

# 基于稀疏表示和词袋模型的高光谱图像分类

任越美<sup>1,2</sup> 张艳宁<sup>1</sup> 魏巍<sup>1</sup> 张秀伟<sup>1</sup>

(西北工业大学计算机学院 西安 710072)<sup>1</sup> (河南工业职业技术学院计算机工程系 南阳 473000)<sup>2</sup>

**摘要** 为增强高光谱图像稀疏表示分类方法中稀疏字典的表征能力并充分利用高光谱图像的光谱信息和空间信息,提出了一种新的基于稀疏表示和词袋模型的高光谱遥感图像分类方法。首先利用词袋模型算法结合高光谱遥感图像数据集生成各类别专业码本,作为字典中对应的原子构造稀疏表示字典。在计算每个像元的对应稀疏表示字典中的稀疏表示特征时,利用空间连续性约束对像元的稀疏表示系数进行空间维的约束。最后根据最小重构误差实现高光谱图像分类。高光谱遥感数据实验结果表明:所提方法能有效提高分类效果,并且其分类精度和 Kappa 系数都优于其他稀疏表示方法以及单独使用光谱信息的方法。

**关键词** 图像处理,高光谱图像,稀疏表示,词袋模型,空间连续性

中图分类号 TP751 文献标识码 A DOI 10.11896/j.issn.1002-137X.2014.10.026

## Classification of Hyperspectral Image Based on Sparse Representation and Bag of Words

REN Yue-mei<sup>1,2</sup> ZHANG Yan-ning<sup>1</sup> WEI Wei<sup>1</sup> ZHANG Xiu-wei<sup>1</sup>

(School of Computer Science, Northwestern Polytechnical University, Xi'an 710072, China)<sup>1</sup>

(Department of Computer Engineering, Henan Polytechnic Institute, Nanyang 473000, China)<sup>2</sup>

**Abstract** To enhance representation ability of the sparse dictionary for hyperspectral image classification using sparse representation and make full use of spectral information and spatial information of hyperspectral image, a novel hyperspectral image classification method based on sparse representation and bag of words was proposed. First, some professional dictionaries of each class are generated by bag of words algorithm based on the hyperspectral remote sensing image dataset, and the sparse representation dictionary is obtained by merging these professional dictionaries. Then, the sparse coefficient of each pixel is calculated according to the sparse representation dictionary, and the spatial continuity is used to constraint the coefficient by using the information of its neighborhoods. Finally, the classification of the objects is determined by computing the minimum reconstruction error of it on each professional dictionary. Experiments on hyperspectral remote sensing images indicate that the proposed method has better performance, a higher overall accuracy and Kappa coefficient than other sparse representation methods and the method based on spectral information respectively.

**Keywords** Image processing, Hyperspectral image, Sparse representation, Bag of words, Spatial continuity

## 1 引言

近年来,国内外高分辨率遥感传感器得到飞速发展,高分辨率对地观测系统成为世界高科技发展的前沿领域,高分辨率遥感图像分析也成为军事和民用等诸多领域研究的重点。高光谱遥感图像是由几十乃至数百个连续波段图像组成的 3 维数据立方体,具有很高的光谱分辨率,为精确的地物分类提供了可能<sup>[1]</sup>,但仅用光谱信息进行分类,往往达不到更好的分类精度。高光谱图像空间分辨率很低,像素间具有很强的空间关系<sup>[2]</sup>。因此,综合且有效地利用高光谱图像的光谱信息以及空间信息进行地物分类,将有助于提升分类器的性能,进而提高分类精度<sup>[3-6]</sup>。

图像稀疏表示(SR)就是把图像投影到由字典中基原子组成的特征子空间时,只有少数表示系数不为零,从而把图像表示为字典中原子的稀疏线性组合。高光谱遥感图像在光谱维满足稀疏性条件,遥感领域的学者已提出利用稀疏表示模型解决高光谱图像处理问题<sup>[7-11]</sup>。文献[7]将稀疏表示分类方法应用于高光谱分类,直接将训练样本组成字典,在该字典上将测试样本进行稀疏分解,通过最小重构误差进行分类。该方法在分类精度上较神经网络等传统方法有了明显提高,但仅利用了光谱信息,没有考虑空间信息的约束。文献[8]在稀疏表示高光谱图像分类中引入拉普拉斯约束,同时利用了光谱信息和空间信息,但只考虑了 4 邻域的空间约束关系。上述稀疏表示分类方法中,稀疏表示字典是由训练样本直接

