

## AdaBoost 图像到类距离学习的图像分类方法

李子龙<sup>1,2,3\*</sup>, 周 勇<sup>2</sup>, 鲍 蓉<sup>1</sup>

(1. 徐州工程学院信电工程学院, 徐州, 221018; 2. 中国矿业大学计算机科学与技术学院, 徐州, 221116;

3. 江苏昂内斯电力科技股份有限公司博士后科研工作站, 徐州, 221003)

**摘 要:**近年来,距离度量学习已经成为图像分类领域的研究热点之一,图像到类距离的度量作为其中的一种方法,取得了不错的分类效果.该方法是一种非参数方法,但由于缺少训练学习,其分类性能很容易受干扰因素的影响,为此提出一种基于 AdaBoost 算法的图像到类距离学习的图像分类方法.首先将图像到类的距离进行阈值化处理,并使用线性分段函数作为图像到类距离的评价函数,然后将该评价函数作为弱分类器加入到 AdaBoost 算法中生成一个强分类器.为了选择最优的弱分类器,使用粒子群优化算法确定图像的相似性阈值,再基于权重错误误差最小化原则得到距离评价函数的两个评价.最后通过实验验证,该方法在 Scene-15 和 Caltech-101 图像数据集上比其他方法有更好的分类效果.

**关键词:**图像分类,图像到类距离,AdaBoost,粒子群优化算法

**中图分类号:**TP301.4

**文献标识码:**A

## AdaBoost image-to-class distance learning for image classification

Li Zilong<sup>1,2,3\*</sup>, Zhou Yong<sup>2</sup>, Bao Rong<sup>1</sup>

(1. School of Information and Electrical Engineering, Xuzhou Institute of Technology, Xuzhou, 221018, China;

2. School of Computer Science and Technology, China University of Mining and Technology, Xuzhou, 221116, China;

3. Post Doctoral Research Center, Onnes Power Technology Co., Ltd, Xuzhou, 221003, China)

**Abstract:** Recently, distance metric learning has become one of the most attractive research areas in image classification. The image-to-class distance metric is a non-parametric method for image classification and achieves a impressive result. However, due to the lack of a training phase, the classification accuracy of it is easily affected by irrelevant factors. In this paper, we propose a novel image-to-class distance learning method for image classification by using the AdaBoost algorithm. We first deal with image-to-class distance through the threshold, and a piecewise linear discriminator function is used as the evaluation function of image-to-class distance. Then, the evaluation function is added to the AdaBoost algorithm as a weak classifier to generate a strong classifier. In order to select the optimal weak classifier, the particle swarm optimization algorithm is used to determine the similarity threshold of the image, and the two evaluation values of the distance evaluation function are obtained based on the principle of weight error minimization. The experimental results on datasets of Scene-15 and Caltech-101 verified that our proposed method can significantly outperform other methods in image classification.

**Key words:** image classification, image-to-class distance, AdaBoost, particle swarm optimization algorithm

图像分类一直是计算机视觉领域的研究热点,近年来各种图像分类方法层出不穷<sup>[1-5]</sup>. 2008 年 Boiman et al<sup>[6]</sup>提出一种新颖的图像分类方法,这是一种非参数的方法,使用图像到类(image-to-class)的距离对图像进行分类,在一定程度上能克服局部特征量化带来的信息损失问题. 然而,由于实际场景中的图像受到光照、遮挡、类内界限模糊等干扰因素的影响,该图像到类距离的分类性能仍需提升,为此许多学者对其进行了改进.

2010 年 Wang et al<sup>[7]</sup>提出使用 Mahalanobis 距离作为图像到类距离的度量公式,并给出了距离度量的学习过程,该方法在特征数量不足时仍有较好的性能. 为了体现不同特征在分类中的不同重要性,2011 年 Wang et al<sup>[8]</sup>提出对图像特征到类距离的权值进行学习,进一步提高图像分类的抗干扰能力. 2012 年 McCann and Lowe<sup>[9]</sup>提出在整个图像集中找到最相关的类,并计算局部邻域的图像特征到类的距离,该方法缩小了查找空间,加快了分类速度. 2013 年 Wang et al<sup>[10]</sup>提出通过线性距离编码得到图像到类的距离矢量,该方法提升了泛化能力. 2015 年 Wang et al<sup>[11]</sup>提出在线性距离编码中考虑局部特征之间的相关性以抑制编码过程中噪声特征的干扰. 2015 年 Qiao and Li<sup>[12]</sup>提出在正则化逻辑回归框架下学习图像到类距离的参数,该方法的性能优于其他图像到类的方法. 2015 年 Tan et al<sup>[13]</sup>将图像到类的距离用于特征子集的选择,提出基于欧式距离的图像到类距离比去选择特征子集的方法,提高了图像分类效果. 2017 年 Peng<sup>[14]</sup>将视觉显著性应用于图像到类距离的计算,通过计算前景和背景的图像到类的距离实现图像分类.

