

基于社会标签的图像情感自动分类标注研究*

■ 陆泉 陈静 丁恒

[摘要] 以现存的大量图像社会标签为基础,设计基于社会标签的图像情感自动分类标注模型,提出图像社会标签对情感的贡献度、区分度以及标签在情感维度上的权重3个因子及计算方法,实现图像5种基本情感语义的分类标注。在实际社会网络数据集上进行验证实验,结果表明建立的模型方法在自动标注图像情感语义方面具有较高的准确性和应用性。同时还发现,实际数据集中情感分布差异与复合情感对图像自动标注具有影响。

[关键词] 情感标注 自动图像标注 社会标签 文本分类 语义鸿沟

[分类号] G353

DOI: 10.13266/j.issn.0252-3116.2014.12.018

1 引言

随着图像信息资源的迅速增长,图像标注与检索成为一项研究热点。从基于内容的图像检索、基于语义的图像检索、图像自动标注到基于情感的图像检索,学者对图像的底层视觉特征以及图像的空间关系、对象、场景、行为和情感等高层语义进行了广泛的研究。目前,利用情感信息进行图像检索已成为图像研究的前沿,而图像情感标注是其核心基础问题。然而,图像视觉特征与图像高层语义之间普遍存在语义鸿沟,影响了标注效果,需要寻找更准确高效的图像情感标注方法。

目前,社会标签方法已经广泛应用到图像标注中。由于标签本身就是用户对图像内容精炼的语言表达,反映了图像的语义内容,同时,图像较低层次的语义内容与图像蕴含的高层情感语义之间存在着本质上的关联关系,因此,在社会图像分享网络(如 flickr.com)中,大量存在的社会标签为研究图像情感语义的自动标注提供了内容基础。本文以此作为研究出发点,利用图像的社会标签,提出一种高效准确的图像情感自动分类标注方法,有效实现图像情感语义标注。

2 文献回顾

图像情感自动标注的基础问题是图像情感模型。其实质是确定图片情感基调及分类,建立一个合理的

图像情感描述框架,一般表示为表达情感的形容词或词对以及这些词之间的关系^[1]。目前,研究人员主要以扩展心理学中的基本情绪理论或情绪维度理论为基本思路。如 S. Stefanie 等利用 P. Ekman 的基本情绪理论建立情感描述框架“scroll bar”^[2],该模型包含“高兴、愤怒、悲伤、厌恶、恐惧”5种基本情感。J. A. Russell 提出的二维情绪理论^[3]被许多学者扩展为包含各种情感概念的二维情感模型;王上飞等则借鉴了 A. Mehrabian 建立的二维情绪理论构建了基于“维量”思想的人工情感模型^[4]。学者们进行理论扩展的重点在于建立情感语义与图像其他信息之间的关系,如王上飞等提出的人工情感模型直接建立图像视觉特征与图像情感语义之间的映射关系^[4];而 J. P. Eakins 提出的图像三层语义模型则认为,图像情感语义是通过场景语义和对象语义进行推理得到的最高层次语义^[5]。

另有更多学者致力于研究图像情感模型之上的图像情感自动标注方法,利用图像视觉特征是其重要特点。研究思路是通过分析图像特定视觉特征给人们带来的情感反应,为图像分配合理的语义标签,建立特定视觉特征到情感语义词汇的直接映射。图像的颜色、形状、纹理、空间关系等视觉特征均可用于建立映射。如 T. Hayashi 等以颜色直方图为图像情感特征,建立了色彩与情感词之间的映射^[6];Dai Ying 通过分析图像的 HSV 颜色分量的纹理参量,发现了纹理对5种情感

* 本文系国家自然科学基金面上项目“图像信息资源可视化协同语义标注及实现研究”(项目编号:71273195)和国家自然科学基金青年项目“图书层次主题自动标引研究”(项目编号:71303089)研究成果之一。

[作者简介] 陆泉,武汉大学信息管理学院副教授,硕士生导师,博士,E-mail: dancinlulu@sina.com;陈静,华中师范大学信息管理学院副教授,硕士生导师;丁恒,武汉大学信息管理学院博士研究生。

