

基于跨领域主动学习的图像分类方法

邵 忻*

(天津外国语大学 基础课教学部, 天津 300204)

(*通信作者电子邮箱 shaoxin1979@126.com)

摘 要:针对基于单一领域主动学习的图像分类方法不能利用不同领域图像共同特征导致标记效率低下的问题,提出一种基于跨领域主动学习的图像分类方法。由不同领域图像学习出含共同隐特征的子空间,综合考虑共同特征和领域相关特征,将数据实例引起的模型损失减少量分解到一个共同部分和领域相关部分,从而领域间的共同信息可以编码到模型损失减少的共同部分并用来进行查询。实验结果显示该方法相对于单一模型学习和混合模型学习方法可以减少将近30%的标记工作,并且可以获得更高的精度,表明该方法可以更高效地运用于各种图像分类任务。

关键词:跨领域;迁移学习;主动学习;图像分类;模式识别

中图分类号: TP391.41 **文献标志码:** A

Cross-domain active learning algorithm for image classification

SHAO Xin*

(Faculty of Fundamental Courses, Tianjin Foreign Studies University, Tianjin 300204, China)

Abstract: To solve the problem of ignoring common information in different domains in traditional active learning image classification, a new multi-domain active learning image classification method was proposed to reduce the labeling effort of image classification. It learned a subspace to represent common features among different images. Considering the common features and domain-specific features, the model loss due to each data instance could be divided into two parts, so that the common information could be queried from the common part. The experimental results show that the new method has some precise increase and has 30% less labeling efforts than the single model method and mixture model method. The results reveal that the new method can be widely used in all kinds of image classification tasks with higher precise and efficiency.

Key words: cross-domain; transfer learning; active learning; image classification; pattern recognition

0 引言

图像分类^[1]是图像检索、模式识别中的一个重要应用,其方法通常是对大量的标记样本训练分类模型,然后将一幅图像按其所描述的场景或者包含的物体进行分类。由于图像数目巨大、种类繁多,且图像存在光照、角度、量级等各种变化,使得对图像的标记耗费大量的人力物力,这直接导致分类任务的困难。

主动学习^[2]是一种有效减少图像标记的学习方法,作为构造有效训练集的方法,其目标是通过迭代抽样,寻找有利于提升分类效果的样本,进而减少分类训练集的大小,在有限的时间和资源的前提下,提高分类算法的效率。主动学习已成为模式识别、机器学习和数据挖掘领域的研究热点问题^[3]。陈荣等^[4]提出了基于主动学习和半监督学习的多类图像分类方法,将基于最优标号和次优标号(Best vs Second-best, BvSB)的主动学习和带约束条件的自学习(Constrained Self-Training, CST)引入到基于支持向量机(Support Vector Machine, SVM)分类器^[5]的图像分类算法中,提出了一种新的图像分类方法。通过BvSB主动学习去挖掘那些对当前分类器模型最有价值的样本进行人工标注,并借助CST半监督学习进一步利用样本集中大量的未标注样本,使得在花费较小

标注代价情况下,能够获得良好的分类性能。陈建杰等^[6]提出了一种基于多分类SVM的主动学习方法,在少量标记类别的初始训练样本集基础上,通过反复迭代主动学习的方式,得到最有利SVM分类器性能的样本为支持向量,有效地减少了样本训练时需要标记样本的数目,并取得了较为理想的分类效果。

然而,当前图像分类中的主动学习主要关注单一领域的分类模型。例如根据人脸图像进行表情分类时只关注人脸领域图像的表情特征,而没有关注动物脸图像的一些表情特征。尽管不同领域的图像有它们独特的特征空间,但它们可能会有共同的隐特征,比如都有高兴、悲伤的表情。如果分别对不同领域的图像进行主动学习,会有很多重复知识,从而造成大量的重复标记。反过来,我们可以考虑用这些共同特征对另一未知领域的图像进行标记,从而降低图像标记的工作量,提高分类效率。

本文提出一种基于跨领域^[7]主动学习的图像分类方法,它的主要思想是:为多领域图像分类任务建立基于全局优化的主动学习模型,所提出的样本选择准则用来抽取未标记样本,使得在标记完后可以最大化地减少所有分类器的模型损失^[8]。在该方法中,首先学习一个表示多个领域的共同隐特征的子空间,再把全局特征空间分割为共同特征部分和领域

