

DOI: 10.11992/tis.201601001

网络出版地址: <http://www.cnki.net/kcms/detail/23.1538.TP.20160718.1521.004.html>

基于感受野学习的特征词袋模型简化算法

赵 骞, 李 敏, 赵晓杰, 陈雪勇

(电子科技大学 计算机科学与工程学院, 四川 成都 611731)

摘 要: 本文研究了在图像识别任务中, 感受野学习对于特征词袋模型的影响。在特征词袋模型中, 一个特征的感受野主要取决于视觉词典中的视觉单词和池化过程中所使用的区域。视觉单词决定了特征的选择性, 池化区域则影响特征的局部性。文中提出了一种改进的感受野学习算法, 用于寻找针对具体的图像识别任务最有效性的感受野, 同时考虑到了视觉单词数量增长所带来的冗余问题。通过学习, 低效、冗余的视觉单词和池化区域会被发现, 并从特征词袋模型中移除, 从而产生一个针对具体分类任务更精简的、更具可分性的图像表达。最后, 通过实验显示了该算法的有效性, 学习到的模型除了结构精简, 在识别精度上相比原有方法也能有一定提升。

关键词: 视觉词袋模型; 感受野学习; 目标识别; 图像分类; 特征学习

中图分类号: TP391.4 **文献标志码:** A **文章编号:** 1673-4785(2016)05-0663-07

中文引用格式: 赵骞, 李敏, 赵晓杰, 等. 基于感受野学习的特征词袋模型简化算法[J]. 智能系统学报, 2016, 11(5): 663-669.

英文引用格式: ZHAO Qian, LI Min, ZHAO Xiaojie, et al. Learning receptive fields for compact bag-of-feature model[J]. CAAI transactions on intelligent systems, 2016, 11(5): 663-669.

Learning receptive fields for compact bag-of-feature model

ZHAO Qian, LI Min, ZHAO Xiaojie, CHEN Xueyong

(School of Computer Science and Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China)

Abstract: In this work, the effects of receptive field learning on a bag-of-features pipeline were investigated for an image identification task. In a bag-of-features model, the receptive field of a feature depends mostly on use of visual words in a visual dictionary and the region used during the pooling process. Codewords make the feature respond to specific image patches and the pooling regions determine the spatial scope of the features. A modified graft feature selecting algorithm was proposed to find the most efficient receptive fields for identification purposes; this considers the redundancy problem created by simultaneously increasing visual words. Using learning receptive fields, inefficient and redundant codewords and pooling regions were found and subsequently eliminated from the pooling region, this made the pipeline more compact and separable for the specified classification task. The experiments show that the modified learning algorithm is effective and the learned pipeline useful for both structural simplification and improving classification accuracy compared with the baseline method.

Keywords: bag-of-features model; receptive field learning; object recognition; image classification; feature learning

在计算机视觉和模式识别领域, 图像识别是一类相当常见的问题。它的作用是预测一幅图像的分类标签, 或者标注出图像内容的属性。在使用分类器对图像所属类别进行预测之前, 一般会使用合适的特征对图像进行描述。特征词袋(bag-of-features,

BoF)是一种常用的简洁而高效的图像中级特征(mid-level feature)学习模型。一个基本的BoF模型通常包含5个部分, 提取图像块、描述图像块、视觉词典学习编码(coding)和池化(pooling)^[1]。其中, 视觉词典学习作为BoF模型的核心, 集中了大量的研究工作, 有不少基于监督^[2]和无监督^[3]的视觉词典学习算法被提出, 比如K-means^[4]、稀疏编码^[5]等都可作为词典学习算法被集成在BoF模型中。对

收稿日期: 2016-01-01. 网络出版日期: 2016-07-18.

基金项目: 国家自然科学基金项目(61371182).

通信作者: 赵骞. E-mail: zhokya@gmail.com.