收稿日期:2014-04-16 修回日期:2014-05-26 本文起止页码:118-123 本文责任编辑:杜杏叶

的影响^[7]。

在映射方法上,也有多种理论方法可以选取。如王上飞等采用支持向量机算法建立图像内容和其所表达的情感语义之间的联系,实现了服装和风景图像的情感检索^[8]; M. Dellagiacoma 等利用支持向量机作为分类器,对自然风景图像进行情感识别^[9]; 刘增荣等进而提出了一种加权值的图像特征融合算法来实现情感语义的识别^[10]。

近年来,学者们采用的方法呈现出多元化趋势。如赵娟娟等综合利用粗糙集理论构建了图像情感语义规则简化的方法^[11]; C. Colombo 等人借鉴 J. Itten 的语义理论,建立了图像的底层视觉特征与图像情感语义之间的映射推导机制^[12]; S. B. Cho^[13-14]、R. Datta^[15]等利用机器学习和软计算等先进算法,通过挖掘图像视觉特征与情感的对应关系进行图像情感自动标注。

综合图像视觉内容和图像相关文本进行图像自动标注得到了许多学者关注。P. Duygulu 等采用了机器翻译模型通过聚类图像区域将 blob token 集合转化为关键词集合^[16]; J. Jeon 等提出了一个跨媒体相关性模型 CMRM,寻找未标注图像的相似图像,并利用相似图像的标签来标注未标注图像^[17]; 此后在 CMRM 的基础上又衍生出连续空间相关性模型^[18]和多重伯努利相关性模型^[19]。但是,相关性模型存在先验概率难以估计及训练集要求较高等局限性。随后, Liu Jing 等提出了一个二元跨媒体相关性模型,将互联网文本信息检索技术融入自动标注技术中,弥补了传统相关性模型的先验概率方面的缺陷^[20]。

基于图像视觉特征的方法也有不足之处,主要在于图像视觉特征与图像高层语义之间存在普遍语义鸿沟^[21]。而且,由于情感语义是图像最高层语义,其自动标注研究中的“维数灾难”现象特别明显^[22],即图像情感语义的相关图像底层视觉特征维数非常高,导致自动标注的性能显著下降,并主要体现为图像情感标注准确率非常低^[21]。另外,图像视觉特征处理的算法时间复杂度较高。张志武等分析了现有大部分研究将标注和排序均在检索中完成导致的效率低下问题,建议将标注与检索研究分离,采取离线标注处理图像语义标注^[22]。

值得注意的是,以图像中层语义为特点的图像社会标签与图像高层情感语义之间有相对较小的语义鸿沟^[5],同时,文本分析技术是社会化媒体中情感分类的一种科学方法和研究手段。余传明等分析了情感词汇共现网络的特性,发现情感词汇共现网络的定点度、定

点强度和定点交互系数之间具有正相关性,从统计学证据角度支持了利用文本共现关系分析情感语义的可能性^[23]。在文本情感分析的应用方面,李纲等利用文本分类技术,开发出比较完整的汽车情感倾向识别系统^[24]; B. Schuller 等利用文本情感分析方法较好地抽取了演讲视频信息中的情感^[25]。结合上述研究及相关综述分析^[26]可以认为,图像的高层情感语义可由图像多个语义标签所传递情感经某种合并计算获得;另外,较之于图像视觉特征分析,语义标签分析具有处理效率上的优势。

3 模型构建

3.1 图像情感模型

本模型的构建主要依据基本情绪理论。该理论认为,基本情绪或分立情绪在人类中是普遍存在的,包括兴趣、高兴等基本积极情绪,以及悲伤、愤怒、厌恶和恐惧等基本消极情绪;与基本情绪相对的是复合情绪,包括基本情绪的各种变化及混合情绪,分类标注较困难。因此,本研究采用情报学专家 G. S. Wolfgang 等^[2]建立的“高兴”、“愤怒”、“悲伤”、“厌恶”、“恐惧”共 5 种基本情感的图像情感模型。