到稿日期:2013-06-30 返修日期:2013-08-09 本文受中国自然科学基金重点项目(61231016),中国自然科学基金(61272288,61201291,61303123),河南省科技攻关计划(142102210557),西工大校基础研究基金(JCT20130108,JCT20130109)资助。

任越美(1984-),女,博士生,讲师,主要研究方向为数字图像处理、模式识别,E-mail:renym2008@163.com;张艳宁(1967-),女,博士,教授,主要研究方向为计算机图像与视频处理、模式识别以及智能信息处理等。

组成的,如果样本选择不当,字典的表征能力就比较弱,分类性能也会下降。若能设计适合高光谱遥感图像地物分类的稀疏表示字典,将有效提高地物分类的精度。文献[9]通过学习的方式得到学习字典,将像元在学习得到的字典上的投影作为像元的稀疏表示特征,然后结合稀疏表示特征及光谱信息,使用随机森林方法对高光谱遥感图像进行分类,但是,高光谱数据维数高、数据量大,其字典学习过程往往比较慢,影响了高光谱图像的地物分类的效率。

针对上述问题,本文提出了一种新的基于稀疏表示和词袋模型的高光谱图像分类方法,通过词袋模型得到的码本生成稀疏表示字典,根据稀疏表示字典进行稀疏分解得到每个像元的稀疏表示特征,并利用高光谱遥感图像中地物的空间连续性约束得到中心像元最终的稀疏表示系数,从而有效地结合图像中的光谱信息和空间结构信息,提高了高光谱图像分类的精度。

## 2 本文算法

稀疏表示技术因其符合生物视觉系统特性,在模式识别领域得到广泛应用<sup>[12,13]</sup>。高光谱遥感图像光谱信息具有内在稀疏性,且这种稀疏特性具有很强的鲁棒性,使得像元在稀疏表示字典下的稀疏表示系数可以作为提取的特征直接用于高光谱图像分类。其中,稀疏表示字典的设计直接影响分类的效率及精度,因此,本文利用词袋模型生成用于分类的稀疏表示字典,并结合高光谱遥感图像地物的空间特性在空间维上对稀疏表示系数特征进行约束,提出了一种新的基于稀疏表示和词袋模型的高光谱图像分类方法。

### 2.1 稀疏表示字典生成

在基于稀疏表示的图像分类中,字典一般由训练样本直接组成,但这种方式组成的字典一般要有较多的原子才能张成较好的特征子空间,对于数据量较大的高光谱图像不太适合,且这种构造方法生成的字典的可区分性不强,不利于地物分类。相比之下,通过学习得到的字典原子更能有效地表示地物像元的光谱特征。令高光谱图像数据集为  $X = \{x_i | x_i \in \mathbb{R}^b, 1 \leq i \leq n\}$ ,字典为  $D \in \mathbb{R}^{b \times m}$ ,则字典学习可转化为求解如下优化问题:

$$\begin{aligned} \min_{D,A} & \frac{1}{2} \|X - DA\|_F^2 + \lambda \|A\|_{1,1} \\ \text{s. t.} & \|D_j\|_2 \leq 1, \forall j \in 1, \dots, m \end{aligned} \quad (1)$$

但上述问题通常不是凸优化问题,不能直接求解,常采用迭代的方法进行求解,导致整个字典学习过程比较费时,而且高光谱图像中存在大量混合像元,致使采用最小均方差字典学习算法得到的超完备字典对于分类可能不是最优的。