于如何找到更好的池化区域,来生成包含更多局部性信息的图像特征,同样吸引了众多研究者的目光。其中,空间金字塔模型(spatial pyramid matching)是一种常用的区域选取策略,他将特征图分割成为一个一个逐层细化的空间子区域,之后在这些子区域上分别进行池化操作^[6]。Coates 等^[7]通过对特征图之间相似性的分析,使用贪婪方法,将相似性较高的特征图分为一组,从而使得特征表达获得了一定的旋转不变性,Jia 等^[8]对于空间金字塔模型进行了进一步的扩展,定义了一组超完备(over-complete)的池化区域,该区域是由所有在特征图上可能存在的矩形区域组成,并在此基础上提出了感受野学习(receptive field learning)的概念,通过在超完备词典生成的特征图上的超完备区域中选择那些最为有效区域。在这些被选择的池化区域上计算图像的特征表达。

本文的工作正是基于 Jia 等^[9]提出的超完备区域的感受野学习方法的进一步改进,在感受野学习的同时考虑对视觉词典规模进行限制,同时实现 BoF 模型中的视觉单词和池化区域的精简,从而到达优化 BoF 模型结构的目的。另外,与其他的词典简化方法相比,该方法充分考虑到了池化的作用和分类任务对视觉单词的需求。

1 Bag-of-features 模型

BoF 模型的核心思想是将图像表示成为一组无序特征的集合,通过统计每一种特征的响应强度来构造一个描述图像的直方图,即图像的中级特征表达。利用图像特征响应的直方图统计,可以训练分类器对图像中所表达的内容进行识别。BoF 模型具有简洁、高效等特点,但同时 BoF 模型在最初提出时是为了解决图像检索问题,对每种特征在整幅图像中的出现次数进行统计,无法对结构信息和空间信息进行有效表示^[9]。

BoF 模型图像分类框架如图 1 所示。

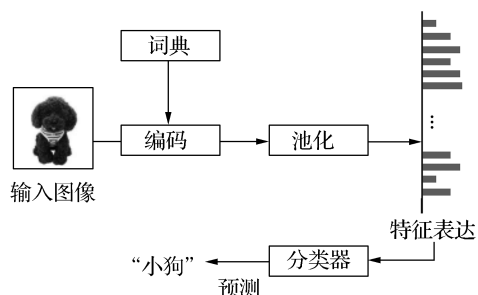


图 1 面向图像识别的 Bag-of-features 模型框架图

Fig.1 Bag-of-features pipeline for image classification

通常将输入图像(或低级特征,比如 SIFT、HOG 等)使用视觉词典中的视觉单词进行编码,从而得到一组特征响应的分布图。相对于 BoF 模型在图像检索领域的应用,局部空间的特征直方图对于图像识别具有更重要的作用。因此对于特征图进行分块池化操作,可以使特征表达中包含多个区域、多个尺度上的特征统计信息。

1.1 词典学习

建立 BoF 模型的第一步就是使用学习算法获得一个视觉词典(codebook)。为此,要从训练样本中随机提取图像块(patch)。然后选择一个学习算法,使用这些图像块作为算法的训练样本。常用的词典学习算法有很多,总体上可分为有监督和无监督 2 类。在这里介绍两种高效的无监督词典学习方案,K-means 和 OMP-K 学习算法^[10]。

$$\min_D \sum_{i=1}^K \min_j \| \mathbf{d}^{(j)} - \mathbf{x}^{(i)} \| \quad (1)$$

K-means 聚类算法^[11]是一种快速而且易实现的学习算法,其通过最小化式(1)中的平方距离来搜索训练数据的 K 个聚类中心,从而可以得到一个大小为 K 的词典。在式(1)中, $\mathbf{x}^{(i)}$ 是由输入的图像块或者初级特征组成的向量, $\mathbf{d}^{(j)}$ 是所求的聚类中心,每一个聚类中心可以被视为一个视觉单词,他们共同构成了一个完整的大小为 K 的词典 \mathbf{D} 。

OMP-K 是另外一种无监督学习算法,该算法通过交叉迭代公式(2)中的损失函数,在服从约束条件的前提下最小化重构误差,来获得词典 \mathbf{D} 。它与 K-means 算法的主要区别在于,在学习阶段每个图像块最多使用 K 个视觉单词计算最小误差,而 K-means 算法每个图像块仅属于一个视觉单词(聚类中心)。

$$\begin{aligned} \min_{\mathbf{D}, \mathbf{s}^{(i)}} \sum_i \| \mathbf{D} \mathbf{s}^{(i)} - \mathbf{x}^{(i)} \|^2 \\ \text{subject to } \| \mathbf{d}^{(j)} \|^2 = 1, \forall j \\ \text{and } \| \mathbf{s}^{(i)} \|_0 \leq k, \forall i \end{aligned} \quad (2)$$

