th 的维度, 则对所有粒子此维的数值重新随机初始化; ⑥设置当前迭代次数 $t=t+1$, 若 $t > MAXITER$, $h=Gbest$, 迭代结束, 输出 h ; 否则转①。

2) 利用 h 对关键字 w_i 的概率预测

对于未标注图像 J , 令 X_j 为分割区域对应的视觉特征, $j=1, 2, \dots, m$, 用式(4)估计 J 与 w_i 的后验概率。

4 实验结果与分析

4.1 实验图像库

为了验证 QPSO-MIL 算法的有效性, 采用图像标注的基准数据集 ECCV 2002 进行了对比实验。该数据集包括 5000 幅图像, 来自 50 个 Corel Photo CDs, 每个 CD 目录下包含同一主题的 100 幅图像。为了确保对比的公平性, 本文直接采用文献[5]处理后的数据进行实验。每幅图像采用 Normalized Cut 方法分割为 1~10 个区域, 每一个区域用一个 36 维的特征向量来表示(其中包括颜色、形状、位置等特征), 并且每幅图像手工标注 1~5 个关键字。把数据集分成两个部分, 其中 4500 幅图像作为训练集, 另 500 幅图像作为测试集。整

个标注词汇集共含有 374 个关键词, 而出现在测试集中的关键词为 263 个。

4.2 实验方法

图 2 给出在含有 4500 幅图像的训练集中各标注关键词出现的次数统计。可以看出, 出现次数大于 100 次的词有 36 个, 出现次数大于 50 次的词有 74 个, 出现次数大于 20 次的词有 141 个。也就是说, 有二分之一以上的词只给定了相当少的训练图像。这样的训练集在 QPSO-MIL 训练过程中存在以下两个问题: (1) 训练集太大, 极大地影响训练效率; (2) 几乎对于每个关键字, 正包与负包的数量严重不平衡, 即负包数量远远大于正包数量。为了解决这两个问题, 在本文实验中, 对于给定关键字, 均随机选择与正包数相同的负包, 构造一个规模更小但数量平衡的训练集。实验中, 粒子规模 $M=500$, 最大迭代次数 $MAXITER=100$, 拥挤度阈值 $Th=0.001$ 。为每幅测试图像选取后验概率最大的 5 个关键字作为图像的标注信息, 采用 Recall(查全率)与 Precision(查准率)来衡量算法的性能。因为涉及到随机选择负包, 在后面的实验结果中, QPSO-MIL 算法记录的是 5 次重复实验的平均查全率与查准率。

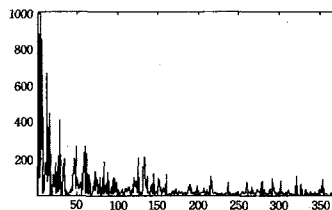


图 2 ECCV2002 训练集标注关键字出现的次数统计

4.3 实验结果与分析

首先, 分别选择出现频率最高的 30 个和 70 个关键字, 与其它基于 MIL 的图像标注方法进行对比, 实验结果如表 1 所列。

表 1 QPSO-MIL 与其它 MIL 算法的性能比较

	70 Keyword		30 Keyword	
	Precision	Recall	Precision	Recall
QPSO-MIL	33.52 %	42.63 %	40.71 %	46.25 %
SPWDD [15]	27.31 %	35.66 %	33.86 %	36.88 %
ASVM-MIL [16]	31.19 %	39.73 %	38.69 %	42.70 %
DD [10]	28.61 %	36.72 %	34.58 %	37.89 %

如表 1 所列, QPSO-MIL 方法性能最好, 主要原因是 QPSO 算法具有很好的全局寻优能力, 往往得到目标概念点是全局最优的, 而不是局部最优的。例如, 分别用 QPSO-MIL 方法与 DD 方法搜索关键字“flower”的概念点, 然后在训练集中列出概念点的 5 个最近邻的示例, 结果如图 3 所示。可以看出, QPSO 方法得到的概念点的 5 个最近邻示例全是正确的“flower”图像块, 而 DD 方法得到的概念点的 5 个最近邻示例均为“grass, tree, leaf”等区域。这是因为“flower”图像中, 往往都伴随着“grass, tree, leaf”等区域, 在相应的 DD 函数中产生一个很大的伪峰, 造成 DD 方法收敛到这个局部最优解, 没有得到正确的概念点。

