

Image Auto-Annotation Using Semantic Link Network

Haijiao Xu, Jialei He, Dehui Ou, Shifeng Peng

Department of Computer Science, Guangdong University of Education, Guangzhou Guangdong
Email: guesskkk99@163.com, 947440047@qq.com, 695470223@qq.com, 2498597150@qq.com

Received: Jul. 30th, 2019; accepted: Aug. 19th, 2019; published: Aug. 26th, 2019

Abstract

Image Auto-annotation remains as a challenge in machine vision. Most conventional approaches concentrate on the relations between visual images and labeled concepts, and neglect the correlations between semantic concepts. In this paper, we present a novel approach called Image Auto-annotation using Semantic Link Network (IA-SLN), which can effectively capture semantic correlations to boost image auto-annotation. Specifically, we first construct a semantic link network based on the global concept correlations and the local concept correlations. Second, we utilize it to annotate unlabeled images. When a concept and its associated concepts frequently occur in images, the prediction of this concept will be boosted in our IA-SLN. Finally, we obtain the most relevant annotations for each of unlabeled images by using a new strategy of annotation promotion. The experimental results on the publicly available dataset IAPR have shown that our approach performs favourably compared with several other state-of-the-art methods.

Keywords

Semantic Link Network, Semantic Estimation, Kolmogorov Complexity, Automatic Image Annotation

基于语义链接网的图像自动标注

徐海蛟, 何佳蕾, 区德辉, 彭世锋

广东第二师范学院, 计算机科学系, 广东 广州
Email: guesskkk99@163.com, 947440047@qq.com, 695470223@qq.com, 2498597150@qq.com

收稿日期: 2019年7月30日; 录用日期: 2019年8月19日; 发布日期: 2019年8月26日

摘要

自动图像标注是机器视觉领域中的一项具有挑战性的课题。大多数传统方法聚集在视觉图像与标注概念

间的建模,而忽略了语义概念之间的关系。本文提出了一种新的标注方法IA-SLN,它能有效捕获语义概念关系从而提升图像标注性能。首先,构建了一个基于全局关联语义与局部关联语义学习的语义链接网。其次,利用它标注未标注图像。当一个标注概念与其关联概念们频繁共现于图像中,IA-SLN方法将提升该标注概念的语义预测值。最后,通过一个标注增强新策略获得未标注图像的最相关语义标注。在IAPR公共数据集上对比了其他方法,实验结果表明我们的图像标注方法IA-SLN性能更优。

关键词

语义链接网, 语义估算, Kolmogorov复杂性, 图像自动标注

Copyright © 2019 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

图像的自动语义标注是多媒体内容智能管理的重要环节。作为一种解决方案,基于概念分类器的图像自动标注方法被提出来了,其识别的语义标注概念可能包括自然(例如,地平线、倒影和瀑布)或者场景(例如,城市、非洲和港口),它们皆来自于一个语义概念单词本。很多标注方法[1]基于一种假设:标注概念是彼此独立无关的。然而,语义概念并非孤立存在。相反地,它们是高度相关的,或者说,在这些语义概念之间存在着语义关联。例如,单概念“树”和“山”往往与“天空”共同出现在图像中,表明它们之间是有强语义依赖的。反过来,三个单概念“海洋”、“桌子”与“沙漠”则没有什么语义依赖关系,它们不太可能同时出现在一副图像中。除此之外,这种语义依赖还可以延伸到场景多概念范围。例如,对于以城市为主题的图像库,一个单概念“城市”与一个场景多概念“街道,房子”很可能频繁地共现在图像中。可以考虑利用这些语义关联去更准确地推断图像的语义标注,尤其是当一副图像的底层视觉特征不足以判断出正确的语义标注的时候。

语义概念及其相互依赖关系可以被组织起来构成一个语义链接网,它由语义节点和带权重边组成。不同于已有的语义网,提出的语义链接网的每一个节点划分为两种类型:语义单概念(如“鲜花”)与场景多概念(如“网球,球场,球手”),它们皆抽取于图像集。每一个场景多概念节点被视为一个整体,即一个整体场景概念,它描绘了一副图像的场景或者某个语义主题。带权重边代表语义概念之间的语义依赖关系,含有较强语义依赖关系的节点边则具有较高的权重。对于上述这两类语义节点,分别使用单概念分类器和多概念分类器去学习它们。为了估算语义边的权重,采用了两个数据资源:全局图像语料库与本地图像语料库。使用这两类语料库去抽取语义是基于如下的考虑。

