

# 基于深度学习的鱼类分类算法研究

顾郑平 朱敏\*

(华东师范大学计算机科学与软件工程学院计算中心 上海 200062)

**摘要** 回顾近年来国内外对鱼类分类的研究进展,指出传统方法存在的缺陷。深度学习是目前图像分类的主流方法。研究基于卷积神经网络 CNN( Convolutional Neural Network) 的鱼类分类模型,并以该模型为基础,进一步提出利用迁移学习,以预训练网络的特征结合 SVM 算法( PreCNN + SVM) 的混合分类模型。实验以 Fish4-Knowledge ( F4K) 作为数据集,使用 TensorFlow 训练网络模型。实验结果表明,利用 PreCNN + SVM 算法,取得了 98.6% 的准确率,较传统方法有显著提高。对于小规模数据集,有效解决了需要人工提取特征的不可迁移性。

**关键词** 深度学习 卷积神经网络 迁移学习 支持向量机

中图分类号 TP3

文献标识码 A

DOI: 10.3969/j.issn.1000-386x.2018.01.035

## FISH CLASSIFICATION ALGORITHM BASED ON DEPTH LEARNING

Gu Zhengping Zhu Min\*

( Computer Center, School of Computer Science and Software Engineering, East China Normal University, Shanghai 200062, China)

**Abstract** Reviewing the research progress of fish classification at home and abroad in recent years, the shortcomings of the traditional methods are pointed out. Deep learning is the mainstream method of image classification. This paper studied the fish classification model based on Convolutional Neural Network ( CNN) . Based on this model, we proposed a hybrid model based on transfer learning algorithm using pre-training neural network and SVM algorithm ( PreCNN + SVM) . The experimental results showed that using PreCNN + SVM algorithm with Fish4-Knowledge ( F4K) data sets and TensorFlow model, the accuracy of 98.6% was achieved, which was significantly higher than the traditional method. For small-scale data sets, it effectively solved the immutability that need to extract features manually.

**Keywords** Deep learning Convolutional neural network Transfer learning SVM

## 0 引言

鱼类在人类生活中扮演着非常重要的角色,也是人类重要的资源之一。根据加拿大学者 Nelson<sup>[1]</sup> 统计,全球已知鱼类约有 28 000 种,占已命名脊椎动物一半以上,且新种鱼类不断被发现。目前全球已命名的鱼种约在 32 100 种。鱼类识别应用广泛,可用于水产、生物、海洋、环境的研究、开发、管理等。对各种鱼类分门别类地建立数据库,利用人工智能方法自动识别鱼类,不仅能够更好地开发利用和保护鱼类资源,也为发展海洋渔业生产发挥了积极的作用,对学术研究

和经济价值具有重大意义。

鱼的形状各种各样,有时相差大,但总的来说大多数鱼呈细长的流线形状,分析起来比较复杂。基于计算机视觉的鱼类分类研究已长达二十多年。多年来,已经有许多学者对其进行了深入研究。1990 年,Strachan 等<sup>[2]</sup> 利用形状特征对鱼进行分类。Larsen 等<sup>[3]</sup> 提取了三个鱼种的形状和纹理特征,使用 LDA 方法对 108 张图像数据集进行分类,准确率为 76%。2013 年,Huang 等<sup>[4]</sup> 提出了 Balance-Guaranteed Optimized Tree( BGOT) 算法,采集了 10 个鱼种 3 179 张图像数据进行实验,并得到 95% 的准确率。吴一全等<sup>[5]</sup> 于 2014 年提出了基于 LS-SVM 的识别方法,识别率达

95.83% 以上。2015 年,杜伟东等<sup>[6]</sup>提出了基于 SVM 的决策融合鱼类识别方法,识别率在 90% 以上。

尽管经过长期的发展,基于计算机视觉的鱼类分类方法研究已经取得了重大进展,但是依然存在各种各样的问题。传统的鱼类分类方法一般采用两步法:

- (1) 从输入的鱼图像中计算出人为设置的特征。
- (2) 根据得出的特征去训练一个分类器,用于测试数据的分类。这种方法表现的好坏很大程度上取决于人为选择的特征是否合理,而人在选择特征时往往都是靠经验,具有很大的盲目性。虽然现在利用人为设置的特征进行分类也取得了较好的结果,但是这些特征都是针对特定数据设计的,如果用同样的特征来处理不同的数据集,结果可能大相径庭,因此这种特征具有不可迁移性。

