

基于二次聚类弱监督学习的图像语义分割<sup>\*</sup>

文笃石

(西安邮电大学计算机学院 西安 710121)

**摘 要:**针对在给定大量用户标注图像基础上,将图像分割成一系列具有单一语义的完整区域,并同时对各区域实现语义标注的问题,提出了一种基于弱监督学习的二次聚类的图像语义分割算法。将谱聚类和判别式聚类相结合,用谱聚类学到的类标指示函数来指导判别式聚类,学习特征的潜在数据结构,利用弱监督信息对聚类分配标签。该方法能够充分利用区域上下文信息,为每个类别选择判别式特征,并且输出鲁棒的多类分类器,对外来没有标签的图像也可以进行有效的区域标注。通过在公用数据集上的充分实验,证明了本方法的有效性。

**关键词:**弱监督学习;谱聚类;判别式聚类

**中图分类号:** TP301    **文献标识码:** A    **国家标准学科分类代码:** 520.1040

## Image semantic segmentation based on weakly supervised learning of two clustering

Wen Dushi

(School of Computer Science and Technology, Xi'an University of Posts and Telecommunications, Xi'an 710121, China)

**Abstract:** According to a large number of users in a given image annotation based on image is segmented into a series of single semantic integrity area, and at the same time to the area to realize the semantic annotation problem, proposes a segmentation algorithm of semantic image two clustering based on weakly supervised learning. The combination of spectral clustering and discriminant clustering is used to guide the discriminant clustering and the latent data structure of learning features. The proposed method can make full use of the regional context information, select discriminative features for each class, and output robust multi class classifier. The effectiveness of the proposed method is demonstrated by experiments on a common data set.

**Keywords:** weakly supervised learning; spectral clustering; discriminant clustering

## 0 引 言

随着互联网进入 Web2.0 时代、数字移动设备的普及,网络图像资源呈现爆炸式的增长,给图像检索带来了巨大挑战。图像语义分割技术能够深层次地理解图像语义内容,把语言标签映射到图像中相应的区域上,能够实现更准确的图像理解效果。虽然有很多文献已经对图像语义分割提出了不同的理论和算法,但是大部分机器学习方法都依赖于精确标注到像素的训练图像。在大数据时代,图像数据迅速增长,这种手工标注图像到像素级别非常耗时耗力。同时这些网络图像数据不仅视觉上变化多端,而且语义上也分散各异,标注结果因标注人员主观理解习惯而不同。另外,对于精确标注的训练集过分依赖,

训练出来的模型很可能存在过拟合现象,缺乏扩展性,不能面向大规模的数据应用。目前,图像分享网站如 Facebook、Picasa、Flickr 等开始涌现,并且有大量的用户标注图片可以提供下载,这为图像语义分割提供了一个新的发展思路,即是否能够利用这些丰富的社会标签(用户对图片添加的分类标签)来学习图像区域和语言之间的映射,这种弱监督(weakly-supervised)方式已经为解决图像语义分割问题展示了惊人的潜力。本文针对弱监督图像语义分割分析,提出了一种基于弱监督的聚类算法。为了同时最大化同一聚类内的视觉一致性和不同聚类之间的可分性,将谱聚类和判别式聚类相结合。用谱聚类学到的类标指示函数来指导判别式聚类,学习特征的潜在结构,并为每个类别选择出具有判别力的特征。利用图像级别标

收稿日期:2017-04

<sup>\*</sup> 基金项目:陕西省教育厅科研计划项目(15JK1679)、陕西省科技统筹创新工程计划项目(2015KTCQ01-14)资助

签作为弱监督信息,为每个聚类分配语义标签。本方法能够充分挖掘区域上下文信息并且进行有效的特征选择,还能输出鲁棒的多类分类器,对没有标签的测试图像也可以进行有效的语义解析。

## 1 弱监督图像语义分割

在计算机视觉领域,图像语义解析通常被称作语义分割。在多媒体领域,常被叫做标注到区域。图像语义分割试图使计算机能够自动地给图像区域分配语义标签,是计算机视觉领域的一个基本任务。图像语义分割通常涉及3个方面:图像分割、感知聚类、物体识别。因此,图像语义分割是将这3个问题在一个框架中联合求解的过程。