3.2 基于社会标签的图像情感自动分类标注模型

由图像社会标签计算图像情感的原理可做如下定义:对于任一图像 P,其社会标签集合为 $T = (t_1, t_2, \dots, t_n)$, T 中的元素属于特征语义、空间关系语义、对象语义、行为语义、场景语义及元数据等类型,每个标签 t_i 携带的情感信息为 E_{t_i} ,那么该图像的情感语义 $|E_T|$ 可由这些标签中的情感信息 $E_{t_i} (i = 1, \dots, n)$ 的某种合并运算 F 获得,表达如式(1)。需要指出,有一些图像标签属于元数据类型,如“canon”(设备品牌元数据)、“closeup”(拍摄手法元数据,意为特写镜头)等,本文认为元数据也对图像情感有贡献,不剔除元数据。

$$|E_T| = F(E_{t_1}, E_{t_2}, \dots, E_{t_n}) \quad (1)$$

式(1)中 $|E_T|$ 的取值为 5 种基本情感之一。若将社会标签的集合看作文档,将 5 种基本情感看作 5 个类别,那么该问题就转化为文档分类问题。本文应用朴素贝叶斯模型(Naive Bayesian Model, NBC)进行标签集文档分类,构建基于社会标签的图像情感自动分类标注模型,如式(2)。

$$\text{score}(P, E) = \prod_{i=1}^n PS(t_i, E) * D(t_i, E) * W(t_i, E) \quad (2)$$

其中 $\text{score}(P, E)$ 为图像 P 在某类情感 E 上的得分,由标签 3 项得分乘积的代数和计算获得。当 5 种

基本情感中某一情感 E 的得分最高(在贝叶斯模型中概率最大)时,模型将图像标注为具有情感语义 E。式(2)中 $PS(t_i, E)$ 为标签 t_i 对情感 E 的贡献度因子(简称“贡献度”),如训练集中标签 t_i 与代表情感 E 的标签共现,则可认为标签 t_i 对情感 E 有贡献,贡献度的值为贝叶斯模型中标签 t_i 出现时情感 E 的概率; $D(t_i, E)$ 是标签 t_i 对情感 E 的区分度因子(简称“区分度”),这是由于将模型所有标签都认为是任何情感类型 E 的特征项,某些情况下会导致模型不能够很好地区分多种情感类型,特别是当多种情感类型之间存在相关关系时,部分标签 t 对多种情感的区分度是很低的,只使用 $PS(t_i, E)$ 计算情感得分并不能得到良好的结果,往往导致多种情感类型的得分很相近; $W(t_i, E)$ 为标签在情感 E 维度上的权重因子(简称“标签权重”),由于在上述算法中高频词比低频词在语义距离上更容易接近某些情感,为了抵消词频分布对标注的影响,模型加入标签权重因子进行调节。

4 因子的计算方法

4.1 标签对情感的贡献度 $PS(t_i, E)$ 的计算方法

语义标签 t_i 对情感 E 的贡献度 $PS(t_i, E)$ 可以通过多种语义距离计算方法获得。本文采用谷歌距离(Normalized Google Distance)^[20] 计算标签 t_i 对情感概念 E 的语义距离 $SD(t_i, E)$,如公式:

$$SD(t_i, E) = \frac{\max(\log f(t_i), \log f(E)) - \log f(t_i, E)}{\log N - \min(\log f(t_i), \log f(E))} \quad (3)$$

其中 t_i 代表某一标签, E 为一种类型的情感概念, $f(t_i)$ 为包含标签 t_i 的图片数目, $f(E)$ 表示属于情感类型 E 的图片数目, $f(t_i, E)$ 表示包含有标签 t_i 且属于情感类型 E 的图片数目, N 为数据库中所有的图片数目。 $SD(t_i, E)$ 越小代表标签 t_i 与情感概念 E 在语义上越接近,即标签 t_i 对情感概念 E 的贡献度 $PS(t_i, E)$ 越大,由公式知 $SD(t_i, E)$ 是一个属于区间 $(0, +\infty)$ 的数,而 $PS(t_i, E)$ 应是一个属于区间 $(0, 1)$ 的数,因此对公式(3)进行如下变型得 $PS(t_i, E)$ 的计算公式:

$$PS(t_i, E) = \exp(-\gamma * SD(t_i, E)) \quad (4)$$

其中 γ 是一个可调参数,表达了标签对情感的贡献度在整个模型 3 个部分中的权重比例。

4.2 标签对情感的区分度 $D(t_i, E)$ 的计算方法

语义标签 t_i 对情感 E 的区分度 $D(t_i, E)$ 可以通过特征选择计算方法获得。本文采用互信息方法。 $D(t_i, E)$ 的计算方法为:在图像总数为 N 的图像社会标

