

融入类别信息的图像标注概率主题模型

曹洁^{1,2}, 罗菊香¹, 李晓旭¹

CAO Jie^{1,2}, LUO Juxiang¹, LI Xiaoxu¹

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

1.College of Computer and Communication, Lanzhou University of Technology, Lanzhou 730050, China

2.Gansu Manufacturing Informatization Engineering Research Center, Lanzhou 730050, China

CAO Jie, LUO Juxiang, LI Xiaoxu. Image annotation probabilistic topic model fusing class information. *Computer Engineering and Applications*, 2017, 53(10): 187-192.

Abstract: The image annotation method based on the probabilistic topic model annotates images by learning the semantic of the image, and researchers pay more and more attention to it in recent years. Class label information can provide the valuable information for image annotation, for example, for images in “tall building” class, annotating “sky”, “skyscraper” is more possible than annotating “sea” and “beach”. However, for images in “coast” class, annotating “sea”, “beach” is more possible than annotating “sky” and “skyscraper”. This paper proposes an image annotation probabilistic topic model fusing class information which uses class information to promote image annotation. And it derives a parameters estimation algorithm based on the variational EM algorithm, as well as gives the method annotating the new images. The experimental results on LableMe and UIUC-Sport datasets show that the image annotation performance of the proposed model is better than other contrastive models.

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

摘要: 基于概率主题模型的图像标注方法旨在通过学习图像语义进行图像标注, 近年来倍受研究人员关注。考虑到类别对图像标注可提供有价值的信息, 例如, “高楼”类图像, 出现“天空”、“摩天楼”的可能性大于“海水”和“沙滩”。而“海岸”类图像出现“海水”、“沙滩”的可能性要大于“天空”和“摩天楼”。在Corr-LDA模型的基础上利用图像类别来改进图像的标注性能, 提出了一个融入类别信息的图像标注概率主题模型。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他相比较模型。

关键词: 图像标注; 图像类别; 变分EM; Corr-LDA模型

文献标志码: A **中图分类号:** TP391 **doi:** 10.3778/j.issn.1002-8331.1512-0234

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^[1-5]的研究也越来越受到学术界和工业界的关注。概率主题模型^[6-7]的图像标注研究作为图像标注研究领域的一个重要分支, 近年来已经

做了大量工作。

Blei^[8-9]等人基于LDA^[10]模型提出了Corr-LDA模型和Mm-LDA模型。Corr-LDA是图像标注的经典模型, Mm-LDA中文本模态和图像模态的主题条件独立。Yakhnenko^[11]提出了MoM-HDP模型, 模型建立在Mm-LDA模型基础上。Chen^[12]提出了HBP模型, 模型考虑整合了背景信息和上下文信息来学习主题特征。

基金项目: 国家自然科学基金(No.61263031); 甘肃省自然科学基金(No.1310RJZA034)。

作者简介: 曹洁(1966—), 女, 教授, 博士生导师, 研究方向: 信息融合、智能交通, E-mail: caoj2976016@qq.com; 罗菊香(1987—), 女, 硕士研究生, 研究方向: 智能信息处理; 李晓旭(1982—), 女, 副教授, 研究方向: 计算机视觉与机器学习。

收稿日期: 2015-12-18 **修回日期:** 2016-03-15 **文章编号:** 1002-8331(2017)10-0187-06

CNKI网络优先出版: 2016-09-28, <http://www.cnki.net/kcms/detail/11.2127.TP.20160928.1609.022.html>