虽然上面的方法从不同的角度对图像到类的距离进行改进,但仍然存在不足:有些特征可能属于多个类,不能简单地以距离哪个类最近就将其划分为哪个类这么简单地进行处理(即不是非此即彼就能明确其类别的);真实场景下图像中往往包含大量噪声,这导致所提取的特征距离与其本类中特征的距离会出现偏差,影响分类效果. 而每个局部特征与某一类中的特征是否相似,取决于这两个局部特征的相近程度,如果两个特征的距离足够接近,就有理由相信它们是属于同一类;

相反,如果它们的特征距离较远,也有充足理由认为它们不属于同一类. 因此,本文对图像到类的距离进行阈值化处理,并使用线性分段函数作为图像特征到类的距离评价函数;同时,借助 AdaBoost 算法组合弱分类器(距离评价函数),进一步提高分类器的分类性能.

## 1 图像到类距离

图像到类距离的方法主要是计算测试图像中的局部特征与候选类特征集合中距离其最近的特征的距离之和,并用其作为测试图像到该类的距离. 最后,根据测试图像距离哪个候选类集合最近决定将其划分到哪一类.

图 1 展示了图像到类距离的图像分类过程. 图中,长方形代表一幅图像,三角形代表从图像中提取的局部特征,椭圆形代表不同类别的候选图像集. 图中的测试图像为  $X$ ,  $Dist(X, c)$  代表测试图像到类别  $c$  的距离.

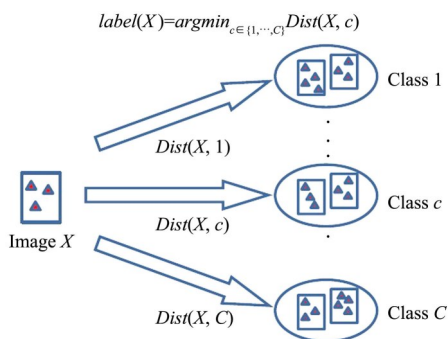


图 1 图像到类距离的分类过程

Fig. 1 The flow of image to classification

## 2 图像到类距离

先给出相关定义,  $\{X_1, \dots, X_M\}$  表示训练样本集,  $M$  是训练样本的数量, 每一个训练样本  $X_i$  是从相应的单幅样本图像提取的特征点集合, 且每个训练样本都有一个对应于类别  $c$  的二进制类别标志  $Y_i^c \in \{-1, +1\}$ .

使用 AdaBoost 算法<sup>[15]</sup>来学习局部特征到类的距离. 根据 AdaBoost 算法的原理, 定义图像  $X_i$  到类  $c$  距离的强分类器如式(1)所示:

$$Dist(X_i, c) = \sum_{k=1}^N \varphi_k(X_i, c) \quad (1)$$

其中,  $N$  为迭代的次数, 每次迭代都会产生一个最优的弱分类器  $\varphi_k$  来更新强分类器  $Dist$ . 在这里, 弱分类器  $\varphi_k$  表示图像到类的距离评价函数. 评价图像到类的距离常选用欧氏距离或 Mahalanobis 距离, 欧式距离的计算简单, 而 Mahalanobis 距离的度量矩阵学习起来较为复杂. 实际场境中的图像类内变化突出、类间界限模糊, 且图像中往往包含大量噪声, 这些都会影响距离评价函数的效果. 为此, 本文将图像到类的距离进行阈值化处理以便于产生一个适应能力更好的距离评价函数,  $\varphi_k$  的定义如式(2)所示:

$$\varphi_k(X_i, c) = \begin{cases} \alpha_k^c & \text{if } \bar{d}(X_i, c) < t_k^c \\ \beta_k^c & \text{otherwise} \end{cases} \quad (2)$$