签数据集中,对任意标签 t_i ,记包含标签 t_i 的图片数目为 N_{ii} ,属于情感类型 E 的图片数目为 N_E ,包含标签 t_i 且属于情感类型 E 的图片数目记为 $N_{(t_i, E)}$,包含标签 t_i 且不属于情感类型 E 的图片数目为 $N_{(t_i, E')}$,不包含标签 t_i 且属于情感类型 E 的图片数目为 $N_{(t_i', E)}$,不包含标签 t_i 且不属于情感类型 E 的图片数目为 $N_{(t_i', E')}$,显然, $N_{ii} = N_{(t_i, E)} + N_{(t_i, E')}$, $N_{ii'} = N_{(t_i', E)} + N_{(t_i', E')}$,则标签 t_i 对情感 E 的区分度 $D(t_i, E)$ 为:

$$D(t_i, E) = \frac{N_{(t_i, E)}}{N} \log_2 \frac{N * N_{(t_i, E)}}{N_{t_i} * N_t} + \frac{N_{(t_i', E)}}{N} \log_2 \frac{N * N_{(t_i', E)}}{N_{t_i'} * N_t} + \frac{N_{(t_i, E')}}{N} \log_2 \frac{N * N_{(t_i, E')}}{N_{t_i} * N_{t_i'}} + \frac{N_{(t_i', E')}}{N} \log_2 \frac{N * N_{(t_i', E')}}{N_{t_i'} * N_{t_i'}} \quad (5)$$

采用互信息法时,多个类的数据规模差异对词项与类的互信息数值大小有很大影响,在本文中体现为情感标签的词频差异以及其他标签的词频分布差异。为了抵消该影响,本文通过调整标签权重参数进行优化。

4.3 标签在情感维度上的权重 $W(t_i, E)$ 的计算方法

语义标签 t_i 情感 E 维度上的权重 $W(t_i, E)$ 可以通过多种先验概率计算方法估计。本研究中采用最大似然估计法,计算方法为:若对于任意标签 t_i ,当整个数据库中包含图片数目为 N,属于情感类型 E 的相关图片数目为 S,且其中包含标签 t_i 的图片数目为 df_{t_i} ,其中词项 t 出现时属于情感类型 E 的相关图片数目为 s 时,该标签 t 属于情感类型 E 的最大似然估计如下式:

$$W(t_i, E) = \log \frac{s / (S - s)}{(df_{t_i} - s) / ((N - df_{t_i}) - (S - s))} \quad (6)$$

由于公式(5)所得的 $W(t_i, E)$ 属于 $(-\infty, +\infty)$ 区间,但模型计算需要将其变换至 $(0, +\infty)$ 区间,变化后如下式:

$$W(t_i, E) = \exp(\psi * \log \frac{s / (S - s)}{(df_{t_i} - s) / ((N - df_{t_i}) - (S - s))}) \quad (7)$$

其中 ψ 是一个可调参数,可调整标签权重在整个模型三部分因子中的权重比例,用以消除词频差异对分类标注的影响。

5 实验与数据分析

5.1 实验过程

由于上述计算所需数据均可利用社会图片分享网站 flickr.com 的 api 获得,因此,本实验于 2013 年 8 月 25 日从 flickr.com 采集包含 happy、sad、fear、anger、dis-

gust 共 5 种情感标签的图片各 100 张(共 500 张)及其标签,剔除这 5 种情感标签后,单幅图像的标签数量最少者为 38 个,500 幅图像里共有 8 912 个不同标签,建立起测试集;同时,以 flickr.com 上剔除了上述 500 幅图像的其余所有图像及其标签为训练集,包括 happy 1 514 902 张、sad 251 972 张、anger 35 736 张、fear 98 811 张、disgust 3 077 张;其次,为获得本文第 4 节计算所需的训练数据,再次通过 java 调用 flickr.api 采集测试集中的 8 912 个标签在 flickr.com 上的各种图片数目数据,并扣除测试集中的对应图片数目数据,得到训练数据;最后,训练模型,并用训练后的模型自动分类标注测试集中图像的情感,通过与剔除情感标签前的数据比较来评估模型。