式中: $\| \mathbf{s}^{(i)} \|_0$ 使用 0 范数表示编码 $\mathbf{s}^{(i)}$ 中非零元素的数量,每个视觉单词 $\mathbf{d}^{(j)}$ 为词典 \mathbf{D} 中一列。使用 OMP-K 算法计算式(2),可以得到词典 \mathbf{D} 。

1.2 特征编码

在完成词典学习之后,需要对输入的数据(图像或者初级特征)进行编码。软阈值(soft thresholding)编码通过计算数据与视觉单词之间的内积对数据进行编码,如式(3)。其中,阈值 α 的引入,为编码后的特征带来一定的稀疏性^[12]。

$$f_i = \max_{d^{(j)} \in D} \{0, d^{(j)T} \mathbf{x} - \alpha\} \quad (3)$$

但是,软阈值方法需要手工指定阈值 α ,为了克服手工指定参数对编码效果的影响。Coates 等^[4]提出了一种三角编码算法 (triangle)。该方法中,当数据 $\mathbf{x}^{(i)}$ 到视觉单词 $\mathbf{d}^{(k)}$ 的距离大于数据 $\mathbf{x}^{(i)}$ 到所有视觉单词 $\mathbf{d}^{(k)}$ 的平均距离 $\mu(z)$ 时,特征对应的值为 0,

$$f_k = \max \{0, \mu(z) - z_k\} \quad (4)$$

式中: $z_k = \|\mathbf{x}^{(i)} - \mathbf{d}^{(k)}\|_2$ 表示数据到视觉单词的欧式距离, $\mu(z)$ 是 z_k 的平均值。

每一个样本在进行编码后,会得到 K 个特征图。这里, K 为词典中视觉单词的个数。

1.3 池化

池化作为 BoF 模型中的一个关键步骤,它通过聚合运算(比如:计算最大值或均值)将一个矩形区域内的特征转换成为一个标量值,从而减少特征的数量。除了降低特征表达的维数之外,池化能够为特征表达带来一些非常有用的特性,比如平移不变性。

另外,为获得更加丰富的特征空间分布的统计信息,规则网格结构、空间金字塔、超完备(如图 2)等池化区域方案被相继提出,使 BoF 特征保留了更多的局部统计信息。

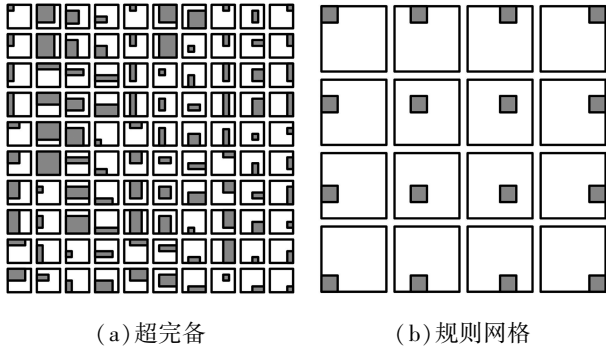


图 2 Pooling 区域

Fig.2 Pooling regions

通常,池化可以用如式(5)的形式进行表示^[13]

$$\mathbf{x}_i^c = \text{op}(A_{R_i}^c) \quad (5)$$

式中: \mathbf{x}_i^c 是池化后得到的特征, c 和 R_i 分别表示特征图对应的视觉单词的下标,以及特征图上的池化的区域, $A_{R_i}^c$ 表示池化区域内的一组特征的集合, $\text{op}(\cdot)$ 表示在 $A_{R_i}^c$ 之上进行的聚合。

2 改进的感受野学习算法

本节的算法是在 Jia 等^[8]的基础上,增加对于视觉词汇数量增长的限制,从而实现感受野学习和选择有效视觉单词的双重效果。算法中使用了 Perkins 等^[14]提出的 grafting 方法对特征的重要性进行

判断, grafting 全称梯度特征测试 (gradient feature testing),需要候选特征的梯度值作为重要性判断的依据,因此要首先定义分类器及其目标函数。