词袋模型算法灵活性强,且鲁棒性好,得到的字典辨别能力比较好<sup>[14]</sup>。本文结合词袋模型,提出了一种新的适合高光谱图像地物分类的稀疏表示字典构造方法。对于有  $C$  类地物的高光谱图像,每类地物的样本数为  $n_c$ ,分别针对每类地物生成一个样本集合  $D_c$ ,可表示为:

$$D_c = [d_1^c, d_2^c, \dots, d_{n_c}^c], c = 1, 2, \dots, C \quad (2)$$

其中,  $d_i^c$  表示第  $c$  类的第  $i$  个样本的光谱向量。

根据词袋模型,利用 K-means 聚类为每类地物分别生成相应的“专业字典” $\tilde{D}_c$ ,再由“专业字典”合并得到最终用于高光谱遥感图像地物分类的稀疏表示字典  $\tilde{D}$ 。

$$\tilde{D} = [\tilde{D}_1, \tilde{D}_2, \dots, \tilde{D}_C] \quad (3)$$

### 2.2 稀疏系数分解及高光谱图像分类

生成字典稀疏表示字典  $\tilde{D}$  后,可依据稀疏表示理论,由字典中的原子对高光谱图像中的测试像元  $x$  进行稀疏线性表示,表示系数可通过求解如下的优化问题而得到:

$$\tilde{\alpha}(x) = \arg \min \|x - \tilde{D}\alpha\|_2^2 + \lambda \|\alpha\|_1 \quad (4)$$

式(4)在数学上是一个凸最优问题,可以转化为线性规划问题加以求解。

得到稀疏表示系数  $\tilde{\alpha}(x)$  后,可根据各“专业字典”近似地重构出像元  $\tilde{x}_c = \tilde{D}_c * \tilde{\alpha}_c(x)$ ,其中,  $\tilde{\alpha}_c(x)$  为系数向量  $\tilde{\alpha}(x)$  中对应第  $c$  类地物的向量子集。

最后,再以最小重构误差为准则得到该像元的地物类别:

$$\text{Class}(x) = \arg \min_{c=1, \dots, C} \|x - \tilde{x}_c\|_2 \quad (5)$$

上述分类过程只考虑了地物的光谱信息差异,忽视了高光谱地物目标在空间维上的信息特性,最终的地物分类精度会受到一定的影响。

### 2.3 空间连续性约束

高光谱图像的成像特性使其地物分布存在空间连续性。因此,可将空域信息引入到我们的稀疏表示分类算法中,综合利用谱域和空域信息,由中心点地物的邻域信息对其光谱信息进行分类约束,以提高基于高光谱遥感图像的地物分类精度。常用的连通域约束有 8 邻域和 4 邻域约束,由高光谱的成像原理可知,8 邻域对中心像素的影响大于 4 邻域的作用,因此,本文算法利用中心像元的 8 邻域像元对光谱信息进行约束。

对于像元  $x$ ,其 8 邻域像元集合可表示为  $x_{8\#} = \{\hat{x}_1, \hat{x}_2, \dots, \hat{x}_8\}$ ,对邻域集合  $x_{8\#}$  中的每个像元分别求解其在稀疏表示字典  $\tilde{D}$  上的稀疏表示系数,记为:

$$\tilde{\alpha}(x_{8\#}) = \{\tilde{\alpha}(\hat{x}_1), \tilde{\alpha}(\hat{x}_2), \dots, \tilde{\alpha}(\hat{x}_8)\} \quad (6)$$

其中,  $\tilde{\alpha}(\hat{x}_i)$  ( $i=1, \dots, 8$ ) 表示像元  $\hat{x}_i$  在字典  $\tilde{D}$  上的稀疏分解系数。

则像元  $x$  最终的稀疏分解系数可由其本身及其邻域像元的稀疏分解系数决定,即

$$\tilde{\alpha}_{\text{new}}(x) = \omega_0 * \tilde{\alpha}(x) + \sum_{i=1}^8 \omega_i * \tilde{\alpha}(\hat{x}_i) \quad (7)$$

其中,  $\omega_0$  和  $\omega_i$  ( $i=1, \dots, 8$ ) 代表权重因子,方便起见,为每个系数赋予相同的权值,取为 1/9。

## 3 实验结果及分析

为了验证本文算法的地物分类效果,对两组高光谱遥感图像数据集进行了实验。为了验证本文字典获取方法的有效性,在相同原子数的情况下,基于稀疏表示图像分类方法,将本文基于词袋模型生成的字典(SRC-Bow),与随机选取训练样本作为字典原子生成的字典(SRC-Random)以及通过学习方式得到的学习字典(SRC-Learning),进行了分类对比实验。同时,在采用本文方法生成的字典基础上,进一步加入空间特性约束(SRC-SC)对高光谱遥感图像地物进行了分类实验。在实验中采用常用的高光谱遥感图像地物分类评价指标,即总精度(Overall Accuracy, OA)、Kappa 系数(Kappa Coefficient, KC)。所有算法和实验都是运行 10 次,对结果求平均值,最优结果用粗体表示。