5.2 实验数据分析

5.2.1 区分度数值分布分析

上文已指出,区分度的数值分布会对模型计算结果产生较大影响。本实验区分度数值具体分布情况如图 1 所示:

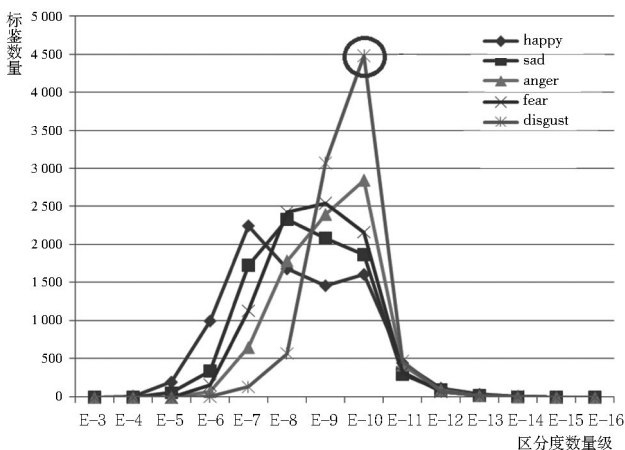


图 1 区分度的数值分布

图 1 中横轴表示区分度数值的数量级,纵轴表示对应某种情感的标签数量,如被圈起的点 1 表示有 4 500 个标签对 disgust 的区分度的值处于 10^{-10} 数量级上。从图 1 可看出,5 种情感的区分度在数量级分布上存在较大差异,这将导致在具有三部分相乘特性的整个模型中,区分度对计算结果影响过大。另外,happy 情感的区分度折线更靠近纵轴,具体而言,区分度最大的前 100 个分类特征中,与 happy 对应的分布在 10^{-3} 至 10^{-5} 数量级上,而与 sad、anger、fear 对应的分布在 10^{-5} 至 10^{-6} 数量级上,与 disgust 对应的分布在 10^{-6} 至 10^{-7} 数量级上,这将使模型标注 happy 的概率远高于其他情感。其原因是在训练集中存在情感分布不均现象,拥有 happy 标签的图像远多于拥有其他 4 种情感标签的图像,所以利用互信息方法计算 $D(t_i, E)$ 时

会导致这样一个极端数值分布。

5.2.2 参数 ψ 调整与标注结果评价分析

为了消除 $D(t_i, E)$ 的数值分布影响,本文进行了调整标签权重参数 ψ 的多组实验,目的是平衡各部分在整个模型中对结果产生影响的权重。考虑可通过调节参数 ψ 来调整模型,实验将 γ 参数取值为 1。实验结果表明,当 $\psi = 15$ 时,本模型标注图像情感的效果达到最佳,其微平均正确率为 0.899 4,微平均召回率为 0.757 6,微平均 F1 值为 0.822 4。

在不同 ψ 值情况下,5 种情感及总体的微平均正确率的变化见图 2。可以看出,模型的微平均正确率接近 0.9;对于 happy 情感,标签权重参数 $10 < \psi < 20$ 时平均正确率达到峰值;标签权重参数 ψ 的变化对于 anger、sad 两种情感的微平均正确率变化影响较小;对于 fear 情感,当 $\psi > 10$ 之后其微平均正确率趋于稳定,对于 disgust 情感,当 $10 < \psi < 20$ 时平均正确率达到峰值。