Putthividhya^[13]提出了 Tr-mmLDA 模型,模型捕捉到了图像和文本主题的不同关联度。Xu^[14]等人引入主题之间的相关性提出了图像标注的 Corr-CTM 模型。这些工作都取得了相对较好的标注效果,但却没有用到图像类别信息,然而类别对图像标注可提供有价值的信息。如图 1 所示,一旦知道图像类别为“coast”,那么图像中出现“sea”和“rock”的可能性一定大于出现“trees”和“grass”;同理一旦知道图像类别为“county”,那么图像中出现“trees”和“grass”的可能性一定大于出现“sea”和“rock”。因此,本文拟利用图像类别来提高图像标注性能,提出了一个融入类别信息的图像标注概率主题模型。值得一提的是,该模型中标注词的生成与 Corr-LDA 模型不同,是依照图像主题比例来生成,即占画面比例的事物越大越有可能被标注。此外,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高于其他相比较模型。本文的主要创新之处在于:(1)提出了一个融入类别信息的图像标注概率主题模型。(2)给出了一个图像和其标注词之间的生成关系。

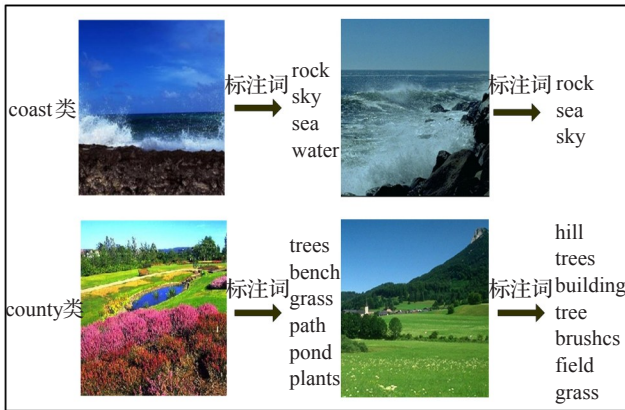


图1 LabelMe 图像集样例图片

本文在第二部分介绍了本文的图像标注模型,第三部分介绍了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了本文的工作。

2 图像标注概率主题模型

2.1 Corr-LDA 模型

Corr-LDA 模型是基于 LDA 模型做图像标注的经典模型,概率图模型表示如图 2 所示。模型标注词主题的选择来自图像主题空间,从而使得文本主题不超出图像主题的范围。模型用到的符号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 $v_m, m \in \{1, 2, \dots, M\}$;

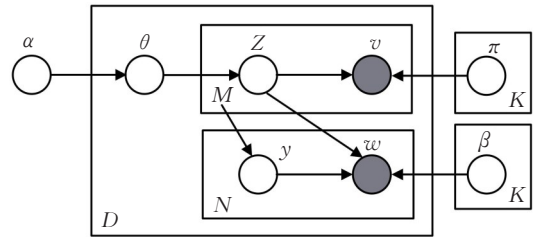


图2 Corr-LDA 模型的概率图模型

表1 符号及其意义

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
L	类别种类数	C	图像类别

- ① 抽取主题分派 $Z_m | \theta \sim Mult(\theta)$;
- ② 抽取图像词 $v_m | Z_m \sim Mult(\pi_{z_m})$;
- (3)对每一个文本词 $w_n, n \in \{1, 2, \dots, N\}$;
- ① 抽取离散索引变量 $y_n \sim Unif(1, 2, \dots, M)$;
- ② 抽取文本词 $w_n \sim Mult(\beta_{z_{y_n}})$;

模型潜变量和观测变量的联合分布为:

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

其中,

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

$$B = P(y_n | M) P(w_n | y_n, Z, \beta)$$

2.2 提出的模型

本文利用图像类别来改进图像的标注性能,提出了一个融入类别信息的图像标注概率主题模型,该模型中标注词的生成与 Corr-LDA 不同,是依照图像主题比例来生成,模型简称为 Corr*-LDA-L 模型。提出模型以词袋(bag of words)特征^[15]为基础,用到的符号及其所表示意义如表 1 所示。