近年来,深度学习在人工智能领域取得重要突破。本文主要研究利用深度学习算法直接在图像上进行自动特征学习与分类。首先研究了基于卷积神经网络的鱼图像分类算法,并在此基础上,进一步提出了以预训练网络的特征结合 SVM 算法(PreCNN + SVM)进行分类的混合模型。利用深度学习算法很好地解决了传统鱼类分类方法存在的缺陷且分类效果优于传统方法。

## 1 相关工作

深度学习是目前图像分类的主流方法,取得比较好的效果,其核心是卷积神经网络。迁移学习能够迁移已有的知识来解决目标领域中仅有少量有标签的样本数据甚至没有的学习问题。

### 1.1 卷积神经网络

卷积神经网络最早由 Le Cun 等在 1998 年提出来<sup>[7]</sup>,并应用在手写数字识别上。随着计算能力的提高,使得大型复杂的模型计算变得可能。AlexNet 模型<sup>[8]</sup>的提出,开启了计算机视觉深度网络的新时代。卷积神经网络在结构上具有局部连接,权重共享及空间或时间上的子采样三个特性,这些特性使得卷积神经网络具有一定程度上的平移、放缩和扭曲不变性。

在图像处理中,图像是以二维矩阵的形式输入到神经网络中。对于二维卷积,假设  $x^l \in (w_l \times h_l)$  和  $x^{l-1} \in (w_{l-1} \times h_{l-1})$  分别是第  $l$  层和第  $l-1$  层的神经元活性。 $X^{(l)}$  的每一个元素为:

$$X_{s,t}^{(l)} = f\left(\sum_{i=1}^u \sum_{j=1}^v w_{i,j}^{(l)} \cdot X_{s-i+u, t-j+v}^{l-1} + b^{(l)}\right) \quad (1)$$

式中:  $W^{(l)} \in (u \times v)$  为二维的滤波器,  $b^{(l)}$  为第  $l$  层的偏置向量。每个滤波器大小为  $u \times v$ 。

卷积层的作用是提取一个局部区域的特征,每一个滤波器相当于一个特征提取器。对于卷积层得到的一个特征映射  $X^{(l)}$ ,可以将其划分为很多区域  $R_k, k = 1, 2, \dots, K$ 。

一个子采样函数  $Sub(\cdot)$  定义为:

$$X_k^{(l+1)} = f(w^{(l+1)} \cdot Sub(R_k) + b^{(l+1)}) \quad (2)$$

式中:  $w^{(l+1)}$  和  $b^{(l+1)}$  分别是可训练的权重和偏置参数。

$$X^{(l+1)} = f(w^{(l+1)} \cdot Sub(X^{(l)}) + b^{(l+1)}) \quad (3)$$

其中,  $Sub(X^{(l)})$  是子采样后的特征映射。

子采样层可以来大大降低特征的维数,避免过拟合,同时还可以使得下一层的神经元对一些小的形态改变保持不变。

### 1.2 迁移学习

将知识迁移到新环境中的能力通常被称为迁移学习。迁移学习<sup>[9]</sup>涉及到域和任务的概念。一个域  $D$  由一个特征空间  $\chi$  和特征空间上的边际概率分布  $P(X)$  组成,即  $D = \{\chi, P(X)\}$ , 其中  $X = x_1, x_2, \dots, x_n \in \chi$ 。对于给定的一个域,一个学习任务  $T$ ,由两部分组成,即标签  $Y$  和目标预测函数  $f(\cdot)$ ,  $T = \{Y, f(\cdot)\}$ 。

给定一个源域  $D_s$  和一个学习任务  $T_s$ ,一个目标域  $D_t$  和一个目标学习任务  $T_t$ 。迁移学习的目的是利用  $D_s$  和  $T_s$  中的知识,来提高目标预测函数  $f(\cdot)$  在  $D_t$  中的性能,其中  $D_s \neq T_s$  或  $D_t \neq T_t$ 。

Maxime 等<sup>[10]</sup>提出迁移卷积神经网络的权重值的方法在 VOC2007 和 VOC2012 上取得了很好的效果。

## 2 基于深度学习的鱼类分类算法

本章研究了基于卷积神经网络的鱼图像分类算法,并在此基础上,研究利用迁移学习,提出了以预训练网络的特征结合 SVM 算法(PreCNN + SVM)的混合模型,能够将适用于大数据的模型迁移到小数据上,实现个性化迁移。