使用全局语料库可以抽取通用语义,具备一定的准确性,然而这种全局语义的抽取仅仅考虑了全局语料库,其是独立于本地图像库的外来异质语料库,并未完全反映本地图像集的独有特性,这可能会影响图像自动标注的准确度。另一方面,对于带有大单词表 $V = \{c_1, \dots, c_i, \dots, c_L\}$ 的本地图像语料库,其语义概念 c_i 的分布很有可能是高度不平衡的。假若仅以本地图像语料库为数据源,抽取的局部语义依赖很可能不能反映真实的语义关系,极有可能会产生过度拟合的现象,这会损失图像语义标注的鲁棒性。因此,仅考虑单个数据源的语义抽取方法具有一些局限性,而结合全局语义与局部语义的方法或许能够引起这种局限的弱化。

借助于语义链接网完成图像的语义标注,主要基于这样几个考虑。1) 在包含大单词表的图像库里,

语义概念的不平衡性极有可能出现,传统的图像标注方法在频繁概念上性能较好,而在稀疏概念上的表现不尽人意。通过语义链接网技术缓解语义概念的不平衡性问题。2) 当图像视觉特征难以辨认的时候,分类器映射出的语义结果很可能会偏离正确的语义概念。例如,考虑抽象场景概念“城市”,相较于有形概念如“桌子”、“房子”、“路灯”和“街道”,它比较难以被语义概念分类器识别。一副包含有“城市”场景概念的图像,极有可能也出现它的依赖概念如“街道”、“街道,房子”和“房子,路灯”,可以利用这种概念依赖关系来推断图像的语义标注。对于一副图像 I ,假若发现了这些易识别的依赖概念,则可以增大图像 I 中存在“城市”标注概念的概率,于是,减少了误判语义概念“城市”的几率。

再考虑两个具有相近视觉外观的语义概念,如“天空”和“海洋”,因为二者具有近似的蓝色形状,概念分类器常常误判,难以分辨它们俩。由于语义概念“树”和“山”常常与“天空”共现于同一幅图像中,因此,可以借助于这种语义依赖去进一步区分辨识正确的语义标注。假如语义概念分类器有很强的视觉证据:图片 I 含有语义标注“树”和“山”,则 I 包含“天空”的概率应该大于“海洋”概率。

我们的语义链接网是一种语义数据模型,但又不同于传统的语义模型如本体模型:1) 本体是领域概念、概念属性、属性约束等知识描述模型,其独立于待标注本地图像库且并未完全反映本地图像的特性,所以可能影响图像语义标注的性能。2) 传统的语义模型仅仅考虑单概念节点,而适于标注的场景多概念被忽视了。我们的语义链接网既考虑了传统的单概念节点,也考虑了场景多概念节点。一个多概念节点被看作一个整体参与机器学习。3) 已有的语义数据模型仅考虑了单一语义信息源,而我们的语义链接网结合了全局语义与局部语义,弱化了其局限性。4) 我们的语义链接网可看作是一种特殊的轻量级本体语义模型,它包含两类节点即单概念节点与场景多概念节点以及含权语义边,移除了一些复杂而无用的通用本体特性如推理机制。总之,它是专用于图像标注的轻语义模型。

2. 基于局部语义与全局语义学习的语义链接网

假设 $A = \{I_1, \dots, I_T\}$ 与 $B = \{I_1, \dots, I_S\}$ 分别是含有 T 个样例的训练图像集与含有 S 个样例的测试图像集。 $V = \{c_1, \dots, c_l, \dots, c_L\}$ 表示单词表,包含 L 个语义概念,每个概念 $c_i \in V$ 是一个单概念,例如“stadium”。在训练集 A 中的每个图像被赋予几个语义单概念 c_i ,而在测试集 B 中的图像没有标记任何语义标注。每个语义多概念 $C_i = \{c_1, \dots, c_l\}$ (例如,“tennis, court, player”)是幂集 V 中的一个元素,即 $C_i \in 2^V$ 或者 $C_i \subseteq V$, l 是 C_i 的长度。如果 $l = 1$,那么 C_i 就退化成一个传统的语义单概念 c_i 。