- (1)抽取一个主题分派 $\theta_l \sim P_{dir}(\alpha)$;
 - (2)对每一个图像词汇 $v_m, m \in \{1, 2, \dots, M\}$;
 - ① 抽取主题分派 $Z_m | \theta, c \sim Mult(\theta, c)$;
 - ② 抽取图像词 $v_m | Z_m \sim Mult(\pi_{z_m})$;
 - (3)对每一个文本词 $w_n, n \in \{1, 2, \dots, N\}$;
 - ① 抽取主题分派 $T_n | Z \sim Mult(Z)$;
- 这里

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

② 抽取文本词 $w_n|T_n \sim \text{Mult}(\beta_{T_n})$;

提出模型的概率图模型表示如图3所示, 其潜变量和观测变量的联合分布为:

$$P(v, w, \theta, Z, T, C|\alpha, \pi, \beta) = \left(\prod_{i=1}^L P(\theta_i|\alpha) \right) \left(\prod_{m=1}^M A \right) \left(\prod_{n=1}^N B \right) \quad (2)$$

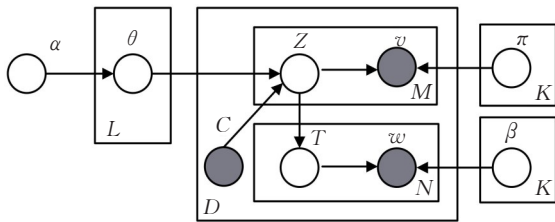


图3 模型的概率图模型

其中,

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

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

模型的生成可总结为两个阶段, 第一阶段: 首先, 按照 Dirichlet 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。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。

提出模型相较于 Corr-LDA 不同主要体现在: 提出模型利用到了图像的分类信息, 同时也用到了图像主题分布的信息; 而 Corr-LDA 这些信息都没有利用到。

3 参数求解与图像标注

对于提出模型, 本文采用变分 EM 算法求解模型参数。EM 算法^[3]是 Dempster, Laird, 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, 在 E 步骤中求解变分参数, 在 M 步骤中求解模型参数。

为了确定模型的参数, 引入潜变量 $H = \{Z, \theta, T\}$ 的变分分布:

$$q(H|\omega) = q(\theta|\gamma) \left(\prod_{m=1}^M q(z_m|\phi_m) \right) \left(\prod_{n=1}^N q(T_n|\rho_n) \right) \quad (3)$$

令 $E = \{v, w\}$ 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:

$$\text{lb} P(v, w|\Delta) = \text{lb} E_q \left(\frac{P(v, w, \theta, Z, T|\Delta)}{q(\theta, Z, T|\omega)} \right) \geq E_q(\text{lb} P(E, H|\Delta)) - E_q(\text{lb} q(H|\omega)) = L(\omega; \Delta)$$

3.1 变分 E 步骤

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:

$$\frac{\partial L(\gamma_{li})}{\gamma_{li}} = \sum_{l=1}^L \sum_{i=1}^K \left(\alpha_i - \gamma_{li} + \sum_{d=1}^D \sum_{m=1}^{M_d} \phi_{dmi} C_d^l \right) \times \left(\Psi'(\gamma_{li}) - \sum_{j=1}^K \Psi' \left(\sum_{j=1}^K \gamma_{lj} \right) \right)$$

令其为零可得:

$$\gamma_{li} = \alpha_i + \sum_{d=1}^D \sum_{m=1}^{M_d} \phi_{dmi} \cdot C_d^l \quad (4)$$

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

$$L(\phi) = + \sum_{d=1}^D \sum_{l=1}^L \sum_{m=1}^{M_d} \sum_{i=1}^K \left(\Psi(\gamma_{li}) - \Psi \left(\sum_{j=1}^K \gamma_{lj} \right) \right) \cdot$$

$$\phi_{dmi} \cdot C_d^l + \sum_{d=1}^D \sum_{m=1}^{M_d} \sum_{i=1}^K \sum_{j=1}^{V_s} \phi_{dmi} V_{dm}^j \text{lb} \pi_{ij} -$$

$$\sum_{d=1}^D \sum_{m=1}^{M_d} \sum_{i=1}^K \phi_{dmi} \text{lb} \phi_{dmi} +$$

$$\sum_{d=1}^D \sum_{n=1}^{N_d} \sum_{i=1}^K \rho_{dni} \left(\left[\frac{1}{M_d} \sum_{m=1}^{M_d} H \right] + \text{lb} M_d \right)$$