2.1 分类器

本文定义了一个多类线性分类器,使用 BoF 模型特征对标签进行预测,

$$\mathbf{y} = \mathbf{W}^T \mathbf{x} + \mathbf{b} \quad (6)$$

式中: \mathbf{x} 是输入图像的特征表达向量, \mathbf{W} 和 \mathbf{b} 分别为权值矩阵和偏置, \mathbf{y} 为分类器的预测结果。式(7)定义了分类器的目标函数,通过最优化算法求出目标函数的最优解,来确定参数 \mathbf{W} 和 \mathbf{b}

$$\min_{\mathbf{W}, \mathbf{b}} \frac{1}{N} \sum_{n=1}^N l(\mathbf{W}^T \mathbf{x}_n + \mathbf{b}, \mathbf{y}_n) + \lambda \|\mathbf{W}\|_{\text{Fro}}^2 \quad (7)$$

式中: \mathbf{x}_n 和 \mathbf{y}_n 分别表示训练样本和对应的标签, N 是用于训练的样本总数。 $\|\mathbf{W}\|_{\text{Fro}}^2$ 是权值矩阵 \mathbf{W} 的 F-范数的平方,用于限制权值矩阵中元素值的大小,起到了降低结构复杂度和抑制过拟合的作用。上式第一项中的 $l(\cdot)$ 表示多类扩展二项负对数似然函数 (BNLL),用于衡量线性分类器对于训练数据的适应程度,其具体定义为

$$l(\mathbf{W}^T \mathbf{x} + \mathbf{b}, \mathbf{y}) = \sum_{i=1}^P \ln(1 + e^{-y_i(\mathbf{W}_{:,i}^T + b_i)}) \quad (8)$$

式(8)中的 $\ln(\cdot)$ 表示自然对数函数, P 是指标签的类别总数, $\mathbf{W}_{:,i}$ 表示矩阵 \mathbf{W} 中的一列。

2.2 感受野学习

为了找到最为合适的感受野,将通过计算目标函数相对于每个特征对应的权值矩阵元素的梯度值来判断特征的重要性。首先,初始化一个空集合 S 来保存已选择的特征,所有没有被选择的特征组成一个候选集合 S_c 。在每次迭代中,针对候选集合 S_c 中每个特征计算一个分值,如式(9)所示。选择分值最大的特征加入集合 S ,并将其从 S_c 中移除。直到集合 S 中特征的数量达到一个指定的值,迭代停止。此时集合 S 中特征所对应的感受野既是针对该分类任务最为有效的感受野。

与 Jia 等人提出的感受野学习方法^[8]不同,本方法在计算特征分值的基础上进一步对特征图进行了区分,将未被选入的特征图所包含特征的分值进行衰减,使感受野学习的过程更加倾向于选取已被选择的特征图上的特征,如式(9)

$$\text{score}(j) = \left\| \frac{\partial L(\mathbf{W}, \mathbf{b})}{\partial \mathbf{W}_{j,:}} \right\|_{\text{Fro}}^2 \cdot (1 - \eta M(j)) \quad (9)$$

式中: $L(\mathbf{W}, \mathbf{b})$ 是式(6)中所定义的目标函数, $\mathbf{W}_{j,:}$ 表示矩阵 \mathbf{W} 中的一行, η 是针对未选中特征图中候选

特征的衰减因子,其取值范围为 $[0, 1)$, $M(j)$ 用于表示当前第 j 个特征是否在选中的特征图之上,其表示为

$$M(j) = \begin{cases} 1, & \text{特征 } j \text{ 在已选特征图} \\ 0, & \text{其他} \end{cases} \quad (10)$$

式中:如果特征图上含有已选中特征,则认为该图被选取,其 $M(j)$ 值为 1。

在每一次迭代后,新的特征被加入集合 S 。在此特征集合之上,需要对分类器参数 W 和 b 进行重新计算。将上次迭代计算出的 W 和 b 作为重新计算时的初始状态,以有效地减少计算参数所消耗的时间。

3 实验与分析

3.1 实现过程

实验中主要使用了 CIFAR-10 数据库^[19]对文中提出的算法进行训练和验证。CIFAR-10 是由加拿大多伦多大学教授 Hinton 等发布的一个用于图像识别算法评估的数据库,该数据库中包含了 50 000 个训练样本和 10 000 个测试样本,共分为 10 类。