其中,  $\bar{d}(X_i, c) = \frac{\sum_{j=1}^{m_i} d(f_{ij}, f_{ij}^c)}{m_i}$ ,  $m_i$  表示从图像  $X_i$

提取的特征数量,  $f_{ij}$  表示提取的第  $j$  个特征,  $f_{ij}^c$  表示类  $c$  中距离  $f_{ij}$  的最近邻特征,  $d(f_{ij}, f_{ij}^c)$  是一个距离度量函数, 这里选用简单的欧式距离进行度量.  $t_k^c$  是一个图像相似性阈值,  $\alpha_k^c$  表示两幅图像在足够相似的情况下的评价值, 而  $\beta_k^c$  则表示两幅图像在不足够相似情况下的评价值. 它们都是要优化学习的参数, 但在确定  $\alpha_k^c$  和  $\beta_k^c$  的值之前, 需要先确定  $t_k^c$  的值.

本文在 AdaBoost 算法的每次迭代过程中都使用粒子群优化算法优化  $t_k^c$  的值, 该算法中每个粒子的位置都是  $t_k^c$  的可能取值, 而整个粒子群的全局最优位置就是  $t_k^c$  的最优值. 在粒子移动的迭代过程中, 粒子会通过适应值来展现其目前的优劣程度, 本文使用权重错误误差  $e_k^c$  作为算法学习过程中的粒子适应值. 第  $k$  次迭代过程中, 参数  $t_k^c$  值的学习过程如下所示:

(1) 初始化粒子群, 确定  $t_k^c$  的初始值集合.

(2) 计算每个粒子的适应值, 获得粒子的个体最优位置和群体最优位置.

(3) 更新第  $k$  个弱分类器, 并更新粒子群中粒子的速度和位置.

(4) 判断是否终止, 否则转向第(2)步执行.

确定  $t_k^c$  值之后, 给出  $\alpha_k^c$  和  $\beta_k^c$  优化的过程.  $W_{++}^c$

表示由弱分类器  $\varphi_k$  所标记为正类的正训练样本的权重和, 而  $W_{+-}^c$  表示由弱分类器  $\varphi_k$  所标记为负类的正训练样本的权重和; 同样的,  $W_{-+}^c$  和  $W_{--}^c$  分别表示真负和假正训练样本的权重和. 也就是说, 这四个权重可由式(3)计算得到:

$$\begin{aligned} W_{++}^c &= \sum_{Y_i=1 \wedge \bar{d}(X_i, c) < t_k^c} w_i \\ W_{+-}^c &= \sum_{Y_i=1 \wedge \bar{d}(X_i, c) \geq t_k^c} w_i \\ W_{-+}^c &= \sum_{Y_i=-1 \wedge \bar{d}(X_i, c) < t_k^c} w_i \\ W_{--}^c &= \sum_{Y_i=-1 \wedge \bar{d}(X_i, c) \geq t_k^c} w_i \end{aligned} \quad (3)$$

在 AdaBoost 算法的每次迭代过程中, 希望选择合适的  $t_k^c$  值能最小化训练样本的权重错误, 如式(4)所示:

$$\begin{aligned} e_k^c &= \sum_{Y_i=1 \wedge \bar{d}(X_i, c) \geq t_k^c} w_i + \sum_{Y_i=-1 \wedge \bar{d}(X_i, c) < t_k^c} w_i = \\ &W_{+-}^c + W_{-+}^c \end{aligned} \quad (4)$$

为了得到最佳的  $\alpha_k^c$  和  $\beta_k^c$  值, 对式(5)<sup>[16]</sup>进行最优化:

$$Z_k^c = \sum_{i=1}^M w_i e^{-Y_i \varphi_k(X_i, c)} \quad (5)$$

为了最优化式(5), 对其进行如下分解:

$$\begin{aligned} Z_k^c &= \sum_{Y_i=1} w_i e^{-\varphi_k(X_i, c)} + \sum_{Y_i=-1} w_i e^{\varphi_k(X_i, c)} = \\ &\sum_{Y_i=1 \wedge \bar{d}(X_i, c) < t_k^c} w_i e^{-\alpha_k^c} + \sum_{Y_i=1 \wedge \bar{d}(X_i, c) \geq t_k^c} w_{ij} e^{-\beta_k^c} + \\ &\sum_{Y_i=-1 \wedge \bar{d}(X_i, c) < t_k^c} w_{ij} e^{\alpha_k^c} + \sum_{Y_i=-1 \wedge \bar{d}(X_i, c) \geq t_k^c} w_{ij} e^{\beta_k^c} = \\ &W_{++}^c e^{-\alpha_k^c} + W_{+-}^c e^{-\beta_k^c} + W_{-+}^c e^{\alpha_k^c} + W_{--}^c e^{\beta_k^c} \end{aligned} \quad (6)$$