相关特征部分。这样,每个数据样本带来的模型损失减少量可以被分解为两部分:领域相关损失减少量和全领域分类器共同损失的减少量。通过联合的选取数据样本,所有分类器的公用模型损失被减少,从而降低很多重复标记。

1 用于跨领域分类的主动学习

本章介绍与本文所讨论问题相关的若干定义:领域、跨领域分类、用于跨领域分类的主动学习。

定义1 领域(Domain)。一个领域是由满足相同数据分布 $P(x)$ 的数据实例组成的集合,这里 $x \in \mathbf{X}$, \mathbf{X} 是一个特征空间。

例如 一系列人脸的图像可以组成一个领域,因为它们有共同的特征空间(如皮肤颜色、眼睛部位、嘴唇形状等)。同样的,一些猫脸的图像也可以组成一个领域。同一领域内的图像特征是独立同分布,不同领域的图像特征不同。例如不能直接用猫脸的表情特征去形容人脸的表情。因此,跨领域分类问题可以定义如下:

定义2 跨领域分类。给定 K 个不同领域中收集的数据样本集合,若 \mathbf{X} 为一个特征空间, Y 为标记的集合。跨领域分类的任务是为 K 个领域训练 K 个分类器: $f^l: \mathbf{X} \rightarrow Y (l = 1, 2, \dots, K)$ 。

基于以上定义,将本文要解决的问题定义如下:

定义3 用于跨领域分类的主动学习。设 $P = \{P^1, P^2, \dots, P^K\}$ 是从 K 个不同领域数据组成的未标记的数据集合,这里 $P^l = \{x_1^l, x_2^l, \dots, x_{N^l}^l\}$ 包含从第 l 个领域中选取的 N^l 个数据样本。用于跨领域分类的主动学习任务为:通过尽可能少的标记选取数据样本建立 K 个分类器 $f^l: \mathbf{X} \rightarrow Y (l = 1, 2, \dots, K)$ 。

本文提出的主动学习框架则是以样本池(Pool)为基础^[9]。在基于样本池的采样中,对未标记的数据池进行迭代地主动学习,每次迭代中扫描最有信息量的数据进行标记。

与传统的单领域主动学习问题不同,跨领域主动学习不仅要降低单个领域的分类器模型损失,还要考虑所有领域全局的模型损失最小。

设 $L(f_D^l)$ 为分类器的模型损失,则全局模型的损失为

$$L_D = \sum_{l=1}^K L(f_D^l) \quad (1)$$

其中 f_D^l 为第 l 个领域的预测函数。

此时,主动学习的目标就是选择一个未标记样例 x^* ,当用它标记后可以最大限度降低全局模型损失,该样本选择准则可以形式化地表示为:

$$x^* = \arg \max_{x^*} (L_D - L_{D+(x^*, y^*)}) = \arg \max_{x^*} \sum_{l=1}^K (L(f_D^l) - L(f_{D+(x^*, y^*)}^l)) \quad (2)$$

其中 $D+(x^*, y^*)$ 代表将样本 (x^*, y^*) 加入域 D 后的扩展训练集。

实际使用中, y^* 在查询之前难以获取准确值,因此采用估计期望的方式对加入样本前后分类器性能损失的期望值进行估算:

$$x^* = \arg \max_{x^*} \sum_{y \in Y} \hat{P}(y|x^*) \sum_{l=1}^K (L(f_D^l) - L(f_{D+(x^*, y^*)}^l)) \quad (3)$$

其中 $\hat{P}(y|x^*)$ 代表给定实例 x^* 后标记为 y 的条件概率。

本文采用 SVM 作为图像分类器,与传统的单领域图像分

类问题不同的是,为了表示多个领域的共同特征,引入一个表示公共隐特征的子空间,对于第 l 个领域的预测函数 f_D^l 可以定义如下:

$$f_D^l(x^l) = \mathbf{w}^l \cdot \Phi(x^l) + \mathbf{v} \cdot \Phi(\theta x^l) \quad (4)$$