2.1. 语义链接网的构建

语义概念及其相互依赖被组织起来形成一个语义链接网(如图 1),可以离线训练,由语义节点与含权边构成。每个语义节点表示为一个语义单概念 $c_i \in V$ 或者语义多概念 $C_i \in 2^V$, C_i 被当作一个整体或者一个场景概念。节点之间以语义边连接,表示两概念之间的语义依赖,边权重 $w_i = p(c|C_i)$ 可以基于全局语料库与本地语料库学习。自连接概率值设置为 1,即 $p(c|c) = 1$ 。为避免产生无意义的场景多概念 C_i ,如:“tennis, waterfall, ship”,在训练集 A 上通过如下共现规则产生出有语境含义的场景多概念 C_i ,于是生成一个语义节点集 V^* :

$$|C_i| \leq u \quad (1)$$

$$N_{\#}(c_1, \dots, c_l) \geq v \quad (2)$$

其中 $N_{\#}(c_1, \dots, c_l)$ 表示所有 l 个语义概念 c_i 在训练集 A 中的共现总数,即多概念频率。 $u \geq 1$ 用来控制多概念 C_i 的最大长度($1 \leq l \leq u$)。假若节点集 V^* 太大,可以调节公式(2)中的共现数 v 以减小计算开销。于是,语义节点集 V^* 生成了。

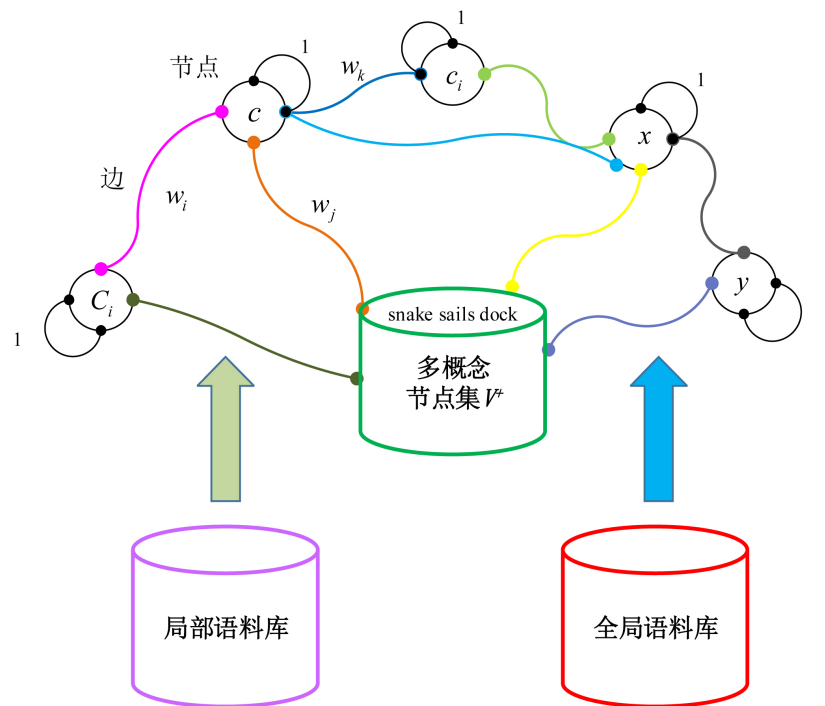


Figure 1. The system framework of semantic link network
图 1. 语义链接网的系统框架

语义边用以连接语义节点，边权重表示两概念节点 x 与 y 的语义依赖概率 $p(x|y)$ ，其值介于 0 (完全不相关)与 1 (完全相等)之间。语义依赖值是对称的，所以该语义链接网是无向语义网络。

2.2. 边权重的估算

使用 Web 图像语料库作为全局语料库，使用图像训练集 A 作为局部语料库，综合利用(全局) Web 图像语料库与(局部)训练图像语料库来计算边权重 w_i ，即概念依赖概率 $p(c|C_i)$ 。边权重采用如下两个语义度量方法计算。