为了同其它经典的图像标注方法进行比较, 将出现在测试集中的 263 个关键字同翻译模型^[5]、跨媒体相关模型^[6]和连续相关模型^[7]等方法进行了比较, 实验结果如图 4 所示。可以看出, QPSO-MIL 算法性能稍优于这些方法, 这说明融

合 MIL 与 QPSO 的图像标注是有效的。

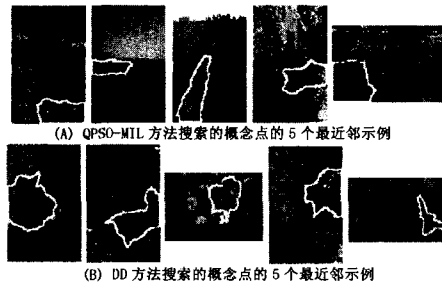


图3 QPSO-MIL与DD方法搜索关键字flower的概念点的性能对比

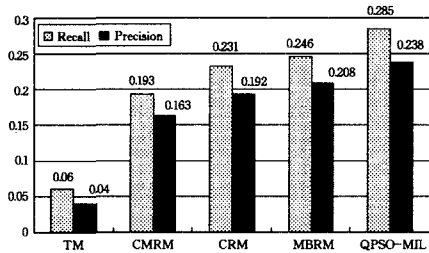


图4 ECCV 2002 图像库上多种图像标注算法的性能对比

4.4 算法效率分析

执行 QPSO-MIL 算法的主体部分是在每次迭代中,要计算每个粒子的适应度向量,其时间复杂度为 $T_{Fit} = O(BN * AIN * D)$,其中 BN 表示训练包的个数, AIN 表示每个包中示例的平均个数, D 表示数据的维数,如果 M 表示粒子规模, N 表示迭代次数,则总时间复杂度 $T = O(N * M * T_{Fit})$ 。在基于 ECCV 2002 图像库的实验中,当训练包数 $BN=200$ 时,计算每个粒子的适应度向量耗时 T_{Fit} 约为 0.008s。迭代过程中,还涉及到粒子的进化与更新,QPSO-MIL 算法总耗时约 430s(其中 $M=500, N=100$,实验平台:AMD4200+处理器,1G DDR400 内存,WindowsXP 操作系统,Matlab7.01 仿真环境)。而 DD 方法要以所有正包中的所有示例为搜索起点,利用梯度下降法进行多次迭代才能得到一次搜索结果。在相同的实验条件下,它完成一次搜索耗时约 6.272s。总体上来说,QPSO-MIL 算法效率高于 DD 方法。

结束语 在基于区域的图像标注问题中,本文的主要创新点在于,将 QPSO 算法与 MIL 问题相结合,提出了一种新的 MIL 算法——QPSO-MIL 算法,为求解 MIL 问题探索了一个新的思路与方向。相对于传统的 DD 算法来说,QPSO 算法不但具有更强的全局寻优能力,而且效率更高。基于 ECCV 2002 图像库的实验结果表明,QPSO-MIL 算法相对其它 MIL 算法,标注准确率更高,是一种有效的图像标注方法,且相对于传统的有监督机器学习算法,不需要用户对训练样本进行精确的手工标注,提高了图像标注的效率。

本文的工作也存在一些不足,主要表现在没有将免疫或遗传等思想引入到 QPSO 算法中,克服量子(粒子)在进化过程中缺乏很好方向指导的缺陷。总之,QPSO-MIL 算法还有待在更大的图像集或其它应用领域进一步验证。由于 MIL 训练样本的特殊性,具有很多探索的空间,是一个值得进一步研究的课题。

参考文献

[1] 张元清,包骏杰,况秀,等. 基于贝叶斯理论的图像标注与检索

[J]. 计算机科学,2008,35(8):229-331

[2] Dietterich T G, Lathrop R H, Lozano-perez T. Solving the multiple instance problem with axis-parallel rectangles [J]. Artificial Intelligence, 1997, 89(12):31-71

[3] Cusano C, Ciocca G, Schettini R. Image annotation using SVM [C] // Proceedings of the SPIE. California, USA, 2004, 5304: 330-338

[4] Fan Jianping, Gao Yuli, Luo Hangzai. Integrating Concept Ontology and Multitask Learning to Achieve More Effective Classifier Training for Multilevel Image Annotation [J]. IEEE Transactions on Image Processing, 2008, 17(3):407-426