对式(6)分别求  $\alpha_k^c$  和  $\beta_k^c$  的偏导数并都设置其值为零, 则可确定这两个参数的最优值, 如式(7)所示:

$$\alpha_k^c = \frac{1}{2} \lg \left( \frac{W_{++}^c}{W_{-+}^c} \right), \quad \beta_k^c = \frac{1}{2} \lg \left( \frac{W_{+-}^c}{W_{--}^c} \right) \quad (7)$$

跟经典的 AdaBoost 算法一样, 训练样本的权重在每次迭代中都要被更新, 如式(8)所示:

$$w_i^{k+1} = w_i^k e^{-Y_i \varphi_k(X_i, c)} \quad (8)$$

### 3 实验结果及分析

为了验证本文所提方法的有效性,在 Scene-15 和 Caltech-101 两个不同的数据集上进行测试,它们分别包含 15 类和 101 类各种不同场景下具有代表性的图像集合. 对于 Caltech-101 数据集,只选取 butterfly, carside, goat, face, bear 和 rifle 六类图像进行实验. 对于这两个图像数据集,随机抽取每一类的 100 幅图像用于训练和测试,并使用 SIFT 特征作为图像所提取的局部特征描述符.

在本文所使用的粒子群优化算法中,粒子数量设置为 50,算法的终止条件为权重错误误差  $e_k^c$  为 0.1. 为了融合空间信息, Wang et al<sup>[7]</sup> 结合空间金字塔和图像到类的距离,以逐渐精细的方式将图像递归地划分成子区域,进一步提高分类性能. 在实验过程中,本文对所有参与对比的方法也都采用空间金字塔,所使用的空间层次及区域块大小为  $1 \times 1, 2 \times 2, 4 \times 4$ , 且文中所有的实验都在 i5-3210 2.5 GHz CPU 6 G 内存 64 位 Windows 10 操作系统下进行.

在 AdaBoost 算法中,弱分类器的数量会对图像分类的准确性产生作用. 为此,本文在 Scene-15 和 Caltech-101 图像数据集上观察弱分类器数量对分类器准确性的影响,如图 2 和图 3 所示. 通过对这两个图的观察可以看出,随着迭代次数的增加,在 Scene-15 图像集上的分类准确率提升较快,而在 Caltech-101 上则提升较慢,在这也符合 Caltech-101 数据集中的图像类内和类间的差异大、分类难度更高的特征. 为了能更好地适应在不同图像数据集上的分类,在进行实验对比时,在 Scene-15 数据集上的 AdaBoost 算法的迭代次数设置为 70,而将在 Caltech-101 数据集上的 AdaBoost 算法的迭代次数设置为 80.

在 Scene-15 和 Caltech-101 这两个图像集上使用本文方法和 NBNN (Naive Bayes Nearest Neighbor)<sup>[6]</sup>, LI2C (Learning Image to Class)<sup>[8]</sup> 方法在每一类别图像上进行分类性能对比,如图 4 和图 5 所示. 可以看出,本文方法对较难识别的图像(如 suburb, carside 等)跟其他方法相差不大,但对于那些难以识别图像(如 living room, butterfly 等)的分类准确性仍然较高. 总之,本文方法能更

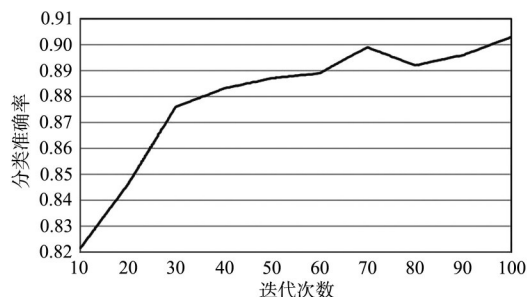


图 2 Scene-15 图像集中 AdaBoost 算法迭代次数与分类准确率的关系

Fig. 2 The relationship between the number of iterations for AdaBoost algorithm and classification accuracy on Scene-15 dataset

