改进 Corr-LDA 的图像标注概率主题模型

曹 洁[1,2],罗菊香[1],李晓旭[1]

[1](兰州理工大学 计算机与通信学院,兰州 730050)
[2](甘肃省制造业信息化工程研究中心,兰州 730050)
E-mail: 894080040@qq.com

摘 要: 对标有文本词的图像数据进行调研,发现真实图像中被标注的事物往往占整个画面的部分较大,而现有的图像标注工作没有考虑这一信息,本文基于此,在 Corr-LDA 的基础上提出一种文本主题选择是依照图像主题分布的图像标注概率主题模型. 同时,为该模型推导了一个基于变分 EM 的参数估计算法,并给出使用该模型标注图像的方法. 在 LabelMe 和 UIUC-Sport 两个真实数据集上验证了提出模型的标注性能要高于其它相比较模型.

关 键 词: 图像标注; 概率主题模型; 变分 EM; Corr-LDA 模型
中图分类号: TP391 **文献标识码:** A **文 章 编 号:** 1000-1220(2017)03-0615-05

Image Annotation Probabilistic Topic Model Improving Corr-LDA Model

CAO Jie[1,2], LUO Ju-xiang[1], LI Xiao-xu[1]
[1](College of Computer and Communication, Lanzhou University of Technology, Lanzhou 730050, China)
[2](Gansu Manufacturing Informatization Engineering Research Center, Lanzhou 730050, China)

Abstract: In this paper, we reach the image data with the text keywords, and find that the things which are marked frequently occupy a large part of the whole picture, while the existing image annotation work did not consider this information, in view of this, the paper proposes a probabilistic topic model of image annotation based on Corr-LDA model, the text topics of this model are from image topic distribution, then derives a parameters estimation algorithm based on the variational EM algorithm, as well as gives the method of annotating new images according to the model. The experimental results on LableMe and UIUC-Sport datasets show that the image annotation performance of the proposed model is better than other compared models.

Key words: image annotation; probabilistic topic model; variational expectation maximization; Corr-LDA model

1 引 言

海量图像数据在互联网上的涌现给图像检索带来了巨大挑战. 在众多检索方式中,基于文本的检索是人们偏爱的检索方式. 标注文本正确与否直接影响图像的检索性能,因此,图像标注[1,2]的研究也越来越受到学术界和工业界的关注.

基于概率主题模型的图像标注研究是图像标注研究领域的一个重要分支,近年来受到研究人员和学者的广泛关注,并已做了大量的工作. 这些工作大致可分为两类. 一类是把标注看作是分类问题[3-6],即把标注词作为类标,建模给定标注词的条件下图像特征的分布: Putthividhya[3] 将图像标注看作是多分类问题,提出了 sLDA-bin 模型; Nguyen[4] 提出了基于多示例多标签多模态标注方法,利用多模型信息促进标注; Li[6] 等人使用类别信息来促进图像标注. 第二类是将图像和标注文本放置在同一水平上[7-13,15-17],来学习图像和文本这两个模态之间的关联: Blei 等人基于 LDA[11] 提出了图像标注的经典模型 Corr-LDA 模型[13]; Xu 等人引入主题之间的相关性提出图像标注的 Corr-CTM 模型[15]; Wang 等人提出同时做图像分类和标注的 Mca-sLDA 模型[16]; Du 等人提出实现同时聚类、标注、分割图像的无参数模型[17]. 这些工作都取得了相对较好的标注效果,但是都没有考虑图像中被标注的事物往往占整个画面的部分较大这一信息,基于此,本文在 Corr-LDA 的基础上提出了一种文本主题选择是依照图像主题分布的图像标注概率主题模型,并为该模型推导了一个基于变分 EM 的参数估计算法,给出了使用该模型标注图像的方法. 最后,在 LabelMe 和 UIUC-Sport 两个真实数据集上验证了提出模型的标注性能要高于其它相比较模型.