[5] Duygulu P, Barnard K, de Freitas JFG, et al. Object recognition as machine translation; learning a lexicon for a fixed image vocabulary [C] // Heyden A, ed. Proc. of the European Conf. on Computer Vision. Berlin:Spring-Verlag, 2002:97-112

[6] Jeon J, Lavrenko V, Manmatha R. Automatic image annotation and retrieval using cross-media relevance models [C] // Proc. of the Int'l ACM SIGIR. Toronto: ACM Press, 2003:119-126

[7] Lavrenko V, Manmatha R, Jeon J. A model for learning the semantics of pictures [C] // Sebastian T, Lawrence K S, Bernhard S, eds. Proc. of the Neural Information Processing Systems (NIPS). Vancouver and Whistler: MIT Press, 2004:553-560

[8] Feng S L, Manmatha R, Lavrenko V. Multiple bernoulli relevance models for image and video annotation [C] // Proc. of the IEEE Conf. Computer Vision and Pattern Recognition. Washington: IEEE Computer Society, 2004:1002-1009

[9] 卢汉清,刘静. 基于图学习的自动图像标注 [J]. 计算机学报, 2008, 31(9):1629-1639

[10] Oded M, Lakshmi R A. Multiple-instance learning for natural scene classification [C] // Proceedings of the 15th International Conference on Machine Learning. Madison, WI, 1998:341-349

[11] Stuart A, Thomas H, Ioannis T. Multiple instance learning with generalized support vector machines [C] // Proceedings of the 18th National Conference on Artificial Intelligence. Edmonton, Canada, 2002:943-944

[12] Kwok J T, Cheung Pak-Ming. Marginalized Multi-instance Kernels [C] // Proceedings of the 20th International Joint Conference on Artificial Intelligence. Hyderabad, India, 2007:901-906

[13] Chen Yixin, Wang Jame Ze. Image categorization by learning and reasoning with regions [J]. Journal of Machine Learning Research, 2004, 5(8):913-939

[14] Chen Yixin, Bi Jinbo, Wang James Ze. MILES: Multiple-instance Learning via Embedded Instance Selection [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2006, 28(12):1931-1947

[15] Yang C, Dong M, Fotouhi F. Region-based image annotation through multiple-instance learning [C] // Proc. of ACM Multimedia. Singapore, 2005:435-438

[16] Yang Chang-bo, Dong Ming, Hua Jing. Region-based Image Annotation Using Asymmetrical Support Vector Machine-based Multiple-instance Learning [C] // Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06). 2006:2057-2063

[17] Zhou Zhi-hua, Xu Jun-ming. On the Relation Between Multi-instance Learning and Semi-supervised Learning [C] // Proceedings of the 24th ICML. Corvallis, Oregon, June 2007:1167-1174

(下转第 296 页)

最高提高了 40%(当 MTU=16bytes 时)。当 MTU=32kbytes 时,改进后 RCIBR 的数据吞吐量达到最高 490Mbytes/sec,是方案一实现的 RCIBR 理想传输速度的 89.4%。根据这组测试结果,笔者最终实现的 RCIBR 底层每次允许传输的最大数据包大小为 32kbytes。

结束语 参考国内外相关文献,笔者首次在 RapidIO 互联协议下实现了 CIBR。本方案实现的 RCIBR 在发送端和接收端均只有一次用户态和内核态之间的数据拷贝,而传统的通信接口在接收端至少有两次内存拷贝。通过采用将 RDMA 操作和内存拷贝等其它操作并行的方法,本方案实现的 RCIBR 表现出了良好的性能优势。参考文献[20],本方案实现的 RapidIO 通信接口比 Linux 下用以太网模拟器实现的 RapidIO 通信接口的最高数据吞吐量提高了 390%,比文献[20]中作者用 RapidIO 消息模式实现的通信接口其数据吞吐量最高提高了 53%。但是这种实现方案尚存在以下缺点。

(1)本方案实现的 RCIBR 占用 CPU 资源最高可达 80%以上。

(2)目前,限制 RCIBR 数据吞吐量的瓶颈因素依然是内存拷贝的速度。

在将来,笔者考虑用双通道内存、优化内存拷贝函数等方式提高内存拷贝的速度,从而进一步提高 RCIBR 的数据吞吐量。

参 考 文 献

[1] RapidIO Trade Association. RapidIO Specification 1.3 [EB/OL]. www.rapidio.org/specs/current,2001