### 2.1 基于卷积神经网络的分类模型

针对传统的鱼类分类算法,本节研究并实现一种基于卷积神经网络的鱼类分类算法。

#### 2.1.1 CNN 模型架构

模型由两个卷积层、两个池化层、一个全连接层和一个 Softmax 回归层组成。使用 Adam<sup>[11]</sup> 基于一阶梯度的随机目标函数算法来优化整个神经网络。模型架构如图 1 所示。

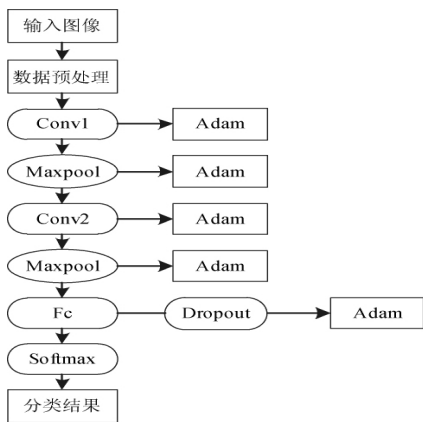


图1 多层架构的卷积神经网络模型

### 2.1.2 权值初始化

为了创建这个模型,需要创建大量的权重和偏置项。这个模型中的权重在初始化时加入少量的噪声来打破对称性以及避免0梯度。模型使用 ReLU<sup>[12]</sup> 激活函数,用一个较小的正数来初始化偏置项,以避免神经元节点输出恒为0的问题。ReLU 激活函数定义为:

$$f(x) = \max(0, x) \quad (4)$$

其中,  $x$  为输入的神经元。

### 2.1.3 卷积和池化

卷积层的任务就是从不同的角度来选择前一层特征图各角度的特征。在获取到卷积特征后,把卷积特征划分到不相交区域上,然后用这些区域的最大特征来获取池化后的卷积特征。第一层由一个卷积接一个 max pooling 完成。卷积在每个  $5 \times 5$  的卷积核中得到 32 特征,对于每一个输出通道都有一个对应的偏置量。同样,第二层中,每个  $5 \times 5$  的卷积核得到 64 个特征。

### 2.1.4 全连接层

在这一层,加入一个有 1 024 个神经元的全连接层,用于处理整个图片。同时将池化层输出的向量转换成一维向量,乘上权重矩阵,加上偏置,并使用 ReLU 激活函数。

### 2.1.5 输出层

为了减少过拟合,在输出层之前加入 Dropout<sup>[13]</sup>。Dropout 可以有效提高大型神经网络的泛化能力。

最后添加一个 Softmax 层。在 Softmax 回归中,假设类标签  $y$  可以取  $k$  个不同的值。对于给定的测试输入  $x$ ,假设函数  $p(y = j | x; w)$  是  $x$  属于类别  $j$  的概率值:

$$p(y = j | x; w) = \frac{\exp(w_k^T x)}{\sum_{j=1}^k \exp(w_j^T x)} \quad (5)$$

式中:  $w$  为模型的参数。

## 2.2 PreCNN + SVM 混合分类模型

深度学习算法能够有效的关键是大规模的数据,

而对于小规模的数据集,为了提高其识别的准确率,往往需要不断调整训练参数以及构建更深层次的模型,这样会增加参数的规模和计算复杂度。

由图1可以看到,卷积神经网络的结构主要由两部分构成:(1) 一系列卷积层。(2) 一个或多个全连接神经网络。卷积层用于特征学习,然后将学习好的特征传入全连接层进行分类。

本小节研究利用迁移学习,将适用于大数据的模型迁移到小数据上,并将 SVM 分类器代替 Softmax 分类器。

### 2.2.1 PreCNN + SVM 模型架构

本文使用的预训练深层模型为 Inception-V3<sup>[14]</sup>。该模型由 Google 2014 年发布,利用该网络在 ImageNet 上预先训练好的参数权重,最后将学习特征输入到 SVM 分类器中进行分类。PreCNN + SVM 模型结构如图2所示。