本文在第二部分介绍了本文的改进模型,第三部分介绍了改进模型的参数求解过程和标注预测方法,第四部分介绍了在 LabelMe 和 UIUC-Sport 两个真实数据集上的实验结果,第五部分总结和分析本文的工作.

2 基于 Corr-LDA 的图像标注

2.1 Corr-LDA 模型

Corr-LDA 模型是基于 LDA 做图像标注的经典模型,其概率图模型表示如下页图 1 所示. 该模型文本主题的选择是在已抽取的图像主题中按均匀分布抽取的. 模型用到的符号

收稿日期: 2016-01-14 **收修改稿日期:** 2016-03-01 **基金项目:** 国家自然科学基金项目(61263031)资助; 甘肃省自然科学基金项目(1310RJZA034)资助. **作者简介:** 曹 洁,女,1966 年生,硕士,教授,博士生导师,研究方向为信息融合、智能交通; 罗菊香,女,1987 年生,硕士研究生,研究方向为智能信息处理; 李晓旭,女,1982 年生,博士,讲师,研究方向为计算机视觉与机器学习.

(C)1994-2022 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

及其所表示意义如表 1 所示.

具体生成过程如下:

1) 抽取一个主题分派 $\theta \sim P_{dir}(\alpha)$;

2) 对每一个图像词汇 $v_m, m \in \{1,2,3,...,M\}$;

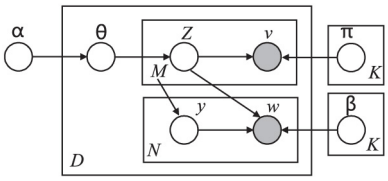

图 1 Corr-LDA 模型的概率图模型

Fig. 1 Probabilistic graphical model of the Corr-LDA model

a) 抽取主题分派 $Z_m | \theta \sim Mult(\theta)$;

b) 抽取图像词 $v_m | Z_m \sim Mult(\pi_{z_m})$;

3) 对每一个文本词 $w_n, n \in \{1,2,3,...,N\}$;

a) 抽取离散索引变量 $y_n \sim Unif(1,2,...,M)$;

b) 抽取文本词 $w_n \sim Mult(\beta_{z_{y_n}})$;

2.2 改进 Corr-LDA 模型

考虑到标有文本词的图像数据,被标注的事物往往占整个画面的部分较大,即一个事物占图像的比例越大,越有可能被标注. 本文在 Corr-LDA 的基础上提出了一种文本主题选择是依照图像主题分布的图像标注概率主题模型,称模型为 Corr-LDA-Image-Topic-Distribution 模型,简称为 Corr-LDA-ITD 模型. 该模型中文本主题的选择是在已抽取的图像主题中按图像主题分布抽取. 模型以词袋(bag of words)特征[18]为基础,用到的的符号及其所表示意义如表 1 所示. 具体生成过程如下:

表 1 符号及其意义

Table 1 Symbol and its meaning

符号	符号意义	符号	符号意义
D	训练集大小	α	狄利克雷参数
K	主题数	β	文本主题
M	图像词个数	π	图像主题
N	文本词个数	ϕ	Z 的变分参数
V_s	图像码书大小	ρ	T 的变分参数
V_t	文本码书大小	λ	y 的变分参数
w	图像词	γ	狄利克雷参数
v	文本词	T	标注潜主题
y	索引变量	P_{dir}	狄利克雷分布
Z	潜主题	$Unif$	均匀分布
θ	主题比例	$Mult$	多项式分布

1) 抽取一个主题分派 $\theta \sim P_{dir}(\alpha)$;

2) 对每一个图像词汇 $v_m, m \in \{1,2,3,...,M\}$;

a) 抽取主题分派 $Z_m | \theta \sim Mult(\theta)$;

b) 抽取图像词 $v_m | Z_m \sim Mult(\pi_{z_m})$;

3) 对每一个文本词 $w_n, n \in \{1,2,3,...,N\}$;

a) 抽取主题分派 $T_n | Z \sim Mult(Z)$;