1) Google 距离

在 Web 级图像语料库上考虑使用 Google 距离[2]，它是一种新的基于信息距离与 Kolmogorov 复杂性[3]的语义相关性度量，且使用了由 Google 搜索引擎返回的 Web 点击页面数来计算两语义概念之间的语义距离。给出两个语义概念 x 与 y ，例如 $x = \text{"city"}$ 和 $y = \text{"street, building"}$ ，相应的 Google 距离 $d(x,y)$ 计算可以参见文献[2]。于是，语义依赖概率 $p(x|y)$ 可以按照公式(3)以 $d(x,y)$ 的负幂来计算出：

$$p(x|y) = \exp(-10 * d(x,y)). \quad (3)$$

可使用任意的全局语料库计算边权重，不失一般性，使用 Google 图像搜索引擎索引的海量 Web 数据作为全局语料库，按照公式(3)去估算全局语义依赖概率 $p_G(c|C_i)$ (即 $x = c$ 和 $y = C_i$)。

2) 共现距离

在训练图像语料库上考虑使用对称共现距离来计算局部语义依赖概率 $p_L(c|C_i)$ (即 $x = c$ 和 $y = C_i$)，可以按照如下对称共现度量方法计算：

$$p(x|y) = \frac{2 * f(x,y)}{f(x) + f(y)}, \quad (4)$$

该方法描述了概念 y 与概念 x 共现的频率，然后被语义概念 x 与 y 的总频率归一化。它可以被理解为给定含有标注 y 的图像 I ，其包含语义概念 x 的概率有多大，其值介于 0.0 和 1.0 范围之内。最后，如下定义语义依赖概率 $p(c|C_i)$ ：

$$p(c|C_i) = \lambda_1 p_G(c|C_i) + \lambda_2 p_L(c|C_i), \quad (5)$$

其中 $\lambda_1 \geq 0$ 和 $\lambda_2 \geq 0$ 是待估算的模型参数且服从约束条件： $\lambda_1 + \lambda_2 = 1$ 。该参数可以用来调节全局语义依赖与局部语义依赖的合适比例，语义自依赖概率设置为 1，即 $p(c|c) = 1$ 。很显然，上述局部语义依赖概率 $p(c|C_i)$ 是对称的，即 $p(C_i|c) = p(c|C_i)$ 。

3. 基于语义链接网的图像标注方法

IA-SLN 考虑了两类线索：语义线索和视觉线索，包含三个步骤：标注上下文生成、语义映射和语义标注增强方法。为了估算语义概念 c 与未标注图像 I 的相关性 $r(c, I)$ ，图中 I 被表示为低层视觉特征和标注上下文 $R(c)$ 。集合 $R(c)$ 中的每一个语义节点 C_i 被表示为一个圆圈(如图 1)，它可以是单概念节点或者场景多概念节点。概念分类器执行从未标注图像 I 到语义概念 C_i 的映射，并且输出图像 I 含有语义概念 C_i 的带权视觉证据 $p(C_i|I)$ 。最后由概念标注增强方法方法精炼语义标注。

3.1. 标注上下文生成

给定任意标注概念 $c \in V$ ，从语义链接网中按照边权重 $w_i = p(c|C_i)$ 选择 K 个最近邻语义节点 C_i 以产生语义标注上下文 $R(c) \subset V^+$ 。其中每个元素 $C_i \in R(c)$ 可以是一个单概念或者一个场景多概念，很显然，因为 $p(c|c) = 1$ ，所以每个语义概念 c 与其本身是相关的(即 $c \in R(c)$)。

为了保持概率的属性，标注上下文 $R(c)$ 中的所有元素值 $p(c|C_i)$ 应该执行如下所示的规范化操作，使得元素和为 1。

$$p(c|C_i) = \begin{cases} \frac{p(c|C_i)}{\sum_{r=1}^K p(c|C_r)} & \text{if } C_i, C_r \in R(c) \\ 0 & \text{elsewhere.} \end{cases} \quad (6)$$