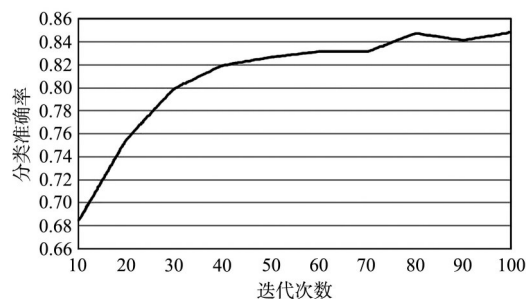


图 3 Caltech-101 图像集中 AdaBoost 算法迭代次数与分类准确率的关系

Fig. 3 The relationship between the number of iterations for AdaBoost algorithm and classification accuracy on Caltech-101 dataset

有效地提高基于图像到类距离的图像分类性能.

图 6 给出本文方法、LI2C 和 NBNN 三种方法分别在 Scene-15 数据集上进行实验所得到的 ROC (Receiver Operating Characteristic) 曲线.

由图 6 可见,本文方法的 ROC 曲线下的区域面积比其他两个方法的面积大一些,这也证明本文方法的分类性能要高于其他算法. 表 1 列出了 NBNN, LI2C, Local NBNN<sup>[9]</sup>, LDC (Linear Distance Coding)<sup>[10]</sup>, SI2C (Saliency-aware Image to Class)<sup>[14]</sup> 和本文方法在 Scene-15 和 Caltech-101 图像集上的平均分类准确性的情况. 可以看出: NBNN 的准确性最低; LI2C 对图像到类距离上的权重进行学习,提高了图像分类的抗干扰能力; Local NBNN 利用局部邻域图像特征到类的距离,加快了分类速度; LDC 通过对图像到类的距离进行稀疏编码,增强其泛化能力; SI2C 将视觉



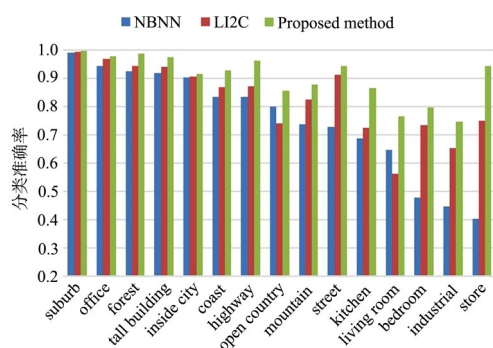


图4 Scene-15图像集中每种类别上不同方法的图像分类性能对比

Fig. 4 The comparison of per - category classification accuracy of three methods on Scene-15 dataset

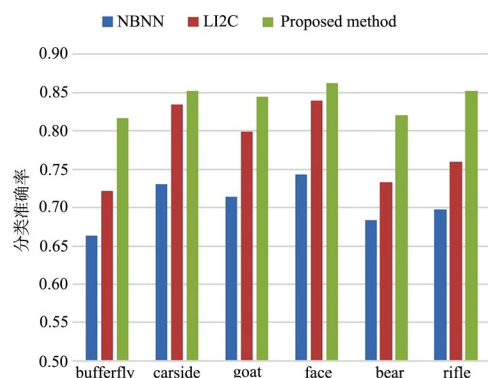


图5 Caltech-101图像集中每种类别上不同方法的图像分类性能对比

Fig. 5 The comparison of per - category classification accuracy of three methods on Caltech-101 dataset

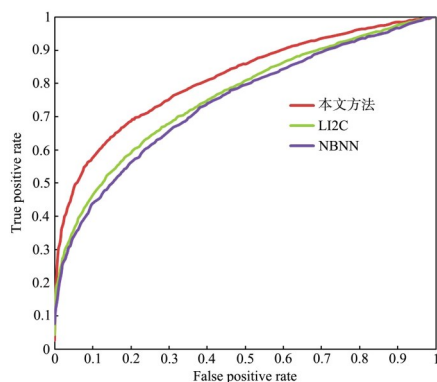


图6 不同方法在Scene-15数据集上的ROC曲线

Fig. 6 The ROC curves of different methods on Scene-15 dataset

显著性应用于图像到类距离的计算,但在类内变化较大、背景复杂的情况下仍表现不好. 本文通过构造一个适应能力强的距离评价函数作为弱分类器,并使用AdaBoost算法生成一个强分类器来获得更好的分类效果,本文方法在两个数据集上的分类准确率比SI2C都高出3%以上.