这里

$$P(T | Z) = \prod_{n}^{N} \prod_{i=1}^{K} \left(\sum_{m=1}^{M} z_{mi} \right)^{T_n^i};$$

b) 抽取文本词 $w_n | T_n \sim Mult(\beta_{T_n})$;

提出模型的概率图模型表示如图 2 所示,其潜变量和观测变量的联合分布为:

$$P(v, w, \theta, Z, T | \alpha, \pi, \beta) = P(\theta | \alpha) \left(\prod_{m=1}^{M} A \right) \left(\prod_{n=1}^{N} B \right) \quad (1)$$

其中,

$$A = P(z_m | \theta) P(v_m | z_m, \pi)$$

$$B = P(T_n | z) P(w_n | T_n, \beta)$$

模型的生成可总结为两个阶段,第一阶段: 首先,按照 Dirchlet 分布抽取潜主题的一个比例; 其次,按照这个比例抽取一个主题标号; 再次,按照相应主题生成一个图像的区块; 重复多次生成整个图像. 该过程与 Corr-LDA 是相同的. 提出模型与 Corr-LDA 不同主要体现在第二阶段.

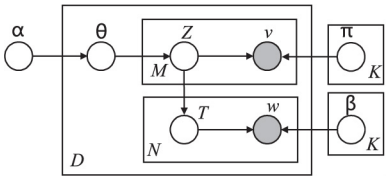

图 2 提出模型的概率图模型

Fig. 2 Probabilistic graphical model of the proposed model

第二阶段: 图像生成完之后,在已抽取的图像主题中按图像主题分布抽取模型的文本主题; 再依照该主题生成相应的标注词; 重复多次就生成了图像的标注词. Corr-LDA 模型文本主题的选择是在已抽取的图像主题中按均匀分布抽取,没有用到图像主题分布的信息,而图像主题分布的信息能够为图像标注提供有价值的信息. 提出模型中文本主题的选择是在已抽取的图像主题中按图像主题分布抽取,用到了图像主题分布的信息.

3 参数求解与图像标注

本文采用变分 EM 算法求解参数. EM 算法[3]是 Dempster,,Laind,,Rubin 于 1977 年提出的参数求解方法. 在变分 EM 算法中,包含 E,M 两个步骤. 为了确定模型的参数,引入潜变量 $H = \{Z, \theta, T\}$的变分分布:

$$q(H | \omega) = q(\theta | \gamma) \left(\prod_{m=1}^{M} q(z_m | \phi_m) \right) \left(\prod_{n=1}^{N} q(T_n | \rho_n) \right) \quad (2)$$

3.1 变分 E 步骤

$E = \{v, w\}$ ‖表示可观测变量,$\Delta = \{\alpha, \pi, \beta\}$ 表示模型参数,$\omega = \{\gamma, \phi, \rho\}$ 表示变分参数,则优化目标为数据集 D 的 log 似然下限 $L(\omega; \Delta)$.

在变分 E 步骤中,对 $E = \{v, w\}$,相对于 $\omega = \{\gamma, \phi, \rho\}$,最大化 $L(\omega; \Delta)$.

计算后验 Dirichlet 参数 γ: 从 $L(\omega; \Delta)$ 中选择包含 γ 的项,并对其求一阶导:

$$\frac{\partial L(\gamma_i)}{\gamma_i} =$$

$$\Psi'(\gamma_i)\left(\alpha_i - \gamma_i + \sum_{m=1}^{M} \phi_{mi}\right) -$$

$$\Psi'\left(\sum_{j=1}^{K} \gamma_j\right) \sum_{j=1}^{K} \left(\alpha_i - \gamma_i + \sum_{m=1}^{M} \phi_{mi}\right)$$