其中, $H = \phi_{dmi} \text{lb}(1+a) + (1-\phi_{dmi}) \text{lb} a$, a 是足够小的常数。

在 $\sum_{i=1}^K \phi_{dmi} = 1$ 约束下最大化 $L(\phi)$ 可得:

$$\phi_{dmi} \propto \pi_{iV_{dm}} \exp \left(\sum_{l=1}^L F \cdot C_d^l + \sum_{n=1}^{N_d} \frac{1}{M_d} \rho_{dni} \text{lb} \left(1 + \frac{1}{a} \right) \right) \quad (5)$$

其中,

$$F = \Psi(\gamma_{li}) - \Psi(\gamma_{lj})$$

计算参数 ρ : 从 $L(\omega; \Delta)$ 中选择包含 ρ 的项, 整理得到 $L(\rho)$:

$$L(\rho) = \sum_{d=1}^D \sum_{n=1}^{N_d} \sum_{i=1}^K \rho_{dni} \left(\left[\frac{1}{M_d} \sum_{m=1}^{M_d} H \right] + \text{lb} M_d \right) +$$

$$\sum_{d=1}^D \sum_{n=1}^{N_d} \sum_{i=1}^K \sum_{j=1}^{V_i} \rho_{dni} \omega_{dn}^j \text{lb} \beta_{ij} - \sum_{d=1}^D \sum_{n=1}^{N_d} \sum_{i=1}^K \rho_{dni} \text{lb} \rho_{dni}$$

在 $\sum_{i=1}^K \rho_{dni} = 1$ 的约束条件下最大化 $L(\rho)$ 可得:

$$\rho_{dni} \propto \beta_{iV_{dn}} \exp \left(\frac{1}{M_d} \sum_{m=1}^{M_d} H + \text{lb} M_d \right) \quad (6)$$

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 $\omega_d = \{\gamma, \phi, \rho\}$, $d \in \{1, 2, \dots, D\}$, 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:

$$L(D) = \sum_{d=1}^D L(\omega_d; \Delta) \quad (7)$$

求解模型参数 π : 从 $L(D)$ 中选择包含 π 的项:

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

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:

$$\pi_{ij} \propto \sum_{d=1}^D \sum_{m=1}^{M_d} \phi_{dmi}^j \nu_{dm}^j \quad (8)$$

求解模型参数 β : 从 $L(D)$ 中选择包含 β 的项:

$$L(\beta) = \sum_{d=1}^D \sum_{n=1}^{N_d} \sum_{i=1}^K \sum_{j=1}^{V_i} \rho_{dmi}^j \omega_{dn}^j \ln \beta_{ij}$$

在 $\sum_{j=1}^{V_i} \beta_{ij} = 1$ 的约束条件下, 相对于 β_{ij} , 最大化

$L(\beta)$ 可得:

$$\beta_{ij} \propto \sum_{d=1}^D \sum_{n=1}^{N_d} \rho_{dmi}^j \omega_{dn}^j \quad (9)$$

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 π 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:

$$P(\omega|v) \approx \sum_{n=1}^N \sum_{z_n} P(\omega|z_n, \beta) q(z_n|\phi) \quad (10)$$

4 实验

本文采用 F -measure 值来评价模型的标注性能, 具体如下所示:

$$F\text{-measure} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

其中, Precision(标注的正确率)是指标注正确的总个数与总的标注个数的比值; Recall(召回率)是指标注正确的总个数与真实的标注词总个数的比值。正确率和召回率都可以作为标注性能的评价指标, 但是由于正确率和召回率有时候会出现矛盾的情况, 而综合评价指标 (F -measure) 是综合两者的评估指标。