### 3.1 AVIRIS Indian Pines 高光谱图像

AVIRIS 高光谱数据<sup>[15]</sup> Indian Pines 采集自美国印第安纳州印度松树测试地,共 220 个波段,波长覆盖范围为 0.4~2.5 $\mu\text{m}$ ,去掉噪声和水汽吸收较明显的波段后,对剩下的 200 个波段数据进行实验。该数据包含 16 类地物类型,共 10366 个有标签的样本点。其中 7 类地物由于样本点太少一般不采用,选用最常用的 9 类典型地物进行分类实验<sup>[7]</sup>。实验时从各类地物样本中随机选取一半作为训练样本,剩余一半作为测试样本。

表 1 给出了不同方法在对应各地物原子数为相应训练样本 10% 时的分类结果,对应的分类结果图如图 1(a)~(d) 所示。针对不同的分类方法,还采用不同大小的稀疏表示字典进行了实验,对稀疏表示字典中对应各类地物的原子数为其训练样本数的 {10%, 20%, 30%, 40%, 50%} 时的情况,分别进行了分类实验,结果如图 2 所示。

表 1 不同分类方法在 Indian Pine 图像上的分类结果对比

方法	SRC-Random	SRC-Learning	SRC-Bow	SRC-SC
总精度(%)	66.49	76.68	80.43	86.87
Kappa 系数	0.605	0.725	0.759	0.835

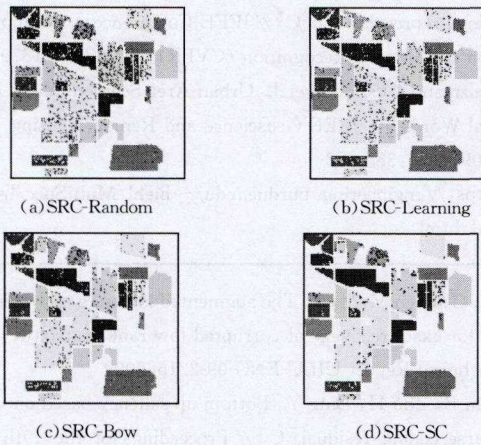


图 1 Indian Pine 高光谱遥感图像地物的分类结果

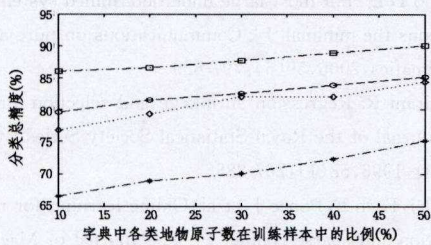


图 2 各类地物选取不同原子数目时(不同大小字典)的分类结果

从表 1 可以看出,在字典大小相同的情况下,本文算法相对采用随机字典和学习字典的稀疏表示方法的地物分类精度有了明显的提高,在未引入空间信息时,(1)总精度分别由基于随机字典方法的 66.49%,基于学习字典方法的 76.68% 提高到了词袋模型生成字典方法的 80.43%;(2)Kappa 系数分别由基于随机字典方法的 0.605,基于学习字典方法的 0.725 提高到了本文方法的 0.759。加入空间结构信息约束后,分类精度和 Kappa 系数分别提高到了 86.87% 和 0.835。

从图 2 可以看出,不同大小的稀疏表示字典对分类性能也有所影响,随着稀疏表示字典中相应类别原子数目的增加,分类精度也有了明显提高。

### 3.2 University of Pavia 高光谱图像

University of Pavia 高光谱图像是由 Reflective Optics System Imaging Spectrometer (ROSIS) 传感器获取的意大利 Pavia 大学数据,光谱范围 0.43~0.86 $\mu\text{m}$ ,共 115 个波段,去掉噪声明显的波段后,对剩下的 103 个波段数据进行实验。意大利 Pavia 大学 Paolo Gamba 教授提供了该数据,并且提供了同步实测的感兴趣区域和测试区域。表 2 给出了不同方法在对应各地物原子数为相应训练样本 10% 时的分类结果,对应的分类结果如图 3(a)~(d) 所示。