(C)1994-2022 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

令其为零可得:

$$\gamma_i = \alpha_i + \sum_{m=1}^{M} \phi_{mi} \tag{3}$$

计算参数 ϕ: 从 $L(\omega;\Delta)$ 中选择包含 ϕ 的项:

$$L(\phi) = \sum_{n=1}^{N}\sum_{i=1}^{K}\rho_{ni}E_q\left(\log\left(\sum_{m=1}^{M} z_{mi}\right)\right) + \sum_{i=1}^{K}\left(\Psi(\gamma_i) - \Psi\left(\sum_{j=1}^{K}\gamma_j\right)\phi_{mi}\right) + \sum_{m=1}^{M}\sum_{i=1}^{K}\sum_{j=1}^{V_s}\phi_{mi}v_m^j\log\pi_{ij} - \sum_{m=1}^{M}\sum_{i=1}^{K}\phi_{mi}\log\phi_{mi}$$

由于 $E_q\left(\log\left(\sum_{m=1}^{M} z_{mi}\right)\right)$直接按定义计算比较复杂,这里首先对其进行变形处理:

$$E_q\left(\log\left(\sum_{m=1}^{M} z_{mi}\right)\right) = E_q\left(\log\left(\frac{1}{M}\sum_{m=1}^{M} z_{mi}\right)\right) + \log M \tag{4}$$

然后,基于 Jensen 不等式对 $\log\left(\frac{1}{M}\sum_{m=1}^{M} z_{mi}\right)$ 进行缩放:

$$E_q\left(\log\left(\frac{1}{M}\sum_{m=1}^{M} z_{mi}\right)\right) \geqslant E_q\left(\frac{1}{M}\sum_{m=1}^{M}(\log(z_{mi}))\right) \tag{5}$$

为了使 $\log z_{mi}$ 有意义,引入一个足够小的常数 a,将 $\log z_{mi}$ 项用 $\log(z_{mi}+a)$ 替代.

同时,根据期望的定义有:

$$E_q[\log(z_{mi}+a)] = \phi_{mi}\log(1+a) + (1-\phi_{mi})\log a \tag{6}$$

基于公式(4)、(5)、(6),可得:

$$E_q\left(\log\left(\sum_{m=1}^{M} z_{mi}\right)\right) \geqslant \left(\frac{1}{M}\sum_{m=1}^{M} H\right) + \log M \tag{7}$$

其中,$H = \phi_{mi}\log(1+a) + (1-\phi_{mi})\log a$

最后将公式(7)代入 $L(\phi)$ 中,在 $\sum_{i=1}^{K}\phi_{mi} = 1$ 的约束下,相对于 ϕ_{mi},最大化 $L(\phi)$ 可得:

$$\phi_{mi} \propto \exp\left(\Psi(\gamma_i) - \Psi(\gamma_j) + \sum_{n=1}^{N}\frac{1}{M}\rho_{ni}\log\left(1+\frac{1}{a}\right)\right) \tag{8}$$

计算参数 ρ: 从 $L(\omega;\Delta)$ 中选择包含 ρ 的项,$L(\rho)$ 中同样含有 $E_q\left(\log\left(\sum_{m=1}^{M} z_{mi}\right)\right)$ 项,采用上述相同的处理方法,在 $\sum_{i=1}^{K}\rho_{ni} = 1$ 的约束条件下,相对于 ρ_{ni},最大化 $L(\rho)$ 可得:

$$\rho_{ni} \propto \beta_{iw_n}\exp\left(\frac{1}{M}\sum_{m=1}^{M} H + \log M\right) \tag{9}$$

3.2　变分 M 步骤

经过 E 步骤,可求得变分参数 $\omega = \{\gamma,\phi,\rho\}$ 的值. 在 M 步骤中,固定变分参数 $\omega_d = \{\gamma,\phi,\rho\}$,$d \in \{1,2,3,\cdots,D\}$,相对于模型参数 $\Delta = \{\alpha,\pi,\beta\}$ 来最大化集合 D 上的 log 似然. 即最大化:

$$L(D) = \sum_{d=1}^{D} L(\omega_d;\Delta) \tag{10}$$

求解模型参数 π: 从 $L(D)$ 中选择包含 π 的项:

$$L(\pi) = \sum_{d=1}^{D}\sum_{m=1}^{M_d}\sum_{i=1}^{K}\sum_{j=1}^{V_s}\phi_{dmi}v_{dm}^j\log\pi_{ij}$$