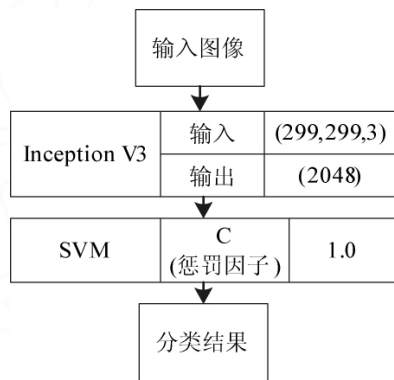


图2 PreCNN + SVM 模型架构

其中 Inception-V3 的网络结构如图3所示。

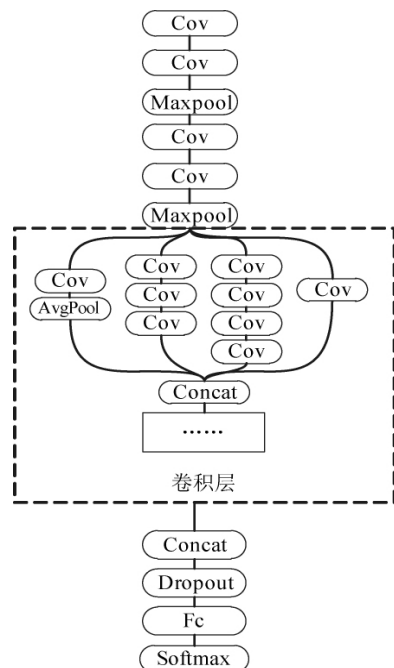


图3 Inception-V3 网络结构图

本文研究的分类数目远小于 ImageNet 上的分类数目,为了更好地得到分类结果,改进 Softmax 分类器为 SVM 分类器。

### 2.2.2 Bottleneck 特征

输入图像经过训练后,提取出了学习特征,如图 4 所示,预训练模型中,第一层卷积完成后得到的特征图像。最后一层全连接层完成后得到的学习特征称为“Bottleneck 特征”,如图 5 所示,但肉眼难以区别这些特征是否能够进行有效分类。

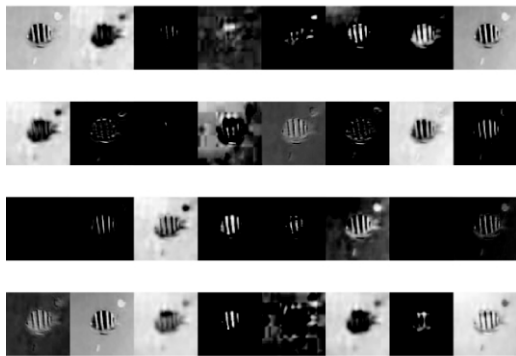


图4 第一层卷积完成后得到的特征图像

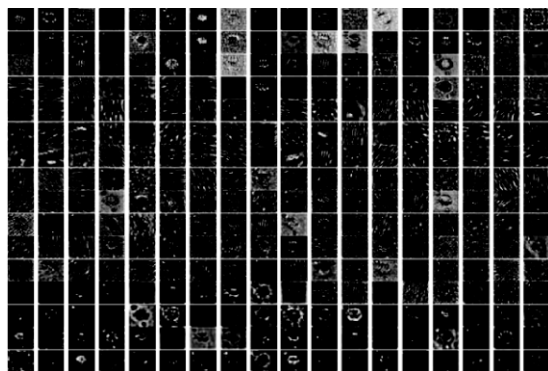


图5 最后一层全连接层完成后得到的特征图像

为了证明利用迁移学习构建的预训练模型的有效性。本文对提取的部分种类的鱼图像的 Bottleneck 特征进行聚类分析,如图 6 所示。图中,不同的颜色(以灰度表示)代表了不同的鱼种。

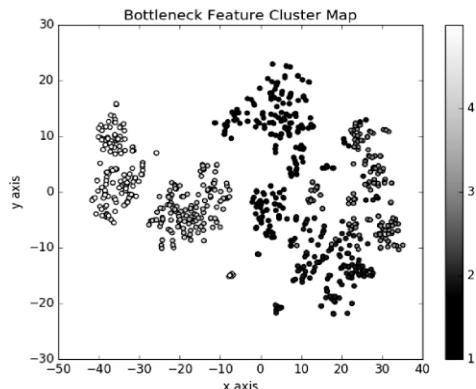


图6 Bottleneck 特征聚类图

可以看到,相同的色点大多聚集在一起。因此,可以使用 Bottleneck 特征有效地进行分类。

### 2.2.3 SVM 分类器

支持向量机(SVM)是一种依赖核函数的机器学习分类算法。在高维空间中最优分类函数为:

$$f(x) = \left[ \sum_{i=1}^n a_i y_i k(x_i, x) + b \right] \quad (6)$$

式中:  $a_i \geq 0$  是拉格朗日因子,  $b$  是阈值。

通过对训练样本交叉验证来调节核函数的相关参数,优化拟合问题,产生的支持向量有的在超平上,有的在超平面之间。本文选用径向向量(RBF)<sup>[15]</sup>形式核函数如下:

$$k(x_i, x) = \exp \left\{ - \frac{\|x_i - x\|^2}{\sigma^2} \right\} \quad (7)$$

式中:  $\sigma$  为可调参数,  $i = 1, 2, \dots, n$ 。

## 3 实验分析

本文实验以在 Fish4-Knowledge(F4K)<sup>[16]</sup>数据集上鱼类图像作为研究对象。该数据集是台湾电力公司、台湾海洋研究所和垦丁国家公园,分享他们 2010 年 10 月 1 日至 2013 年 9 月 30 日期间,在台湾南湾、兰屿和胡比湖的水下观景台收集的影像数据。最后实验结果与前人的多项研究成果进行对比。

### 3.1 数据预处理

本实验数据集是从 Fish4-Knowledge(F4K)<sup>[16]</sup>获取的,有 23 类鱼种数据集,共 27 370 张鱼的图像。图 7 给出了 23 种鱼类的图像。

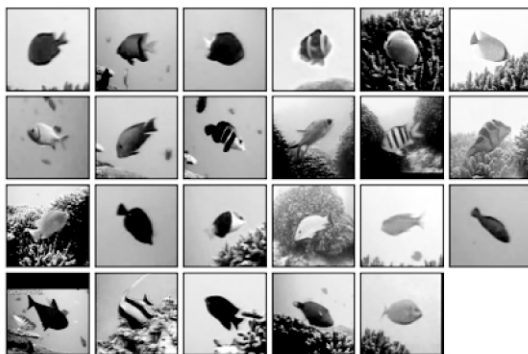


图7 23 种鱼类的图像

在数据预处理中,由于这些 RGB 鱼图像大小不一,范围约为  $20 \times 20$  至约  $200 \times 200$  像素。考虑到图像的平均大小,调整所有的图像为  $28 \times 28$  像素,使其减少计算资源的消耗。整个数据集被划分为两个子集:4/5 用于训练,1/5 用于测试。同时,由于不同鱼类物种的数量不平衡,针对每个种类按相同的比例分配。

对于训练集还采用了一系列随机的变换,人为地增加了数据集的大小:(1) 随机翻转图像由左到右。(2) 随机变换图像的亮度。(3) 随机变换图像的对比度。如图 8 所示,一张图像经过变换后得到的一系列图像。

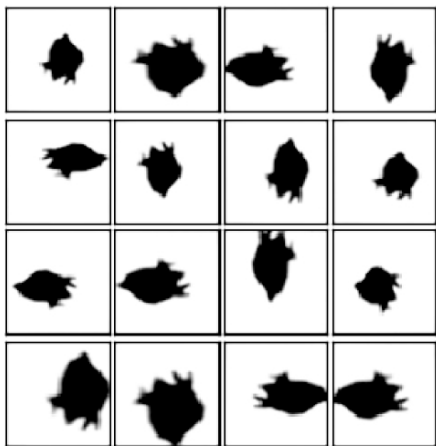


图 8 一张图像变换后得到的一系列图像

### 3.2 模型训练

实验在配备 Intel Core i7 处理器,内存 8 GB 的 OS X Yosemite 操作系统下进行,卷积神经网络的训练使用 Google 的深度学习框架 TensorFlow<sup>[17]</sup>。

在训练过程中,使用交叉熵作为优化的目标函数。利用 TensorBoard<sup>[17]</sup> 得到损失函数变化图,如图 9 所示。

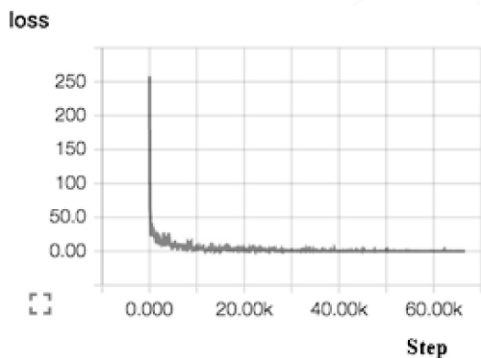


图 9 迭代步数与损失值变化曲线

图 9 表示,当学习速率为 0.001 时,得到平均损失值变化曲线。由于训练中使用的数据批量比较小,损失值中夹杂了相当对的噪声。相比原始值,损失值的移动平均值显得更有意义。