式(4) 将一个域的预测函数划分为两个部分:一部分是在公共特征子空间上的预测,另一部分是在原始特征空间上的预测。这里的 Φ 是特征映射, \mathbf{w}^l 和 \mathbf{v} 是权值向量, θ 是学习得到的从原始特征空间到公共特征子空间的转换矩阵。共享参数 \mathbf{v} 和 θ 用于从不同域提取公共隐特征。

参数的学习由两步组成:第一步用谱特征校准(Spectral Feature Alignment, SFA)^[7] 估计参数 θ 。第二步用 SVM 优化来估计参数 \mathbf{w}^l 和 \mathbf{v} , SVM 优化方法形式化的表示如下:

$$\min_{\{\mathbf{w}^l, \mathbf{v}, \theta\}} \frac{1}{2} \|\mathbf{v}\|^2 + \frac{1}{2} \sum_{l=1}^K \|\mathbf{w}^l\|^2 \quad (5)$$

s. t. $y_i^l(\mathbf{w}^l \cdot \Phi(x_i^l) + \mathbf{v} \cdot \Phi(\theta x_i^l)) \geq 1; l = 1, 2, \dots, K$

式(5)可以看成标准的 SVM 优化问题^[11],因此可以用 SVM 优化方法进行优化。

2 基于跨领域主动学习的图像分类

跨领域主动学习问题即解决式(3)的标记样本选取问题。为了估计式(3),需要对式中的模型损失函数 $L(f_D^l)$ 进行合理有效的估计。本文为了估计 $L(f_D^l)$,用变形空间(Version Space)的大小来估计模型的损失值。对于第 l 个域,变形空间 V^l 定义为

$$V^l = \{\mathbf{u}^l / \|\mathbf{u}^l\| \mid \mathbf{u}^l \in \mathbf{W}^l; \forall i, y_i^l(\mathbf{w}^l \Phi(x_i^l) + \mathbf{v} \Phi(\theta x_i^l)) > 0\} \quad (6)$$

其中: $\mathbf{u}^l = [\mathbf{w}^l, \mathbf{v}]$, \mathbf{W}^l 是参数空间。在 SVM 中,可以用边缘距离(Margin)^[10] 来估计变形空间的大小,即越靠近分离平面的样本认为最有信息量。用 $\|\mathbf{V}_D^l\|$ 表示当前变形空间的大小,用当前变形空间大小来估计新的变形空间大小,可以近似表示如下:

$$\|\mathbf{V}_{D+(x^*, y^*)}^l\| \approx \frac{1 + y f_D^l(x^*)}{2} \|\mathbf{V}_D^l\| \quad (7)$$

从而式(3)中的模型损失减少量可以表示如下:

$$L(f_D^l) - L(f_{D+(x^*, y^*)}^l) = \|\mathbf{V}_D^l\| - \|\mathbf{V}_{D+(x^*, y^*)}^l\| \approx \frac{1 - y f_D^l(x^*)}{2} \|\mathbf{V}_D^l\| \quad (8)$$

代入式(3),有

$$x^* = \arg \max_{x^*} \sum_{y \in Y} \hat{P}(y|x^*) \sum_{l=1}^K \frac{1 - y f_D^l(x^*)}{\|\mathbf{u}^l\|} \quad (9)$$

这样整个跨领域主动学习的图像分类方法就可以用式(9)所示的样本选择准则来概括。整个跨领域主动学习算法流程见算法1。

算法1 跨领域主动学习算法。

输入 从 K 个领域取的样例集合 P , 每个域的初始训练样本数 M , 未标记样本集 U , 迭代次数 T , 每次迭代查询的样例数 S ;

输出 K 个域的分类器。

- 1) 随机对每个域的 M 个数据样例进行标记,形成初始训练集;
- 2) 采用 SFA 训练低维共享子空间;
- 3) for $t \leftarrow 1$ to T do
 - 在训练集 D 上用(5)式训练 K 个分类器;
 - for each $x \in P$ do
 - 用式(9)评估全局模型损失减少量;

end
 从 U 中找出 S 个使有最大的全局模型损失减少量的样本进行标记,记作 Y ;
 通过 $D \leftarrow D \cup (U^*, Y^*)$ 更新训练集,并从 P 中移除 U ;
 end