在 $\sum_{j=1}^{V_s}\pi_{ij} = 1$ 的约束条件下,相对于 π_{ij},最大化 $L(\pi)$ 可得:

$$\pi_{ij} \propto \sum_{d=1}^{D}\sum_{m=1}^{M_d}\phi_{dmi}v_{dm}^j \tag{11}$$

求解模型参数 β: 从 $L(D)$ 中选择包含 β 的项:

$$L(\beta) = \sum_{d=1}^{D}\sum_{n=1}^{N_d}\sum_{i=1}^{K}\sum_{j=1}^{V_t}\rho_{dni}w_{dn}^j\log\beta_{ij}$$

在 $\sum_{j=1}^{V_t}\beta_{ij} = 1$ 的约束条件下,相对于 β_{ij},最大化 $L(\beta)$ 可得:

$$\beta_{ij} \propto \sum_{d=1}^{D}\sum_{n=1}^{N_d}\rho_{dni}w_{dn}^j \tag{12}$$

参数 π_{ij},β_{ij} 分别表示第 i 个主题下第 j 个图像词汇、文本词汇的概率.

确定模型参数 α: 本文没有对 α 进行优化,多次实验发现,将 α 设置成全为 1 的向量,模型性能较好.

重复执行 E,M 步骤,直到 $L(D)$ 收敛.

3.3　图像标注

模型参数确定之后,可以使用该模型对新图像预测标注词. 首先,使用训练好的模型参数计算图像主题分布; 然后,再依据该分布选择文本主题; 最后,按照这个选中的主题生成标注词. 具体标注方法为:

$$P(w \mid v) \approx \sum_{n=1}^{N}\sum_{z_n} P(w \mid z_n,\beta)q(z_n \mid \phi) \tag{13}$$

4　实　验

4.1　数据和预处理

为了评估模型的性能,本文选择一个场景图像集-LabelMe 图像集和一个运动图像集-UIUC-Sport 图像集. 数据集 LabelMe 包含 8 类 "海岸", "森林", "高速公路", "城市", "高山", "乡村", "街道", "高楼". 每类包含 200 幅图像,共 1600 幅图像. 数据集 UIUC-Sport 也包含 8 类 "羽毛球", "马球", "门球", "地滚球", "攀岩", "帆船", "划船", "滑雪". 其中每类图像数目不等,共 1791 幅图像.

本文提出模型是基于词袋(bag of words)表示,对于 LabelMe 图像集特征提取过程:

1) 对于图像数据集,利用网格抽样技术(网格大小为 5 × 5),从每一个网格中心抽取一个大小为 16 × 16 的区块,然后用 128 维的 sift 描述子进行描述. 采用 k-means 算法对所有的 sift 描述子进行聚类. 所有聚类中心(240 个)构成图像码书,码书长度 240.

2) 对于文本词,移除出现次数少于 3 次的标注词,剩余所有文本词(294 个)构成文本码书,其长度为 294.

对于 UIUC-Sport 图像集,特征提取过程和 LabelMe 图像集大致相同,唯一不同是对每幅图像都一致的抽取大小为 32 × 32 的 2500 个区块.

4.2　标注性能比较

本文采用 *F-measure* 值来评价模型的标注性能,将提出模型与 mm-LDA 模型、Corr-LDA 模型、Multi-class sLDA with annotation(Mca-sLDA)[16] 模型进行比较. mm-LDA 模型是 Blei[13] 提出的图像标注模型,Corr-LDA 模型是基于概率主题模型做图像标注的经典模型,Mca-sLDA 模型是 Wong[16] 提出的同时做图像分类和标注的模型.