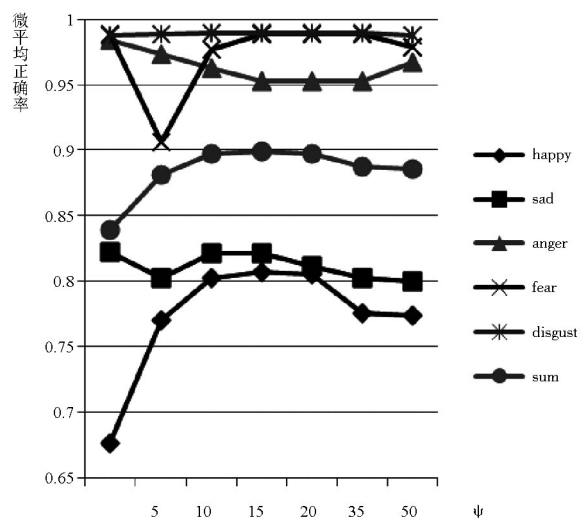


图 2 微平均正确率变化

5 种情感及总体的微平均召回率的变化见图 3。可以看出,模型的微平均召回率在 0.75 左右;标签权重参数 ψ 的增大会使 happy 情感的微平均召回率不断下降,同时会使 sad 情感的微平均召回率不断上升,其原因是:500 幅测试图像中存在复合情感图像,因此具有多种基本情感标签,其中 happy 与 sad 标签共现 26 次,而其他基本情感标签共现次数均小于 8。由于本文采用单类分类器设计思路,因此标签权重参数 ψ 的增大实际上是将测试数据中部分既属于 happy 情感又属于 sad 情感的图像从 happy 标注结果集中向 sad 标注结果集中转移,从而导致了上述变化。

当 $\psi > 35$ 之后, happy 情感的平均召回率基本不变,这说明模型已经能够基本区分积极情感(happy)和

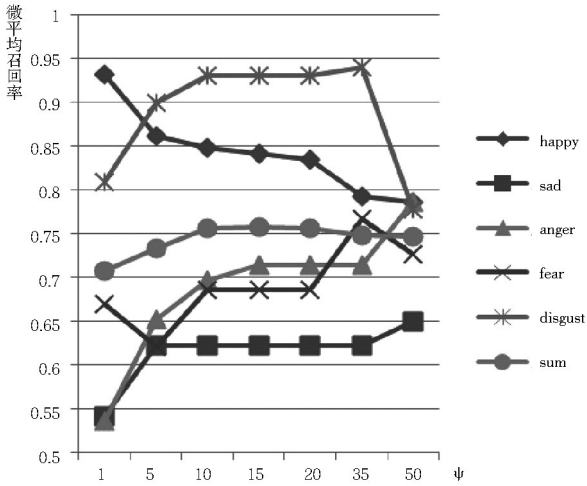


图3 微平均召回率变化

消极情感(sad、anger、fear、disgust);另外,disgust与fear情感的平均召回率大幅下降,同时anger情感的平均召回率大幅上升,sad情感的平均召回率小幅上升,说明anger、fear、disgust三者之间的相关性要高于anger、fear、disgust和sad之间的相关性,体现为anger、fear、disgust三种情感的交集图像大部分随着 ψ 的变化不断从disgust、fear的标注结果集中向anger标注结果集转移,极少数随着 ψ 的变化不断从disgust、fear的标注结果集中向sad标注结果集转移。

5种情感及总体的微平均F1值的变化见图4。可以看出,模型的微平均F1值在0.8左右,disgust情感的F1值一直处于较大值0.9左右,而sad情感的F1值始终处于较小值0.7左右,说明模型对于disgust情感的分辨能力最强,对于sad情感的分辨能力最弱。happy情感的F1值基本稳定在0.8左右,说明模型对happy情感的分辨能力最稳定;fear和anger情感的F1值有明显变化(全距在1.5左右),说明模型对fear和anger情感的分辨能力较不稳定。

从实验结果来看,本文提出的基于社会标签的情感自动标注模型及方法的标注正确率总体较高,当 $10 < \psi < 20$ 时各种情感标注正确率都大于0.8,其中对anger、disgust、fear3种情感的标注正确率在0.95以上;其次,有4种情感标注召回率都大于0.65,而sad召回率较低的原因是,采自实际网络的测试数据集含有复合情感类型图像,其中sad和happy图像交集较大,受到happy情感类型的影响;再次,从F1值上看,当 $10 < \psi < 20$ 时,模型的分类标注能力将处于稳定状态;最后,综合来看,5种基本情感中,自动标注disgust的准确性最高。