3.2. 语义标注映射

标注映射的任务是计算标注概念 c 与测试图像 I 的相关性 $r(c, I)$ 。考虑由 K 个标注上下文元素 $C_i \in R(c)$ 共同参与相关性计算，获得如下相关性估值 $r(c, I)$ ：

$$r(c, I) = \sum_{i=1}^K w_i v_i = \sum_{i=1}^K p(c|C_i) p(C_i|I), \quad (7)$$

该后验概率 $p(C_i|I)$ 指出了图像 I 与语义多概念 C_i 的相关性，它可以视为图像 I 包含有语义概念 c 的视觉证据 v_i ，而两语义概念 c 和 C_i 之间的概念依赖 $p(c|C_i)$ 可视为视觉证据 v_i 的权重 w_i ，高依赖关系的语义概念具有较高权重。对于给定的标注概念 $c \in R(c)$ (例如“city”)，假若有很强的视觉证据 $p(C_i|I)$ 表明 I 包含有该概念 c 及其多个强依赖概念 $C_i \in R(c)$ (例如“street, building”和“building, light”等)，则 $p(c|C_i) p(C_i|I)$ 项变得很大，因此相关性 $r(c, I)$ 也变大。反之，则 $r(c, I)$ 会变得很小。

概率 $p(C_i|I)$ 可以通过分类器来估算。由于近年来卷积神经网络在图像分类领域取得了巨大成功[4] [5]，因此 IA-SLN 选用了一种高效的卷积神经网络 DenseNet [5]。假若 $|C_i| = 1$ (即传统的单概念 c_i)，使用传统的单概念分类器训练方法：把标记有单概念 c_i 的样例作为正例，而剩余图像样例作为反例。假若 $|C_i| > 1$ (即一个语义场景多概念 C_i)，正例集 $S^+(C_i)$ 与反例集 $S^-(C_i)$ 可以如下方法构建：

$$\begin{aligned} S^+(C_i) &= \{I_i | C_i \subseteq A_s(I_i)\}, \\ S^-(C_i) &= \{I_i | C_i \not\subseteq S^+(C_i)\}, \end{aligned} \quad (8)$$

其中 $A_s(I_i)$ 是训练图像 $I_i \in A$ 的标注概念集。基于上述两个集合 $S^+(C_i)$ 和 $S^-(C_i)$ ，训练出分类器作为多概念场景检测器。

为了找到最合适的模型参数值 $\Delta = (\lambda_1, \lambda_2)$ ，在训练数据上最大化概念预测值的对数似然函数。按照公式(3)与(5)，相关性估值 $r(c, I)$ 可以表达为如下形式：

$$r(c, I) = \sum_{i=1}^K [\lambda_1 p_G(c | C_i) + \lambda_2 p_L(c | C_i)] \times p(C_i | I), \quad (9)$$

$$r(c, I) = \lambda_1 \sum_{i=1}^K p_G(c | C_i) p(C_i | I) + \lambda_2 \sum_{i=1}^K p_L(c | C_i) p(C_i | I), \quad (10)$$

IA-SLN 使用 $y_{ci} \in \{0, 1\}$ 表示概念 c 不存在/存在于训练图像 I 中，因此概念 c 的预测值 $p(y_{ci})$ 可以如下形式给出：

$$\begin{aligned} p(y_{ci} = 1) &= r(c, I), \\ p(y_{ci} = 0) &= 1 - r(c, I), \\ p(y_{ci}) &= r(c, I)^{y_{ci}} (1 - r(c, I))^{1 - y_{ci}}, \end{aligned} \quad (11)$$

于是，语义概念 c 的对数似然函数改写为如下形式：

$$J_c = \sum_{i=1}^T n_{ci} \log p(y_{ci}), \quad (12)$$

其中 n_{ci} 是代价因子，它考虑了每一个概念 c 的正例数 N^+ 与反例数 N^- 的不平衡性。如果 $y_{ci} = 1$ 则设定 $n_{ci} = 1/N^+$ ，否则设定 $n_{ci} = 1/N^-$ 。将公式(10)和(11)带入(12)，得到了下列对数似然函数：

$$J_c = \sum_{i=1}^T n_{ci} \log [(\Delta\Theta)^{y_{ci}} (1 - \Delta\Theta)^{1 - y_{ci}}], \quad (13)$$

其中 $\Theta = [\sum_{i=1}^K p_G(c | C_i) p(C_i | I), \sum_{i=1}^K p_L(c | C_i) p(C_i | I)]^T$ 。对于每个对数似然函数 $J_c (c \in V)$ 分别采用梯度下降法[6]进行最大化求解。