(C)1994-2022 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

由于主题模型中主题的数目会影响模型的性能，本文均匀的从 10 到 100 中选择 10 组主题. 对真实数据集 UIUC-Sport 和 LabelMe 中的 8 类图像，从每类图像中分别随机抽取总数的一半作为训练集，剩下的作为测试集. 随机抽取 5 次，得到 5 组训练测试集，计算 *F-measure* 的平均值. 实验根据公式(13) 标注图像，选取概率大的前 5 个词作为标注词.

UIUC-Sport 数据集标注性能

UIUC-Sport 数据集上三个模型的标注性能比较如表 2 所示，表格每行的粗体是此行的最高值.

表 2 UIUC-Sport 数据集上 *F-measure* 比较

Table 2 Comparison of *F-measure* on UIUC-Sport datasets

主题 \ 模型	mm-LDA	Corr-LDA	Mca-sLDA	Corr-LDA-ITD
K = 10	0.354	0.356	0.359	0.374
K = 20	0.359	0.365	0.366	0.386
K = 30	0.354	0.371	0.375	0.398
K = 40	0.353	0.369	0.374	0.400
K = 50	0.360	0.369	0.376	0.405
K = 60	0.368	0.370	0.374	0.388
K = 70	0.370	0.371	0.373	0.389
K = 80	0.372	0.372	0.376	0.393
K = 90	0.371	0.373	0.376	0.388
K = 100	0.370	0.372	0.375	0.383

表 3 LabelMe 数据集上 *F-measure* 比较

Table 3 Comparison of *F-measure* on LabelMe datasets

主题 \ 模型	mm-LDA	Corr-LDA	Mca-sLDA	Corr-LDA-ITD
K = 10	0.416	0.411	0.409	0.418
K = 20	0.418	0.422	0.423	0.428
K = 30	0.417	0.414	0.416	0.435
K = 40	0.416	0.416	0.417	0.432
K = 50	0.417	0.419	0.421	0.429
K = 60	0.419	0.419	0.420	0.423
K = 70	0.420	0.419	0.420	0.422
K = 80	0.419	0.416	0.418	0.424
K = 90	0.420	0.418	0.420	0.422
K = 100	0.420	0.416	0.416	0.421

从表 2 可知，提出模型在 UIUC-Sport 真实图像集上，在不同主题数目时都表现出了提出模型效果要高于 mm-LDA 模型、Corr-LDA 模型、Mca-sLDA 模型，并且提出模型在主题数目为 50 的时候表现了最好的标注性能 40.5%，比 mm-LDA 模型、Corr-LDA 模型、Mca-sLDA 模型超过了约 3.0%. 可见，在 UIUC-Sport 数据集上，提出模型的标注性能要高于其它相比较模型.

LabelMe 数据集标注性能

LabelMe 数据集上三个模型的标注性能比较如表 3 所示，表格每行的粗体是此行的最高值.

从表 3 可知，提出模型在 LabelMe 真实图像集上，在不同的主题数目时都表现出了提出模型效果要高于 mm-LDA 模型、Corr-LDA 模型、Mca-sLDA 模型，并且提出模型在主题数目为 30 的时候表现了最好的标注性能 43.5%，比 mm-LDA 模型、Corr-LDA 模型、Mca-sLDA 模型超过了约 2%. 可见，在 LabelMe 数据集上，提出模型的标注性能要高于其它相比较模型.

4.3 log 似然值比较

表 4、表 5 是 UIUC-Sport 和 LabelMe 两个真实数据集的 log 似然值比较，在主题数目为 20、40、60、80 都表现出提出模型的 log 似然值相较 Corr-LDA 模型要大，说明提出模型的数据拟合能力较强，更适合真实数据.