根据监督信息设置的不同,图像语义分割方法大致可以分为3类:全监督、半监督、弱监督方法。全监督算法需要人工准备图像及相应的区域标签,用于训练模型。其优点是可从训练样本提取模型或者知识,往往有较高性能。然而像素级别的标注需要耗费大量的人力和时间,并且模型难以扩展,同时优化和推断也存在困难。为了能够在模型精度和上述问题中折中,半监督方法应运而生,但是半监督的方法还是不能从根本上摆脱对精细标注数据的依赖。较早期的图像语义分割方法大部分着重于底层特征设计和模型设计,出于精度的考虑,大部分是全监督或者半监督方法。近年来,随着网络的发展和数码便携设备的增加,计算机视觉的任务变得更加复杂。如果还是沿用传统的手工标注方法准备训练数据,已经不能满足任务的需要。因此就需要另辟蹊径,寻找新的方法,用尽可能少的人力资源获得更多有用数据并加以有效利用。弱监督方法仅仅需要图像级别的标注数据进行学习,就能够把标签标注到相应的图像区域上去。即在训练数据中,每个样本都仅仅被标注了一个候选标签集合,在这个标签集合中只有一个标签是这个样本的正确标签。

基于弱监督的图像语义分割主要包括3个部分:图像预处理、弱监督学习算法设计、算法性能评价。图像预处理主要包括两个步骤:图像分割和特征提取。图像分割<sup>[1]</sup>是从输入图像中提取重要对象的过程,或将图像分解为包含相似像素部分和区域的过程。大多数图像分割算法基于像素灰度级的两个基本特征:边界上的不连续性和区域内的相似性。图像分割中比较经典的算法有数据驱动的马尔科夫蒙特卡洛<sup>[2]</sup>(DDMCMC)算法、Mean Shift<sup>[3]</sup>算法、Normalized Cuts<sup>[4]</sup>算法等。其中,DDMCMC具有较高的性能,但计算开销相对比较大。而Mean Shift和Normalized Cuts算法易于实现,也能达到一定精度,但是容易把语义区域分成过于细小的区域。图像特征提取是视觉信息分析的必要条件和关键环节,提取的特征是否能够有效描述目标区域和场景并具有明显的区分性对图像理解十分重要。根据通用的视觉特征描述图像共有的特征,通常分为全局特征和局部特征两类。全局特征反映的是图像的整体特征,无法反映图像中所包含的对象或者物体,

如颜色、纹理、形状等。局部特征使用能够反映图像内容的显著点描述图像,能灵活描述图像的局部信息和细节内容。显著点检测算子有Harris算子<sup>[5]</sup>、高斯拉布拉斯(Laplacian of Gaussian)算子<sup>[6]</sup>、SURF(speeded up robust features)<sup>[7]</sup>和SIFT(scale-invariant feature transform)<sup>[8]</sup>算子等。目前弱监督学习大部分使用了Expectation-maximization(EM)算法来评估模型的参数和示例的真实类标。

## 2 二次聚类弱监督图像语义分割

### 2.1 方法

图像语义分割的目的是将图像划分为多个语义上相互独立的区域,并进一步预测区域所属的类别,它的本质是聚类问题,即把包括亮度、颜色和纹理等特征相似的图像块聚在一起,并将正确的标签赋予到这些图像块,使得类内的图像像素在语义上一致,空间上相连。在本次实验中,采用了基于弱监督二次聚类(weakly-supervised twice clustering for image semantic segmentation, WSTC)的方法,将数据集的所有子区域进行聚类,使得属于同一语义类的子区域被聚到一类。同时用图像级别标签作为弱监督信息,为聚类分配标签。如图1所示。