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。图像集 LabelMe 包含 8 类: “coast”, “forest”, “highway”, “inside city”, “mountain”, “open country”, “street”, “tall building”。每类包含 600 幅图像, 共 4 800 幅图像。图像集 UIUC-Sport 也包含 8 类: “badminton”, “bocce”, “croquet”, “polo”, “rock climbing”, “rowing”, “sailing”, “snowboarding”。每类图像数不等, 共 1 791 幅图像。

对 LabelMe 和 UIUC-Sport 图像集分别提取图像的词袋特征^[15], LabelMe 图像集具体过程如下:

(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。聚类中心构成图像码书, 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 2 500 个区块。

4.2 实验结果与分析

为了评估标注性能, 本文将提出模型与 Corr-LDA 模型、Multi-class sLDA with annotation (Mca-sLDA)^[16] 模型和 smmLDA^[17] 模型进行比较。Corr-LDA 是基于概率主题模型做图像标注的经典模型, Mca-sLDA^[16] 模型是同时做图像分类和标注的模型, smmLDA^[17] 模型是 2014 年 Tran 提出的做图像标注的概率主题模型。

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(10)标注图像, 选取概率大的前 5 个词作为标注词。

4.2.1 UIUC-Sport 数据集标注性能

UIUC-Sport 数据集上四个模型的标注性能比较如图 4 所示。

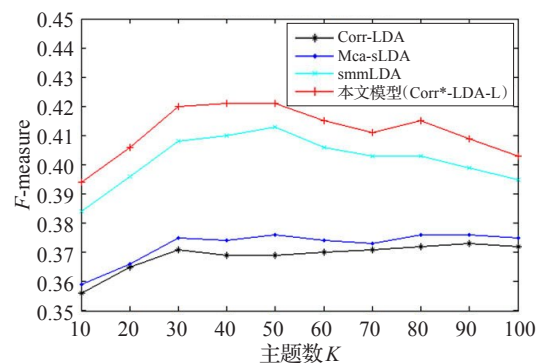


图4 UIUC-Sport 数据集上 F -measure 值性能比较

从图4可知, 本文提出模型在UIUC-Sport 真实图像集上, 在主题数目为10、20、30、40、50、60、70、80、90、100都表现出了提出模型效果要高于Corr-LDA模型、Mca-sLDA模型和smmLDA模型, 并且提出模型在30个主题的时候表现了最好的标注性能42.1%, 比Corr-LDA模型和Mca-sLDA模型超过了约4.5%, 比smmLDA模型超过约1.3%。可见, 在UIUC-Sport数据集上, 提出模型的标注效果要高于其他相比较模型。

4.2.2 LabelMe数据集标注性能

LabelMe数据集上四个模型的标注性能比较如图5所示。

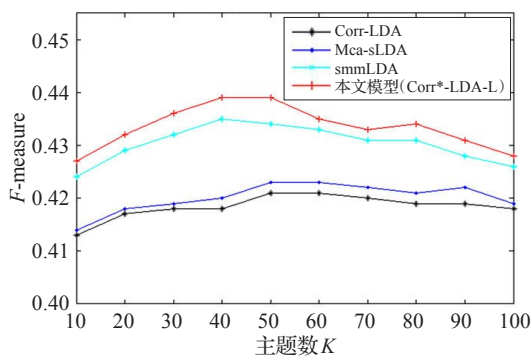


图5 LabelMe数据集上F-measure值性能比较

从图5可知, 提出模型在LabelMe真实图像集上, 在主题数目为10、20、30、40、50、60、70、80、90、100都表现

出了提出模型效果要高于Corr-LDA模型、Mca-sLDA模型和smmLDA模型, 并且提出模型在40个主题的时候表现了最好的标注性能44.0%, 比Corr-LDA和Mca-sLDA超过了约2.5%, 比smmLDA模型超过约1.0%。可见, 在LabelMe数据集上, 提出模型的标注效果要高于其他相比较模型。