通过算法 1,可以对多个领域图像数据集中进行自动标记,标记后训练分类器,如此迭代,直到模型稳定,则可以对图像进行分类。该算法避免了对各个相关领域的图像分别标记带来的额外开销,从而能提高图像分类的效率。

3 实验分析

3.1 实验设置

为了验证本文算法的分类性能,在 UCI 数据库^[12]中选取了两个具有相似信息的数据集进行实验。一个是美国邮政手写体数字图像集 (USPS),另外一个为手写数字图像集 (Pendig-its)。这是两个不同风格的图像,因此各自的字体等特征有差别,但表达的主题类似,因此可以看成具有相似特征的不同领域图像。为了测试本文算法,一方面分别对两个数据集进行单一模型学习,统计总体的平均标记量和平均分类精度,称这种方式为 SepAL;另一方面,对这两个数据集进行混合,并整体训练分类模型,统计标记数量和分类准确率,称这种分类为 MixAL。最后用本文提出的算法对两个数据集进行主动学习和分类,记为 MultiAL。比较在这三种学习方式下算法需要的人工标记量和分类精度。采用线性 SVM 进行分类,为了获得较可靠的结果,将实验重复 10 次,每次对训练集和测试集作不同的随机划分,最终得到一个分类精度的平均结果,如图 1 所示,其中 SepAL 方法的迭代次数是两个模型的迭代次数的总和,精度是两个模型的平均。

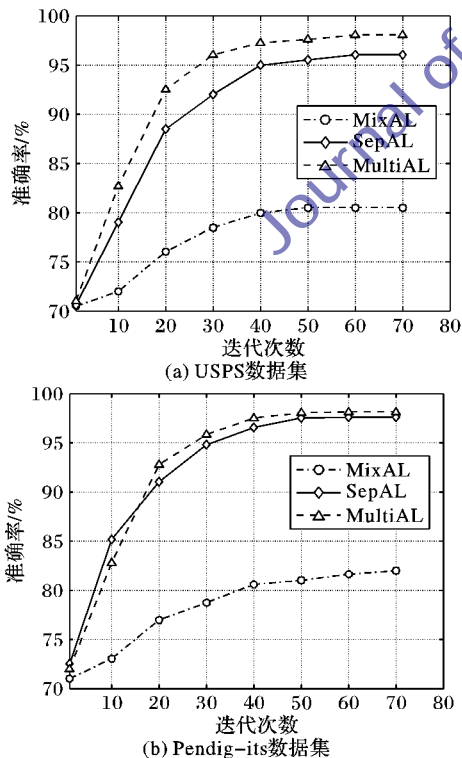


图1 算法的分类准确率对比

此外,考虑分类精度达到一定量的情况下需要的最少标记量,结果如表 1 所示,其中 SepAL 是在平均精度达到一定值

时两者标注负担的总和。表 1 中“—”表示无论怎样标记都无法达到相应精度要求。

表 1 达到相同的最大分类精度需要的标注数

算法	USPS 数据集 最大分类精度/%					Pendig-its 数据集 最大分类精度/%				
	75	80	85	90	95	75	80	85	90	95
SepAL	8	12	18	28	41	7	11	16	27	40
MixAL	15	11	—	—	—	14	9	—	—	—
MultiAL	7	9	14	20	27	6	8	13	20	25

3.2 实验结果与分析

从图 1 所示的实验结果可以看到,简单将样本混合学习模型的 MixAL 方法始终不能得到全局的较优模型,在各个阶段分类准确度都不及对比方法。本文的算法的分类准确率在实验初期不及普通的 SepAL 算法,这是因为在实验初期样本较少,全局模型没有完成学习导致精确度不及原始方法;而在 Pendig-its 数据集上当迭代次数达到 18 以后本文的 MultiAL 方法在精确度上要超过 SepAL 方法,在 USPS 数据集上迭代 8 次精确度就已超过 SepAL 方法。因为不同类别的模型的公共特征被学习出来,使得分类器有更精确的分类能力。该结果表明本文算法的准确率超过其他针对单领域图像分类的算法。