表 4 UIUC-Sport 数据集上 log 似然值比较

Table 4 Comparison of log-likelihood on UIUC-Sport datasets

模型 \ 主题	K = 20	K = 40	K = 60	K = 80
Corr-LDA	−13742.8	−13690.5	−13648.9	−13605.5
Corr-LDA-ITD	−12754.4	−12587.5	−12645.3	−12627.2

表 5 LabelMe 数据集上 log 似然值性能比较

Table 5 Comparison of log -likelihood on LabelMe datasets

模型 \ 主题	K = 20	K = 40	K = 60	K = 80
Corr-LDA	−12892.3	−12739.2	−12702.6	−12684.2
Corr-LDA-ITD	−11680.2	−11517.4	−11798.0	−11697.9

综合 4.2、4.3 节，在真实图像集 UIUC-Sport 和 LabelMe 上的实验结果表明：提出模型的标注效果较高，模型标注结果更加准确，模型拟合数据的能力更强. 原因在于，提出模型利用了真实图像中被标注事物往往占整个画面的部分较大这一信息，实验标注结果样例展示如图 3 所示. 例如第一张图像，

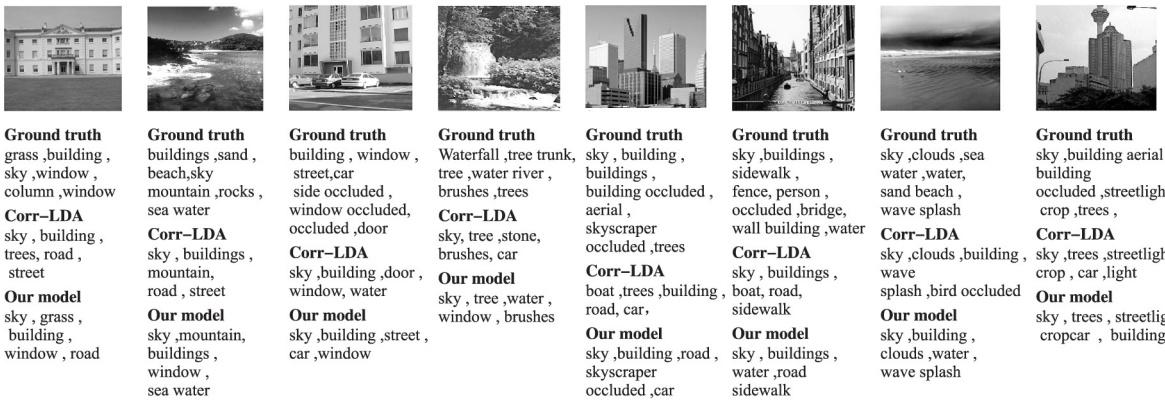

图 3 LabelMe 数据集上标注结果性能比较

Fig. 3 Comparison of annotation performance on LabelMe datasets

(C)1994-2022 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

Corr-LDA 模型标注了“sky”、“building”、“trees”、“road”、“street”5 个文本词,其中“trees”、“road”、“street”3 个错误标注词. 而提出模型标注了“sky”、“building”、“window”、“grass”、“road”5 个文本词,其中“road”1 个错误标注词. 并且 Corr-LDA 对图像中所占面积较大的事物“grass”没有标注,而提出模型对图像中所占面积较大的事物“sky”、“building”、“grass”都进行了正确标注.

5 结　论

本文利用真实图像中被标注事物往往占整个画面部分较大这一信息,在 Corr-LDA 模型的基础上提出了一种改进的图像标注概率主题模型(Corr-LDA-ITD 模型). 同时,为该模型推导了一个基于变分 EM 的参数估计算法,并给出了使用该模型标注图像的方法. 在两个真实数据集 UIUC-Sport 和 LabelMe 上与 mm-LDA、Corr-LDA、Mca-sLDA 模型进行比较,实验结果表明提出模型的标注性能相较于其它相比较模型有所提高.

另外,本文没有考虑文本之间的词间相关性对图像标注的影响. 例如,假设图像标注了“街道”,那么它标注“汽车”的概率一定高于标注“沙滩”的概率. 可见,词间相关性可提高图像标注的性能. 因此,本文后续将致力于利用词间相关性改善图像标注性能.