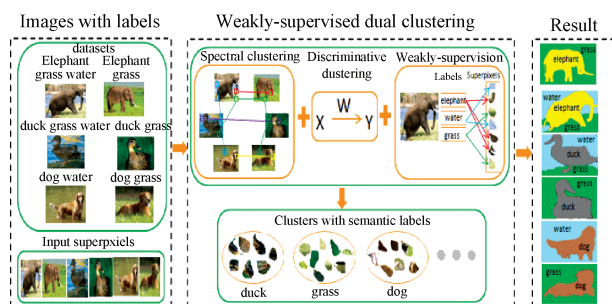


图1 弱监督二次聚类的图像语义分割方法

### 2.2 公式标记

假设训练集是一个包含  $I$  幅图像的数据集  $\mathbf{X} = [\mathbf{X}_1, \dots, \mathbf{X}_i, \dots, \mathbf{X}_I]$ 。其中,  $\mathbf{X}_i = [x_i^1, \dots, x_i^{n_i}]$ ,  $x_i^k \in R^d$  是第  $i$  幅图像中第  $k$  个子区域的特征表示,  $d$  为特征的维数,  $n_i$  代表第  $i$  幅图像中子区域的个数。将数据矩阵  $\mathbf{X}$  简写为  $\mathbf{X} = [x_1, \dots, x_i, \dots, x_N]$ , 其中  $N = \sum_{i=1}^I n_i$  是数据集中所有子区域的总和。假设数据集中出现的标签为  $C$  个,  $\mathbf{G} = [g_1, \dots, g_i, \dots, g_I] \in \{0, 1\}^{C \times I}$  代表与图像相对应的标签信息矩阵,  $g_i \in \{0, 1\}^{C \times I}$  是第  $i$  幅图像的标签向量。如果第  $i$  幅图像被打上了标签  $c$ , 则  $g_i^c = 1$ , 否则为 0。定义子区域类标矩阵  $\mathbf{Y}$ :

$$y_n^c = \begin{cases} 1, & \text{如果第 } n \text{ 个子区域属于第 } c \text{ 类} \\ 0, & \text{如果第 } n \text{ 个子区域不属于第 } c \text{ 类} \end{cases} \quad (1)$$

直接推断二值的类标是非常困难的,因此在实际求解中将类标松弛为  $[0, 1]$  的变量。

### 2.3 谱聚类

谱聚类是一种基于图论的聚类方法,其基本思想是利用样本数据的相似矩阵(拉普拉斯矩阵)进行特征分解后得到的特征向量进行聚类。在本次实验中,采用谱聚类挖掘子区域之间的这种一致性上下文关系。首先需要说明如何构建谱聚类所需要的相似度图。图的结点是过分割的图像区域,图的边的权重根据图像区域的一致性关系得到。视觉相似的子区域对和空间近邻的子区域对都应该有边相连。子区域候选标签是它所在的图像标签,两个候选标签集合没有交集的子区域是不可能具有相同类标的,它们之间不应该有相连边。因此,仅有那些所属图像具有共同标签的子区域才被计算相似度。子区域之间的相似度矩阵  $S$  定义如下:

$$S_{ij} = \begin{cases} \exp\left(-\frac{\|x_i - x_j\|^2}{\sigma^2}\right) \\ 0 \quad \text{otherwise.} \end{cases} \quad (2)$$

$x_i \in N_k(x_j) \text{ or } x_j \in N_k(x_i)$

式中:  $N_k(x)$  表示子区域  $x$  的  $k$  近邻,  $k$  近邻是从满足上述条件的子区域中产生,  $\sigma$  是参数。为了保持标注的空间平滑性,还将与子区域  $x$  在同一幅图像中空间近邻的子区域与  $x$  相连。计算出相似度矩阵  $S$  后,谱聚类的目标函数为最小化下列公式:

$$I(Y) = \frac{1}{2} \sum_{i,j=1}^N S_{ij} \left\| \frac{y_i}{\sqrt{A_{ii}}} - \frac{y_j}{\sqrt{A_{jj}}} \right\|_2^2 = \text{Tr}[Y^T L Y] \quad (3)$$

式中:  $y_i$  和  $y_j$  分别为数据集中第  $i$  个子区域和第  $j$  个子区域的标签向量,即  $Y$  中的第  $i$  行和第  $j$  行。上式中,  $A$  是一个对角矩阵,  $A_{ii} = \sum_{j=1}^N S_{ij}$ ,  $L = A^{-1/2}(A - S)A^{-1/2}$ , 是归一化的拉普拉斯矩阵。  $\text{Tr}(\cdot)$  是矩阵积的运算。