3.3. 标注增强及算法

为了获得图像 I 的最佳语义标注 c ，需要对所有概念 $c \in V$ 的语义相关性分数 $r(c, I)$ 从大到小排序，从中选择出最好的 w 个语义概念 c 作为候选标注集 Y 。此外，一个图像的标注概念数量 w 并非是固定值，换句话说，一些图像的语义标注概念数可能小于 w 个。由于候选概念集 Y 可能包含有不相关的噪音标签，为了提升标注效果，IA-SLN 基于阈值 t_i 过滤掉一些噪音候选标签。给出图像 I ，选择最高 $\alpha\%$ 个候选概念相关分数的算术均值作为该标注阈值 t_i 。因为单词表中相关的概念标注总是远远少于不相关的语义概念，所以 α 被赋予一个较小的值，例如 5, 6 或者 7。

基于上述方法，如下给出 IA-SLN 自动图像标注算法。

输入：训练集 A ，标注字典 V ，测试图像 $I \in B$ ；

输出：标注结果集 $O = \{c_1, c_2, \dots, c_k\}$ 。

- 1 构建多概念节点集 V^+ ；
- 2 计算任意两节点之间的边权重 w_i ，生成语义链接网；
- 3 for each $c \in V$

- 4 构建标注上下文 $R(c)$;
- 5 映射语义标注, 获得相关性估值 $r(c, I)$;
- 6 end for
- 7 执行标注增强方法;
- 8 返回标注结果集 O 。

4. 实验和评价

4.1. 数据集

IAPR 数据集[7]包含有 19,627 副自然风景图像, 包括人物、动物、城市、地标以及现代生活的方方面面, 其语义标注更接近真实场景, 每幅图像含有 1~23 个语义标注, 单词表 V 含有 291 个语义概念(即 $|V|=291$)。IAPR 图像集被随机分割为两部分: 17,665 副图像作为带标注的训练集且余下图像作为未标注测试集。

4.2. 实验结果与分析

考虑使用四个标准性能指标来评估图像标注方法, 即文献[7]使用的平均准确率、平均召回率、F1 分数以及 N^+ 值, 指标值越高表示标注性能越好。测试图像 I 的最相关 w 个语义概念的选择, 受到了已有经典工作[7]的启发。按照这个思路, 初步考虑 w 的近似值是 5 或者 6。由于在单词表中, 相关标注概念数远远小于不相关的概念数, 所以考虑 α 赋予一个较小的值, 例如 5, 6 或者 7。关于参数 (w, α) 的选择也得到交叉验证实验的证实。随机地把整个训练集划分为 10 份, 一份做验证集而剩余 9 份做训练集, 基于该验证集划分, 在参数 $w = [5, 6]$ 与 $\alpha = [3, 4, 5, 6, 7, 8]$ 上做了一组交叉验证实验。设置 $w = 6$, $\alpha = 7$, 因为它们表现出最好的 F1 性能。为确定语义标注上下文 $R(c)$ 的大小 K , 即语义链接网节点 c 的近邻节点数, 再次执行 10-路交叉验证实验。发现当 K 值到达 20 时获得最佳的 F1 分数, 因此设置 $K = 20$ 。

Table 1. Performance comparisons of automatic image annotation

表 1. 自动图像标注性能比较

标注方法	平均准确率	平均召回率	F1 分数	N^+ 值
SKL-CRM [8]	0.47	0.32	0.381	274
FastTag [9]	0.47	0.26	0.335	280
KSVM-VT [10]	0.47	0.29	0.359	268
RNN [11]	0.33	0.31	0.320	255
DLD-FCM [1]	0.34	0.30	0.320	251
CNN-MLL [12]	0.49	0.42	0.45	-
CNN-MLSU [13]	0.44	0.38	0.41	-
CNN-LDA [14]	0.44	0.40	0.42	-
GAN-W [15]	0.55	0.34	0.42	-
IA-SLN (Ours)	0.55	0.44	0.489	281

表 1 列出了与最新图像标注方法的对比实验结果。提出的 IA-SLN 方法超越了所有比较方法, 相较于两种较好的已有标注方法即 CNN-MLL 和 CNN-LDA 方法, IA-SLN 方法分别获得了 12.2%或 25.0%, 4.8%或 10.0%, 8.6%或 16.4%的性能提升(平均准确率, 平均召回率, F1 值)。IA-SLN 不仅考虑了单概念,