在使用样本训练和测试前,都要对样本数据进行归一化和白化处理,这样能够提高 BoF 模型特征的预测性能。

从训练样本中随机提取 5×5 的图像块用于词典学习,选择 K-means 作为词典学习的算法,以及 triangle 作为编码算法。输入图像经编码后会形成 28×28 的特征图,然后最大值池化会将每张特征图上 7×7 相邻但不重叠的空间区域聚合成为一个特征,因此每张特征图被降维到 4×4 的大小。超完备的池化区域以及对照实验都将基于这些池化后的特征图。

在训练分类器时,式(7)中的参数 λ 固定为 0.01,用于超完备池化区域和感受野学习实验。对于其他预定义的池化区域(比如网格结构和空间金字塔),在实验中使用 L2-SVM 作为分类器,其参数通过五折交叉验证的方式搜索确定。

3.2 算法性能测试

表 1 将本文提出的改进的感受野学习算法,同原有方法^[8]以及几种常用的池化策略进行了比较。表中包括网格结构的池化区域(2×2 , 4×4),空间金字塔(spatial pyramid matching, SPM),超完备感受野(over-complete, OC),基于超完备感受野的感受野学习(OC+学习),本文方法是在超完备感受野上的实验结果(OC+改进算法)。

表 1 多种 pooling 策略的比较

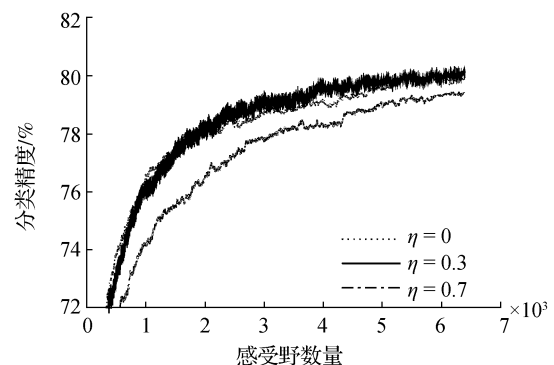
Table 1 Comparison between different pooling strategies

池化区域	特征数	词典规模	衰减因子	精度/%
2×2	800	200	—	65.73
4×4	3 200	200	—	73.25
SPM	4 800	200	—	74.01
OC ^[8]	20 000	200	—	76.44
OC+学习 ^[8]	6 400	200	—	76.72
OC+改进算法	4 408	200(400)	0.7	76.95
OC+改进算法	5 370	200(800)	0.8	77.89

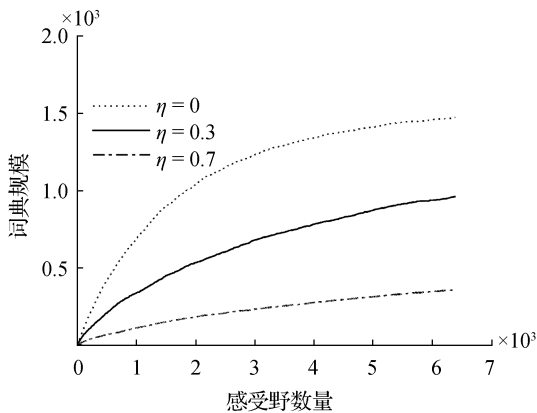
在未使用学习策略的池化方案中,由于超完备感受野能够较好地捕获特征图中的空间信息,因此得到了较高的分类精度。在增加学习策略之后冗余被消除,超完备感受野又获得了 0.28% 的提升。在此基础上使用本文方法,以一个较大的词典作为初始词典,同时在学习过程中对词典规模的增长进行了限制。例如在表 1 中,分别以 400 和 800 为初始词典大小,将衰减因子 η 设为 0.7 和 0.8,在词典规模为 200 时,学习到了 4 408 和 5 370 个特征。尽管使用的特征数有所减少,但是分类结果却比原有算法提高了 0.23% 和 1.27%。该结果说明了改进算法的有效性。