### 2.4 判别式聚类

采用线性分类器作为一种判别式的聚类方法,使其与谱聚类方法共享统一的类标指示向量。考虑到不同标签具有不同的判别性特征,对分类器施以  $l_{2,1}$  范数约束,起到稀疏特征选择的作用。假设特征和预测的标签之间存在线性变换  $W \in R^{d \times C}$ , 判别式聚类的损失为最小化下列目标函数

$$\min_{Y, W} \mathcal{L}(Y, W) = \alpha \sum_{i=1}^N \text{loss}(y_i, W^T x_i) + \beta \|W\|_{2,1} \quad (4)$$

式中:  $\text{loss}$  是一个损失函数,  $\alpha$  和  $\beta$  是两个非负参数。  $l_{2,1}$  范数定义为  $\|W\|_{2,1} = \sum_{i=1}^d \sqrt{\sum_{j=1}^C W_{ij}^2}$ , 这样的定义能够确保  $W$  的行稀疏,从而能够在有效地选择判别性特征的同时去除冗余和噪声特征。采用最小平方损失来定义损失函数,那么线性分类器的具体形式为:

$$\mathcal{L}(Y, W) = \alpha \|X^T W - Y\|_F^2 + \beta \|W\|_{2,1} \quad (5)$$

### 2.5 弱监督约束

弱监督信息是指图像级标签与图像子区域标签的对应约束关系,包括 3 个方面:

1) 如果图像具有某一个标签,那么该图像中一定有至少一个子区域能够解释这个标签;

2) 如果图像没有某个标签,图像中所有的图像子区域都不能解释这个标签;

3) 一个图像子区域只能对应一个标签。

为了满足最后一个条件,提出一个弱监督正则项,

$$Q(Y) = \gamma \sum_{i=1}^I \sum_{c=1}^C \left| \max_{x_i \in X_i} y_{ij}^c - g_i^c \right| \quad (6)$$

式中:  $\gamma$  是一个控制权重的参数,  $y_{ij}^c$  代表第  $i$  幅图像的第  $j$  个子区域在第  $c$  个标签的取值。直接处理上面带有绝对值符号的正则项是非常困难的,因此首先要去掉绝对值。对公式进行松弛优化,得到统一的目标函数:

$$\begin{aligned} \min_{Y, W} & \text{Tr}[Y^T L Y] + \alpha \|X^T W - Y\|_F^2 + \beta \|W\|_{2,1} + \\ & \gamma \sum_{i=1}^I \sum_{c=1}^C [(1 - g_i^c) h_c^T Y^T q_i + g_i^c (1 - \max_{x_i \in X_i} p_{ij}^T Y h_c)] \\ \text{s. t. } & Y^T Y = I_C, Y \geq 0. \end{aligned} \quad (7)$$

### 2.6 CCCP 算法

由于  $l_{2,1}$  范数项是非光滑的,  $\max$  项是非凸的,因此整个目标函数是非光滑且非凸的。在求解过程中,采用了凸凹过程(convex concave programming, CCCP)与非负乘子法的迭代更新对目标函数进行优化,求得问题的次优解。在每一轮的迭代过程中,通过凸凹过程,用一阶泰勒展开近似非凸函数,将原问题转化为一个凸的优化问题。引入指示向量  $B$  和  $U$  后将统一目标函数的最后一项进一步改写,非凸项被线性项所取代,那么可以用迭代的方法来求解问题了。将正交性约束写入统一目标函数,得到

$$\begin{aligned} \min_{Y, W} \mathcal{L}(Y, W) &= \text{Tr}(Y^T L Y) + \alpha \|X^T W - Y\|_F^2 + \beta \|W\|_{2,1} + \\ & \gamma \sum_{i=1}^I \sum_{c=1}^C [(1 - g_i^c) h_c^T Y^T q_i + g_i^c (1 - h_c^T B U_i Y h_c^T)] + \\ & \frac{\mu}{2} \|Y^T Y - I_C\|_F^2 \\ \text{s. t. } & Y \geq 0. \end{aligned} \quad (8)$$

式中:  $\mu \geq 0$  是控制正交性约束的参数。采用迭代优化过程来解这个问题: 固定  $W$ , 更新  $Y$ ; 固定  $Y$ , 更新  $W$ 。将目标函数对  $W$  求导为 0, 得到  $W$  的迭代公式:

$$\begin{aligned} \frac{\partial \mathcal{L}(Y, W)}{\partial W} &= 2(\alpha X(X^T W - Y) + \beta D W) = 0 \\ \Rightarrow W &= \alpha(\alpha X X^T + \beta D)^{-1} X Y \end{aligned} \quad (9)$$