### 3.3 实验结果

对于 CNN 模型,经过 60 000 步迭代计算,训练数据集的准确率为 97.57%,测试数据集得到的准确率为 96.67%。平均训练时间为 5 小时 32 分。

对于 PreCNN + SVM 模型,经过训练后,训练数据集的准确率为 98.90%,测试数据集的准确率为

98.6%。平均训练时间为 2 小时 45 分。

与前人研究比较结果如表 1 所示。

表 1 各个算法比较结果

算法	准确率/%
LDA + SVM <sup>[18]</sup>	80.4
VLFeat Dense-SIFT <sup>[18]</sup>	93.58
CNN	96.67
PreCNN + SVM	98.60

对比前人算法,利用深度学习算法对鱼类图像进行分类,其准确率均有提高,并且方法更灵活,不需要人工提取特征值。对比 CNN 模型,PreCNN + SVM 模型利用预先训练好的参数权重,减少了训练时间,并且准确率也有一定提高。

通过实验数据表明,本文提出的基于深度学习的鱼类分类算法,能够有效解决鱼类图像分类问题,该方法对传统的分类准确率均有提高,对于小规模数据集有效解决了需要人工提取特征的不可迁移性。

## 4 结 语

本文研究了基于深度学习的鱼类分类算法,提出了基于卷积神经网络的分类模型和基于迁移学习的 PreCNN + SVM 分类模型。这些方法提取的特征都是从训练数据中提取的,不需要关于鱼类领域的知识。

通过在深海活鱼图像数据集上的实验证明,对比传统鱼类识别算法,利用深度学习算法识别鱼类准确度有一定的提高。

下一步将进一步优化基于深度学习的鱼类识别模型,以 GPU 代替 CPU 加快训练速度。同时采集更多的数据来训练本文的模型,使得能够识别更多的鱼类。最终期望能够将本文的方法和应用进行推广,形成一个实用的自动鱼类识别产品,推进水下活鱼识别研究,探索水下物体识别问题,并使海洋生物学家,生态学家以及鱼类养殖等商业应用受益。

## 参 考 文 献

- [1] Froese R, Pauly D. FishBase. World Wide Web electronic publication [OL]. (2017-02). www.fishbase.org.
- [2] Strachan N J C, Nesvadba P, Allen A R. Fish species recognition by shape analysis of images [J]. Pattern Recognition, 1990, 23(5): 539-544.
- [3] Larsen R, Olafsdottir H, Ersbøll B. Shape and texture based classification of fish species [J]. Image Analysis, 2009, 5575: 745-749.
- [4] Huang P X, Boom B J, Fisher R B. Underwater live fish rec-