而且考虑了场景多概念以及结合全局语义与局部语义的边权重学习。全局语义侧重于语义概念的通用属性，而局部语义侧重于本地特定的图像属性，这两种类型的语义具有一定的互补性，因此，它们的结合可以较好地提高图像语义标注的效果。

参考文献

- [1] 李长磊, 吕学强, 张凯, 董志安. 一种改进模糊 C 均值聚类的图像标注方法[J]. 小型微型计算机系统, 2018, 39(8): 1860-1864.
- [2] 杨惠荣, 尹宝才, 付鹏斌, 曲亮. 基于 Google 距离的语义 Web 服务发现[J]. 北京工业大学学报, 2012, 38(11): 1670-1675.
- [3] 王有华, 陈笑蓉. 基于 Kolmogorov 复杂性的文本聚类算法改进[J]. 计算机科学, 2016, 43(5): 243-246.
- [4] 周非, 李阳, 范馨月. 图像分类卷积神经网络的反馈损失计算方法改进[J]. 小型微型计算机系统, 2019, 40(7): 1532-1537.
- [5] Huang, G., Liu, Z., Laurens, V.D.M. and Weinberger, K.Q. (2017) Densely Connected Convolutional Networks. *IEEE Conference on Computer Vision and Pattern Recognition*, Honolulu, 21-26 July 2017, 2261-2269. <https://doi.org/10.1109/CVPR.2017.243>
- [6] 常永虎, 李虎阳. 基于梯度的优化算法研究[J]. 现代计算机, 2019(17): 3-8+15.
- [7] Makadia, A., Pavlovic, V. and Kumar, S. (2010) Baselines for Image Annotation. *International Journal of Computer Vision*, **90**, 88-105. <https://doi.org/10.1007/s11263-010-0338-6>
- [8] Moran, S. and Lavrenko, V. (2014) Sparse Kernel Learning for Image Annotation. *International Conference on Multimedia Retrieval*, Glasgow, 1-4 April 2014, 113-120. <https://doi.org/10.1145/2578726.2578734>
- [9] Chen, M., Zheng, A. and Weinberger, K.Q. (2013) Fast Image Tagging. *International Conference on International Conference on Machine Learning*, Atlanta, 16-21 June 2013, 1274-1282.
- [10] Verma, Y. and Jawahar, C. (2013) Exploring SVM for Image Annotation in Presence of Confusing Labels. *British Machine Vision Conference*, Bristol, 9-13 September 2013, 1-25. <https://doi.org/10.5244/C.27.25>
- [11] Cui, C., Ma, J. and Lian, T. (2013) Ranking-Oriented Nearest-Neighbor Based Method for Automatic Image Annotation. *ACM SIGIR Conference on Research and Development in Information Retrieval*, Dublin, 28 July-1 August 2013, 957-960. <https://doi.org/10.1145/2484028.2484113>
- [12] 高耀东, 侯凌燕, 杨大利. 基于多标签学习的卷积神经网络的图像标注方法[J]. 计算机应用, 2017, 37(1): 228-232.
- [13] 汪鹏, 张奥帆, 王利琴, 董永峰. 基于迁移学习与多标签平滑策略的图像自动标注[J]. 计算机应用, 2018, 38(11): 3199-3203+3210.
- [14] 张蕾, 蔡明. 融合卷积神经网络与主题模型的图像标注[J]. 激光与光电子学进展, 2019, 56(20): 1-12.
- [15] 税留成, 刘卫忠, 冯卓明. 基于生成式对抗网络的图像自动标注[J]. 计算机应用, 2019, 39(7): 2129-2133.

知网检索的两种方式:

1. 打开知网首页: <http://cnki.net/>, 点击页面中“外文资源总库 CNKI SCHOLAR”, 跳转至: <http://scholar.cnki.net/new>, 搜索框内直接输入文章标题, 即可查询;
或点击“高级检索”, 下拉列表框选择: [ISSN], 输入期刊 ISSN: 2326-3415, 即可查询。
2. 通过知网首页 <http://cnki.net/>顶部“旧版入口”进入知网旧版: <http://www.cnki.net/old/>, 左侧选择“国际文献总库”进入, 搜索框直接输入文章标题, 即可查询。

投稿请点击: <http://www.hanspub.org/Submission.aspx>

期刊邮箱: airr@hanspub.org