3.3 参数 η 的影响

本节实验中,在 CIFAR-10 数据库上比较了 3 种衰减因子 η 值下,学习算法所受到的影响(如图 3)。首先,使用 K-means 算法得到一个大小为 1 600 的词典。然后,以该词典生成的特征图为基础,分别将 η 值设为 0、0.3 和 0.7,观察和比较在学习过程中,精度和词典规模的增长。其中,当 η 值设为 0 时(即不对词典的增长进行衰减),则本方法与 Jia 等提出的学习方法^[8]等价。



(a) 样本数量对分类精度的影响



(b) 样本数量对词典规模的影响

图 3 不同 η 值对应的分类精度及词典规模增长曲线

Fig.3 The curve of classification accuracy and codebook scale with different η

在图 3(a) 是分类精度跟随感受野数量(即感受野对应的特征数)增长的曲线,图 3(b) 是视觉单词量随感受野数量增长的曲线。可以看出,在这 3 种 η 值情况下,精度的增长受到的影响不大,而视觉单词数量却受到了明显抑制,说明算法能够对于分类贡献较低的视觉单词进行识别。

在表 2 中,池化区域为 OC,比较了图 3 中的实验在学习 6 400 个感受野时的状态。从表中可以看出,随着 η 值的增加,所需的视觉单词的数量不断下降。而且对于特征图数量增长一定程度上的限制,对分类结果也会有一些提高。现象也说明了基于梯度的贪婪算法并不能确保针对分类任务获得一个最优的特征集合。

表 2 CIFAR-10 数据库的实验结果

Table 2 Results on CIFAR-10

词典规模	衰减因子	精度/%
1 477(1 600)	0	79.80
963(1 600)	0.3	80.13
357(1 600)	0.7	79.39

使用 CIFAR-100 数据库重复上述实验。这里,由于数据库有 100 类对象,更大的权值矩阵 W 使得算法对 η 值变得更敏感,因此实验中选择 η 值为 0、0.1 和 0.3。在表 3 中,可以看到与 CIFAR-10 上类似的实验结果。

3.4 样本数量的影响

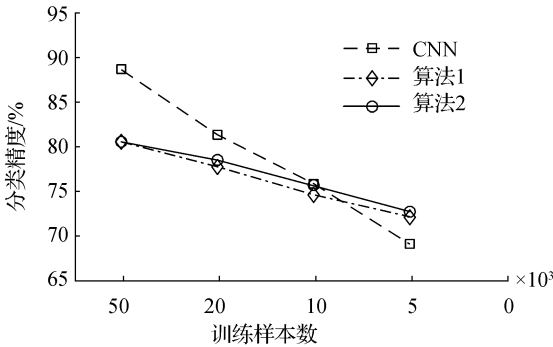
为了研究训练样本数量对本文所描述的算法的影响程度,随机从 CIFAR-10 数据库中选择出多个样本子集作为训练数据,使用这些子集学习词典和感受野。除此之外,使用同样样本子集训练卷积神

经网络(convolutional neural networks, CNN)^[15],与本文中提出的算法进行对比。

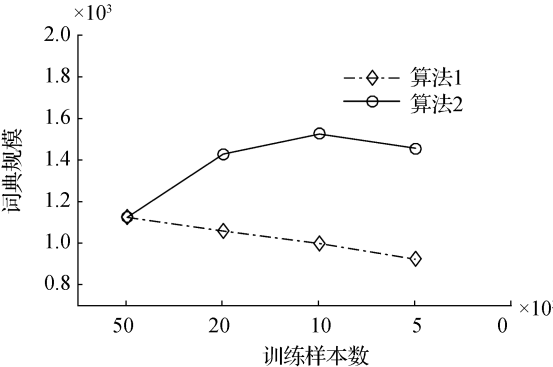
表 3 CIFAR-100 数据库的实验结果

Table 3 Results on CIFAR-100

词典规模	衰减因子	精度/%
1 273(1 600)	0	54.88
821(1 600)	0.1	55.04
313(6 400)	0.3	54.28



(a) 样本数量对分类精度的影响



(b) 样本数量对词典规模的影响

图 4 训练样本数量对分类精度和词典规模的影响

Fig.4 Effects of training sample size on classification accuracy and codebook scale