4.3 标注结果

考虑到Mca-sLDA模型和Corr-LDA模型性能相近, 仅展示提出模型(Corr*-LDA-L模型)和Corr-LDA的标注结果, 如图6所示(斜体字表示错误的标注词)。从图6不难发现本文的标注结果更加准确。在第一排图像中, Corr-LDA的错误标注词要多于提出模型, 例如, 第一张图像中Corr-LDA标注的错误词有“buiding”和“car”, 而提出模型的错误标注词只有“stone”, 剩余图像也如此。在第二排图像中, 提出模型的标注误差要小于Corr-LDA, 例如, 在第一张图像中Corr-LDA标注的“building”和“car”, 不属于“mountain”类的标注词汇, 而提出模型标注的“ground”属于“mountain”类中的标注词汇, 离正确的标注词更接近。这也说明提出模型在标注词的生成过程中, 利用图像类别来促进标注这一思想是有效的, 利用图像类别信息确实可以改善图像的标注效果。

综上, 从UIUC-Sport和LabelMe两个真实数据集上的实验结果看, 提出模型的标注效果要高于Mca-sLDA、

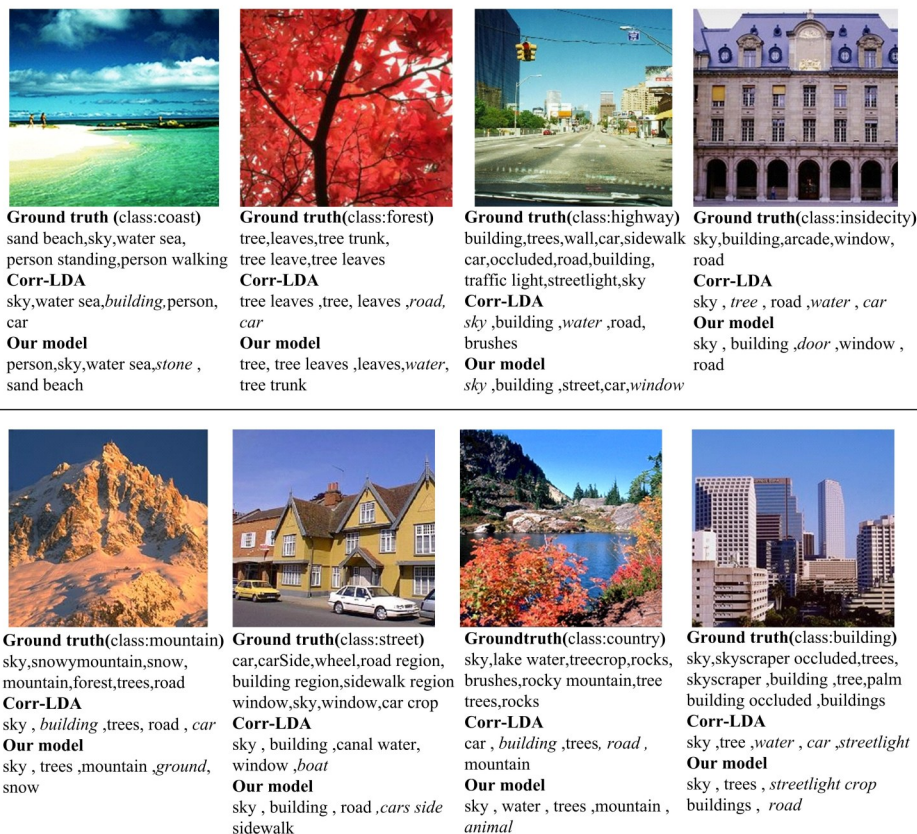


图6 LabelMe数据集上标注结果性能比较

Corr-LDA 模型和 smmLDA 模型。

5 结论

考虑到类别对图像标注可提供有价值的信息,本文在 Corr-LDA 模型的基础上利用图像类别来改进图像的标注性能,提出了一个融入类别信息的图像标注概率主题模型。该模型中标注词的生成与 Corr-LDA 模型不同,是依照图像主题比例来生成,即占画面比例的事物越大越有可能被标注。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 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他相比较模型。此外,在现实的应用中获得数据的类别往往是不准确、不完整的;弱类别图像大量存在于标准评测集和互联网图像集中,为了更好的利用这些弱类别的样本,所以本文后续将致力于弱类别图像的标注问题。

参考文献:

- [1] 杨栋,周秀玲,郭平.基于贝叶斯通用背景模型的图像标注[J].自动化学报,2013(10):1674-1680.
- [2] Cam-Tu C T N, De-Chuan Z, Zhihua Z. Multi-modal image annotation with multi-instance multi-label LDA[C]// Twenty-Third International Joint Conference on Artificial Intelligence(IJCAI),2013:1558-1564.
- [3] Gao Yanyu, Yin Yixin, Uozumi T. A hierarchical image annotation method based on SVM and semi-supervised EM[J]. Acta Automatica Sinica, 2010, 36(7): 960-967.
- [4] Xing X, Shimada A, Taniguchi R. Latent topic model for image annotation by modeling topic correlation[C]// 2013 IEEE International Conference on Multimedia and Expo(ICME). IEEE, 2013: 1-6.
- [5] Zhu S, Li X, Li Z, et al. An improved approach for image annotation[C]// 2014 33rd Chinese, Control Conference(CCC). IEEE, 2014: 4673-4678.
- [6] 李晓旭. 基于概率主题模型的图像分类和标注研究[D]. 北京:北京邮电大学,2012.