然后再把  $W$  带入统一目标函数中,得到新的目标函数

$$\begin{aligned} \min_Y \mathcal{L} &= \text{Tr}[Y^T M Y] + \gamma \left[ \sum_i \sum_c (1 - g_i^c) h_c^T Y^T q_i + \right. \\ & \left. \sum_i \sum_c g_i^c (1 - h_c^T B U_i Y h_c^T) \right] + \frac{\mu}{2} \|Y^T Y - I_C\|_F^2 \\ \text{s. t. } & Y \geq 0. \end{aligned} \quad (10)$$

式中:  $M = L + \alpha(I_N - \alpha X X^T + \beta D)^{-1} X$ ,  $I_N \in R^{N \times N}$  是单位矩阵。对于新的目标函数,未知变量为  $Y$ ,引入非负乘子法进行求解。设  $\phi_{ij}$  是对于  $Y_{i,j} \geq 0$  的拉格朗日乘子,且  $\Phi = [\phi_{ij}]$ ,则拉格朗日函数形式为:

$$\mathcal{L} + f_r(\Phi Y^T)$$

令它对  $\mathbf{Y}$  求导数为 0, 得到:

$$2\mathbf{M}\mathbf{Y} + \mathbf{P} + 2\mu\mathbf{Y}\mathbf{Y}^T\mathbf{Y} - 2\mu\mathbf{Y} + \Phi = 0 \quad (11)$$

$$\mathbf{P} = \gamma \sum_{i=1}^I \sum_{c=1}^C [(1 - g_i^c) q_i h_c - g_i^c \mathbf{U}_i^T \mathbf{B}^T h_c^T h_c]$$

利用 KKT 条件令  $\phi_{ij} Y_{ij} = 0$ , 得到对  $\mathbf{Y}$  的更新公式为:

$$Y_{ij} \leftarrow Y_{ij} \frac{2(\mu Y)_{ij}}{(2\mathbf{M}\mathbf{Y} + \mathbf{P} + 2\mu\mathbf{Y}\mathbf{Y}^T\mathbf{Y})_{ij}} \quad (12)$$

最后, 将  $\mathbf{Y}$  归一化使得  $(\mathbf{Y}^T \mathbf{Y})_{ii} = 1, i = 1, \dots, C$ . 至此, 一轮迭代结束了。反复迭代优化凸的子问题直到收敛, 可以得到原始问题的次优解。整个算法的详细过程如表 1 所示。

表 1 弱监督二次聚类(WSTC)算法流程

输入: 数据矩阵 $\mathbf{X} \in \mathbf{R}^{d \times N}$ ;
标签矩阵 $\mathbf{G} \in \mathbf{R}^{C \times I}$ ;
参数 $\alpha, \beta, \gamma, \mu$ .
① 构建 $k$ 近邻图计算拉普拉斯矩阵 $\mathbf{L}$
② 设置迭代轮数 $t = 1$ ;
初始化 $\mathbf{Y} \in \mathbf{R}^{N \times C}$ ;
设置 $\mathbf{D}' \in \mathbf{R}^{d \times d}$ 是一个单位阵.
③ 循环
④ 根据 $\mathbf{W}$ 的迭代公式计算
$\mathbf{W}' = \alpha(\alpha \mathbf{X} \mathbf{X}^T + \beta \mathbf{D}')^{-1} \mathbf{X} \mathbf{Y}'$
$\mathbf{M}' = \mathbf{L} + \alpha(\mathbf{I}_N - \alpha \mathbf{X}^T (\alpha \mathbf{X} \mathbf{X}^T + \beta \mathbf{D}')^{-1} \mathbf{X})$
⑤ 计算 $\mathbf{B}'$
⑥ 根据 $\mathbf{P}$ 的公式计算 $\mathbf{P}'$
⑦ 根据 $\mathbf{Y}$ 的公式更新
$Y_{ij}^{t+1} \leftarrow Y_{ij}^t \frac{2(\mu Y^t)_{ij}}{(2\mathbf{M}' \mathbf{Y}^t + \mathbf{P}' + 2\mu \mathbf{Y}^t (\mathbf{Y}^t)^T \mathbf{Y}^t)_{ij}}$
⑧ 更新矩阵 $\mathbf{D}^{t+1}, D_{ii}^t = \frac{1}{2\ \mathbf{W}_i^t\ _2}$
⑨ $t = t + 1$
⑩ until 满足收敛条件
输出: 标签矩阵 $\mathbf{Y}$ ;
分类器 $\mathbf{W}$ .