作为参照对象的 CNN 使用了 Lin 等人提出的 Network in Network 模型^[16]。该模型曾在 CIFAR-10 数据库上获得了准确率 89.21% 的最好成绩。在实验中,逐渐的减少训练样本的数量,使用了完整的 50 000 个样本,以及随机抽取的 20 000、10 000 和 5 000 个样本分别进行试验。在本文算法的实现过程中,一组实验采用相同的初始词典大小,对于不同大小的训练集均取 1 600 个视觉单词,称为算法 1。另一组随着样本数量的减少逐步扩大初始词典的规模,分别使用 1 600、3 200、4 800 和 6 400 个视觉单词,称为算法 2。实验中的 η 被固定为 0.3。

在图 4(a) 中, 可以发现在训练集规模较大的情况下, CNN 模型具有更高的分类精度。伴随着训练样本数量的减少 CNN 模型的性能逐渐下降。当训练样本数量为 10 000 时, CNN 模型同本文中的 BoF 模型性能相当。另外, 算法 2 由于逐步提高了初始词典的大小, 对于分类性能的损失产生了一定的补偿作用, 因此略好于使用固定初始词典大小的算法 1 的结果。

在图 4(b) 中, 学习后的词典规模稳定在一定范围内。尽管算法 2 的初始词典大小在不断增加, 但是实际使用的词典并没有随着初始词典而膨胀。因此, 在计算能力允许的情况下, 可以使用较大的初始字典, 本文提出的方法可以学习到一个大小合适的模型。

通过以上分析, 认为针对小样本集的目标识别问题, BoF 模型依然是一种十分有效的工具。

3.5 分析讨论

实验验证了方法的有效性, 可将感受野学习和选择有效视觉单词相结合, 用于对 BoF 模型的精简。尽管对于词典进行了大量的删减, 但是对于精度仅产生了较小影响, 从侧面说明了模型中包含了很多与分类无关的冗余参数。

然而, 本文提出的方法仍有几点不足需要进一步改进。首先, 通常对于参数 η 的调节缺少指导性的原则, 只能通过尝试多个 η 值来估计合适的取值。其次, 如何指定一个合适的特征数, 同样缺乏指导原则。一般来说, 可以指定一个较大的特征数, 观测算法在验证集上的效果, 从而做出判断。

4 结束语

本文通过对现有的感受野学习方法的分析和研究的基础上, 提出了一种改进型的感受野学习算法。相比于原有算法, 本文提出的方法除了考虑生成特征的空间区域和词典对于分类的影响之外, 还将去除词典中低效的视觉单词纳入学习过程。使用此方法, 从词典大小和池化区域两方面对 BoF 模型的规模进行了精简。实验结果表明, 本文提出的方法能够利用更少的词典规模形成更有效的特征表达, 即使在使用相同初始词典的情况下, 也可以较少冗余成分对于分类任务的干扰, 提高 BoF 模型生成的特征表达的表达能力, 从而提高分类精度。特别是在训练样本较小的情况下, 具有一定的应用价值。

参考文献:

- [1] HUANG Yongzhen, WU Zifeng, WANG Liang, et al. Feature coding in image classification: a comprehensive study [J]. IEEE transactions on pattern analysis and machine intelligence, 2013, 36(3): 493-506.
- [2] YANG Jianchao, YU Kai, GONG Yihong, et al. Linear spatial pyramid matching using sparse coding for image classification[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Miami, FL, 2009: 1794-1801.
- [3] YU Kai, ZHANG Tong, GONG Yihong. Nonlinear learning using local coordinate coding[C]//Advances in Neural Information Processing Systems 22: 23rd Annual Conference on Neural Information Processing Systems. Vancouver, British Columbia, Canada, 2009: 2223-2231.
- [4] COATES A, NG A, LEE H. An analysis of single-layer networks in unsupervised feature learning[J]. Journal of machine learning research, 2011, 15: 215-223.
- [5] GREGOR K, LECUN Y. Learning fast approximations of sparse coding[C]//Proceedings of the 27th International Conference on Machine Learning. Haifa, Israel, 2010.
- [6] LAZEBNIK S, SCHMID C, Ponce J. Beyond bags of features: spatial pyramid matching for recognizing natural scene categories [C]//Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. New York, NY, USA, 2006, 2: 2169-2178.
- [7] COATES A, NG A Y. Selecting receptive fields in deep networks[C]//Advances in Neural Information Processing Systems 24: 25th Annual Conference on Neural Information Processing Systems. Granada, Spain, 2011: 2528-2536.
- [8] JIA Yangqing, HUANG Chang, DARRELL T. Beyond spatial pyramids: receptive field learning for pooled image features [C]//Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition. Providence, RI, 2012: 3370-3377.
- [9] SIVIC J, ZISSERMAN A. Video google: a text retrieval approach to object matching in videos[C]//Proceedings of the Ninth IEEE International Conference on Computer Vision. Nice, France, 2003: 1470-1477.
- [10] COATES A, NG A Y. The importance of encoding versus training with sparse coding and vector quantization [C]//Proceedings of the 28th International Conference on Machine Learning. Bellevue, WA, USA, 2011.
- [11] JAIN A K. Data clustering: 50 years beyond K-means[J]. Pattern recognition letters, 2010, 31(8): 651-666.
- [12] 汪启伟. 图像直方图特征及其应用研究[D]. 合肥: 中