从表 1 中对达到各个精度所需标记量的对比结果可见,本文提出的方法在各个精度要求下都极大地降低了额外的标记工作,在较高精度要求下对标记量的要求相比单领域算法降低了将近 30%,从而使得算法性能得到很大提升。

4 结语

传统的主动学习图像分类方法有效地降低了大量的样本标记工作,但在多个领域的图像分类中,这类方法忽略了多个领域图像的信息关联性,造成很多不必要的冗余标记。本文提出一种基于跨领域主动学习的图像分类方法,利用多个领域的图像直接的关联性,学习既有共同特征的信息又有领域相关信息的训练样本,从而降低样本标记的工作量,提高分类效率。实验结果表明本文方法将图像标记效率提高了近 30%,从而有效说明了算法的有效性和优越性。利用更多的领域数据例如文本标记数据辅助图像分类将成为下一步研究的方向。

参考文献:

- [1] WANG J, YANG J, YU K, *et al.* Learning locality-constrained linear coding for image classification [C]// CVPR2010: Proceedings of the 2010 IEEE Conference on Computer Vision and Pattern Recognition. Washington, DC: IEEE Computer Society, 2010: 3360 – 3367.
- [2] LIU K, QIAN X, WANG Z. Survey on active learning algorithms [J]. Computer Engineering and Applications, 2012, 48(34): 1 – 6. (刘康, 钱旭, 王自强. 主动学习算法研究综述 [J]. 计算机工程与应用, 2012, 48 (34): 1 – 6.)
- [3] HOI S C H, JIN R, LYU M R. Batch mode active learning with applications to text categorization and image retrieval [J]. IEEE Transactions on Knowledge and Data Engineering, 2009, 21(9): 1233 – 1248.

从而保证细节在不同亮度下的真实再现。经过测试证明,本文映射算法在细节再现方面有一定的成效,且对不同场景的HDR图像均具有良好的视觉效果。但是由于BLF分层算法自身的问题,不能得到理想的分层效果,容易产生光晕现象,下一步工作将设计新的边缘保护滤波器以杜绝光晕现象。

参考文献:

- [1] TUMBLIN J, RUSHMEIER H. Tone reproduction for realistic images [J]. *IEEE Computer Graphics & Applications*, 1993, 13(6): 42–48.
- [2] LARSON G W, RUSHMEIER H, PIATKO C. A visibility matching tone reproduction operator for high dynamic range scenes [J]. *IEEE Transactions on Visualization and Computer Graphics*, 1997, 3(4): 291–306.
- [3] REINHARD E, DEVLIN K. Dynamic range reduction inspired by photoreceptor physiology [J]. *IEEE Transactions on Visualization and Computer Graphics*, 2005, 11(1): 13–24.
- [4] CHIU K, HERF M, SHIRLEY P, *et al.* Spatially nonuniform scaling functions for high contrast images [C]// *Proceedings of Graphics Interface 93*. Toronto: Canadian Human-Computer Communications Society, 1993: 245–253.
- [5] MEYLAN L, SUSSTRUNK S. High dynamic range image rendering using a retinex-based adaptive filter [J]. *IEEE Transactions on Image Processing*, 2006, 15(9): 2820–2830.
- [6] REINHARD E, STARK M, SHIRLEY P, *et al.* Photographic tone reproduction for digital images [J]. *ACM Transactions on Graphics*, 2002, 21(3): 267–276.
- [7] TUMBLIN J E, HODGINS J K, GUENTER B K. Two methods for display of high contrast image [J]. *ACM Transactions on Graphics*, 1999, 18(1): 56–94.
- [8] TUMBLIN J, TURK G. LCIS: a boundary hierarchy for detail-preserving contrast reduction [C]// *Proceedings of the 1999 International Conference on Computer Graphics and Interactive Techniques*. New York: ACM, 1999: 83–90.
- [9] DURAND F, DORSEY J. Fast bilateral filtering for the display of high-dynamic-range images [J]. *ACM Transactions on Graphics*, 2002, 21(3): 257–266.
- [10] CHOUDHURY A, MEDIONI G. Hierarchy of nonlocal means for preferred automatic sharpness enhancement and tone mapping [J]. *Journal of the Optical Society of America A: Optics Image Science and Vision*, 2013, 30(3): 353–366.
- [11] STVENS J C, STEVENS S S. Brightness function: effects of adaptation [J]. *Journal of the Optical Society of America*, 1963, 53(3): 375–384.
- [12] KUANG J T, JOHNSON G M, FAIRCHILD M D. iCAM06: a refined image appearance model for HDR image rendering [J]. *Journal of Visual Communication and Image Representation*, 2007, 18(5): 406–414.
- [13] JOHNSON G M, FAIRCHILD M D. Rendering HDR images [C]// *CIC 2003: Proceedings of the Eleventh Color Imaging Conference: Color Science and Engineering Systems, Technologies, Applications*. Scottsdale, Arizona: The Society for Imaging Science and Technology, 2003: 36–41.
- [14] AYDIN T O, MANTIUK R, MYSZKOWSKI K, *et al.* Dynamic range independent image quality assessment [J]. *ACM Transactions on Graphics*, 2008, 27(3): 1–10.
- [15] HEBER S, STOYE J. Finding all common intervals of k permutations [C]// *CPM '01: Proceedings of the 12th Annual Symposium on Combinatorial Pattern Matching*, LNCS 2089. Berlin: Springer-Verlag, 2001: 207–218.
- [16] XIAO F, DICARLO J, WANDELL B. HDR data [EB/OL]. [2013–06–15]. <http://white.stanford.edu/~brian/hdr/hdr.html>.
- [17] CADIK M. Index of /home/cadikm/tmo/hdr_images [EB/OL]. (2006–12–03) [2013–06–03]. http://dcgi.felk.cvut.cz/home/cadikm/tmo/hdr_images/.
- [18] DEBEVEC P. Recovering high dynamic range radiance maps from photographs [EB/OL]. [2013–06–03]. <http://www.pauldebevec.com/Research/HDR/>.
- [8] TONG S, KOLLER D. Support vector machine active learning with applications to text classification [J]. *Journal of Machine Learning*, 2002, 14(2): 45–66.
- [9] DAGAN I, ENGELSON S P. Committee-based sampling for training probabilistic classifiers [C]// *ICML1995: Proceedings of the 12th International Conference on Machine Learning*. San Francisco: Morgan Kaufmann Publishers Inc., 1995: 150–157.
- [10] LEWIS D D, GALE W A. A sequential algorithm for training text classifiers [C]// *SIGIR'94: Proceedings of the 17th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, LNCS 1513. Berlin: Springer-Verlag, 1994: 3–12.
- [11] TUIA D, RATLE F, PACIFICI F, *et al.* Active learning of very-high resolution optical imagery with SVM: entropy vs margin sampling [C]// *IGARSS 2008: Proceedings of the 2008 IEEE International Conference on Geoscience and Remote Sensing Symposium*. Washington, DC: IEEE Computer Society, 2008, 4: 73–76.
- [12] ASUNCION A, NEWMAN D J. UCI machine learning repository [EB/OL]. (2010–01–10) [2013–01–15]. <http://archive.ics.uci.edu/ml/datasets.html>.

(上接第1171页)

- [4] CHEN R, CAO Y. Multi-class image classification with active learning and semi-supervised learning [J]. *Acta Automatica Sinica*, 2011, 37(8): 954–962. (陈荣, 曹永锋, 孙洪. 基于主动学习和半监督学习的多类图像分类[J]. *自动化学报*, 2011, 37(8): 954–962.)
- [5] MAJI S, BERG A, MALIK J. Classification using intersection kernel support vector machines is efficient [C]// *CVPR2008: Proceedings of the 2008 IEEE Conference on Computer Vision and Pattern Recognition*. Washington, DC: IEEE Computer Society, 2008: 1–8.
- [6] CHEN J, YE Z. Remote sensing image classification using active learning with multiclass support vector machine [J]. *Science of Surveying and Mapping*, 2009, 34(4): 34–39. (陈建杰, 叶智宣. 多分类SVM主动学习及其在遥感图像分类中的应用[J]. *测绘科学*, 2009, 34(4): 34–39.)
- [7] PAN S, NI X, SUN J, *et al.* Cross-domain sentiment classification via spectral feature alignment [C]// *WWW2010: Proceedings of the 19th International Conference on World Wide Web*. New York: ACM, 2010: 751–760.