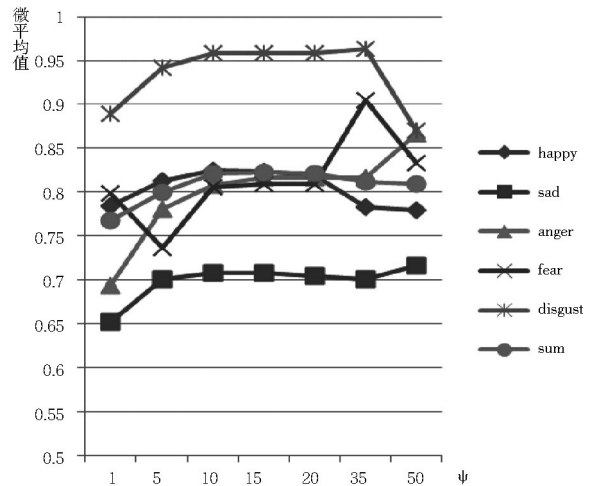


图4 微平均F1值变化

6 结语

本文提出基于社会标签的图像情感自动分类标注模型及方法,并在实际社会网络数据集flickr上进行了实验。结果表明:其在图像情感标注中具有较高的准确性和应用性,同时还发现了实际数据集中情感分布差异与复合情感对自动标注的影响。

本文的主要特色在于:首先,模型考虑了图像社会标签对情感的贡献度、标签对情感的区分度及标签在情感维度上的权重三个因子,其中,采用语义距离计算贡献度比统计概率更加适合表达情感与语义标签之间的关系,区分度的提出精简了传统特征选择的繁琐环节,权重抵消了词频分布对模型的不良影响。其次,可通过调整标签权重参数 ψ 优化分类标注效果。再者,由于完全基于社会标签中包含的语义信息,所以不存在图像视觉特征处理导致的语义鸿沟问题,分类标注准确性高,且算法性能优越。最后,利用flickr上的实际数据集进行了实证,并有所发现。今后的工作将进一步探索复合情感图像的识别与自动标注,以及学习实际数据集中情感分布差异来自动优化参数设置的方法,最终建立一个比较完善的基于社会标签的图像情感自动分类标注系统。

参考文献:

- [1] 王上飞,王煦法.图像情感检索研究的进展与展望[J].电路与系统学报,2005,10(4):102-110.
- [2] Stefanie S,Wolfgang G S. Collective indexing of emotions in images: A study in emotional information retrieval[J]. Journal of the American Society for Information Science and Technology,2009,60(5):863-876.
- [3] Russell J A. A circumplex model of affect[J]. Journal of Personality and Social Psychology,1980,39(6):1161-1178.