### 3 数据分析对比

#### 3.1 数据集的选择

本次实验是在通用数据集 MSRC 上进行了大量的实验。MSRC 是语义分割任务中应用最为广泛的一个数据集, 它包含了来自 23 个类 532 幅图像, 平均每幅图有 3 个标签。其中 270 幅作为训练图像, 262 幅作为测试图像。

#### 3.2 图像分割和特征提取

采用了 SLIC<sup>[9]</sup> 分割算法来对每幅图进行过分割, 平均每幅图像被分割成 50 个子区域, 提取图像的 SIFT 点用文本中分类词袋 BOW (Bag-of-Words) 模型进行 500 维向量表示作为特征。特殊说明: 该方法是普遍适用的, 并不针对某一种特定的分割算法。

#### 3.3 模型评估

从两个角度来评价 WSTC 方法的效果: 语义分割和图像协同分割。对于语义分割的效果采用两种不同的定量指标来评价: 1) 总准确率 (T\_Pre), 即数据集中所有分类正确的像素之和比上数据集像素总和; 2) 分类准确率 (C\_Pre), 即统计每类分类的准确率, 最后在所有类上做平均。对于协同分割采用在 PASCAL 比赛中的通用指标 intersection-over-union<sup>[10]</sup> (IOU score) 来衡量。

#### 3.4 在 MSRC 数据集上的实验结果

在 MSRC 数据集上, 将 LAS<sup>[11]</sup>、MTL-RF<sup>[12]</sup>、MIM<sup>[13]</sup> 作为对比方法。从监督信息 (supervision, Sup)、标签先验 (image label prior, ILP)、特征 (multiple of feature, MOF) 这 3 个方面来简单总结一下这些方法, 具体内容如表 2 所示。

表 2 WSTC 与其他方法对比的实验参数设置

Method	MTL-RF	LAS	MIM	WSTC
Sup	Weak	Weak	Weak	Weak
ILP	Without	Without	With	With
MOF	No	No	Yes	No

其中, “Full” 监督代表每个像素都被人工打标用于训练模型, “Weak” 代表仅仅有图像级别的标签用于训练模型。 “With” 代表在测试阶段图像带有标签, 仅仅从图像的标签中选择每个子区域概率最大的标签作为预测的标签。 “Without” 代表在测试阶段图像的标签完全未知, 从整个数据集的标签集中选择每个子区域概率最大的标签。 “Yes” 代表此方法使用了多种特征。 “No” 代表此方法使用了单一特征。基于几种实验方法的语义分割准确率如图 2 和表 3 所示。

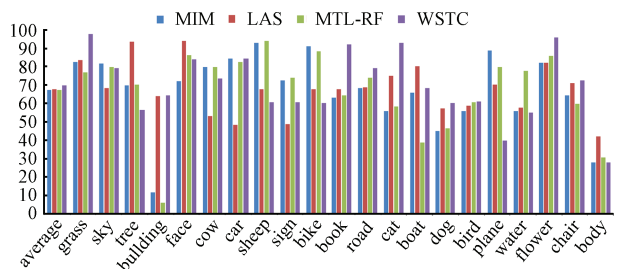


图 2 WSTC 等几种方法的语义分割分类准确率

表 3 WSTC 等几种方法的语义分割总准确率

Method	MTL-RF	LAS	MIM	WSTC
T_Pre	53	61	65	70

由实验结果分析可以看出 WSTC 方法较其他几种对比方法有较好的效果, 证明 WSTC 方法有效。另外, 本方法在 21 个类中的 9 个类上取得了最好的结果, 特别是在一些非常难分的类上, 比如狗、猫、鸟等也有好的表现。这