- ognition using a balance-guaranteed optimized tree [C]//Asian Conference on Computer Vision. Springer Berlin Heidelberg, 2012: 422-433.
- [5] 吴一全,殷骏,戴一冕,等. 基于蜂群优化多核支持向量机的淡水鱼种类识别 [J]. 农业工程学报, 2014, 30(16): 312-319.
- [6] 杜伟东,李海森,魏玉阔,等. 基于 SVM 的决策融合鱼类识别方法 [J]. 哈尔滨工程大学学报, 2015(5): 623-627.
- [7] Lécun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition [J]. Proceedings of the IEEE, 1998, 86(11): 2278-2324.
- [8] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks [C]//International Conference on Neural Information Processing Systems. Curran Associates Inc. 2012: 1097-1105.
- [9] Pan S J, Yang Q. A Survey on Transfer Learning [J]. Knowledge & Data Engineering IEEE Transactions on, 2010, 22(10): 1345-1359.
- [10] Oquab M, Bottou L, Laptev I, et al. Learning and Transferring Mid-level Image Representations Using Convolutional Neural Networks [C]//IEEE Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, 2014: 1717-1724.
- [11] Kingma D P, Ba J. Adam: A Method for Stochastic Optimization [C]//The third International Conference for Learning Representations, San Diego, 2015.
- [12] Hahnloser R H, Sarpeshkar R, Mahowald M A, et al. Digital selection and analogue amplification coexist in a cortex-inspired silicon circuit [J]. Nature, 2000, 405(6789): 947-951.
- [13] Srivastava N, Hinton G, Krizhevsky A, et al. Dropout: a simple way to prevent neural networks from overfitting [J]. Journal of Machine Learning Research, 2014, 15(1): 1929-1958.
- [14] Szegedy C, Vanhoucke V, Ioffe S, et al. Rethinking the inception architecture for computer vision [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016: 2818-2826.
- [15] 张学工. 关于统计学习理论与支持向量机 [J]. 自动化学报, 2000, 26(1): 32-42.
- [16] Boom B J, Huang P X, He J, et al. Supporting ground-truth annotation of image datasets using clustering [C]//International Conference on Pattern Recognition. IEEE, 2012: 1542-1545.
- [17] Abadi M, Agarwal A, Barham P, et al. TensorFlow: Large-scale machine learning on heterogeneous systems [EB/OL]. (2015-12-12). <http://download.tensorflow.org/paper/whitepaper2015.pdf>.
- [18] Qin H, Li X, Liang J, et al. DeepFish: Accurate underwater live fish recognition with a deep architecture [J]. Neurocomputing, 2015, 187: 49-58.
- (上接第 136 页)
- [3] Fawcett T. "In vivo" spam filtering: a challenge problem for KDD [J]. Acm Sigkdd Explorations Newsletter, 2003, 5(2): 140-148.
- [4] 李勇,刘战东,张海军. 不平衡数据的集成分类算法综述 [J]. 计算机应用研究, 2014, 31(5): 1287-1291.
- [5] Zhang X, Song Q, Wang G, et al. A dissimilarity-based imbalance data classification algorithm [J]. Applied Intelligence, 2015, 42(3): 544-565.
- [6] Elkan C. The Foundations of Cost-Sensitive Learning [C]//Seventeenth International Joint Conference on Artificial Intelligence. 2001: 973-978.
- [7] Chawla N V, Bowyer K W, Hall L O, et al. SMOTE: synthetic minority over-sampling technique [J]. Journal of Artificial Intelligence Research, 2002, 16(1): 321-357.
- [8] 张泉山,罗强. 一种基于聚类融合欠抽样的不平衡数据分类方法 [J]. 计算机科学, 2015, 42(S2): 63-66.
- [9] Sun Z, Song Q, Zhu X, et al. A novel ensemble method for classifying imbalanced data [J]. Pattern Recognition, 2015, 48(5): 1623-1637.
- [10] Weiss G M. Mining with rarity [J]. Acm Sigkdd Explorations Newsletter, 2004, 6(1): .
- [11] Weiss G M. Mining with rarity: a unifying framework [J]. Acm Sigkdd Explorations Newsletter, 2004, 6(1): 7-19.
- [12] Weiss G M, Hirsh H. A Quantitative Study of Small Disjuncts [C]//Seventeenth National Conference on Artificial Intelligence and Twelfth Conference on Innovative Applications of Artificial Intelligence. AAAI Press, 1970: 665-670.
- [13] Provost F, Fawcett T. Robust Classification for Imprecise Environments [J]. Machine Learning, 2001, 42(3): 203-231.
- [14] Ng W W Y, Zeng G, Zhang J, et al. Dual autoencoders features for imbalance classification problem [J]. Pattern Recognition, 2016, 60: 875-889.
- [15] 刘学,张素伟. 基于二次随机森林的不平衡数据分类算法 [J]. 软件, 2016, 37(7): 75-79.
- [16] 胡晓星,李辉. 胸外科肿瘤患者术前医疗风险评估表在病案中的应用 [J]. 中国病案, 2014(11): 15-17.
- [17] 王丹丹,陈情,毕平. 肺癌左全肺切除术后心肺并发症的发生与术前低肺功能的相关性 [J]. 中国肿瘤临床, 2015, 42(7): 397-400.
- [18] Zieba M, Tomczak J M, Lubicz M, et al. Boosted SVM for extracting rules from imbalanced data in application to prediction of the post-operative life expectancy in the lung cancer patients [J]. Applied Soft Computing, 2014, 14(1): 99-108.