References:

[1] Bao Hong, Xu Guang-mei, Feng Song-he, et al. Advances in automatic image annotation [J]. Computer Science, 2011, 38(7): 35-40.

[2] Guo Qiao-jin, Ding Yi, Li Ning. Image annotation review based on keyword [J]. Computer Engineering and Applications, 2011, 47(30): 155-158.

[3] Gao Yan-yu, Yin Yi-xin, UOZUMI Takashi. A hierarchical image annotation method based on SVM and semi-supervised EM [J]. Acta Automatica Sinica, 2010, 36(7): 960-967.

[4] Qi X, Han Y. Incorporating multiple SVMs for automatic image annotation [J]. Pattern Recognition, 2007, 40(2): 728-741.

[5] Zhao Y, Zhao Y, Zhu Z, et al. A novel image annotation scheme based on neural network [C]. Eighth International Conference on Intelligent Systems Design and Applications. IEEE Computer Society, 2008: 644-647.

[6] Li X, Wang X, Wu C, et al. Correspondence with category latent dirichlet allocation for image annotation [C]. Multimedia Technology (ICMT), 2011 International Conference on. IEEE, 2011: 4879-4882.

[7] Jeon J, Manmatha R. Using aximum entropy for automatic image annotation [J]. Lecture Notes in Computer Science, 2004, 3115: 24-32.

[8] Blei D M, Ng A Y, Jordan M I. Latent dirichlet allocation [J]. Journal of Machine Learning Research, 2003, 3: 993-1022.

[9] Blei D M, Mcauliffe J D. Supervised topic models [J]. Advances in Neural Information Processing Systems, 2010, 3: 327-332.

[10] Blei D M, Lafferty J D. A correlated topic model of Science [J]. Annals of Applied Statistics, 2007, 1(1): 17-35.

[11] Monay F, Gatica-Perez D. PLSA-based image auto-annotation: constraining the latent space [C]. Proc. acm Int. conf. on Multimedia, 2004: 348-351.

[12] Li Xiao-xu. Study of image classification and annotation based on probabilistic topic model [D]. Beijing: Beijing University of Posts and Telecommunications, 2012.

[13] Blei D M, Jordan M I. Modeling annotated data [C]. International Acm Sigir Conference on Research & Development in Informaion Retrieval, ACM, 2003: 127-134.

[14] Yuan Liu, Zhang Long-bo. Tags predicted based on probability topic model [J]. Computer Science, 2011, 38(7): 175-180.

[15] Xing X, Shimada A, Taniguchi R. Latent topic model for image annotation by modeling topic correlation [C]. Multimedia and Expo (ICME), 2013 IEEE International Conference on. IEEE, 2013: 1-6.

[16] Wang C, Blei D, Li F F. Simultaneous image classification and annotation [C]. Computer Vision and Pattern Recognition, 2009, CVPR 2009. IEEE Conference on. 2009: 1903-1910.

[17] Du L, Ren L, Dunson D B, et al. A bayesian model for simultaneous image clustering, annotation and object segmentation [J]. Advances in Neural Information Processing Systems, 2009, 2009: 486-494.

[18] Lei W, Hoi S C H, Nenghai Y. Semantics-preserving bag-of-words models and applications [J]. IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society, 2010, 19(7): 1908-1920.

附中文参考文献:

[1] 鲍　泓,徐光美,冯松鹤. 自动图像标注技术研究进展 [J]. 计算机科学,2011,38(7):35-40.

[2] 郭乔进,丁　轶,李　宁. 基于关键词的图像标注综述 [J]. 计算机工程与应用,2011,47(30):155-158.

[12] 李晓旭. 基于概率主题模型的图像分类和标注的研究 [D]. 北京: 北京邮电大学,2012.

[14] 袁　柳,张龙波. 基于概率主题模型的标签预测 [J]. 计算机科学,2011,38(7):175-180.

(C)1994-2022 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net