- [7] Morales-González A, García-Reyes E, Sucar L E. Unsupervised segmentation evaluation for image annotation[C]// International Conference on Computer Vision Theory and Applications, 2015.
- [8] Blei D M, McAuliffe J D. Supervised topic models[J]. Advances in Neural Information Processing Systems, 2010, 3: 327-332.
- [9] Blei D M, Jordan M I. Modeling annotated data[C]// Proceedings of the 26th annual international ACM SIGIR Conference on Research and Development in Information Retrieval. ACM, 2003: 127-134.
- [10] Blei D M, Ng A Y, Jordan M I. Latent dirichlet allocation[J]. the Journal of Machine Learning Research, 2003, 3: 993-1022.
- [11] Yakhnenko O, Honavar V. Annotating images and image objects using a hierarchical Dirichlet process model[C]// Proceedings of the 9th International Workshop on Multimedia Data Mining: Held in Conjunction with the ACM SIGKDD, 2008: 1-7.
- [12] Chen X, Lu C, An Y, et al. Probabilistic models for topic learning from images and captions in online biomedical literatures[C]// Proceedings of the 18th ACM Conference on Information and Knowledge Management, 2009: 495-504.
- [13] Putthividhya D, Attias H T, Nagarajan S S. Topic regression multi-modal Latent Dirichlet Allocation for image annotation[C]// The IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Francisco, CA, United States, 2010: 3408-3415.
- [14] Blei D M, Lafferty J D. A correlated topic model of science[J]. The Annals of Applied Statistics, 2007: 17-35.
- [15] Wu L, Hoi S C H, Yu N. Semantics-preserving bag-of-words models and applications[J]. IEEE Transactions on Image Processing, 2010, 19(7): 1908-1920.
- [16] Wang C, Blei D, Li F F. Simultaneous image classification and annotation[C]// IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2009: 1903-1910.
- [17] Tran T H, Choi S. Supervised multi-modal topic model for image annotation[C]// 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2014: 5979-5983.