表 2 不同分类方法在 ROSIS University of Pavia 图像上的分类结果对比

方法	SRC-Random	SRC-Learning	SRC-Bow	SRC-SC
总精度(%)	76.56	79.43	80.86	83.74
Kappa 系数	0.701	0.736	0.753	0.789

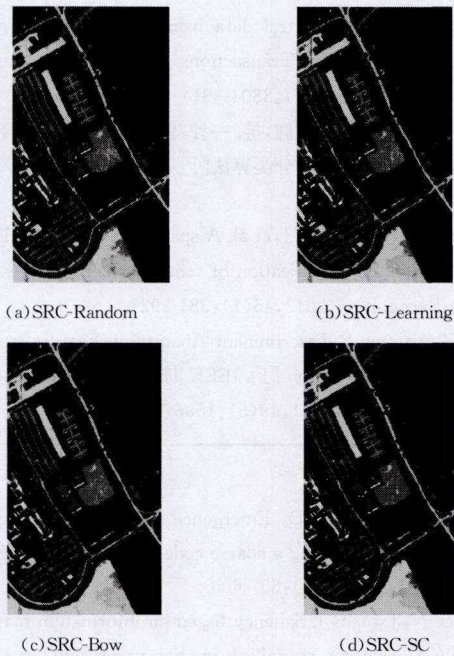


图 3 ROSIS University of Pavia 高光谱遥感图像地物的分类结果

从表 2 和图 3 同样可以看出,本文提出方法的分类精度要高于其他稀疏表示特征方法,且同时利用空-谱信息的分类结果优于单独使用光谱信息的分类结果。

从上述两个高光谱遥感图像数据集上的实验结果对比可以看出,本文算法利用词袋模型生成的各类的“专业字典”能更好地表征本类地物,合成生成的字典原子间的“高内聚、低耦合”性能比较好。另外,空间结构信息的引入,弥补了单从光谱信息进行分类时对其周围像元考虑不足的缺陷。因此,本文算法在分类精度和视觉效果上都要优于光谱信息方法和稀疏表示特征方法,最终达到了较好的地物分类效果。

**结束语** 基于高光谱遥感图像,利用稀疏表示理论完成地物分类,并利用词袋模型为每类地物分别生成“专业字典”,再将其合并成用于分类的整体字典。针对只利用光谱信息造成的分类精度不高的问题,根据地物像元的空间连续性约束像元在字典上的稀疏分解系数,完成综合光谱信息和空间信息的地物分类。在高光谱遥感图像数据集上的实验结果表明,本文所提方法提高了地物分类效果,在分类精度和 Kappa 系数方面都优于基于随机生成字典和学习字典的稀疏表示方

法,并且在本文字典基础上,加入空间信息约束的分类效果也优于单独使用光谱信息的分类效果。

本文算法提高了分类效果,但也有许多值得进一步改进的地方,如在高光谱图像的分类方法中的邻域约束仅采用了固定的8邻域信息,若能根据像元的分布采用随机窗口约束或是空间金字塔进行约束,则可以提高分类精度。