国科学技术大学, 2014.

WANG Qiwei. Study on image histogram feature and application[D]. Hefei, China: University of Science and Technology of China, 2014.

[13]BOUREAU Y L, ROUX N L, BACH F, et al. Ask the locals: multi-way local pooling for image recognition[C]//Proceedings of the 2011 International Conference on Computer Vision. Barcelona, Spain, 2011: 2651-2658.

[14]PERKINS S, LACKER K, THEILER J. Grafting: fast, incremental feature selection by gradient descent in function space[J]. Journal of machine learning research, 2003, 3: 1333-1356.

[15]LECUN Y, BOSER B E, DENKER J S, et al. Handwritten digit recognition with a back-propagation network[C]//Advances in Neural Information Processing Systems 2: 3rd Annual Conference on Neural Information Processing Systems. Vancouver, British Columbia, Canada. San Francisco, CA, USA, 1989: 396-404.

[16]KRIZHEVSKY A. Learning multiple layers of features from tiny images[D]. Toronto, Canada: University of Toronto, 2009.

作者简介:



赵骞,男,1986年生,博士研究生,主要研究方向为计算机视觉、神经网络。参与“863”项目1项,国家自然科学基金项目1项。



李敏,男,1981年生,讲师,博士,主要研究方向为仿生机器人、外骨骼机器人。参与“863”项目2项。曾获得教育部技术发明奖一等奖1项,授权国家发明专利5项,发表学术论文7篇。



赵晓杰,男,1972年生,博士研究生,主要研究方向为航迹规划、传感器网络,参与“973”项目1项。

2017 第2届 IEEE 大数据分析国际会议 2017 the 2nd IEEE International Conference on Big Data Analysis (ICBDA 2017)

我国已全面迈入4G时代,大数据、云计算等产业关键技术演进和应用创新加速发展,产业政策规划纷纷出台,数据开放共享机制不断推进,产业发展基础环境初步具备,产业生态体系逐步完善建立,大数据产业将进入理性发展、落地应用阶段,并有望在通信、互联网、金融、医疗等重点行业率先取得突破,将成为全球最重要的大数据市场。大数据带来了全新的变革与机遇,但我国大数据产业还处于发展初期,总体战略规划仍有待部署,行业标准化和规模化应用还未实现,核心技术和数据资源开放有待突破,行业协作和商业模式瓶颈尚待解决,数据安全与隐私保护形势严峻等诸多挑战。

为了迎接大数据时代的到来,更好地交流大数据的实践经验,进一步推进大数据技术创新与应用,展示国内外大数据领域的最新科研成果。2017年第二届大数据分析国际会议将于2017年3月10日-12日在中国北京召开。本届ICBDA 2017由IEEE和西交利物浦大学大数据分析研究所联合主办,University of Texas at Dallas, USA 协办,由西交利物浦大学的关圣威教授,和 University of Texas at Dallas 张康教授担任联合大会主席。会议旨在促进大数据分析等领域的学术交流与合作,热忱欢迎从事相关技术研究的专家、学者和专业技术人员踊跃投稿并参加大会。

会议网站:<http://www.icbda.org/index.html>