- [4] 王上飞, 王煦法. 基于“维量”思想的人工情感模型(英文)[J]. 中国科学技术大学学报, 2004(1): 83-91.
- [5] Eakins J P. Automatic image content retrieval - Are we getting anywhere? [C] // Proceedings of the Third International Conference on Electronic Library and Visual Information Research. Milton Keynes: Aslib, 1996: 123-135.
- [6] Hayashi T, Hagiwara M. Image query by impression words - The IQI System[J]. IEEE Trans On Consumer Electronics, 1998, 44(2): 347-352.
- [7] Dai Ying. Intention-based image retrieval with or without a query image [C] // Proceedings of the 10th International Multimedia Modeling Conference. Washington, DC: IEEE Computer Society, 2004: 26-32.
- [8] 王上飞, 陈恩红, 汪祖媛, 等. 基于支持向量机的图像情感语义注释和检索算法的研究[J]. 模式识别与人工智能, 2004(1): 28-31.
- [9] Dellagiacoma M, Zontone P, Boato G. Emotion based classification of natural images [C] // Proceedings of the International Workshop on Detecting and Exploiting Cultural diversity on the Social Web. New York: ACM, 2011: 17-22.
- [10] 刘增容, 余雪丽, 李志. 基于特征融合的图像情感语义识别研究[J]. 太原理工大学学报, 2012, 43(5): 553-557.
- [11] 赵娟娟, 陈俊杰, 李国庆. 图像情感语义规则简化学习算法[J]. 计算机科学, 2009, 36(10): 240-243.
- [12] Colombo C, Bimbo D, Pala P. Semantics in visual information retrieval [J]. IEEE Multimedia, 1999, 6(3): 38-53.
- [13] Cho S B. Emotional image and musical information retrieval with interactive genetic algorithm [J]. Proceedings of the IEEE, 2004, 92(4): 702-711.
- [14] Cho S B, Lee J Y. Advanced signal processing technology by soft computing [M]. World Scientific Pub Co. Inc. 2001: 225-239.
- [15] Datta R, Li Jia, Wang J Z. Algorithmic inferencing of aesthetics and emotion in natural images: An exposition [C] // Proceedings of 15th IEEE International Conference. San Diego: IEEE, 2008: 105-108.
- [16] Duygulu P, Barnard K, Freitas J, et al. Object recognition as machine translation: Learning a lexicon for a fixed Image vocabulary [C] // Proceedings of the 7th European Conference on Computer Vision. London: Lecture Notes in Computer Science, 2002: 97-112.
- [17] Jeon J, Lavrenko V, Manmatha R. Automatic image annotation and retrieval using cross-media relevance model [C] // Proceedings of the 26th annual international ACM SIGIR. New York: ACM, 2003: 119-126.
- [18] Lavrenko V, Manmatha R, Jeon J. A model for learning the semantics of pictures [EB/OL]. [2014-02-16]. <http://ciir.cs.umass.edu/pubfiles/mm-46.pdf>.
- [19] Duygulu P, Barnard K, Freitas J, et al. Object recognition as machine translation: Learning a lexicon for a fixed Image vocabulary [C] // Proceedings of the 7th European Conference on Computer Vision. London: Lecture Notes in Computer Science, 2002: 97-112.
- [20] Liu Jing, Wang Bin, Li Mingjing, et al. Dual cross-media relevance model for image annotation [C] // Proceedings of the 15th International Conference on Multimedia. New York: ACM, 2007: 605-614.
- [21] Enser P G B, Sandom C J, Hare J S, et al. Facing the reality of semantic image retrieval [J]. Journal of Documentation, 2007, 63(4): 465-481.
- [22] 张志武, 阚德涛. 基于语义学习的自动图像标注技术研究述评 [J]. 情报学报, 2013, 32(10): 1112-1120.
- [23] 余传明, 周丹. 情感词汇共现网络的复杂网络特性分析 [J]. 情报学报, 2010, 25(5): 906-914.
- [24] 李纲, 程明结, 寇广增. 基于情感倾向识别的汽车评论挖掘系统构建 [J]. 情报学报, 2011, 30(2): 204-211.
- [25] Schuller B, Reiter S, Muller R, et al. Speaker independent speech emotion recognition by ensemble classification [C] // 2005 IEEE International Conference on Multimedia and Expo. Amsterdam: IEEE, 2005: 864-867.
- [26] 杨立公, 朱俊, 汤世平. 文本情感分析综述 [J]. 计算机应用, 2013, 33(6): 1574-1578.

Automatic Image Emotion Categorization and Annotation Based on Social Tags

Lu Quan¹ Chen Jing² Ding Heng¹

¹School of Information Management, Wuhan University, Wuhan 430072

²School of Information Management, Central China Normal University, Wuhan 430079

[Abstract] This paper, based on the existing massive social tagged images, designs an automatic image emotion categorization and annotation model based on these social tags, introduces three factors and their calculation methods, including weight to emotion, discrimination to emotion and contribution to emotion of a social tag, which can categorize images into five basic emotions and annotate each image with the corresponding emotion. Finally experiments are designed, and detailed data analysis is presented on a real social tagging website.

[Keywords] emotion annotation automatic image annotation social tagging text categorization semantic gap