### 参考文献

- [1] Plaza A, Benediktsson J A, Boardman J W, et al. Recent advances in techniques for hyperspectral image processing [J]. *Remote Sensing of Environment*, 2009, 113(1): 110-122
- [2] Geng Xiu-rui, Zhang Xia, Chen Zheng-chao, et al. Classification algorithm based on spatial continuity for hyperspectral image [J]. *Journal of Infrared Millimeter Waves*, 2004, 23(4): 209-302
- [3] Mathieu F, Jón A B, Jocelyn C, et al. Spectral and spatial classification of hyperspectral data using SVMs and morphological profiles [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2008, 46( 11): 3804-3814
- [4] 高恒振,万建伟,粘永健,等.一种基于谱域-空域组合特征支持向量机的高光谱图像分类算法[J]. *宇航学报*, 2011, 32(4): 917-921
- [5] Fauvel M, Chanussot J, et al. A spatial-spectral kernel-based approach for the classification of remote-sensing images [J]. *Pattern Recognition*, 2012, 45(1): 381-392
- [6] Fu Z, Antonio R. Discriminant Absorption-Feature Learning for Material Classification [J]. *IEEE Transaction of Geoscience and Remote Sensing*, 2011, 49(5): 1536-1556
- [7] Olshausen B, Field D. Emergence of simple-cell receptive field properties by learning a sparse code for natural images[J]. *Natural*, 1996, 381(6583): 607-609
- [8] Bruce N, Tsotsos J. Saliency based on information maximization [C]// *Advances in neural information processing systems*. 2006, 18: 155-162
- [9] Hou X, Zhang L. Dynamic visual attention, Searching for coding length increments[C]// *Advances in Neural Information Processing Systems*. 2008, 21: 681-688
- [10] Yan J, Zhu M, Liu H. Visual saliency detection via sparsity pursuit[J]. *IEEE Signal Processing Letters*, 2010, 17(8): 739-742
- [11] Sun X, Yao H, Ji R. Saliency detection based on short-term sparse representation[C]// *Proceedings of 2010 IEEE 17th International Conference on Image Processing*. 2010: 1101-1104
- [12] Harel J, Koch C, Perona P. Graph-based visual saliency[C]// *Advances in Neural Information Processing System*. 2006: 545-552
- [13] Vinje W, Gallant J. Receptive fields and functional architecture in two nonstriate visual areas[J]. *Journal of Neurophysiol*, 1965, 28: 229-289
- [14] Liu G, Lin Z, Yu Y. Robust subspace segmentation by low-rank representation[C]// *International Conference on machine learning*. 2010
- [7] Haq Q S, Shi Li-xin, Tao Lin-mi, et al. A L1-minimization based approach for hyperspectral data classification[C]// *Proceedings of 2010 International Conference on Remote Sensing (ICRS)*. 2010, 2: 139-142
- [8] Yi Chen, Nasrabadi N M, Tran T D. Hyperspectral image classification using dictionary-based sparse representation [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2011, 49(10): 3973-3985
- [9] 宋相法, 焦李成. 基于稀疏表示及光谱信息的高光谱遥感图像分类[J]. *电子与信息学报*, 2012, 34(2): 268-272
- [10] Haq S, Tao Q L, et al. A Fast and Robust Sparse Approach for Hyperspectral Data Classification Using a Few Labeled Samples [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2012, 99: 1-16
- [11] Iordache M D, Dias J M B, Plaza A. Sparse unmixing of hyperspectral data[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2011, 49(6): 2014-2039
- [12] Wright J, Yang A, Ganesh A, et al. Robust face recognition via sparse representation[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2009, 31(2): 201-227
- [13] Elhamifar E, Vidal R. Robust classification using structured sparse representation[C]// *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2011: 1873-1879
- [14] Weizman L, Goldberger J. Urban-Area Segmentation Using Visual Words[J]. *IEEE Geoscience and Remote Sensing Letters*, 2009, 6(3): 388-392
- [15] <https://engineering.purdue.edu/~biehl/MultiSpec/hyperspectral.html>
- [15] Liu G, Chen M, Ma Y. The augmented lagrange multiplier method for exact recovery of corrupted low-rank matrix[R]. *UTUC Technical Report UILU-ENG-09-2215*. 2009
- [16] Han B, Zhu H, Ding Y. Bottom up saliency based on weighted sparse coding residual[C]// *Proceedings of the 19th Annual ACM International Conference on Multimedia*. 2011: 1117-1120
- [17] Donoho D L. For most large underdetermined systems of linear equations the minimal[J]. *Communications on pure and applied mathematics*, 2006, 59(6): 797-829
- [18] Tibshirani R. Regression shrinkage and selection via the lasso [J]. *Journal of the Royal Statistical Society Series B (Methodological)*, 1996, 58(1): 267-288
- [19] Mairal J, Bach F, Ponce J, et al. Online learning for matrix factorization and sparse coding[J]. *The Journal of Machine Learning Research*, 2010, 11: 19-60
- [20] Borji A, Itti L. Exploiting local and global patch rarities for saliency detection[C]// *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*. 2012: 478-485
- [21] Davis J, Goadrich M. The relationship between Precision-Recall and ROC curves[C]// *Proceedings of the 23th International Conference on Machine Learning*. 2006: 233-240
- [22] Hou X, Zhang L. Saliency detection: A spectral residual approach[C]// *IEEE International conference on Computer Vision and Pattern Recognition*. 2007: 1-8