表1 不同方法的平均分类准确性(%)对比

Table 1 The average classification ( % ) of different methods

方法	Scene-15	Caltech-101
NBNN	$78.7 \pm 0.74$	$71.6 \pm 1.12$
LI2C	$81.7 \pm 0.54$	$77.8 \pm 0.98$
Local NBNN	$83.3 \pm 0.59$	$79.4 \pm 1.10$
LDC	$84.4 \pm 0.52$	$80.3 \pm 0.95$
SI2C	$86.5 \pm 0.50$	$81.2 \pm 0.81$
本文方法	$89.8 \pm 0.44$	$84.7 \pm 0.78$

## 4 结 论

本研究对基于图像到类距离的图像分类方法进行改进,提出在AdaBoost算法框架下学习图像到类距离的分类方法. 首先,为构造一个适应能力强的距离评价函数,对图像到类的距离进行阈值化处理,并将其作为弱分类器加入到AdaBoost算法中;其次,为学习距离评价函数的参数,分别使用粒子群优化算法和误差最小化原理去选择最佳的弱分类器. 实验证明,本文提出的方法能有效地提高基于图像到类距离的分类性能.

### 参考文献

- [1] Zhang L, Zhen X T, Shao L. Learning object-to-class kernels for scene classification. IEEE Transactions on Image Processing, 2014, 23(8): 3241—3253.
- [2] Moutafis P, Leng M J, Kakadiaris I A. An overview and empirical comparison of distance metric learning methods. IEEE Transactions on Cybernetics, 2016, 47(3): 612—625.
- [3] Rasche C. Rapid contour detection for image classification. IET Image Processing, 2017, 12(4): 532—538.
- [4] Tang P, Wang X G, Feng B, et al. Learning multi-instance deep discriminative patterns for image

- classification. IEEE Transactions on Image Processing, 2017, 26(7): 3385–3396.
- [5] Lu Y W, Lai Z H, Li X L, et al. Learning parts-based and global representation for image classification. IEEE Transactions on Circuits and Systems for Video Technology, 2018, 28(12): 3345–3360.
- [6] Boiman O, Shechtman E, Irani M. In defense of nearest-neighbor based image classification//IEEE Conference on Computer Vision and Pattern Recognition. Anchorage, AL, USA: IEEE, 2008: 1–8.
- [7] Wang Z X, Hu Y Q, Chia L T. Image-to-class distance metric learning for image classification//European Conference on Computer Vision. Springer Berlin Heidelberg, 2010: 709–719.
- [8] Wang Z X, Hu Y Q, Chia L T. Improved learning of I2C distance and accelerating the neighborhood search for image classification. Pattern Recognition, 2011, 44(10–11): 2384–2394.
- [9] McCann S, Lowe D G. Local naive bayes nearest neighbor for image classification//2012 IEEE Conference on Computer Vision and Pattern Recognition. Providence, RI, USA: IEEE, 2012: 3650–3656.
- [10] Wang Z L, Feng J S, Yan S C, et al. Linear distance coding for image classification. IEEE Transactions on Image Processing, 2013, 22(2): 537–548.
- [11] Wang Z L, Feng J S, Yan S C. Collaborative linear coding for robust image classification. International Journal of Computer Vision, 2015, 114(2–3): 322–333.
- [12] Qiao S Y, Li Z L. Discriminative learning of I2C distance for image classification//The 27<sup>th</sup> Chinese Control and Decision Conference. Qingdao, China: IEEE, 2015: 1371–1375.
- [13] Tan S B, Liu L, Peng C Y, et al. Image-to-class distance ratio: a feature filtering metric for image classification. Neurocomputing, 2015, 165: 211–221.
- [14] Peng P. Context-aware image-to-class distances from image classification. Ph. D Dissertation. Sheffield: University of Sheffield, 2017.
- [15] Wang H J, Hu J N, Deng W H. Face feature extraction: a complete review. IEEE Access, 2017, 6: 6001–6039.
- [16] Fernández-Baldera A, Buenaposada J M, Baumela L. BAdaCost: multi-class boosting with costs. Pattern Recognition, 2018, 79: 467–479.
- (责任编辑 杨可盛)