也证明了弱监督信息对容易混淆的类具有指导作用。

对于图像协同分割,用 LAS、MIM 与 WSTC 方法进行比较,协同分割的结果参如表 4 所示。

表 4 WSTC 与其他图像协同分割方法的结果对比

class	WSTC	LAS	MIM
chair	53.5	40.2	29.8
bike	39.4	43.8	30.7
face	46.4	69.7	34.2
tree	58.7	67.4	60.8
cat	51.9	32.9	26.4
house	48.8	51.1	33.5
car	51.8	59.9	36.1
dog	50.6	42.8	35.1
flower	80.6	50.8	42.2
sheep	67.3	64.2	61.8
bird	48.5	47.7	30.5
cow	43.9	52.5	32.8
sign	58.3	58.8	43.9
plane	36.5	22.4	26.7
Average	52.6	50.3	37.5

从表 4 中可以看到 WSTC 方法与其他方法对比得到了最高的 IOU 分数,这个结果证明了弱监督信息对图像分割是具有指导作用的,可以改善图像分割。同时,WSTC 方法比其他方法的准确率要高一些。

#### 4 结 论

图像语义解析、分割能够深层次地理解图像语义内容,不仅仅要指出图像中“有什么”,还要指出“在哪里”,即把语义标签映射到图像中相应的区域上去,实现准确的图像理解。本文基于弱监督的图像语义分割展开研究,提出了 WSTC 方法,训练图像标注图像级别的语义标签,取得了不错的效果。但是弱监督图像语义分割仍存在很多尚未解决或有待提高的地方,如精度问题、社会标签所带的噪声问题、特征和语义之间的映射问题等。这些工作不仅具有较高的研究价值,而且有广阔的应用前景,其研究成果必将更好地服务于人们的生活。

#### 参 考 文 献

[1] 谭静. 基于半监督聚类的图像分割算法研究[D]. 青

岛:中国海洋大学,2012.

- [2] 周玮. 基于数据驱动的马尔科夫蒙特卡洛视网膜血管分割[D]. 南京:南京航空航天大学,2011.
- [3] 景军锋,赵娟. 基于 Mean Shift 滤波的织物疵点检测方法[J]. 电子测量与仪器学报,2016,30(5):739-749.
- [4] 黄先楼. 基于 Normalized Cut 的图像分割及其 CUDA 并行实现[D]. 北京:北京交通大学,2014.
- [5] 朱力强,王春薇,王耀东,等. 基于特征点集距离描述的裂缝图像匹配算法研究[J]. 仪器仪表学报,2016,37(12):2851-2858.
- [6] 田素云,王小明,赵雪青. 基于拉普拉斯算子和图像修补的图像去噪算法[J]. 计算机应用,2012,32(10):2793-2797.
- [7] 孙东辉,鞠秀亮,冯登超,等. 基于 FAST 检测器和 SURF 描述子的聚合图像人脸识别[J]. 国外电子测量技术,2016,35(1):94-98.
- [8] 张维坤,叶伟,劳国超. 基于 SIFT 特征的 SAR 图像飞机目标匹配分类方法研究[J]. 国外电子测量技术,2016,35(8):19-21.
- [9] ACHANTA R, SHAJI A, SMITH K, et al. Slic superpixels compared to state-of-the-art superpixel methods[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2012,34(11):2274-2282.
- [10] PONCE J. Multi-class cosegmentation [C]. IEEE Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, 2012:542-549.
- [11] LIU X, YAN S, LUO J, et al. Nonparametric label-to-region by search[C]. Computer Vision and Pattern Recognition. IEEE, 2010:3320-3327.
- [12] VEZHNEVETS A, BUHMANN J M. Towards weakly supervised semantic segmentation by means of multiple instance and multitask learning[C]. Computer Vision and Pattern Recognition. IEEE, 2010:3249-3256.
- [13] KIM G, XING E P, LI F F, et al. Distributed cosegmentation via submodular optimization on anisotropic diffusion[C]. International Conference on Computer Vision. IEEE Computer Society, 2011:169-176.

#### 作 者 简 介

文笃石,1978 年出生,硕士研究生,主要研究方向为图像处理、嵌入式系统。  
E-mail:wds3721@sina.com