[2] Callaghan B, Lingutla T, Chiu A. NFS over RDMA[C]//Work-in-Progress Presentation, USENIX File Access and Storage Symposium, Monterey, CA, January 2002

[3] Voruganti K, Sarkar P. An Analysis of Three Gigabit Networking Protocols for Storage Area Networks[C]//Proceeding of International Conference on Performance, Computing, and Communications. 2001;259-265

[4] Regnier G, Makinen S, Illikkal R, et al. TCP Onloading for Data Center Servers[J]. IEEE Computer, 2004, 37(11):48-58

[5] Romanow A, Bailey S. An Overview of RDMA over IP[C]//First International Workshop on Protocols for Fast Long-Distance Networks. 2003

[6] Recio R, Culley P, Garcia D, et al. An RDMA Protocol Specification [EB/OL]. draft-ietf-rddp-rdmap-02 (work in progress), 2004

[7] Hilland J, Culley P, Pinkerton J, et al. RDMA protocol verbs

specification (version 1.0)[EB/OL]. RDMA Consortium, 2003

[8] Boden N J, Cohen D, Felderman R E, et al. Myrinet: A gigabit per second local area network[J]. IEEE-Micro. 1995, 15(1):29-36

[9] Liu J, Panda D K, Banikazemi M. Evaluating the impact of rdma on storage I/O over infiniband[C]//SAN-03 Workshop in conjunction with HPCA. 2004

[10] Tipparaju V, Santhanaraman G, Nieplocha J, et al. Host-assisted zero-copy remote memory access communication on infiniband [C]//Intel Parallel and Distributed Processing Symposium (IP-DPS 04). April 2004

[11] Wu J, Wyckoff P, Panda D K. PVFS over InfiniBand: Design and performance evaluation[C]//International Conference on Parallel Processing. Oct. 2003

[12] Jin H W, Narravula S, Brown G, et al. Performance evaluation of rdma over ip: A case study with the ammasso gigabit ethernet nic[C]//Workshop on High-Performance Interconnects for Distributed Computing (at HPDC'05). July 2005

[13] Magoutis K. Memory Management Support for Multi-Programmed Remote Direct Memory Access Systems[C]//2005 IEEE International Conference on Cluster Computing. 2005

[14] Yoon In-su, Chung Sang-hwa. Implementation and Analysis of TCP/IP Offload Engine and RDMA Transfer Mechanisms on an Embedded System[C]//Lecture Notes in Computer Science. 2005;818-830

[15] Liang S, Noronha R, Panda D K. Swapping to remote memory over infiniband: An approach using a high performance network block device[C]//IEEE International Conference on Cluster Computing (Cluster 2005). September 2005

[16] Ou L, He X, Han J. A Fast Read/Write Process to Reduce RDMA Communication Latency[C]//Proceeding of the 2006 International Workshop on Networking, Architecture, and Storage. August 2006

[17] Huang W, Gao Q, Liu J, et al. Panda. High Performance Virtual Machine Migration with RDMA over Modern Interconnects[C]//Proc. Cluster. Austin, TX. September 2007

[18] Liu J, Jiang W, Wyckoff P, et al. Design and Implementation of MPICH2 over InfiniBand with RDMA Support[C]//Proceeding of IPDPS. April 2004

[19] Corbet J, Rubini A, Kroah-Hartman G. Linux Device Drivers (Third Edition)[M]. O'REILLY, 2005;114-115

[20] 梁基. 基于 RapidIO 的高性能通信接口的设计与实现[D]. 上海:上海大学, 2008

(上接第 282 页)

[18] Zhou Zhi-hua, Zhang Min-ling. Multi-instance multi-label learning with application to scene classification [A]//Advances in Neural Information Processing Systems 19[C]. Cambridge, MA, 2007;1609-1616

[19] Kennedy J, Eberhart R. Particle swarm optimization [C]//Proceedings of IEEE International Conference of Neural Networks,

1995. USA: IEEE, 1995;1942-1948

[20] 段其昌, 张红雷. 基于搜索空间可调的自适应粒子群优化算法与仿真 [J]. 控制与决策, 2008, 23(10):1192-1195

[21] Sun Jun, Feng Bin, Xu Wen-bo. Particle Swarm Optimization with Particles Having Quantum Behavior[C]//IEEE Proceedings of Congress on Evolutionary Computation. USA: IEEE, 2004;326-331